Part IV

Calculus and Ordinary Differential Equations in Banach Spaces

15

Ordinary Differential Equations in a Banach Space

Let X be a Banach space, $U \subset_o X$, $J = (a,b) \ni 0$ and $Z \in C(J \times U, X)$ – Z is to be interpreted as a time dependent vector-field on $U \subset X$. In this section we will consider the ordinary differential equation (ODE for short)

$$\dot{y}(t) = Z(t, y(t)) \text{ with } y(0) = x \in U. \tag{15.1}$$

The reader should check that any solution $y \in C^1(J,U)$ to Eq. (15.1) gives a solution $y \in C(J,U)$ to the integral equation:

$$y(t) = x + \int_0^t Z(\tau, y(\tau))d\tau \tag{15.2}$$

and conversely if $y \in C(J,U)$ solves Eq. (15.2) then $y \in C^1(J,U)$ and y solves Eq. (15.1).

Remark 15.1. For notational simplicity we have assumed that the initial condition for the ODE in Eq. (15.1) is taken at $t = 0$. There is no loss in generality in doing this since if $\tilde{y}$ solves

$$\frac{d\tilde{y}}{dt}(t) = \tilde{Z}(t, \tilde{y}(t)) \text{ with } \tilde{y}(t_0) = x \in U$$

iff $y(t) := \tilde{y}(t + t_0)$ solves Eq. (15.1) with $Z(t,x) = \tilde{Z}(t + t_0, x)$.

15.1 Examples

Let $X = \mathbb{R}$, $Z(x) = x^n$ with $n \in \mathbb{N}$ and consider the ordinary differential equation

$$\dot{y}(t) = Z(y(t)) = y^n(t) \text{ with } y(0) = x \in \mathbb{R}. \tag{15.3}$$

If y solves Eq. (15.3) with $x \neq 0$, then $y(t)$ is not zero for t near 0. Therefore up to the first time y possibly hits 0, we must have

$$t = \int_0^t \frac{\dot{y}(\tau)}{y(\tau)^n} d\tau = \int_0^{y(t)} u^{-n} du = \begin{cases} \frac{[y(t)]^{1-n} - x^{1-n}}{1-n} & \text{if } n > 1 \\ \ln \left| \frac{y(t)}{x} \right| & \text{if } n = 1 \end{cases}$$

and solving these equations for $y(t)$ implies

$$y(t) = y(t,x) = \begin{cases} \frac{x}{\sqrt[n-1]{1-(n-1)tx^{n-1}}} & \text{if } n > 1 \\ e^t x & \text{if } n = 1. \end{cases} \tag{15.4}$$

The reader should verify by direct calculation that $y(t,x)$ defined above does indeed solve Eq. (15.3). The above argument shows that these are the only possible solutions to the Equations in (15.3).

Notice that when $n = 1$, the solution exists for all time while for $n > 1$, we must require

$$1 - (n-1)tx^{n-1} > 0$$

or equivalently that

$$t < \frac{1}{(1-n)x^{n-1}} \text{ if } x^{n-1} > 0 \text{ and}$$
$$t > -\frac{1}{(1-n)|x|^{n-1}} \text{ if } x^{n-1} < 0.$$

Moreover for $n > 1$, $y(t,x)$ blows up as t approaches the value for which $1 - (n-1)tx^{n-1} = 0$. The reader should also observe that, at least for s and t close to 0,

$$y(t, y(s,x)) = y(t+s, x) \tag{15.5}$$

for each of the solutions above. Indeed, if $n = 1$ Eq. (15.5) is equivalent to the well know identity, $e^t e^s = e^{t+s}$ and for $n > 1$,

$$\begin{aligned} y(t, y(s,x)) &= \frac{y(s,x)}{\sqrt[n-1]{1-(n-1)ty(s,x)^{n-1}}} \\ &= \frac{\frