Differential Equations: Dynamical Systems

Joe Erickson

Table of Contents

1

Initial-Value Problems

1.1 – Normed Vector Spaces

A suitable setting for our study of ordinary differential equations and dynamical systems is the normed vector space. First we make clear what precisely is meant by a norm.

Definition 1.1. *A* ***norm*** *on a vector space X over the field $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ is a mapping $\|\cdot\| : X \to [0, \infty)$ having the following properties.*

N1. $\|x\| \geq 0$ *for* $x \in X$.

N2. $\|x\| = 0$ *if and only if* $x = 0$.

N3. $\|\alpha x\| = |\alpha|\|x\|$ *for* $\alpha \in \mathbb{F}$ *and* $x \in X$.

N4. $\|x + y\| \leq \|x\| + \|y\|$ *for* $x, y \in X$.

Property (N4) is the famous **triangle inequality**, and with it one can derive the **inverse triangle inequality** which states that

$$\big|\|x\| - \|y\|\big| \leq \|x - y\|$$

for $x, y \in X$.

A vector space X equipped with a norm $\|\cdot\|$ is called a **normed vector space**, and may be denoted by the symbol $(X, \|\cdot\|)$. The norm naturally induces a **metric** that gives the **distance** between vectors $x, y \in X$ to be $\|x - y\|$. Such a metric, in turn, induces a **topology** on X which defines a set $S \subseteq X$ to be **open** if, for each $s \in S$, there exists some $r > 0$ such that the **open ball**

$$B_r(s) := \big\{x \in X : \|x - s\| < r\big\}$$

is a subset of S. If $(X, \|\cdot\|)$ is a finite-dimensional vector space, then any choice for the norm $\|\cdot\|$ will induce the same topology on X, meaning in particular that properties we define later such as convergence and continuity are preserved.

A set $S \subseteq X$ is **closed** if its complement $X \setminus S$ is open, and the symbol $\overline{B}_r(s)$ denotes a **closed ball** centered at s with radius r; that is,

$$\overline{B}_r(s) := \big\{x \in X : \|x - s\| \leq r\big\}.$$

Thus $\overline{B}_r(s) = B_r(s) \cup \partial B_r(s)$, with $\partial B_r(s)$ being the **boundary** of the ball.

Example 1.2. Suppose $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ are normed vector spaces over a field $\mathbb{F}$. The set $X \times Y$ becomes a vector space if we define vector addition by

$$(x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1, y_2)$$

for $(x_1, y_1), (x_2, y_2) \in X \times Y$, and scalar multiplication by

$$\alpha(x, y) = (\alpha x, \alpha y)$$

for $(x, y) \in X \times Y$ and $\alpha \in \mathbb{F}$. Such a vector space is usually denoted by $X \oplus Y$ and called the **direct sum** of X and Y. There are many choices of norm for $X \oplus Y$. Indeed, defining

$$\|(x, y)\|_p = \left(\|x\|_X^p + \|y\|_Y^p\right)^{1/p} \tag{1.1}$$

for any real constant $p \in [1, \infty)$ satisfies the properties of a norm on $X \oplus Y$, with $p = 1$ resulting in the **rectilinear norm** and $p = 2$ the **Euclidean norm**. For instance, when $p = 1$ the triangle inequality is readily found to hold, with

$$\begin{aligned}\|(x_1, y_1) + (x_2, y_2)\|_1 = \|(x_1 + y_1, x_2 + y_2)\|_1 &= \|x_1 + x_2\|_X + \|y_1 + y_2\|_Y \\ &\le \left(\|x_1\|_X + \|x_2\|_X\right) + \left(\|y_1\|_Y + \|y_2\|_Y\right) \\ &= \left(\|x_1\|_X + \|y_1\|_Y\right) + \left(\|x_2\|_X + \|y_2\|_Y\right) \\ &= \|(x_1, y_1)\|_1 + \|(x_2, y_2)\|_1\end{aligned}$$

for any $(x_1, y_1), (x_2, y_2) \in X \oplus Y$.

Provided that $X \oplus Y$ is a finite-dimensional vector space (which is the case whenever X and Y are both finite-dimensional), the possible norms formulated by (1.1) are all equivalent in the sense that they induce the same topology on $X \oplus Y$, and so properties such as convergence and continuity are unaffected by the choice of value for p. ■

Given a sequence $(x_n)_{n=1}^{\infty}$ in X, we say that x_n **converges** to $x \in X$ if

$$\lim_{n\to\infty} \|x_n - x\| = 0; \tag{1.2}$$

that is, if for each $\epsilon > 0$ there exists an integer n_0 such that $\|x_n - x\| < \epsilon$ for all $n \ge n_0$. We may also write (1.2) as $\lim_{n\to\infty} x_n = x$ or simply $x_n \to x$, and call x the **limit** of the sequence. If there exists no $x \in X$ for which (1.2) holds, then the sequence (x_n) is said to **diverge**. According to Definition 1.1 the norm in $(X, \|\cdot\|)$ is only given to have domain X, and so there is no occasion to ponder the possibility that (x_n) "converges" to something that is not an element of X.

A sequence (x_n) in $(X, \|\cdot\|)$ is called a **Cauchy sequence** if for each $\epsilon > 0$ there exists integer k such that $\|x_m - x_n\| < \epsilon$ for all $m, n \ge k$. We say a normed vector space is **complete** if every Cauchy sequence in the space converges (i.e. has a limit). A complete normed vector space is otherwise known as a **Banach space**.

Some properties of limits involving the vector addition, scalar multiplication, and norm operations on a Banach space are given in the following proposition.

Proposition 1.3. *Let $(X, \|\cdot\|)$ be normed vector spaces over $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$, let (x_n) and (y_n) be sequences in X, and let (α_n) be a sequence in $\mathbb{F}$. Suppose $x, y \in X$ and $\alpha \in \mathbb{F}$. If $x_n \to x$, $y_n \to y$, and $\alpha_n \to \alpha$, then $\|x_n\| \to \|x\|$, $x_n + y_n \to x + y$, and $\alpha_n x_n \to \alpha x$.*

Proof. We show only that $\alpha_n x_n \to \alpha x$, leaving the rest as an exercise. First, with the inverse triangle inequality we find

$$0 \leq \big||\alpha_n| - |\alpha|\big| \leq |\alpha_n - \alpha|,$$

and since $\lim_{n\to\infty} |\alpha_n - \alpha| = 0$ by hypothesis, the squeeze theorem implies that

$$\lim_{n\to\infty} \big||\alpha_n| - |\alpha|\big| = 0,$$

and hence $\lim_{n\to\infty} |\alpha_n| = |\alpha|$. Now, by the triangle inequality followed by use of property (N3),

$$\begin{aligned}
\|\alpha_n x_n - \alpha x\| &= \|\alpha_n x_n - \alpha_n x + \alpha_n x - \alpha x\| \\
&\leq \|\alpha_n x_n - \alpha_n x\| + \|\alpha_n x - \alpha x\| \\
&= |\alpha_n|\|x_n - x\| + |\alpha_n - \alpha|\|x\|,
\end{aligned}$$

and since

$$\begin{aligned}
\lim_{n\to\infty} \big(|\alpha_n|\|x_n - x\| + |\alpha_n - \alpha|\|x\|\big) &= \lim_{n\to\infty} |\alpha_n| \cdot \lim_{n\to\infty} \|x_n - x\| + \|x\| \lim_{n\to\infty} |\alpha_n - \alpha| \\
&= |\alpha| \cdot 0 + \|x\| \cdot 0 = 0
\end{aligned}$$

by the usual laws of limits in $\mathbb{F}$, another application of the squeeze theorem leads us to conclude that $\lim_{n\to\infty} \|\alpha_n x_n - \alpha x\| = 0$. ■

Definition 1.4. *Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be normed vector spaces. A mapping $F : X \to Y$ is* ***continuous*** *at $\hat{x} \in X$ if for each $\epsilon > 0$ there exists some $\delta > 0$ such that, for any $x \in X$, $\|x - \hat{x}\|_X < \delta$ implies $\|F(x) - F(\hat{x})\|_Y < \epsilon$.*

The following proposition offers an equivalent means of characterizing continuity that is often more convenient to work with.

Proposition 1.5. *Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be normed vector spaces. A mapping $F : X \to Y$ is continuous at $\hat{x} \in X$ if and only if $F(x_n) \to F(\hat{x})$ for every sequence (x_n) in X that converges to $\hat{x}$.*

Proof. Suppose F is continuous at $\hat{x} \in X$, and let (x_n) be a sequence in X that converges to $\hat{x}$. Fix $\epsilon > 0$. Then there exists $\delta > 0$ such that $\|x - \hat{x}\|_X < \delta$ implies $\|F(x) - F(\hat{x})\|_Y < \epsilon$. But $x_n \to \hat{x}$ implies there exists integer n_0 such that $\|x_n - \hat{x}\|_X < \delta$ for all $n \geq n_0$, and thus $\|F(x_n) - F(\hat{x})\|_Y < \epsilon$ for all $n \geq n_0$. Therefore $F(x_n) \to F(\hat{x})$.

For the converse we prove the contrapositive. Suppose that F is not continuous at $\hat{x}$. Then there exists some $\epsilon > 0$ such that, for each $\delta > 0$, there can be found some $x_\delta \in X$ for which $\|x_\delta - \hat{x}\|_X < \delta$ and yet $\|F(x_\delta) - F(\hat{x})\|_Y \geq \epsilon$. In particular, for each $n \geq 1$ there exists $x_n \in X$ such that $\|x_n - \hat{x}\|_X < 1/n$ and yet $\|F(x_n) - F(\hat{x})\|_Y \geq \epsilon$. In this manner we construct a sequence $(x_n)_{n=1}^\infty$ such that $x_n \to \hat{x}$ but $(F(x_n))_{n=1}^\infty$ fails to converge to $F(\hat{x})$. ■

Example 1.6. Strictly speaking, given a normed vector space $(X, \|\cdot\|)$, the vector addition operation $+$ has domain $X \oplus X$ and thus is a function of two independent variables. Without bringing to bear any additional machinery it can be shown that $+$ is a **jointly continuous** function on $X \oplus X$, meaning that for fixed x the mapping $X \to X$ given by $(x, y) \mapsto x + y$ for all $y \in X$ is continuous on X, and for fixed y the mapping $(x, y) \mapsto x + y$ for all $x \in X$ is also continuous on X. But is $+$ continuous on $X \oplus X$ in the sense of Definition 1.4? A careful study of the definition shows this question to be meaningless unless $X \oplus X$ is equipped with a norm. What norm should this be?

From Example 1.2 we know that, so long as $X \oplus X$ is a finite-dimensional vector space, the choice of norm will have no impact on continuity. Thus, if $+ : X \oplus X \to X$ is continuous for one choice of norm for $X \oplus X$, then it will be continuous for any choice of norm. Therefore we may as well equip $X \oplus X$ with the rectilinear norm given by (1.1) for $p = 1$.

Fix $(x, y) \in (X \oplus X, \|\cdot\|_1)$, and let (x_n, y_n) be a sequence in $X \oplus X$ that converges to (x, y). Then

$$\lim_{n\to\infty} \big(\|x_n - x\| + \|y_n - y\|\big) = \lim_{n\to\infty} \|(x_n - x, y_n - y)\|_1 = \lim_{n\to\infty} \|(x_n, y_n) - (x, y)\|_1 = 0,$$

and since

$$0 \le \|(x_n + y_n) - (x + y)\| = \|(x_n - x) + (y_n - y)\| \le \|x_n - x\| + \|y_n - y\| \to 0,$$

the squeeze theorem implies that

$$\lim_{n\to\infty} \|(x_n + y_n) - (x + y)\| = 0,$$

or equivalently $x_n + y_n \to x + y$. Therefore $+$ is continuous at (x, y) by Proposition 1.5. ■

Normed vector spaces consisting of collections of continuous functions will be of especial importance in the sequel, enough so to warrant some special notations. Given sets U and V, let $\mathcal{C}(U, V)$ denote the collection of continuous functions $U \to V$, and let $C_b(U, V)$ denote the collection of bounded continuous functions $U \to V$. Quite often $V = \mathbb{R}$, and so we also define $\mathcal{C}(U) = \mathcal{C}(U, \mathbb{R})$ and $\mathcal{C}_b(U) = \mathcal{C}_b(U, \mathbb{R})$.

Sequences of functions will also figure prominently in upcoming theoretical developments, with one mode of convergence for such sequences, called uniform convergence, having particular significance.

Definition 1.7. *Let $(X, \|\cdot\|)$ be a normed vector space and $S \subseteq \mathbb{R}$, and suppose $f_n : S \to X$ for all $n \in \mathbb{N}$. We say $(f_n)_{n\in\mathbb{N}}$* ***converges uniformly to f on S*** *if for every $\epsilon > 0$ there exists some $n_0 \in \mathbb{N}$ such that*

$$\|f_n(t) - f(t)\| < \epsilon$$

for all $n > n_0$ and $t \in S$.

We use the symbol $f_n \xrightarrow{\text{u}} f$ to denote that the sequence (f_n) converges uniformly to f on some set. Definition 1.7 could be generalized to let S be a subset of an arbitrary normed vector space, as opposed to a subset of $\mathbb{R}$, but we shall have no use for such a generalization at this juncture. One reason uniform convergence is desirable is that it preserves the property of continuity. The following proposition reveals precisely what this means.

Proposition 1.8. *Let $(X, \|\cdot\|)$ be a normed vector space and $S \subseteq \mathbb{R}$, and suppose $f_n : S \to X$ is continuous on S for each $n \in \mathbb{N}$. If $(f_n)_{n\in\mathbb{N}}$ converges uniformly to f on S, then f is continuous on S.*

Proof. Suppose (f_n) converges uniformly to f on S. Fix $s \in S$, and let $\epsilon > 0$. There is an integer n_0 such that $\|f_n(t) - f(t)\| < \epsilon/3$ for all $n > n_0$ and $t \in S$, and so $\|f_n(s) - f(s)\| < \epsilon/3$ for any $n > n_0$ as well. Fixing $m > n_0$, the continuity of f_m at s implies there exists some $\delta > 0$ such that $\|f_m(t) - f_m(s)\| < \epsilon/3$ for all $t \in S$ for which $|t - s| < \delta$. Now, for any $t \in (s - \delta, s + \delta) \cap S$, we have

$$\begin{aligned}\|f(t) - f(s)\| &= \big\|[f(t) - f_n(t)] + [f_n(t) - f_n(s)] + [f_n(s) - f(s)]\big\| \\ &\leq \|f(t) - f_m(t)\| + \|f_m(t) - f_m(s)\| + \|f_m(s) - f(s)\| \\ &< \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon.\end{aligned}$$

This shows that f is continuous at s, and since $s \in S$ is arbitrary it follows that f is continuous on S. ■

Example 1.9. Let $I \subseteq \mathbb{R}$ be a compact (i.e. closed and bounded) interval, and define a norm on $\mathcal{C}(I)$ by

$$\|f\|_\infty = \sup_{t\in I} |f(t)|, \tag{1.3}$$

called the **sup norm** or **uniform norm**. That $\|\cdot\|_\infty$ satisfies the four properties of a norm in Definition 1.1 we leave to the reader to verify, but we note here that $\|f\|_\infty \in [0, \infty)$ for any f in the normed vector space $(\mathcal{C}(I), \|\cdot\|_\infty)$ is assured by the compactness of I and the extreme value theorem.

It is straightforward to check that a sequence (f_n) converges to f in the space $(\mathcal{C}(I), \|\cdot\|_\infty)$ if and only if (f_n) converges uniformly to f on I. That is, convergence with respect to the uniform norm is synonymous with uniform convergence.

If I is not compact then $\|\cdot\|_\infty$ will fail to be a norm on $\mathcal{C}(I)$. For instance if $I = (0, 1)$, which is not closed and thus not compact, we find $f(t) = 1/t$ to be an element of $\mathcal{C}(I)$ such that $\|f\|_\infty = \infty$, and therefore $\|\cdot\|_\infty$ is not a norm. A similar difficulty arises if we let $I = [0, \infty)$ and $f(t) = t$. However, for I not compact, we may restrict the domain of $\|\cdot\|_\infty$ to $\mathcal{C}_b(I) \subseteq \mathcal{C}(I)$ in order to construct the normed vector space $(\mathcal{C}_b(I), \|\cdot\|_\infty)$. ■

A fundamental fact from the subject of real analysis is that the normed vector space $(\mathbb{R}, |\cdot|)$ is complete, where the norm $|\cdot|$ is the usual absolute value operation. This fact is used in the proof of the following.

Theorem 1.10. *Let $I \subseteq \mathbb{R}$ be an interval.*

1. *If I is compact, then $(\mathcal{C}(I), \|\cdot\|_\infty)$ is a Banach space.*
2. *If I is arbitrary, then $(\mathcal{C}_b(I), \|\cdot\|_\infty)$ is a Banach space.*

Proof.
Proof of (1). Suppose $I \subseteq \mathbb{R}$ is compact, so that $(\mathcal{C}(I), \|\cdot\|_\infty)$ is properly a normed vector space. We only need to show that $(\mathcal{C}(I), \|\cdot\|_\infty)$ is complete. Let (f_n) be a Cauchy sequence in $\mathcal{C}(I)$, so

for any $\epsilon > 0$ there exists an integer k such that $\|f_m - f_n\|_\infty < \epsilon$ for all $m, n > k$. In light of (1.3) this implies that $|f_m(t) - f_n(t)| < \epsilon$ for all $m, n > k$ and $t \in I$. Thus $(f_n(t))$ is a Cauchy sequence in $(\mathbb{R}, |\cdot|)$ for all $t \in I$, and since $(\mathbb{R}, |\cdot|)$ is a Banach space we conclude that $(f_n(t))$ converges to some point in $\mathbb{R}$ denoted by $f(t)$ for each $t \in I$. That is, (f_n) converges pointwise to $f : I \to \mathbb{R}$.

Next to show is that $f_n \to f$ with respect to the uniform norm $\|\cdot\|_\infty$ on $C(I)$, which will be accomplished by demonstrating that $f_n \xrightarrow{\text{u}} f$ on I. Fix $\epsilon > 0$. Then there exists integer k such that $|f_n(t) - f_m(t)| < \epsilon/2$ for all $m, n > k$ and $t \in I$. From this comes

$$|f_m(t) - f_n(t)| < \frac{\epsilon}{2},$$

whence

$$f_n(t) - \frac{\epsilon}{2} < f_m(t) < f_n(t) + \frac{\epsilon}{2},$$

for all $m, n > k$ and $t \in I$. This implies that

$$f_n(t) - \frac{\epsilon}{2} \leq \lim_{m\to\infty} f_m(t) \leq f_n(t) + \frac{\epsilon}{2},$$

hence

$$f_n(t) - \frac{\epsilon}{2} \leq f(t) \leq f_n(t) + \frac{\epsilon}{2},$$

and finally

$$|f_n(t) - f(t)| \leq \frac{\epsilon}{2} < \epsilon \tag{1.4}$$

for $n > k$ and $t \in I$. Therefore $f_n \xrightarrow{\text{u}} f$ on I, and since $f \in \mathcal{C}(I)$ by Proposition 1.8, we conclude that $(\mathcal{C}(I), \|\cdot\|_\infty)$ is complete.

Proof of (2). Suppose $I \subseteq \mathbb{R}$ is any interval, and let (f_n) be a Cauchy sequence in $\mathcal{C}_b(I)$. Since $(\mathcal{C}_b(I), \|\cdot\|_\infty)$ is a normed vector space the proof that (f_n) converges uniformly to some $f \in C(I)$ is the same as before, and so it only remains to show that $f : I \to \mathbb{R}$ is bounded to place f in $\mathcal{C}_b(I)$.

Let $\epsilon > 0$. Then (1.4) holds for all $n > k$ and $t \in I$, and so for any particular $m > k$ we have $|f_m(t) - f(t)| < \epsilon$ for all $t \in I$. Suppose f is not bounded on I, and let $M > 0$. Then there exists some $\tau \in I$ for which $f(\tau) > M + \epsilon$, and thus we obtain $f_m(\tau) > M$. Since $M > 0$ is arbitrary it follows that f_m is not bounded on I, contradicting the hypothesis that (f_n) is a sequence in $\mathcal{C}_b(I)$. Therefore f is bounded on I. ■

Unless otherwise indicated, we take the norm on $\mathbb{F}^n$ for $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and $n \in \mathbb{N}$ to be the Euclidean norm $|\cdot|$ given by

$$|x| = \left(\sum_{k=1}^{n} |x_k|^2\right)^{1/2} \tag{1.5}$$

for any $x := (x_1, \ldots, x_n) \in \mathbb{F}^n$. This convention is relevant to the following proposition, which generalizes Theorem 1.10.

Proposition 1.11. *Let $I \subseteq \mathbb{R}$ be an interval.*

1. *If I is compact, then $(\mathcal{C}(I, \mathbb{R}^n), \|\cdot\|_\infty)$ is a Banach space.*
2. *If I is arbitrary, then $(\mathcal{C}_b(I, \mathbb{R}^n), \|\cdot\|_\infty)$ is a Banach space.*

Proof.
Proof of (1). Suppose $I \subseteq \mathbb{R}$ is compact, and let $(f_k)_{k\in\mathbb{N}}$ be Cauchy in $(\mathcal{C}(I,\mathbb{R}^n), \|\cdot\|_\infty)$. Thus for each $k \in \mathbb{N}$ we have

$$f_k(t) = \big(f_{k1}(t), \dots, f_{kn}(t)\big) = \big(f_{kj}(t)\big)_{j=1}^n \tag{1.6}$$

for $t \in I$. The continuity of f_k on I implies the continuity of each component function $f_{kj} : I \to \mathbb{R}$ on I. Fix $\epsilon > 0$. Then there exists $\ell \in \mathbb{N}$ such that $\|f_m - f_k\|_\infty < \epsilon$ for all $m, k > \ell$, and in accordance with the Euclidean norm defined by (1.5) it follows that

$$\sqrt{\sum_{j=1}^n |f_{mj}(t) - f_{kj}(t)|^2} < \epsilon \qquad \forall m, k > \ell \ \ \forall t \in I,$$

and hence, for each $1 \le j \le n$,

$$\|f_{mj} - f_{kj}\|_\infty < \epsilon \qquad \forall m, k > \ell.$$

For each j, then, we find $(f_{kj})_{k\in\mathbb{N}}$ to be a Cauchy sequence in $(\mathcal{C}(I), \|\cdot\|_\infty)$, and by Theorem 1.10 there exists $\varphi_j \in \mathcal{C}(I)$ such that $f_{kj} \xrightarrow{\text{u}} \varphi_j$. Hence $f_k = (f_{kj})_{j=1}^n \xrightarrow{\text{u}} (\varphi_j)_{j=1}^n := \varphi$ as $k \to \infty$. The continuity of each $\varphi_j : I \to \mathbb{R}$ implies the continuity of $\varphi : I \to \mathbb{R}^n$, and so every Cauchy sequence in $\mathcal{C}(I,\mathbb{R}^n)$ converges to some element in $\mathcal{C}(I,\mathbb{R}^n)$. This makes $\mathcal{C}(I,\mathbb{R}^n)$ complete, and therefore a Banach space.

Proof of (2). Let $I \subseteq \mathbb{R}$ be arbitrary, and let $(f_k)_{k\in\mathbb{N}}$ be Cauchy in $(\mathcal{C}_b(I,\mathbb{R}^n), \|\cdot\|_\infty)$ for f_k given by (1.6). The proof of part (1) shows (f_k) converges uniformly to some $\varphi \in (\mathcal{C}(I,\mathbb{R}^n), \|\cdot\|_\infty)$, and so it only remains to show that φ is bounded. Also in the proof of part (1) it was shown that $(f_{kj})_{k\in\mathbb{N}}$ is Cauchy in $\mathcal{C}(I)$ for each $1 \le j \le n$. Now, the boundedness of each f_k for $k \in \mathbb{N}$ implies the boundedness of each f_{kj} for $k \in \mathbb{N}$ and $1 \le j \le n$, and hence for each j the sequence $(f_{kj})_{k\in\mathbb{N}}$ is Cauchy in $\mathcal{C}_b(I)$. By part (2) of Theorem 1.10 it follows that each $(f_{kj})_{k\in\mathbb{N}}$ converges uniformly to some $\varphi_j \in \mathcal{C}_b(I)$, and this in turn implies that $f_k \xrightarrow{\text{u}} (\varphi_j)_{j=1}^n = \varphi$. Since the boundedness of each φ_j implies the boundedness of φ, we conclude that $\varphi \in \mathcal{C}_b(I,\mathbb{R}^n)$ and therefore $(\mathcal{C}_b(I,\mathbb{R}^n), \|\cdot\|_\infty)$ is complete. ∎

We end this section with a couple results concerning a sequence that lies in a closed subset of a normed vector space, whether complete or not.

Proposition 1.12. *Let $(X, \|\cdot\|)$ be a normed vector space and $S \subseteq X$ closed. If (x_n) is a sequence in S that converges to x, then $x \in S$.*

Proof. Suppose (x_n) is a sequence in S, so that $x_n \in S$ for all n. Suppose $x \notin S$. Then $x \in X \setminus S$, and since $X \setminus S$ is open there exists some $r > 0$ such that $B_r(x) \subseteq X \setminus S$. It follows that $x_n \notin B_r(x)$ for all n, which is to say $\|x_n - x\| \ge r$ for all n, and thus the sequence (x_n) cannot converge to x. Therefore if (x_n) converges to x, it must be that $x \in S$. ∎

Proposition 1.13. *Let $(X, \|\cdot\|)$ be a Banach space and $S \subseteq X$. Then S is closed if and only if every Cauchy sequence in S converges to a vector in S.*

Proof. Suppose S is closed, and let (x_n) be a Cauchy sequence in S. Then (x_n) converges to some $x \in X$ since $(X, \|\cdot\|)$ is complete, and so $x \in S$ by Proposition 1.12.

We prove the contrapositive of the converse. Suppose S is not closed. Then $X \setminus S$ is not open, and so there exists some $x \in X \setminus S$ such that any open ball centered at x is not a subset of $X \setminus S$. Thus for each $n \in \mathbb{N}$ there exists $x_n \in B_{1/n}(x) \cap S$. Clearly $x_n \to x$, so that $(x_n)_{n\in\mathbb{N}}$ is a convergent—and hence Cauchy—sequence in X; and since $x_n \in S$ for all n, we see that (x_n) is a Cauchy sequence in S that does not converge to a vector in S. ■

The necessity of having $(X, \|\cdot\|)$ be complete in Proposition 1.13 is readily seen by considering the sequence $(1/n)_{n\in\mathbb{N}}$ in the normed vector space $((0,2), |\cdot|)$. In the topology of this space (induced by the Euclidean norm) the interval $(0,1]$ is a closed set, and $(1/n)_{n\in\mathbb{N}}$ is a Cauchy sequence in this closed set that fails to converge to an element of the set.

While any subset of a metric space is a metric space in its own right, it is not generally true that a subset of a vector space (normed or otherwise) is itself a vector space. Thus, the conclusion of Proposition 1.13 notwithstanding, care must be taken not to speak of an arbitrary closed subset of a Banach space as being itself a Banach space. To put it plainly, the set S in Proposition 1.13 may fail to be closed under either the operation of vector addition or scalar multiplication.

1.2 – Calculus on Euclidean Spaces

We now give some calculus results without proof, as they are typically broached in advanced calculus or elementary analysis texts; but nonetheless it will be convenient to have many of the tools most useful for the study of ordinary differential equations collected in one place. The setting will be specialized to Euclidean normed vector spaces $(\mathbb{R}^n, |\cdot|)$, as this will suffice to carry us far into the sequel. Letting $\mathcal{E}_n$ denote the **standard basis** for $\mathbb{R}^n$ consisting of the usual set of n orthonormal vectors $e_i := (\delta_{ij})_{j=1}^n$,[1] the **standard matrix** for a linear map $L : \mathbb{R}^n \to \mathbb{R}^m$ we define to be the $m \times n$ matrix $[L]$ corresponding to L with respect to bases $\mathcal{E}_n$ and $\mathcal{E}_m$. Hence $L(x) = [L]x$ for all $x \in \mathbb{R}^n$, with $x \in \mathbb{R}^{n\times 1}$ specifically (i.e. the components of vector x are presented as an $n \times 1$ matrix) in order for $[L]x$ to be defined.[2]

Given an open set $U \subseteq \mathbb{R}^n$, a function $F : U \to \mathbb{R}^m$ is **differentiable** at $a \in U$ if there exists a linear map $L : \mathbb{R}^n \to \mathbb{R}^m$ such that

$$\lim_{h\to 0} \frac{|F(a+h) - F(a) - L(h)|}{|h|} = 0.$$

The map L is called the **total derivative of *F* at *a***,[3] which herein shall be denoted by $\mathrm{d}F(a)$. Letting $F_1, \ldots, F_m$ denote the real-valued components of F, it is a fact that if

$$F(x_1, \ldots, x_n) = \big(F_1(x_1, \ldots, x_n), \ldots, F_m(x_1, \ldots, x_n)\big)$$

is differentiable at a, then all the first-order partial derivatives of F exist at a, and the standard matrix $[\mathrm{d}F(a)]$ for $\mathrm{d}F(a) : \mathbb{R}^n \to \mathbb{R}^m$ is the $m \times n$ matrix with ij-entry $[\mathrm{d}F(a)]_{ij}$ given by

$$[\mathrm{d}F(a)]_{ij} = \frac{\partial F_i}{\partial x_j}(a)$$

for $1 \le i \le m$ and $1 \le j \le n$, so that

$$[\mathrm{d}F(a)] = \begin{bmatrix} \dfrac{\partial F_1}{\partial x_1}(a) & \cdots & \dfrac{\partial F_1}{\partial x_n}(a) \\ \vdots & \ddots & \vdots \\ \dfrac{\partial F_m}{\partial x_1}(a) & \cdots & \dfrac{\partial F_m}{\partial x_n}(a) \end{bmatrix}. \tag{1.7}$$

This is the **Jacobian** (or **derivative**) **matrix of *F* at *a***. If $\mathrm{d}F(a)$ happens to be a square matrix, then its determinant $\det(\mathrm{d}F(a)) := \det([\mathrm{d}F(a)])$ is the **Jacobian determinant of *F* at *a***. Omitting the argument a, the operator $\mathrm{d}F = [\partial F_i/\partial x_j]_{m,n}$ is the **Jacobian matrix of *F***.

Arguably the most straightforward means of determining that a function is differentiable on some open set is to check that its first-order partial derivatives exist and are continuous on the set.

Proposition 1.14. *If $U \subseteq \mathbb{R}^n$ is open, then $F : U \to \mathbb{R}^m$ is differentiable at each point in U if all the first-order partial derivatives of F exist and are continuous on U.*

---

[1] Here δ_{ij} is the **Kronecker delta** function: $\delta_{ij} = 0$ if $i \ne j$, and $\delta_{ii} = 1$.

[2] It will not be our habit to distinguish between $\mathbb{R}^n$ and $\mathbb{R}^{n\times 1}$ save on occasions when it serves to make definitions precise.

[3] Other common terms are total differential, derivative, or differential of F at a.

Having waded in this far, we would be remiss if we did not at least state the chain rule for total derivatives.

Theorem 1.15 (**Total Derivative Chain Rule**). *Let $U \subseteq \mathbb{R}^n$ and $V \subseteq \mathbb{R}^m$ be open. If $F : U \to \mathbb{R}^m$ is differentiable at $a \in U$, $F(U) \subseteq V$, and $G : V \to \mathbb{R}^\ell$ is differentiable at $F(a)$, then $G \circ F$ is differentiable at a with*

$$\mathrm{d}(G \circ F)(a) = \mathrm{d}G(F(a)) \circ \mathrm{d}F(a).$$

Remark. From elementary linear algebra we know the standard matrix $[\mathrm{d}G(F(a)) \circ \mathrm{d}F(a)]$ for $\mathrm{d}G(F(a)) \circ \mathrm{d}F(a)$ to be the matrix given by the product $[\mathrm{d}G(F(a))][\mathrm{d}F(a)]$. Thus the total derivative chain rule as presented in Theorem 1.15 may be written as

$$[\mathrm{d}(G \circ F)(a)] = [\mathrm{d}G(F(a))][\mathrm{d}F(a)],$$

so that for any $x \in \mathbb{R}^{n\times 1}$ we find $\mathrm{d}(G \circ F)(a)(x)$ to be the vector in $\mathbb{R}^{\ell\times 1}$ resulting from the matrix product $[\mathrm{d}G(F(a))][\mathrm{d}F(a)]x$. ■

A function $\Phi : U \subseteq \mathbb{R}^{n+1} \to \mathbb{R}^m$ may be regarded as depending on (t, x), with $t \in \mathbb{R}$ and $x \in \mathbb{R}^n$, and it may occasion that we are interested in the variation in the value of $\Phi(t, x)$ as t is held constant and x varies. Of course we speak here of what is essentially the partial derivative of Φ with respect to x, but for fixed t the function $\Phi(t, \cdot)$ maps from a subset of $\mathbb{R}^n$ into $\mathbb{R}^m$ just as F does in Theorem 1.15, and so our partial derivative must be formulated in a manner that accords with (1.7). So there is no misunderstanding, for $a \in U$ we employ the special notation $\partial_x\Phi(a)$ to denote the **partial derivative with respect to $\boldsymbol{x}$ of $\boldsymbol{\Phi}$ at $\boldsymbol{a}$**, which for

$$\Phi(t, x_1, \dots, x_n) = \big(\Phi_1(t, x_1, \dots, x_n), \dots, \Phi_m(t, x_1, \dots, x_n)\big)$$

is the linear transformation $\partial_x\Phi(a) : \mathbb{R}^n \to \mathbb{R}^m$ whose standard matrix $[\partial_x\Phi(a)]$ has ij-entry given by

$$[\partial_x\Phi(a)]_{ij} = \frac{\partial\Phi_i}{\partial x_j}(a)$$

for $1 \le i \le m$ and $1 \le j \le n$.

1.3 – Fixed Points and Contractions

A **fixed point** of a mapping $f : X \to Y$ is a point $p \in X$ such that $f(p) = p$. Clearly for such a function to have any chance of possessing a fixed point the sets X and Y must not be disjoint. We shall be especially interested in the fixed points of mappings known as contractions.

Definition 1.16. *Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be normed vector spaces. A function $K : S \subseteq X \to Y$ is a* ***contraction mapping on S*** *if there exists a constant $\theta \in [0, 1)$, called the* ***contraction constant****, such that*

$$\|K(x_1) - K(x_2)\|_Y \leq \theta \|x_1 - x_2\|_X$$

for all $x_1, x_2 \in S$.

A contraction mapping on a subset S of $(X, \|\cdot\|)$ may also be called a **contraction on *S***, or simply a **contraction** if $S = X$. We will be dealing almost exclusively with contractions for which the set Y in Definition 1.16 is equal to X.

Generalizing the notion of a contraction mapping, a function $F : S \subseteq (X, \|\cdot\|_X) \to (Y, \|\cdot\|_Y)$ is **Lipschitz continuous on *S*** if there exists a constant $\theta \in [0, \infty)$ such that

$$\|F(x_1) - F(x_2)\|_Y \leq \theta \|x_2 - x_2\|_X \tag{1.8}$$

for all $x_1, x_2 \in S$. The inequality (1.8) is a **Lipschitz condition** for F on S, and θ is the **Lipschitz constant**. Clearly a contraction mapping on S is also Lipschitz continuous on S.

We say $F : S \subseteq (X, \|\cdot\|_X) \to (Y, \|\cdot\|_Y)$ is **locally Lipschitz continuous on *S*** if for each $x \in S$ there exists some open set O containing x such that F is Lipschitz continuous on $O \cap S$. The following proposition we give without proof.

Proposition 1.17. *If $F : S \subseteq (\mathbb{R}^n, |\cdot|) \to (\mathbb{R}^m, |\cdot|)$ is locally Lipschitz continuous on S and C is a compact set such that $C \subseteq S$, then F is Lipschitz continuous on C.*

The next proposition shows that Lipschitz continuity is a stronger condition on a function than the conventional continuity of Definition 1.4.

Proposition 1.18. *If $F : S \subseteq (X, \|\cdot\|_X) \to (Y, \|\cdot\|_Y)$ is Lipschitz continuous on S, then F is continuous on S.*

Proof. Suppose F is Lipschitz continuous on S, so there is some $\theta \geq 0$ such that (1.8) holds for all $x_1, x_2 \in S$. Fix $x_0 \in S$ and let $\epsilon > 0$. If $\theta = 0$, then (1.8) implies that $\|F(x) - F(x_0)\|_Y = 0 < \epsilon$ for all $x \in S$, and thus F is continuous at x_0. If $\theta > 0$, choose $\delta = \epsilon/\theta$, and suppose $x \in S$ is such that $\|x - x_0\|_X < \delta$. Then we have

$$\big\|F(x) - F(x_0)\big\|_Y \leq \theta \|x - x_0\|_X < \theta \cdot \frac{\epsilon}{\theta} = \epsilon,$$

and again F is continuous at x_0. Since $x_0 \in S$ is arbitrary, we conclude that F is continuous on S. ■

With the proposition above we immediately obtain the following result that will be needed presently to prove the contraction principle.

Corollary 1.19. *A contraction mapping is continuous on its domain.*

A connection between fixed points and contractions is made by the following theorem, called either the contraction principle or the Banach fixed-point theorem.

Theorem 1.20 (Contraction Principle). *Let $(X, \|\cdot\|)$ be a Banach space and $\varnothing \neq S \subseteq X$ closed, and suppose $K : S \to S$ is a contraction mapping with contraction constant θ. Then K has a unique fixed point $x^* \in S$, and moreover*

$$\|K^n(x) - x^*\| \leq \frac{\theta^n}{1-\theta}\|K(x) - x\| \tag{1.9}$$

for all $n \in \mathbb{N}$ and $x \in S$.

Proof. Suppose $x_1, x_2 \in S$ are fixed points for K. Then

$$\|x_1 - x_2\| = \|K(x_1) - K(x_2)\| \leq \theta\|x_1 - x_2\|,$$

and so if $\|x_1 - x_2\| \neq 0$ then the contradiction $\theta \geq 1$ results. Hence $\|x_1 - x_2\| = 0$, so that $x_1 = x_2$. This proves the uniqueness of any fixed point that K may have in S.

Now fix $x_0 \in S$, and define the sequence $x_{n+1} = K(x_n)$ for $n \geq 0$. We prove by induction that

$$\forall\, n \in \mathbb{N}\big(\|x_{n+1} - x_n\| \leq \theta^n\|x_1 - x_0\|\big). \tag{1.10}$$

That K is a contraction implies

$$\|x_2 - x_1\| = \|K(x_1) - K(x_0)\| \leq \theta\|x_1 - x_0\|,$$

thereby verifying the inequality in (1.10) when $n = 1$. Suppose the inequality holds for some arbitrary $n \in \mathbb{N}$. Then

$$\|x_{n+2} - x_{n+1}\| = \|K(x_{n+1}) - K(x_n)\| \leq \theta\|x_{n+1} - x_n\| \leq \theta \cdot \theta^n\|x_1 - x_0\| = \theta^{n+1}\|x_1 - x_0\|,$$

and (1.10) is proven. Since $\theta \in [0, 1)$, one important implication of (1.10) that comes via the squeeze theorem is that

$$\lim_{n\to\infty} \|x_{n+1} - x_n\| = 0. \tag{1.11}$$

Next, for $m > n$ we use the triangle inequality and (1.10) to obtain

$$\begin{aligned}
\|x_m - x_n\| &= \left\|\sum_{j=n+1}^{m}(x_j - x_{j-1})\right\| \leq \sum_{j=n+1}^{m}\|x_j - x_{j-1}\| \\
&\leq \sum_{j=n+1}^{m}\theta^{j-1}\|x_1 - x_0\| = \sum_{j=0}^{m-n-1}\theta^{j+n}\|x_1 - x_0\| \\
&= \theta^n\|x_1 - x_0\|\sum_{j=0}^{m-n-1}\theta^j = \theta^n\|x_1 - x_0\| \cdot \frac{1-\theta^{m-n}}{1-\theta} \\
&= \frac{\theta^n - \theta^m}{1-\theta}\|x_1 - x_0\| \leq \frac{\theta^n}{1-\theta}\|x_1 - x_0\|.
\end{aligned} \tag{1.12}$$

Since $\theta \in [0,1)$, for each $\epsilon > 0$ there exists some $\ell \in \mathbb{N}$ such that

$$\frac{\theta^\ell}{1-\theta}\|x_1 - x_0\| < \epsilon$$

and so $\|x_m - x_n\| < \epsilon$ for all $m, n \geq \ell$ by (1.12) and its counterpart in which the roles of m and n are reversed. This shows that $(x_n)_{n\in\mathbb{N}}$ is a Cauchy sequence in S, and since $(X, \|\cdot\|)$ is a Banach space and $S \subseteq X$ is closed, by Proposition 1.13 there exists $x^* \in S$ such that $x_n \to x^*$. Because K is continuous on S by Corollary 1.19, we also have $K(x_n) \to K(x^*)$.

Now, by Proposition 1.3 and (1.11),

$$\|K(x^*) - x^*\| = \left\| \lim_{n\to\infty} K(x_n) - \lim_{n\to\infty} x_n \right\| = \lim_{n\to\infty} \|x_{n+1} - x_n\| = 0,$$

which shows that $K(x^*) = x^*$ and thus x^* is a fixed point for K.

Finally, since $x_n = K^n(x_0)$ by induction, from (1.12) we have

$$\|K^n(x_0) - x^*\| = \lim_{m\to\infty} \|K^n(x_0) - x_m\| \leq \frac{\theta^n}{1-\theta}\|x_1 - x_0\| = \frac{\theta^n}{1-\theta}\|K(x_0) - x_0\|$$

for any $n \in \mathbb{N}$. The point $x_0 \in S$ being arbitrary, we have proven (1.9) for all $n \in \mathbb{N}$ and $x \in S$. ■

An occasionally useful approach to determining whether a real-valued function of a single real variable is a contraction mapping is the following.

Proposition 1.21. *Suppose f is continuous on $[a,b]$ and differentiable on (a,b). If there exists some $\theta \in [0,1)$ such that $|f'(x)| \leq \theta$ for all $x \in (a,b)$, then f is a contraction mapping on $([a,b], |\cdot|)$.*

Proof. Suppose there exists $\theta \in [0,1)$ such that $|f'(t)| \leq \theta$ for all $t \in (a,b)$. To show is that, for all $t_1, t_2 \in [a,b]$,

$$|f(t_2) - f(t_1)| \leq \theta|t_2 - t_1|. \tag{1.13}$$

Let $t_1, t_2 \in [a,b]$. We assume that $t_1 \neq t_2$ since (1.13) clearly holds whenever $t_1 = t_2$, with $t_1 < t_2$ for definiteness. By the mean value theorem there exists $\tau \in (t_1, t_2)$ such that

$$f'(\tau) = \frac{f(t_2) - f(t_1)}{t_2 - t_1},$$

whence we obtain

$$\left| \frac{f(t_2) - f(t_1)}{t_2 - t_1} \right| = |f'(\tau)| \leq \theta,$$

which in turn yields (1.13). ■

Example 1.22. The **Newton-Raphson method** is an algorithm for approximating a zero ξ of a real-valued function f of a single real variable x to an arbitrary degree of accuracy, with $\xi = \lim_{n\to\infty} x_n$ for sequence $(x_n)_{n\in\mathbb{N}}$ defined by the recurrence relation

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}, \quad x_0 \in I, \tag{1.14}$$

where x_0 is a point chosen "near" the zero ξ.

We seek conditions that are sufficient to ensure that the method will indeed converge to ξ. First we express the sequence (1.14) as $x_{n+1} = K(x_n)$ for

$$K(x) := x - \frac{f(x)}{f'(x)},$$

so that $(x_n)_{n\in\mathbb{N}}$ may be written as $(K^n(x_0))_{n\in\mathbb{N}}$. Provided that $f'(\xi) \neq 0$, we have $K(\xi) = \xi - f(\xi)/f'(\xi) = \xi$ since $f(\xi) = 0$, and so ξ is a fixed point for K. Provided that f'' exists on some closed interval I_1 containing ξ in its interior, we find K to be differentiable on I_1 with

$$K'(x) = \frac{f(x)f''(x)}{[f'(x)]^2}.$$

Let $I_2 \subseteq I_1$ be a compact interval on which f'' is continuous and $\xi \in \operatorname{Int}(I_2)$. Then f' and f are also continuous on I_2, and hence so too is K'. Now, because $K'(\xi) = 0$, there must exist some compact interval $I \subseteq I_2$ such that $|K'(x)| < 1$ for all $x \in I$; indeed, by the extreme value theorem there exists some $\theta \in [0, 1)$ such that $|K'(x)| \leq \theta$ for all $x \in I$. Therefore, since K is continuous on I and differentiable on $\operatorname{Int}(I)$, Proposition 1.21 implies that $K : I \to \mathbb{R}$ is a contraction mapping on I.

Finally, because $(\mathbb{R}, |\cdot|)$ is a Banach space and $\varnothing \neq I \subseteq \mathbb{R}$ is closed, the contraction principle informs us that ξ is a unique fixed point for K on I, with

$$|K^n(x_0) - \xi| \leq \frac{\theta^n}{1-\theta}|K(x_0) - x_0|$$

for all $n \in \mathbb{N}$ and any fixed $x_0 \in I$. Since $\theta^n \to 0$ as $n \to \infty$, it follows by the squeeze theorem that $(K^n(x_0))_{n\in\mathbb{N}}$ converges to ξ as desired.

We note that though the conditions imposed above are indeed sufficient to ensure the Newton-Raphson method converges to the zero ξ for f, they are not altogether necessary. Depending on the properties of the function f, the method may still find zeros for f even if, for instance, I is not compact or $f'(\xi) = 0$. ■

We now give a modest generalization of the contraction principle in which the hypothesis that a mapping K is a contraction is replaced by a rather less strict condition.

Theorem 1.23 (Weissinger's Theorem). *Let $(X, \|\cdot\|)$ be a Banach space and $\varnothing \neq S \subseteq X$ closed, and suppose $K : S \to S$ is such that, for convergent series $\sum_{n=1}^{\infty} \theta_n$,*

$$\|K^n(x_1) - K^n(x_2)\| \leq \theta_n \|x_1 - x_2\| \tag{1.15}$$

for all $n \in \mathbb{N}$ and $x_1, x_2 \in S$. Then K has a unique fixed point $x^* \in S$, and moreover

$$\|K^n(x) - x^*\| \leq \left(\sum_{j=n}^{\infty} \theta_j\right) \|K(x) - x\| \tag{1.16}$$

for all $n \in \mathbb{N}$ and $x \in S$.

Proof. Clearly $\theta_n \geq 0$ for all $n \in \mathbb{N}$, and since $\sum \theta_j$ is convergent there exists $\ell \in \mathbb{N}$ such that $\theta_n \in [0, 1)$ for all $n \geq \ell$. Suppose $x_1, x_2 \in S$ are fixed points for K. Then x_1 and x_2 are fixed points for K^ℓ, so that

$$\|x_1 - x_2\| = \|K^\ell(x_1) - K^\ell(x_2)\| \leq \theta_\ell \|x_1 - x_2\|,$$

and so if $x_1 \neq x_2$ then the contradiction $\theta_\ell \geq 1$ results. Hence $x_1 = x_2$, thereby proving the uniqueness of any fixed point that K may have in S.

Now fix $x_0 \in S$, and define the sequence $x_{n+1} = K(x_n)$ for $n \geq 0$. It can be shown by induction that $x_n = K^n(x_0)$, and with this we obtain

$$\|x_{n+1} - x_n\| = \|K^n(x_1) - K^n(x_0)\| \leq \theta_n \|x_1 - x_0\| \tag{1.17}$$

for $n \in \mathbb{N}$. Thus for $m > n$ we have, by the triangle inequality,

$$\|x_m - x_n\| = \left\| \sum_{j=n}^{m-1} (x_{j+1} - x_j) \right\| \leq \sum_{j=n}^{m-1} \|x_{j+1} - x_j\| \leq \left(\sum_{j=n}^{m-1} \theta_j \right) \|x_1 - x_0\|,$$

and thus

$$\|K^m(x_0) - K^n(x_0)\| \leq \left(\sum_{j=n}^{m-1} \theta_j \right) \|K(x_0) - x_0\| \tag{1.18}$$

for $m > n$. If $K(x_0) = x_0$, then $(K^n(x_0))_{n \in \mathbb{N}}$ is a constant sequence and hence Cauchy in S. Suppose $K(x_0) \neq x_0$, and let $\epsilon > 0$. Since $\sum \theta_j$ converges there is some $n \in \mathbb{N}$ such that $\sum_{j=n}^{\infty} \theta_j < \epsilon / \|K(x_0) - x_0\|$, and so by (1.18) we have

$$\|K^m(x_0) - K^n(x_0)\| \leq \left(\sum_{j=n}^{m-1} \theta_j \right) \|K(x_0) - x_0\| \leq \left(\sum_{j=n}^{\infty} \theta_j \right) \|K(x_0) - x_0\| < \epsilon \tag{1.19}$$

for $m > n$. Again we conclude that $(K^n(x_0))_{n \in \mathbb{N}}$ is Cauchy in S, and so by Proposition 1.13 there exists $x^* \in S$ such that $K^n(x_0) \to x^*$. Letting $m \to \infty$ in (1.19), we find by Proposition 1.3 that

$$\|K^n(x_0) - x^*\| \leq \left(\sum_{j=n}^{\infty} \theta_j \right) \|K(x_0) - x_0\|$$

for all $n \in \mathbb{N}$. Because $x_0 \in S$ is arbitrary we will have verified the estimate (1.16) once it is shown that x^* is a fixed point for K.

From (1.17) and the observation that $\theta_n \to 0$, the squeeze theorem implies $\|x_{n+1} - x_n\| \to 0$. Now, setting $n = 1$ in (1.15) makes clear that K is Lipschitz continuous on S, so K is continuous on S by Proposition 1.18, and hence the fact that $x_n \to x^*$ implies $x_{n+1} = K(x_n) \to K(x^*)$. Now we have

$$0 = \lim_{n \to \infty} \|x_{n+1} - x_n\| = \lim_{n \to \infty} \|K(x_n) - x_n\| = \|K(x^*) - x^*\|,$$

so that $K(x^*) = x^*$ and therefore $x^* \in S$ is a fixed point for K. ∎

If K in Weissinger's theorem were a contraction mapping with contraction constant θ, then each θ_n in (1.15) would specifically become θ^n. However, because $\theta_n \to 0$ as $n \to \infty$, Weissinger's theorem implies that K^n is a contraction mapping on S for sufficiently large n.

The following proposition is notable in that requires neither K^n to be a contraction mapping nor $(X, \|\cdot\|)$ to be complete.

Proposition 1.24. *Given a function $K : (X, \|\cdot\|) \to (X, \|\cdot\|)$, if K^n has unique fixed point x^* for some $n \in \mathbb{N}$, then K also has unique fixed point x^*.*

Proof. Suppose K^n has unique fixed point x^* for some $n \geq 1$, so that $K^n(x^*) = x^*$. If $n = 1$ then there is nothing to prove since $K^1 = K$, so assume that $n \geq 2$. Now, using the commutativity of the function composition operation, we have

$$K^n(K(x^*)) = K(K^n(x^*)) = K(x^*),$$

which shows $K(x^*)$ to be a fixed point for K^n, and therefore $K(x^*) = x^*$ since x^* is given to be a unique fixed point for K^n.

Now suppose that $\hat{x}$ is a fixed point for f, so $K(\hat{x}) = \hat{x}$. Then $\hat{x}$ must be a fixed point for K^n since

$$K^n(\hat{x}) = K^{j-1}(K(\hat{x})) = K^{k-1}(\hat{x})$$

for any $j \geq 1$, and so

$$K^n(\hat{x}) = K^{n-1}(\hat{x}) = K^{n-2}(\hat{x}) = \cdots = K^1(\hat{x}) = K(\hat{x}) = \hat{x}.$$

We conclude that $\hat{x} = x^*$ by the uniqueness of x^* as a fixed point for K^n, and therefore x^* is a unique fixed point for K also. ∎

Theorem 1.20 and Proposition 1.24 taken together deliver a result that is modestly stronger than the contraction principle alone, insofar as we may dispense with the latter's requirement that the mapping $K : S \to S$ be itself a contraction. Specifically we have the following.

Theorem 1.25. *Let $(X, \|\cdot\|)$ be a Banach space and $\varnothing \neq S \subseteq X$ closed, and suppose $K : S \to S$ is such that K^n is a contraction mapping for some $n \in \mathbb{N}$. Then K has a unique fixed point in S.*

Proof. Since K^n is a contraction mapping on the closed set S, the contraction principle implies that K^n possesses a unique fixed point x^* in S. By Proposition 1.24 it follows that x^* is also the unique fixed point for K on S. ∎

1.4 – The Picard-Lindelöf Theorem

Given an open set $U \subseteq \mathbb{R}^{n+2}$ and $F \in \mathcal{C}(U)$, an **ordinary differential equation** (ODE) is an equation of the form

$$F\big(t, x, x^{(1)}, \ldots, x^{(n)}\big) = 0, \tag{1.20}$$

where $x \in \mathcal{C}^n(S)$ for some $S \subseteq \mathbb{R}$, and

$$x^{(k)}(t) := \frac{d^k x}{dt^k}(t)$$

for each $0 \le k \le n$ (so $x^{(0)} := x$ in particular). We say (1.20) is an ***n*th-order** ODE to convey that the highest-order derivative of the dependent variable x in the equation is order n. Given an interval $I \subseteq S$, we say $\varphi \in \mathcal{C}^n(I)$ is a **solution** to (1.20) on I if, for all $t \in I$, we have $(t, \varphi(t), \varphi^{(1)}(t), \ldots, \varphi^{(n)}(t)) \in U$ such that

$$F\big(t, \varphi(t), \varphi^{(1)}(t), \ldots, \varphi^{(n)}(t)\big) = 0.$$

Unless otherwise indicated, all derivatives are considered to be two-sided, including at any endpoints of I!

We are particularly interested in nth-order ODEs for which $x^{(n)}$ may be isolated on one side of the equation, thereby obtaining an **explicit** nth-order ODE having the form

$$x^{(n)} = f\big(t, x, x^{(1)}, \ldots, x^{(n-1)}\big).$$

An explicit first-order ODE thus has the form $dx/dt = f(t, x)$. However, as t is naturally thought of as denoting time, we shall employ the dot notation $\dot{x} := dx/dt$ that customarily denotes differentiation with respect to time, and so write the first-order ODE as $\dot{x} = f(t, x)$. If $x(t)$ and $f(t, x)$ are vector-valued functions, then $\dot{x} = f(t, x)$ in fact represents a system of explicit first-order ODEs. Much of the remainder of this chapter will be occupied with the business of developing many theoretical results concerning such systems.

Letting

$$x(t) := (x_1(t), \ldots, x_n(t)) := (x_k(t))_{k=1}^n,$$

where each x_k is a real-valued function, in this section we state and prove our first existence-uniqueness theorem for a first-order **initial-value problem** (IVP) of the form

$$\dot{x} = f(t, x), \quad x(t_0) = x_0, \tag{1.21}$$

assuming that $f \in \mathcal{C}(U, \mathbb{R}^n)$ for some open set $U \subseteq \mathbb{R}^{n+1}$, with $(t_0, x_0) \in U$. If $n \ge 2$ then $\dot{x} = f(t, x)$ is a vector equation that amounts to a system of at least two first-order ordinary differential equations.

Given a function $\varphi : S \subseteq \mathbb{R} \to \mathbb{R}^n$, the symbol $\Gamma_\varphi(S)$ denotes the graph of φ on S; that is,

$$\Gamma_\varphi(S) := \big\{(t, \varphi(t)) : t \in S\big\},$$

which lies in $\mathbb{R}^{n+1}$.

Proposition 1.26. *Suppose $f(t,x) = (f_k(t,x))_{k=1}^n$ is continuous on an open set $U \subseteq \mathbb{R}^{n+1}$, $(t_0, x_0) \in U$, I is an interval containing t_0, and $\varphi : I \to \mathbb{R}^n$ is such that $\Gamma_\varphi(I) \subseteq U$. Then $\varphi(t)$ is a solution to the IVP* (1.21) *on I if and only if it is a solution to*

$$x(t) = x_0 + \int_{t_0}^{t} f(s, x(s))\, ds \tag{1.22}$$

on I.

Proof. Suppose $\varphi(t) := (\varphi_k(t))_{k=1}^n$ is a solution to (1.21) on I, so that $\dot{\varphi}(t) = f(t, \varphi(t))$ for all $t \in I$ and $\varphi(t_0) = x_0$. Since integration of a vector-valued function is carried out componentwise, by the fundamental theorem of calculus we have

$$\begin{aligned} x_0 + \int_{t_0}^{t} f(s, \varphi(s))\, ds = x_0 + \int_{t_0}^{t} \dot{\varphi}(s)\, ds &= x_0 + \left(\int_{t_0}^{t} \dot{\varphi}_k(s)\, ds \right)_{k=1}^n \\ &= x_0 + \big(\varphi_k(t) - \varphi_k(t_0)\big)_{k=1}^n \\ &= x_0 + \big(\varphi_k(t)\big)_{k=1}^n - \big(\varphi_k(t_0)\big)_{k=1}^n \\ &= x_0 + \varphi(t) - \varphi(t_0) = \varphi(t) \end{aligned}$$

for each $t \in I$, and thus φ satisfies (1.22) on I.

Next suppose φ is a solution to (1.22) on I; that is,

$$\varphi(t) = x_0 + \int_{t_0}^{t} f(s, \varphi(s))\, ds \tag{1.23}$$

for all $t \in I$. For each $t \in I$ the closed interval J with endpoints t_0 and t is a subset of I, so that $\Gamma_\varphi(J) \subseteq \Gamma_\varphi(I) \subseteq U$, and thus the integrand $s \mapsto f(s, \varphi(s))$ in (1.23) is continuous over the interval of integration. From this it follows that the components of $f(s, \varphi(s))$ are themselves continuous over the interval of integration, thereby assuring the existence of the integral in (1.23), and so

$$\dot{\varphi}(t) = \frac{d}{dt} \int_{t_0}^{t} f(s, \varphi(s))\, ds = \left(\frac{d}{dt} \int_{t_0}^{t} f_k(s, \varphi(s))\, ds \right)_{k=1}^n = \big(f_k(t, \varphi(t))\big)_{k=1}^n = f(t, \varphi(t))$$

for all $t \in I$ by the fundamental theorem of calculus, with

$$\varphi(t_0) = x_0 + \int_{t_0}^{t_0} f(s, \varphi(s))\, ds = x_0.$$

Therefore φ satisfies (1.21) on I. ∎

With t_0 as in (1.21) and some suitable $T > 0$ let $I_T := [t_0 - T, t_0 + T]$. To prove an existence-uniqueness theorem for the initial-value problem (1.21) with $f \in \mathcal{C}(U, \mathbb{R}^n)$, where $U \subseteq \mathbb{R}^{n+1}$ is open, we start by defining an operator K with domain some suitable subset of the normed vector space $(\mathcal{C}(I_T, \mathbb{R}^n), \|\cdot\|_\infty)$ by

$$K[\varphi](t) = x_0 + \int_{t_0}^{t} f(s, \varphi(s))\, ds; \tag{1.24}$$

that is, for each suitable $\varphi \in \mathcal{C}(I_T, \mathbb{R}^n)$, $K[\varphi]$ is the function given by (1.24). Proposition 1.11 provides assurance that $(\mathcal{C}(I_T, \mathbb{R}^n), \|\cdot\|_\infty)$ is a Banach space so long as $T < \infty$, but for any

$T > 0$ is may be possible to find some $\varphi \in \mathcal{C}(I_T, \mathbb{R}^n)$ such that $(t, \varphi(t)) \notin U$ for some $t \in I_T$. As f in (1.24) is only given to be continuous on U, it becomes clear that certain restrictions must be placed on our choice for either T or φ, and as it turns out, restrictions must be imposed on both. Indeed, in addition to the foregoing considerations, we seek some $T_0 > 0$ and closed set $S \subseteq (\mathcal{C}(I_{T_0}, \mathbb{R}^n), \|\cdot\|_\infty)$ such that K maps S into S and also K is a contraction on S. Before stating a lemma to address some of these matters, however, we need a definition.

Given normed vector spaces $(X, \|\cdot\|_X)$, $(Y, \|\cdot\|_Y)$, $(Z, \|\cdot\|_Z)$, with $W \subseteq X \times Y$, a function $F(x, y)$ that maps W into Z is said to be **Lipschitz continuous in the second argument, uniformly with respect to the first argument on $\boldsymbol{W}$** if there exists a constant $\theta \in [0, \infty)$ such that

$$\|F(x, y_1) - F(x, y_2)\|_Z \leq \theta \|y_1 - y_2\|_Y$$

for all $(x, y_1), (x, y_2) \in W$. We say $F : W \to Z$ is **locally Lipschitz continuous in the second argument, uniformly with respect to the first argument on $\boldsymbol{W}$** if at each point $(x, y) \in W$ there is an open set O containing (x, y) such that F is Lipschitz continuous in the second argument, uniformly with respect to the first argument on $O \cap W$. If X, Y, and Z are subsets of Euclidean spaces, then Proposition 1.17 may be used to show that if F is locally Lipschitz continuous in the second argument, uniformly with respect to the first argument on W, then F is Lipschitz continuous in the second argument, uniformly with respect to the first argument on any compact set C such that $C \subseteq W$.

Lemma 1.27. *Let $U \subseteq \mathbb{R}^{n+1}$ be open with $(t_0, x_0) \in U$, and suppose $T, \delta > 0$ are such that $V := I_T \times \overline{B}_\delta(x_0) \subseteq U$. For $f \in \mathcal{C}(U, \mathbb{R}^n)$ continuous on U, set*

$$M = \sup_{(t,x) \in V} |f(t, x)|,$$

and define $T_0 = \min\{T, \delta/M\}$, with $T_0 = T$ if $M = 0$. Also define

$$S = \left\{\varphi \in \mathcal{C}(I_{T_0}, \mathbb{R}^n) : \|\varphi - x_0\|_\infty \leq \delta\right\},$$

where $\|\varphi\|_\infty := \sup\{|\varphi(t)| : t \in I_{T_0}\}$.

1. *Let $V_0 = I_{T_0} \times \overline{B}_\delta(x_0)$. If $\varphi \in S$, then $\Gamma_\varphi(I_{T_0}) \subseteq V_0 \subseteq V$.*
2. *S is closed.*
3. *$K : S \to S$.*
4. *Suppose f is Lipschitz continuous in the second argument, uniformly with respect to the first argument on V_0, with Lipschitz constant L. If $T_0 < 1/L$, then K is a contraction on S.*

Proof.
Proof of (1). Fix $\varphi \in S$, so $\varphi : I_{T_0} \to \mathbb{R}^n$ is such that $\|\varphi - x_0\|_\infty \leq \delta$. Then for all $t \in I_{T_0}$ we have $|\varphi(t) - x_0| \leq \delta$, so that

$$\varphi(t) \in \overline{B}_\delta(x_0) = \{x \in \mathbb{R}^n : |x - x_0| \leq \delta\},$$

and therefore $(t, \varphi(t)) \in V_0$.

Proof of (2). Let $(\psi_k)_{k \in \mathbb{N}}$ be a Cauchy sequence in S, so $\psi_k \in \mathcal{C}(I_{T_0}, \mathbb{R}^n)$ is such that $\Gamma_{\psi_k}(I_{T_0}) \subseteq V_0$ for each k by part (1). Now, because (ψ_k) is also a Cauchy sequence in $(\mathcal{C}(I_{T_0}, \mathbb{R}^n), \|\cdot\|_\infty)$, which is complete by Proposition 1.11, there exists $\psi \in \mathcal{C}(I_{T_0}, \mathbb{R}^n)$ such that, with respect to the sup norm $\|\cdot\|_\infty$, $\psi_k \to \psi$ as $k \to \infty$. All that remains is to demonstrate that $\|\psi - x_0\|_\infty \le \delta$.

The sequence (ψ_k) certainly converges pointwise to ψ; that is, for each $t \in I_{T_0}$ we have $\psi_k(t) \to \psi(t)$ as $k \to \infty$, and thus $(t, \psi_k(t))_{k\in\mathbb{N}}$ is a sequence in V_0 that converges to $(t, \psi(t))$. But V_0 is a closed set, implying that $(t, \psi(t)) \in V_0$ for each $t \in I_{T_0}$, or equivalently $\Gamma_\psi(I_{T_0}) \subseteq V_0$, which in turn leads to $\varphi(I_{T_0}) \subseteq \overline{B}_\delta(x_0)$, and finally $\|\varphi - x_0\|_\infty \le \delta$. Therefore $\psi \in S$, and S is closed by Proposition 1.13.

Proof of (3). Fix $\varphi \in S$ and $t \in I_{T_0}$, the latter implying that $|t - t_0| \le T_0$. We have

$$\begin{aligned} |K[\varphi](t) - x_0| &= \left| \int_{t_0}^{t} f(s, \varphi(s))\, ds \right| \le \left| \int_{t_0}^{t} \big|f(s, \varphi(s))\big|\, ds \right| \\ &\le \left| \int_{t_0}^{t} M \right| = M|t - t_0| \le MT_0 \le M \cdot \frac{\delta}{M} = \delta, \end{aligned} \tag{1.25}$$

and thus

$$\|K[\varphi] - x_0\|_\infty = \sup_{t\in I_{T_0}} |K[\varphi](t) - x_0| \le \delta.$$

It remains to show that $K[\varphi] : I_{T_0} \to \mathbb{R}^n$ is continuous. For any $t_1, t_2 \in I_{T_0}$ we have

$$|K[\varphi](t_1) - K[\varphi](t_2)| = \left| \int_{t_1}^{t_2} f(s, \varphi(s))\, ds \right| \le \left| \int_{t_1}^{t_2} \big|f(s, \varphi(s))\big|\, ds \right| \le M|t_1 - t_2|, \tag{1.26}$$

so that $K[\varphi]$ is Lipschitz continuous on I_{T_0}, and hence continuous on I_{T_0} by Proposition 1.18. Therefore $K[\varphi] \in S$.

Proof of (4). Fix $\varphi, \psi \in S$. Now, for any $t \in I_{T_0}$,

$$\begin{aligned} |K[\varphi](t) - K[\varphi](t)| &= \left| \int_{t_0}^{t} \big[f(s, \varphi(s)) - f(s, \psi(s))\big]\, ds \right| \\ &\le \left| \int_{t_0}^{t} \big|f(s, \varphi(s)) - f(s, \psi(s))\big|\, ds \right| \\ &\le \left| \int_{t_0}^{t} L|\varphi(s) - \psi(s)|\, ds \right| \\ &\le L|t - t_0| \|\varphi - \psi\|_\infty \le LT_0\|\varphi - \psi\|_\infty, \end{aligned}$$

so that $\|K[\varphi] - K[\psi]\|_\infty \le LT_0\|\varphi - \psi\|_\infty$. Thus if $T_0 < 1/L$ it follows that K is a contraction on S. ∎

We observe that $M < \infty$ in the lemma since f is given to be continuous on the compact set V_0, and thus $T_0 > 0$ is assured.

In light of the contraction principle, the motivation for working with the operator (1.24) to demonstrate the uniqueness of a solution to (1.21) is perhaps made more manifest by the following proposition.

Proposition 1.28. *Let $f \in \mathcal{C}(U, \mathbb{R}^n)$ for open $U \subseteq \mathbb{R}^{n+1}$ with $(t_0, x_0) \in U$, let I be any interval containing t_0, and suppose $\varphi : I \to \mathbb{R}^n$ is such that $\Gamma_\varphi(I) \subseteq U$. Then φ satisfies the IVP* (1.21) *on I if and only if $K[\varphi] \equiv \varphi$ on I.*

Proof. Comparing (1.22) with (1.24), we see that Proposition 1.28 is an immediate consequence of Proposition 1.26. ∎

Broadly speaking, any solution to the initial-value problem (1.21) is a fixed point for the operator K as defined by (1.24).

For the statement of the Picard-Lindelöf theorem that is the centerpiece of the present section we introduce some notational conventions. For $U \subseteq \mathbb{R}^m$ and $V \subseteq \mathbb{R}^n$, we denote by $\mathcal{C}^k(U, V)$ the collection of all functions $F : U \to V$ that have continuous partial derivatives of order up to and including k on U. With $\mathcal{C}^0(U, V) := \mathcal{C}(U, V)$ and $\mathcal{C}^\infty(U, V) := \bigcap_{k=0}^\infty \mathcal{C}^k(U, V)$, it is clear that

$$\mathcal{C}^\infty(U, V) \subseteq \mathcal{C}^{k+1}(U, V) \subseteq \mathcal{C}^k(U, V) \subseteq \mathcal{C}(U, V)$$

for all $k \geq 0$. Since $V = \mathbb{R}$ quite frequently, we further define $\mathcal{C}^k(U) = \mathcal{C}(U, \mathbb{R})$ for $0 \leq k \leq \infty$. If $U \subseteq \mathbb{R}$, then of course all partial derivatives become ordinary derivatives. Differentiation of vector-valued functions is performed componentwise as usual.

In order to fully and honestly prove the Picard-Lindelöf theorem we shall need the following mean value theorem for vector-valued functions: if $h \in \mathcal{C}([a, b], \mathbb{R}^n) \cap \mathcal{C}^1((a, b), \mathbb{R}^n)$, then there exists $t \in (a, b)$ such that

$$|h(b) - h(a)| \leq (b - a)|h'(t)|.$$

Theorem 1.29 (Picard-Lindelöf). *Let $U \subseteq \mathbb{R}^{n+1}$ be open, with $(t_0, x_0) \in U$ and $T, \delta > 0$ such that $V := I_T \times \overline{B}_\delta(x_0) \subseteq U$. Suppose $f \in \mathcal{C}(U, \mathbb{R}^n)$ is locally Lipschitz continuous in the second argument, uniformly with respect to the first argument on U,*

$$M = \sup_{(t,x) \in V} |f(t, x)|,$$

and $T_0 = \min\{T, \delta/M\}$. For $V_0 := I_{T_0} \times \overline{B}_\delta(x_0)$, if L is a Lipschitz constant for f on V_0 and $T_0 < 1/L$, then the IVP (1.21) *has a unique solution $x^* \in \mathcal{C}^1(I_{T_0}, \mathbb{R}^n)$, and moreover this solution is such that $\Gamma_{x^*}(I_{T_0}) \subseteq V_0$.*

Proof. Suppose L is a Lipschitz constant for f on V_0 and $T_0 < 1/L$. The mapping $K : S \to S$ is a contraction on the closed set S by Lemma 1.27, and since $\mathcal{C}(I_{T_0}, \mathbb{R}^n)$ is a Banach space by Proposition 1.11, the contraction principle implies there exists a unique fixed point $x^* \in S$ for K. Now, $\Gamma_{x^*}(I_{T_0}) \subseteq V_0 \subseteq U$ by Lemma 1.27, and since $x^* : I_{T_0} \to \mathbb{R}^n$ is such that $K[x^*] \equiv x^*$ on I_{T_0}, by Proposition 1.28 we find x^* to be a solution to (1.21) on I_{T_0}, with $x^* \in \mathcal{C}^1(I_{T_0}, \mathbb{R}^n)$ since f is continuous and $\dot{x}^*(t) = f(t, x^*(t))$ for all $t \in I_{T_0}$. This proves the existence in S of a solution to the IVP.

Next suppose that $\varphi \in S$ satisfies the IVP on I_{T_0}. Then $K[\varphi] \equiv \varphi$ on I_{T_0} by Proposition 1.28, so that φ is a fixed point for $K : S \to S$, and therefore $\varphi = x^*$ by the uniqueness of x^* as a fixed point for K in S. This proves the uniqueness in S of a solution to the IVP.

It remains to show that there can exist no function $\varphi : I_{T_0} \to \mathbb{R}^n$ that lies *outside of* S for which $\Gamma_\varphi(I_{T_0}) \subseteq U$ and φ satisfies the IVP.[4] First we show that $\Gamma_{x^*}(\mathrm{Int}(I_{T_0})) \subseteq \mathrm{Int}(V_0)$. Suppose there exists $\tau \in [t_0, t_0 + T_0)$ such that $|x^*(\tau) - x_0| = \delta$, where $\tau > t_0$ since $x^*(t_0) = x_0$. By the mean value theorem there is some $\hat{\tau} \in (t_0, \tau)$ such that

$$|\dot{x}^*(\hat{\tau})| \geq \frac{|x^*(\tau) - x^*(t_0)|}{\tau - t_0} = \frac{|x^*(\tau) - x_0|}{\tau - t_0},$$

and so, since x^* satisfies (1.21) and $T_0 \leq \delta/M$,

$$\big|f(\hat{\tau}, x^*(\hat{\tau}))\big| = |\dot{x}^*(\hat{\tau})| \geq \frac{\delta}{\tau - t_0} > \frac{\delta}{T_0} \geq M. \tag{1.27}$$

However, $(\hat{\tau}, x^*(\hat{\tau})) \in V_0$ since $x^* \in S$, so that $|f(\hat{\tau}, x^*(\hat{\tau}))| \leq M$. This contradicts (1.27), and hence $|x^*(t) - x_0| < \delta$ for all $t \in [t_0, t_0 + T_0)$. This also holds for $t \in (t_0 - T_0, t_0]$, and therefore $|x^*(t) - x_0| < \delta$ for all $t \in \mathrm{Int}(I_{T_0})$.

Now suppose $\varphi : I_{T_0} \to \mathbb{R}^n$ is such that $\Gamma_\varphi(I_{T_0}) \subseteq U$ and φ satisfies (1.21) on I_{T_0}, but $\varphi \notin S$. Being a solution to the IVP certainly requires $\varphi \in \mathcal{C}(I_{T_0}, \mathbb{R}^n)$, so for $h : I_{T_0} \to \mathbb{R}$ given by $h(t) = |\varphi(t) - x_0|$ there exists $\hat{t} \in I_{T_0}$ such that $h(\hat{t}) > \delta$. For the sake of argument we assume $\hat{t} \in I_{T_0}^+ := [t_0, t_0 + T_0]$. Because $h(t_0) = 0$, the intermediate value theorem implies $h(t) = \delta$ for at least one value of t between t_0 and $\hat{t}$. Define

$$T_1 = \min\{t > 0 : h(t_0 + t) = \delta\}, \tag{1.28}$$

so that $t_0 < t_0 + T_1 < \hat{t} \leq t_0 + T_0$; also define $I = [t_0, t_0 + T_1]$. Since $K[\varphi] \equiv \varphi$ and $K[x^*] \equiv x^*$ on I_{T_0} by Proposition 1.28, the same identities hold on I; however, because $I \subseteq \mathrm{Int}(I_{T_0})$ we have $|x^*(t) - x_0| < \delta$ for all $t \in I$, whereas $|\varphi(t_0 + T_1) - x_0| = \delta$, and hence $\varphi \not\equiv x^*$ on I.

For $\|\psi\|_I := \sup_{t \in I} |\psi(t)|$, define

$$S' = \{\psi \in \mathcal{C}(I, \mathbb{R}^n) : \|\psi - x_0\|_I \leq \delta\},$$

which is a closed subset of $\mathcal{C}(I, \mathbb{R}^n)$ by much the same argument that proved part (3) of Lemma 1.27. Certainly $x^* \in S'$, so that $S' \neq \varnothing$. Also from (1.28) it's apparent that $|\varphi(t) - x_0| \leq \delta$ for $t \in I$, so that $\Gamma_\varphi(I) \subseteq I \times \overline{B}_\delta(x_0)$, and hence $\varphi \in S'$ as well. We also have $K : S' \to S'$. To see this, fix $\psi \in S'$ and $t \in I$; then, noting that $\Gamma_\psi(I) \subseteq V$, the find from

$$|K[\psi](t) - x_0| \leq \int_{t_0}^{t} |f(s, \psi(s))|\, ds \leq \int_{t_0}^{t} M\, ds \leq MT_0 \leq \delta$$

that $\|K[\psi] - x_0\|_I \leq \delta$, and the argument that $K[\psi] \in \mathcal{C}(I, \mathbb{R}^n)$ is largely the same as (1.26). Finally, K is a contraction on S', since for any $\psi_1, \psi_2 \in S'$ and $t \in I$ we have[5]

$$|K[\psi_1](t) - K[\psi_2](t)| \leq \int_{t_0}^{t} L|\psi_1(s) - \psi_2(s)|\, ds \leq LT_1 \|\psi_1 - \psi_2\|_I,$$

so that $\|K[\psi_1] - K[\psi_2]\|_I \leq \theta \|\psi_1 - \psi_2\|_I$ for $\theta = LT_1 \leq LT_0 < 1$. The contraction principle now informs us that $K : S' \to S'$ has a unique fixed point in S', whereas we have already discovered

[4] Curiously a significant number of ODE textbooks, both elementary and advanced, wholly ignore this apparent possibility.

[5] Compare with the proof of part (4) of Lemma 1.27.

that $x^*, \varphi \in S'$ are distinct fixed points for K. As this is a contradiction, we conclude that any solution to the IVP on I_{T_0} must lie in S. ■

It is generally desirable to secure the largest interval of validity possible for any solution to an initial-value problem. When applying the Picard-Lindelöf theorem this translates into determining the largest T_0 that satisfies the theorem's hypotheses. Since the theorem requires that $T_0 < 1/L$, we therefore wish to find the smallest Lipschitz constant L for f on V_0, which will be

$$L_0 := \sup_{(t,x)\neq(t,y)\in V_0} \frac{|f(t,x) - f(t,y)|}{|x-y|}. \tag{1.29}$$

Certainly for all $(t,x) \neq (t,y)$ in V_0 the equation (1.29) implies that

$$|f(t,x) - f(t,y)| \leq L_0|x-y|, \tag{1.30}$$

and since (1.30) is trivially true whenever $(t,x) = (t,y)$, we see that L_0 indeed qualifies as a Lipschitz constant for f on V_0.

Another possible means of extending an interval of validity I for a solution to the IVP (1.21) is to not contrive to have I be centered at t_0. Depending on the nature of f it may be possible to have $I = [t_0 - T_0', t_0 + T_0]$ for $T_0' \neq T_0$. For this small loss of symmetry we may find we can satisfy the hypotheses of Theorem 1.29 on a longer interval to one side of t_0 than on the other side; that is, the Picard-Lindelöf theorem may be applied to intervals of the form $[t_0 - T, t_0]$ and $[t_0, t_0 + T]$ separately for different values of T that give rise to different values of T_0.

If f in (1.21) happens to be of class $\mathcal{C}^1$, which is to say $f \in \mathcal{C}^1(U, \mathbb{R}^n)$, then it turns out that f is locally Lipschitz continuous in the second argument, uniformly with respect to the first argument. Another proposition will be necessary in order to prove this, but there are some preliminaries to dispense with. First, a subset C of a vector space is **convex** if, for any $x, y \in C$, the **line segment**

$$[x,y] := \big\{(1-t)x + ty : t \in [0,1]\big\}$$

whose endpoints are x and y is a subset of C. Next, given a field $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$, the **Frobenius norm** of a matrix $A \in \mathbb{F}^{m\times n}$ is defined to be

$$|A| = \sum_{j=1}^{n}\sum_{i=1}^{m} |a_{ij}|,$$

where a_{ij} denotes the ij-entry of A.

Proposition 1.30. *For $U \subseteq \mathbb{R}^n$ open, suppose $F \in \mathcal{C}^1(U, \mathbb{R}^m)$. If $C \subseteq U$ is compact and convex, then F is Lipschitz continuous on C with Lipschitz constant* $\sup_{x\in C} |[\mathrm{d}F(x)]|$.

Proof. Suppose $C \subseteq U$ is both compact and convex. The first-order partial derivatives of F being continuous on U by hypothesis, Proposition 1.14 implies F is differentiable on U, so that the mapping $(U, |\cdot|) \to (\mathbb{R}^{m\times n}, |\cdot|)$ given by

$$x \mapsto [\mathrm{d}F(x)] := \left[\frac{\partial F_i}{\partial x_j}(x)\right]$$

for each $x \in U$ is continuous on U, and hence bounded on C with respect to the Frobenius norm. Let

$$M = \sup_{x \in C} \left|[\mathrm{d}F(x)]\right|,$$

and fix $a, b \in C$. Then $[a, b] \subseteq C$ since C is convex. Defining $h : [0, 1] \to C$ by $h(t) = (1-t)a+tb$, we apply the fundamental theorem of calculus to each of the m components of F to obtain

$$\int_0^1 (F \circ h)'(t)\, dt = (F \circ h)(1) - (F \circ h)(0) = F(b) - F(a). \tag{1.31}$$

Now, because h is continuous, $h([0, 1]) \subseteq C$, and $C \subseteq U$ is closed, there exists an open interval $I \supseteq [0, 1]$ such that $h : I \to U$, whereupon $F \circ h : I \to \mathbb{R}^m$. Equation (1.7) makes clear that $\mathrm{d}h(t) : \mathbb{R} \to \mathbb{R}^n$ exists for each $t \in I$, and if $h_1, \ldots, h_n$ are the components of h, then

$$[\mathrm{d}h(t)] = \begin{bmatrix} h_1'(t) \\ \vdots \\ h_n'(t) \end{bmatrix} = h'(t) = b - a.$$

Similarly we have $[\mathrm{d}(F \circ h)(t)] = (F \circ h)'(t)$ for $t \in I$. In terms of their standard matrices, the linear map $\mathrm{d}F(h(t)) : \mathbb{R}^n \to \mathbb{R}^m$ is an $m \times n$ matrix and $\mathrm{d}h(t) : \mathbb{R} \to \mathbb{R}^n$ is an $n \times 1$ matrix, and thus

$$(F \circ h)'(t) = [\mathrm{d}(F \circ h)(t)] = [\mathrm{d}F(h(t))][\mathrm{d}h(t)] = [\mathrm{d}F(h(t))]h'(t) = [\mathrm{d}F(h(t))](b - a)$$

by the remark following the total derivative chain rule of Theorem 1.15. From (1.31) we now obtain

$$F(b) - F(a) = \int_0^1 [\mathrm{d}F(h(t))](b - a)\, dt.$$

Finally, using a property of the Frobenius norm which holds that $|AB| \le |A||B|$ for any matrices A and B for which the product AB is defined,

$$\begin{aligned} |F(b) - F(a)| &\le \int_0^1 \left|[\mathrm{d}F(h(t))](b - a)\right| dt \le \int_0^1 \left|[\mathrm{d}F(h(t))]\right||b - a|\, dt \\ &\le \int_0^1 M|b - a|\, dt = M|b - a| = \left(\sup_{x \in C} \left|[\mathrm{d}F(x)]\right|\right)|b - a|, \end{aligned}$$

finishing the proof. ■

Corollary 1.31. *If $U \subseteq \mathbb{R}^n$ is open and $F \in \mathcal{C}^1(U, \mathbb{R}^m)$, then F is locally Lipschitz continuous on U.*

Proof. Suppose $U \subseteq \mathbb{R}^n$ is open and $F \in \mathcal{C}^1(U, \mathbb{R}^m)$. For any $x \in U$ there exists $\delta > 0$ such that $B_{2\delta}(x) \subseteq U$, and hence $\overline{B}_\delta(x) \subseteq U$. Now, since the closed ball $\overline{B}_\delta(x)$ is both compact and convex, Proposition 1.30 implies that F is Lipschitz continuous on $\overline{B}_\delta(x)$, and hence on the open ball $B_\delta(x)$. ■

1.5 – Other Existence-Uniqueness Theorems

Given normed vector spaces $(X, \|\cdot\|_X)$, $(Y, \|\cdot\|_Y)$, $(Z, \|\cdot\|_Z)$, let $W \subseteq X \times Y$, and for each $x \in X$ let

$$W_x := \{y \in Y : (x, y) \in W\}.$$

A function $F(x, y)$ that maps W into Z is said to be **Lipschitz continuous in the second argument on W** if, for each fixed $x \in X$ for which $W_x \neq \varnothing$, the function $F(x, \cdot) : W_x \to Z$ is Lipschitz continuous on W_x; that is, for each $x \in X$ there exists a constant $\theta_x \in [0, \infty)$ such that

$$\|F(x, y_1) - F(x, y_2)\|_Z \leq \theta_x \|y_1 - y_2\|_Y$$

for all $y_1, y_2 \in W_x$. We say $F : W \to Z$ is **locally Lipschitz continuous in the second argument on W** if $F(x, \cdot) : W_x \to Z$ is locally Lipschitz continuous on W_x for each $x \in X$ for which $W_x \neq \varnothing$. If X, Y, and Z represent Euclidean spaces (or subsets of same), then Proposition 1.17 informs us that whenever F is given to be locally Lipschitz continuous in the second argument on W, then for each $x \in X$ the function $F(x, \cdot)$ is Lipschitz continuous on compact subsets of W_x.

We start with the statement and proof of a local existence-uniqueness theorem for the initial-value problem

$$\dot{x} = f(t, x), \quad x(t_0) = x_0, \tag{1.32}$$

that dispenses with the Picard-Lindelöf theorem's requirement that $T_0 < 1/L$, and relaxes other conditions. For instance the hypothesis of Theorem 1.29 that $f \in \mathcal{C}(U, \mathbb{R}^n)$ be locally Lipschitz continuous in the second argument, uniformly with respect to the first argument on U will be weakened to local Lipschitz continuity in the second argument on U. Another way in which the following theorem contrasts with Theorem 1.29 is it restricts the time interval on which the unique solution to (1.32) exists to $[t_0, t_0 + T_0]$. There exists an analogous result which considers $t < t_0$. We let $I_T^+ := [t_0, t_0 + T]$ for $T > 0$.

Theorem 1.32. *Let $U \subseteq \mathbb{R}^{n+1}$ be open, with $(t_0, x_0) \in U$ and $T, \delta > 0$ such that $I_T^+ \times \overline{B}_\delta(x_0) \subseteq U$. Suppose $f \in \mathcal{C}(U, \mathbb{R}^n)$ is locally Lipschitz continuous in the second argument on U, define*

$$M(t) = \int_{t_0}^{t} \left(\sup_{x \in \overline{B}_\delta(x_0)} |f(s, x)| \right) ds$$

and

$$L(t) = \sup_{x_1 \neq x_2 \in \overline{B}_\delta(x_0)} \frac{|f(t, x_1) - f(t, x_2)|}{|x_1 - x_2|}$$

for $t \in I_T^+$, and assume

$$\lambda := \int_{t_0}^{t_0 + T_0} L(t)\, dt < \infty$$

for

$$T_0 := \sup\{t : t_0 + t \in I_T^+ \text{ and } M(t_0 + t) \leq \delta\}.$$

Then there is a unique solution $x^* \in \mathcal{C}^1(I_{T_0}^+, \overline{B}_\delta(x_0))$ for the IVP (1.32), with

$$\sup_{t\in I_{T_0}^+} \left|K^m[x_0](t) - x^*(t)\right| \le \frac{\lambda^m e^\lambda}{m!} \int_{t_0}^{t_0+T_0} |f(s,x_0)|\,ds \tag{1.33}$$

for all $m \in \mathbb{N}$, and thus $K^m[x_0] \xrightarrow{\text{u}} x^*$ on $I_{T_0}^+$.

Proof. The normed vector space $(\mathcal{C}(I_{T_0}^+, \mathbb{R}^n), \|\cdot\|_\infty)$ is a Banach space by Proposition 1.11, and

$$S := \left\{\varphi \in \mathcal{C}(I_{T_0}^+, \mathbb{R}^n) : \|\varphi - x_0\|_\infty \le \delta\right\}$$

is a closed subset by the same argument as in the proof of Lemma 1.27.[6] Now, for any $\varphi \in S$ and $t \in I_{T_0}^+$ we have $\varphi(t) \in \overline{B}_\delta(x_0)$, and so

$$|K[\varphi](t) - x_0| \le \int_{t_0}^t |f(s,\varphi(s))|\,ds \le M(t) \le \delta$$

by the definition of T_0 and the fact that $M : I_{T_0}^+ \to \mathbb{R}$ is a monotone increasing function. Hence $\|K[\varphi] - x_0\|_\infty \le \delta$, and the calculation (1.26) for $t_1, t_2 \in I_{T_0}^+$ shows that $K[\varphi] : I_{T_0}^+ \to \mathbb{R}^n$ is continuous and therefore $K : S \to S$.

Define

$$\ell(t) = \int_{t_0}^t L(s)\,ds,$$

where $\ell : I_{T_0}^+ \to \mathbb{R}$ since $\ell(t) \le \lambda < \infty$ for all $t \in I_{T_0}^+$. It can be shown by induction that

$$\forall m \in \mathbb{N} \left[\forall \varphi, \psi \in S \;\; \forall t \in I_{T_0}^+ \left(|K^m[\varphi](t) - K^m[\psi](t)| \le \frac{\ell^m(t)}{m!} \sup_{r\in[t_0,t]} |\varphi(r) - \psi(r)|\right)\right], \tag{1.34}$$

and thus

$$\|K^m[\varphi] - K^m[\psi]\|_\infty \le \frac{\lambda^m}{m!}\|\varphi - \psi\|_\infty.$$

Observing that $\sum_{m=1}^\infty \lambda^m/m! < \infty$, Theorem 1.23 implies that K has a unique fixed point $x^* \in S$ such that

$$\|K^m[\varphi] - x^*\|_\infty \le \left(\sum_{j=m}^\infty \frac{\lambda^j}{j!}\right)\|K[\varphi] - \varphi\|_\infty \tag{1.35}$$

for all $m \in \mathbb{N}$ and $\varphi \in S$. Since the constant function x_0 is in S, we may substitute it for φ in (1.35) to obtain

$$\|K^m[x_0] - x^*\|_\infty \le \left(\sum_{j=m}^\infty \frac{\lambda^j}{j!}\right)\|K[x_0] - x_0\|_\infty. \tag{1.36}$$

However, we find that

$$\|K[x_0] - x_0\|_\infty = \sup_{t\in I_{T_0}^+} |K[x_0](t) - x_0| \le \sup_{t\in I_{T_0}^+} \int_{t_0}^t |f(s,x_0)|\,ds \le \int_{t_0}^{t_0+T_0} |f(s,x_0)|\,ds \tag{1.37}$$

[6] We refrain from citing Lemma 1.27 itself since technically that lemma has a different definition for T_0 than the present setting.

and

$$\sum_{j=m}^{\infty} \frac{\lambda^j}{j!} = \sum_{j=0}^{\infty} \frac{\lambda^{j+m}}{(j+m)!} \leq \sum_{j=0}^{\infty} \frac{\lambda^{j+m}}{j!m!} = \frac{\lambda^m}{m!} \sum_{j=0}^{\infty} \frac{\lambda^j}{j!} = \frac{\lambda^m e^\lambda}{m!}, \tag{1.38}$$

and the inequalities (1.36), (1.37), and (1.38) taken together yield (1.33). Because $x^* \in S$ it is clear that $\Gamma_{x^*}(I_{T_0}^+) \subseteq I_T^+ \times \overline{B}_\delta(x_0) \subseteq U$, and thus by Proposition 1.28 we conclude that x^* is a solution to the IVP (1.32) on $I_{T_0}^+$, with uniqueness assured by the uniqueness of x^* as a fixed point for K in S.

The proof is done once we attend to the detail of affirming the veracity of (1.34) by an inductive argument. The fact that

$$|f(t,x_1) - f(t,x_2)| \leq L(t)|x_1 - x_2| \tag{1.39}$$

for each $t \in I_{T_0}^+$ and $x_1, x_2 \in \overline{B}_\delta(x_0)$ implies

$$\begin{aligned} |K[\varphi](t) - K[\psi](t)| &\leq \int_{t_0}^t \big|f(s,\varphi(s)) - f(s,\psi(s))\big| ds \leq \int_{t_0}^t L(s)|\varphi(s) - \psi(s)|\, ds \\ &\leq \int_{t_0}^t L(s) \sup_{r \in [t_0,t]} |\varphi(r) - \psi(r)|\, ds = \ell(t) \sup_{r \in [t_0,t]} |\varphi(r) - \psi(r)|, \end{aligned}$$

establishing the bracketed statement in (1.34) for $m = 1$. Now suppose that the bracketed statement holds for some $m \in \mathbb{N}$. Recalling (1.39) and observing that $\ell'(t) = L(t)$,

$$\begin{aligned} |K^{m+1}[\varphi](t) - K^{m+1}[\psi](t)| &\leq \int_{t_0}^t \big|f(s, K^m[\varphi](s)) - f(s, K^m[\psi](s))\big| ds \\ &\leq \int_{t_0}^t L(s)|K^m[\varphi](s) - K^m[\psi](s)|\, ds \\ &\leq \int_{t_0}^t L(s) \frac{\ell^m(s)}{m!} \sup_{r \in [t_0,s]} |\varphi(r) - \psi(r)|\, ds \\ &\leq \frac{1}{m!} \sup_{r \in [t_0,t]} |\varphi(r) - \psi(r)| \int_{t_0}^t \ell'(s)\ell^m(s)\, ds. \end{aligned} \tag{1.40}$$

Applying integration by parts with $u = \ell^m(s)$ and $v' = \ell'(s)$ yields

$$\int_{t_0}^t \ell'(s)\ell^m(s)\, ds = \ell^{m+1}(t) - m \int_{t_0}^t \ell'(s)\ell^m(s)\, ds,$$

whence comes

$$\int_{t_0}^t \ell'(s)\ell^m(s)\, ds = \frac{\ell^{m+1}(t)}{m+1},$$

and so, returning to (1.40),

$$\begin{aligned} |K^{m+1}[\varphi](t) - K^{m+1}[\psi](t)| &\leq \frac{1}{m!} \sup_{r \in [t_0,t]} |\varphi(r) - \psi(r)| \frac{\ell^{m+1}(t)}{m+1} \\ &= \frac{\ell^{m+1}(t)}{(m+1)!} \sup_{r \in [t_0,t]} |\varphi(r) - \psi(r)|, \end{aligned}$$

thereby affirming the bracketed statement in (1.34) for m replaced by $m+1$. This finishes the proof. ■

The bound furnished by (1.33) on the error incurred when approximating the unique solution $x^*(t)$ to the IVP (1.32) using the function $K^m[x_0](t)$ makes clear that $K^m[x_0]$ converges uniformly to x^* on $I^+_{T_0}$. The process of finding $K^m[x_0]$ for successive values of $m \in \mathbb{N}$ is known as **Picard iteration**, with $K^m[x_0]$ itself being the ***m*th Picard iterate**. From

$$K^m[x_0](t) = K\big[K^{m-1}[x_0]\big](t) = x_0 + \int_{t_0}^{t} f(s, K^{m-1}[x_0](s))\,ds$$

we have

$$\frac{d}{dt}\big(K^m[x_0](t)\big) = f(t, K^{m-1}[x_0](t)),$$

which is continuous on $I^+_{T_0}$, and therefore $K^m[x_0] \in \mathcal{C}^1(I^+_{T_0}, \overline{B}_\delta(x_0))$ for all m since $K : S \to S$ implies $K^m : S \to S$.

For the sake of completeness we now state the counterpart to Theorem 1.32 that addresses solutions to (1.32) for $t < t_0$. The proof is of course quite similar. We let $I^-_T := [t_0 - T, t_0]$ for $T > 0$.

Theorem 1.33. *Let $U \subseteq \mathbb{R}^{n+1}$ be open, with $(t_0, x_0) \in U$ and $T, \delta > 0$ such that $I^-_T \times \overline{B}_\delta(x_0) \subseteq U$. Suppose $f \in \mathcal{C}(U, \mathbb{R}^n)$ is locally Lipschitz continuous in the second argument on U, define*

$$M(t) = \int_{t}^{t_0} \left(\sup_{x \in \overline{B}_\delta(x_0)} |f(s,x)| \right) ds$$

and

$$L(t) = \sup_{x_1 \neq x_2 \in \overline{B}_\delta(x_0)} \frac{|f(t,x_1) - f(t,x_2)|}{|x_1 - x_2|}$$

for $t \in I^-_T$, and assume

$$\lambda := \int_{t_0 - T_1}^{t_0} L(t)\,dt < \infty$$

for

$$T_1 := \sup\{t : t_0 - t \in I^-_T \text{ and } M(t_0 - t) \leq \delta\}.$$

Then there is a unique solution $x^* \in \mathcal{C}^1(I^-_{T_1}, \overline{B}_\delta(x_0))$ for the IVP (1.32), with

$$\sup_{t \in I^-_{T_1}} \big|K^m[x_0](t) - x^*(t)\big| \leq \frac{\lambda^m e^\lambda}{m!} \int_{t_0 - T_1}^{t_0} |f(s, x_0)|\,ds$$

for all $m \in \mathbb{N}$, and thus $K^m[x_0] \xrightarrow{\text{u}} x^*$ on $I^-_{T_1}$.

In general one can expect the T_0 of Theorem 1.32 and the T_1 of Theorem 1.33 to be different values, and the two theorems taken together assure the existence of a unique solution to the IVP (1.32) on the interval $[t_0 - T_1, t_0 + T_0]$.

Theorem 1.34. *Let $U = I \times \mathbb{R}^n$ for open interval I, with $(t_0, x_0) \in U$. If $f \in \mathcal{C}(U, \mathbb{R}^n)$ is Lipschitz continuous in the second argument on U, and $T > t_0$ is such that $I_T^+ \subseteq I$ and*

$$\lambda := \int_{t_0}^{t_0+T} \left(\sup_{x_1 \neq x_2 \in \mathbb{R}^n} \frac{|f(t,x_1) - f(t,x_2)|}{|x_1 - x_2|} \right) dt < \infty, \tag{1.41}$$

then there is a unique solution $x^* \in \mathcal{C}^1(I_T^+, \mathbb{R}^n)$ to the IVP (1.32).

Proof. Let $S = \mathcal{C}(I_T^+, \mathbb{R}^n)$, which is a closed subset of the Banach space $(\mathcal{C}(I_T^+, \mathbb{R}^n), \|\cdot\|_\infty)$. Fix $\varphi \in S$. The function $t \mapsto |f(t, \varphi(t))|$ is continuous on I_T^+, which is compact and so $\alpha := \sup_{t \in I_T^+} |f(t, \varphi(t))| < \infty$ by the extreme value theorem. Then for any $t_1, t_2 \in I_T^+$,

$$|K[\varphi](t_1) - K[\varphi](t_2)| = \left| \int_{t_1}^{t_2} f(s, \varphi(s))\, ds \right| \leq \left| \int_{t_1}^{t_2} \big|f(s, \varphi(s))\big|\, ds \right| \leq \alpha |t_1 - t_2|,$$

so that $K[\varphi]$ is Lipschitz continuous on I_T^+, hence continuous on I_T^+ by Proposition 1.18, and so $K : S \to S$.

For each $t \in I_T^+$ define

$$L(t) = \sup_{x_1 \neq x_2 \in \mathbb{R}^n} \frac{|f(t,x_1) - f(t,x_2)|}{|x_1 - x_2|}, \tag{1.42}$$

and in light of (1.41) define $\ell : I_T^+ \to [0, \lambda]$ by

$$\ell(t) := \int_{t_0}^{t} L(s)\, ds.$$

Using (1.34) with $I_{T_0}^+$ replaced by I_T^+, we have

$$\|K^m[\varphi] - K^m[\psi]\|_\infty \leq \frac{\lambda^m}{m!} \|\varphi - \psi\|_\infty$$

for all $\varphi, \psi \in S$. Since $\sum_{m=1}^\infty \lambda^m / m! < \infty$, Theorem 1.23 implies K has a unique fixed point $x^* \in S$ such that (1.35) holds for all $m \in \mathbb{N}$ and $\varphi \in S$. The proof of Theorem 1.32 from (1.35) onwards shows that the bound (1.33) applies in the present setting; however, we only need the fact that $x^* : I_T^+ \to \mathbb{R}^n$ possesses the property $K[x^*] \equiv x^*$ on I_T^+. Proposition 1.28 then implies that x^* is a solution to the IVP (1.32) on I_T^+, with uniqueness assured by the uniqueness of x^* as a fixed point for K in S. ■

Corollary 1.35. *If $f \in \mathcal{C}(\mathbb{R}^{n+1}, \mathbb{R}^n)$ is Lipschitz continuous in the second argument on $\mathbb{R}^{n+1}$, and*

$$\int_{t_0 - T}^{t_0 + T} L(t)\, dt < \infty$$

for all $T > 0$ and $L(t)$ given by (1.42), *then the IVP* (1.32) *has a unique solution $x^* \in \mathcal{C}^1(\mathbb{R}, \mathbb{R}^n)$.*

Proof. We find by Theorem 1.34 together with the analogous result pertaining to $[t_0 - T, t_0]$ that the IVP has a unique solution $x^* \in \mathcal{C}^1([t_0 - T, t_0 + T], \mathbb{R}^n)$ for each $T > 0$, and thus $x^* \in \mathcal{C}^1(\mathbb{R}, \mathbb{R}^n)$. The bound (1.33), which applies to the setting of Theorem 1.34, makes clear that if $I, J \subseteq \mathbb{R}$ are intervals containing t_0, with φ the unique solution to the IVP on I and ψ the unique solution to the IVP on J, then $\psi|_J \equiv \varphi$. ■

1.6 – Dependence on Initial Conditions

We now make a study of how the solution to the initial-value problem (1.32) depends on the initial condition $x(t_0) = x_0$. Specifically, what are the bounds on the change in the solution to the IVP when the point (t_0, x_0) is changed? Essential to the development of the theory is Grönwall's inequality, of which there are many variants, but the version that follows will suffice. As ever we define $I_T^+ = [t_0, t_0 + T]$ for $0 < T < \infty$.

Proposition 1.36 (**Grönwall's Inequality**). *Suppose $u, \alpha, \beta \in \mathcal{C}(I_T^+)$ with $\beta \geq 0$ on I_T^+. If*

$$u(t) \leq \alpha(t) + \int_{t_0}^{t} \beta(s)u(s)\,ds \tag{1.43}$$

for all $t \in I_T^+$, then

$$u(t) \leq \alpha(t) + \int_{t_0}^{t} \alpha(s)\beta(s)\exp\left(\int_{s}^{t} \beta(r)\,dr\right)ds \tag{1.44}$$

for all $t \in I_T^+$.

Proof. Define $\varphi : I_T^+ \to [0, \infty)$ by

$$\varphi(t) = \exp\left(-\int_{t_0}^{t} \beta(s)\,ds\right),$$

so $\dot{\varphi}(t) = -\beta(t)\varphi(t)$. Now, supposing (1.43) to be the case, we find for any $t \in I_T^+$ that

$$\begin{aligned}\frac{d}{dt}\left(\varphi(t)\int_{t_0}^{t} \beta(s)u(s)\,ds\right) &= \varphi(t)\beta(t)u(t) - \varphi(t)\beta(t)\int_{t_0}^{t} \beta(s)u(s)\,ds \\ &= \varphi(t)\beta(t)\left(u(t) - \int_{t_0}^{t} \beta(s)u(s)\,ds\right) \\ &\leq \varphi(t)\beta(t)\alpha(t).\end{aligned}$$

Thus for $s \in [t_0, t]$,

$$\frac{d}{ds}\left(\varphi(s)\int_{t_0}^{s} \beta(r)u(r)\,dr\right) \leq \alpha(s)\beta(s)\varphi(s),$$

and hence

$$\varphi(t)\int_{t_0}^{t} \beta(r)u(r)\,dr = \int_{t_0}^{t}\left[\frac{d}{ds}\left(\varphi(s)\int_{t_0}^{s} \beta(r)u(r)\,dr\right)\right]ds \leq \int_{t_0}^{t} \alpha(s)\beta(s)\varphi(s)\,ds.$$

Dividing by $\varphi(t)$ then yields

$$\int_{t_0}^{t} \beta(s)u(s)\,ds \leq \int_{t_0}^{t} \frac{\alpha(s)\beta(s)\varphi(s)}{\varphi(t)}\,ds = \int_{t_0}^{t} \alpha(s)\beta(s)\exp\left(\int_{s}^{t} \beta(r)\,dr\right)ds,$$

whereupon adding $\alpha(t)$ to both sides and making use of (1.43) once more leads to (1.44). ∎

Proposition 1.36 is still valid if the interval I_T^+ is replaced by a non-compact interval I such as $[t_0, T)$ or $[t_0, \infty)$. Also the continuity of α may be replaced by a weaker requirement that the negative part of α (i.e. α^- given by $\alpha^-(t) = \max\{-\alpha(t), 0\}$) is integrable on compact subintervals of I.

Corollary 1.37. *Suppose $u, \alpha, \beta \in \mathcal{C}(I_T^+)$ with β nonnegative and α nondecreasing on I_T^+. If* (1.43) *holds for all $t \in I_T^+$, then*

$$u(t) \leq \alpha(t) \exp\left(\int_{t_0}^{t} \beta(s)\, ds\right)$$

for all $t \in I_T^+$.

Proof. Since $\alpha(t_1) \leq \alpha(t_2)$ whenever $t_1 \leq t_2$, by Grönwall's inequality we have

$$\begin{aligned}
u(t) &\leq \alpha(t) + \int_{t_0}^{t} \alpha(s)\beta(s) \exp\left(\int_{s}^{t} \beta(r)\, dr\right) ds \\
&\leq \alpha(t) + \alpha(t) \int_{t_0}^{t} \beta(s) \exp\left(\int_{s}^{t} \beta(r)\, dr\right) ds \\
&= \alpha(t) - \alpha(t) \int_{t_0}^{t} \frac{d}{ds}\left[\exp\left(\int_{s}^{t} \beta(r)\, dr\right)\right] ds \\
&= \alpha(t) \exp\left(\int_{t_0}^{t} \beta(r)\, dr\right)
\end{aligned}$$

for all $t \in I_T^+$. ■

Corollary 1.38. *Fix $\alpha, \gamma \in \mathbb{R}$ and $\beta > 0$. If $u \in \mathcal{C}(I_T^+)$ is such that*

$$u(t) \leq \alpha + \int_{t_0}^{t} [\beta u(s) + \gamma]\, ds \tag{1.45}$$

for all $t \in I_T^+$, then

$$u(t) \leq \alpha \exp[\beta(t - t_0)] + \frac{\gamma}{\beta}\big(\exp[\beta(t - t_0)] - 1\big) \tag{1.46}$$

for all $t \in I_T^+$.

Proof. From (1.45) we have

$$u(t) + \frac{\gamma}{\beta} \leq \left(\alpha + \frac{\gamma}{\beta}\right) + \int_{t_0}^{t} \beta\left(u(s) + \frac{\gamma}{\beta}\right) ds$$

for all $t \in I_T^+$, so that

$$u(t) + \frac{\gamma}{\beta} \leq \left(\alpha + \frac{\gamma}{\beta}\right) \exp\left(\int_{t_0}^{t} \beta\, ds\right) = \left(\alpha + \frac{\gamma}{\beta}\right) \exp[\beta(t - t_0)]$$

by Corollary 1.37. The inequality (1.46) immediately follows. ■

Theorem 1.39. *Let $U \subseteq \mathbb{R}^{n+1}$ be open, with $(t_0, x_0), (t_0, y_0) \in U$. Suppose $f, g \in \mathcal{C}(U, \mathbb{R}^n)$, with f locally Lipschitz continuous in the second argument, uniformly with respect to the first argument on U. Suppose further that $x(t)$ and $y(t)$ satisfy the initial-value problems*

$$\dot{x} = f(t, x), \;\; x(t_0) = x_0 \quad \text{and} \quad \dot{y} = g(t, y), \;\; y(t_0) = y_0 \tag{1.47}$$

for $t \in I_T := [t_0 - T, t_0 + T]$, respectively. For $V \subseteq U$ a compact set with $\Gamma_x(I_T), \Gamma_y(I_T) \subseteq V$, let $L > 0$ be such that

$$|f(\tau, \xi_1) - f(\tau, \xi_2)| \leq L|\xi_1 - \xi_2|$$

for all $(\tau, \xi_1), (\tau, \xi_2) \in V$, and let

$$M = \sup_{(\tau,\xi) \in V} |f(\tau, \xi) - g(\tau, \xi)|.$$

Then

$$|x(t) - y(t)| \leq |x_0 - y_0| e^{L|t-t_0|} + \frac{M}{L}\big(e^{L|t-t_0|} - 1\big)$$

for all $t \in I_T$.

Proof. We carry out the proof only for $t \in I_T^+$, as the argument is essentially the same for $t \in [t_0 - T, t_0]$. By Proposition 1.26,

$$x(t) = \int_{t_0}^{t} f(s, x(s))\, ds \quad \text{and} \quad y(t) = \int_{t_0}^{t} g(s, y(s))\, ds$$

for $t \in I_T^+$, and so

$$\begin{aligned}
|x(t) - y(t)| &= \left|(x_0 - y_0) + \int_{t_0}^{t} \big[f(s, x(s)) - g(s, y(s))\big]\right| ds \\
&\leq |x_0 - y_0| + \int_{t_0}^{t} \big|f(s, x(s)) - g(s, y(s))\big|\, ds \\
&\leq |x_0 - y_0| + \int_{t_0}^{t} \big|f(s, x(s)) - f(s, y(s))\big|\, ds + \int_{t_0}^{t} \big|f(s, y(s)) - g(s, y(s))\big|\, ds \\
&\leq |x_0 - y_0| + \int_{t_0}^{t} L|x(s) - y(s)|\, ds + \int_{t_0}^{t} M\, ds.
\end{aligned}$$

Using Corollary 1.38 with $u(t) = |x(t) - y(t)|$, we finally obtain

$$\begin{aligned}
|x(t) - y(t)| &\leq |x_0 - y_0| + \int_{t_0}^{t} \big(L|x(s) - y(s)| + M\big)\, ds \\
&\leq |x_0 - y_0| e^{L|t-t_0|} + \frac{M}{L}\big(e^{L|t-t_0|} - 1\big),
\end{aligned}$$

where of course $t - t_0 = |t - t_0|$ for $t \in I_T^+$. ■

Subject to assumptions that assure uniqueness, let $\varphi(t, t_0, x_0)$ denote the solution to the initial-value problem (1.32) on some interval containing t_0. Keeping t_0 fixed, suppose there exists a closed interval I containing t_0 such that, for all ξ sufficiently close to x_0, the IVP $\dot{x} = f(t, x)$, $x(t_0) = \xi$ has a unique solution on I. Then the IVP (1.32) is **well-posed** if φ is continuous in the third argument at each $t \in I$. Thus there exists some $\alpha > 0$ such that, for all $\xi \in \mathbb{R}^n$ for which $|\xi - x_0| < \alpha$, the function $t \mapsto \varphi(t, t_0, \xi)$ is a solution to the IVP

$$\dot{x} = f(t, x), \quad x(t_0) = \xi \tag{1.48}$$

on I, and

$$\lim_{\xi \to x_0} \varphi(t, t_0, \xi) = \varphi(t, t_0, x_0)$$

for each $t \in I$.

Substituting f for g in Theorem 1.39, we take $x(t)$ and $y(t)$ to be solutions on I_T to the initial-value problems

$$\dot{x} = f(t, x),\ \ x(t_0) = x_0 \quad \text{and} \quad \dot{y} = f(t, y),\ \ y(t_0) = y_0,$$

respectively. Assuming the hypotheses of the theorem hold, and noting that $M = 0$ when $f = g$, the theorem concludes that

$$|x(t) - y(t)| \le |x_0 - y_0| e^{L|t - t_0|} \tag{1.49}$$

for each $t \in I_T$. Clearly

$$\varphi(t, t_0, y_0) = y(t) \to x(t) = \varphi(t, t_0, x_0)$$

as $y_0 \to x_0$, and hence the IVP (1.32) is well-posed provided solutions to (1.48) are valid on I_T for all ξ sufficiently close to x_0. The next theorem has something to say about this last point.

Theorem 1.40. *Let $U \subseteq \mathbb{R}^{n+1}$ be open, and suppose $f \in \mathcal{C}(U, \mathbb{R}^n)$ is locally Lipschitz continuous in the second argument, uniformly with respect to the first argument. Fix $(t_0, x_0) \in U$.*

1. *There exists compact interval I and compact set $B \subseteq \mathbb{R}^n$ such that $(t_0, x_0) \in I \times B \subseteq U$, and $\varphi(t, \tau, \xi)$ exists on $I \times I \times B$. Thus for each $\tau \in I$ and $\xi \in B$ the function $\varphi(\cdot, \tau, \xi) : I \to \mathbb{R}^n$ is a solution on I to the IVP*

$$\dot{x} = f(t, x), \quad x(\tau) = \xi.$$

2. *If V is a compact set such that $I \times \varphi(I \times I \times B) \subseteq V \subseteq U$,*

$$M := \sup_{(\tau, \xi) \in V} |f(\tau, \xi)|,$$

and $L > 0$ is such that

$$|f(\tau, \xi_1) - f(\tau, \xi_2)| \le L|\xi_1 - \xi_2|$$

for all $(\tau, \xi_1), (\tau, \xi_2) \in V$, then

$$|\varphi(t_1, \tau_1, \xi_1) - \varphi(t_2, \tau_2, \xi_2)| \le |\xi_1 - \xi_2| e^{L|t_1 - \tau_1|} + \big(|t_1 - t_2| + |\tau_1 - \tau_2| e^{L|t_1 - \tau_2|}\big) M \tag{1.50}$$

for all $(t_1, \tau_1, \xi_1), (t_2, \tau_2, \xi_2) \in I \times I \times B$.

3. *$\varphi(t, \tau, \xi) \in \mathcal{C}(I \times I \times B, \mathbb{R}^n)$.*

Proof.
Proof of (1). Fixing $(t_0, x_0) \in U$, let T, δ, V and M be defined as in Theorem 1.29, and assume L is a Lipschitz constant for f on V. Let $0 < \epsilon < \min\{T, \delta/M, 1/L\}$, and define $V_0 \subseteq V$ by $V_0 = [t_0 - \epsilon, t_0 + \epsilon] \times \overline{B}_\delta(x_0)$. By Theorem 1.29 the IVP (1.32) has a unique solution $\varphi(\cdot, t_0, x_0)$

on $[t_0-\epsilon, t_0+\epsilon]$. Now let $\tau \in I := [t_0-\epsilon/4, t_0+\epsilon/4]$ and $\xi \in B := \overline{B}_{\delta/2}(x_0)$ be arbitrary, and consider the IVP

$$\dot{x} = f(t,x), \quad x(\tau) = \xi. \tag{1.51}$$

Define $J_\tau = [\tau-\epsilon/2, \tau+\epsilon/2]$. Since $V_{\tau\xi} := J_\tau \times \overline{B}_{\delta/2}(\xi) \subseteq V_0 \subseteq V \subseteq U$, Theorem 1.29 implies there is a unique solution $\varphi(\cdot,\tau,\xi)$ to (1.51) on J_τ with $\Gamma_{\varphi(\cdot,\tau,\xi)}(J_\tau) \subseteq V_{\tau\xi}$. But $I \subseteq J_\tau$, so that $\varphi(\cdot,\tau,\xi)$ satisfies (1.51) on I with

$$\Gamma_{\varphi(\cdot,\tau,\xi)}(I) \subseteq \Gamma_{\varphi(\cdot,\tau,\xi)}(J_\tau) \subseteq V_{\tau\xi},$$

and we conclude that $\varphi(t,\tau,\xi)$ is defined for all $(t,\tau,\xi) \in I \times I \times B$.

Proof of (2). Fix $(t_1,\tau_1,\xi_1),(t_2,\tau_2,\xi_2) \in I \times I \times B$. By Theorem 1.39, and more specifically (1.49), we have

$$|\varphi(t_1,\tau_1,\xi_1) - \varphi(t_1,\tau_1,\xi_2)| \le |\xi_1-\xi_2|e^{L|t_1-\tau_1|}. \tag{1.52}$$

Now define $u : I \to \mathbb{R}$ by

$$u(t) = |\varphi(t,\tau_1,\xi_2) - \varphi(t,\tau_2,\xi_2)|.$$

By Proposition 1.26,

$$u(t) = \left|\left(\xi_2 + \int_{\tau_1}^t f(r,\varphi(r,\tau_1,\xi_2))\,dr\right) - \left(\xi_2 + \int_{\tau_2}^t f(r,\varphi(r,\tau_2,\xi_2))\,dr\right)\right|$$

and so

$$\begin{aligned}
u(t) &= \left|\int_{\tau_1}^t f(r,\varphi(r,\tau_1,\xi_2))\,dr - \int_{\tau_2}^t f(r,\varphi(r,\tau_2,\xi_2))\,dr\right| \\
&= \left|\int_{\tau_1}^{\tau_2} f(r,\varphi(r,\tau_1,\xi_2))\,dr + \int_{\tau_2}^t f(r,\varphi(r,\tau_1,\xi_2))\,dr - \int_{\tau_2}^t f(r,\varphi(r,\tau_2,\xi_2))\,dr\right| \\
&\le \left|\int_{\tau_1}^{\tau_2} f(r,\varphi(r,\tau_1,\xi_2))\,dr\right| + \left|\int_{\tau_2}^t |f(r,\varphi(r,\tau_1,\xi_2)) - f(r,\varphi(r,\tau_2,\xi_2))|\,dr\right| \\
&\le M|\tau_1-\tau_2| + \left|\int_{\tau_2}^t L|\varphi(r,\tau_1,\xi_2) - \varphi(r,\tau_2,\xi_2)|\,dr\right| \\
&= M|\tau_1-\tau_2| + \left|\int_{\tau_2}^t Lu(r)\,dr\right|
\end{aligned}$$

for all $t \in I$. If $t \in [\tau_2,\infty) \cap I$ then

$$u(t) \le M|\tau_1-\tau_2| + \int_{\tau_2}^t Lu(r)\,dr,$$

and so

$$u(t) \le M|\tau_1-\tau_2| \exp\left(\int_{\tau_2}^t L\,dr\right) = M|\tau_1-\tau_2|e^{L(t-\tau_2)}$$

by Corollary 1.37; and if $t \in (-\infty,\tau_2] \cap I$ then

$$u(t) \le M|\tau_1-\tau_2| + \int_t^{\tau_2} Lu(r)\,dr,$$

and so

$$u(t) \le M|\tau_1 - \tau_2|e^{L(\tau_2 - t)}$$

by the analogous result to Corollary 1.37. Therefore

$$u(t) \le M|\tau_1 - \tau_2|e^{L|t-\tau_2|} \tag{1.53}$$

for all $t \in I$.

Next, with another application of Proposition 1.26,

$$|\varphi(t_1,\tau_2,\xi_2) - \varphi(t_2,\tau_2,\xi_2)| = \left|\int_{t_2}^{t_1} f(r,\varphi(r,\tau_2,\xi_2))\,dr\right| \le M|t_1 - t_2|. \tag{1.54}$$

Finally, applying the triangle inequality along with (1.52), (1.53), and (1.54), we obtain

$$\begin{aligned}|\varphi(t_1,\tau_1,\xi_1) - \varphi(t_2,\tau_2,\xi_2)| &\le |\varphi(t_1,\tau_1,\xi_1) - \varphi(t_1,\tau_1,\xi_2)| + u(t_1) + |\varphi(t_1,\tau_2,\xi_2) - \varphi(t_2,\tau_2,\xi_2)| \\ &\le |\xi_1 - \xi_2|e^{L|t_1-\tau_1|} + M|\tau_1 - \tau_2|e^{L|t_1-\tau_2|} + M|t_1 - t_2|,\end{aligned}$$

which is (1.50).

Proof of (3). This will follow from part (2) provided that we can find a compact set V such that $I \times \varphi(I \times I \times B) \subseteq V \subseteq U$. Indeed, in the proof of part (1) we constructed the set $V_0 = [t_0 - \epsilon, t_0 + \epsilon] \times \overline{B}_\delta(x_0)$ with the property that

$$\Gamma_{\varphi(\cdot,\tau,\xi)}(I) \subseteq \Gamma_{\varphi(\cdot,\tau,\xi)}([\tau - \epsilon/2, \tau + \epsilon/2]) \subseteq [\tau - \epsilon/2, \tau + \epsilon/2] \times \overline{B}_{\delta/2}(\xi) \subseteq V_0 \subseteq U \tag{1.55}$$

for all $(\tau,\xi) \in I \times B$. We show that V_0 is a suitable choice for V.

Let $(t', x) \in I \times \varphi(I \times I \times B)$. Thus $t' \in [t_0 - \epsilon, t_0 + \epsilon]$, and there exists $(t,\tau,\xi) \in I \times I \times B$ such that $\varphi(t,\tau,\xi) = x$. Since $t \in I$ and

$$\{(s,\varphi(s,\tau,\xi)) : s \in I\} \subseteq [\tau - \epsilon/2, \tau + \epsilon/2] \times \overline{B}_{\delta/2}(\xi)$$

by (1.55), we have $x = \varphi(t,\tau,\xi) \in \overline{B}_{\delta/2}(\xi) \subseteq \overline{B}_\delta(x_0)$, and therefore $(t', x) \in V_0$. Having now shown that $I \times \varphi(I \times I \times B) \subseteq V_0$, the inequality (1.50) and a squeeze theorem argument may be applied to show that φ is continuous on $I \times I \times B$. ■

Fixing $t_0 \in I$, an immediate implication of part (1) of Theorem 1.40 is that the function $\varphi(\cdot, t_0, \xi)$ is a solution on I to the IVP (1.48) for all $\xi \in B$.

In (1.48) we now assume that $f \in \mathcal{C}^k(U, \mathbb{R}^n)$ for some $k \ge 1$, and that the solution $\varphi(\cdot, t_0, \xi)$ to the IVP on I is differentiable in the third argument (i.e. with respect to ξ). We continue to assume $\xi \in B$ for B as defined in Theorem 1.40, and $t_0 \in I$ is fixed. Since $\dot\varphi(t, t_0, \xi) = f(t, \varphi(t, t_0, \xi))$ for all $t \in I$ and $\xi \in B$, it follows that $\dot\varphi(t, t_0, \xi)$ is differentiable in the third argument as well.

Define $\varphi = (\varphi_j)_{j=1}^n$, so that $\varphi_j(\cdot, t_0, \cdot) : I \times \{t_0\} \times B \to \mathbb{R}$ for each $1 \le j \le n$. Also let $\xi = (\xi_j)_{j=1}^n$. Using the notation introduced in §1.2, the linear mapping $\partial_\xi\varphi : \mathbb{R}^n \to \mathbb{R}^n$ has $n \times n$ matrix $[\partial_\xi\varphi]$ given by

$$[\partial_\xi\varphi(t, t_0, \xi)]_{ij} = \frac{\partial\varphi_i}{\partial\xi_j}(t, t_0, \xi)$$

for $1 \leq i, j \leq n$. More explicitly we have

$$[\partial_\xi \varphi] = \begin{bmatrix} \dfrac{\partial \varphi_1}{\partial \xi_1} & \cdots & \dfrac{\partial \varphi_1}{\partial \xi_n} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial \varphi_n}{\partial \xi_1} & \cdots & \dfrac{\partial \varphi_n}{\partial \xi_n} \end{bmatrix}.$$

Now, by (1.48) and the chain rule,

$$\partial_\xi \partial_t \varphi(t, t_0, \xi) = \partial_\xi f(t, \varphi(t, t_0, \xi)) = \partial_x f(t, \varphi(t, t_0, \xi)) \circ \partial_\xi \varphi(t, t_0, \xi),$$

where $\partial_x f$ denotes the partial derivative of f with respect to the second argument. If we adopt the convention of denoting the composition of functions by juxtaposition, then we may write simply

$$\partial_\xi \partial_t \varphi(t, t_0, \xi) = \partial_x f(t, \varphi(t, t_0, \xi)) \partial_\xi \varphi(t, t_0, \xi). \tag{1.56}$$

Now we consider the so-called **first variational equation**

$$\dot{y} = A(t, \xi) y, \tag{1.57}$$

where

$$A(t, \xi) := \partial_x f(t, \varphi(t, t_0, \xi)).$$

The equation is linear in y, and if we suppose $\partial_\xi \partial_t \varphi = \partial_t \partial_\xi \varphi$, then by inspection it's seen that the function $t \mapsto \partial_\xi \varphi(t, t_0, \xi)$ is a solution to (1.57) on I for each $\xi \in B$. Equivalent to (1.57) is the integral equation

$$y(t) = E_n + \int_{t_0}^{t} A(s, \xi) y(s) \, ds, \tag{1.58}$$

where E_n is the $n \times n$ identity matrix. We pass from (1.57) to (1.58) by first integrating the former with respect to t to obtain the family of antiderivatives

$$y(t) = C + \int_{t_0}^{t} A(s, \xi) y(s) \, ds,$$

where C is an arbitrary constant $n \times n$ matrix. Thus $y(t_0) = C$, and to arrange for $\partial_\xi \varphi(t, t_0, \xi)$ to be the solution we next observe that $\varphi(t_0, t_0, \xi) = \xi$, so that $\partial_\xi \varphi(t_0, t_0, \xi) = E_n$, and thus we must choose C to be E_n.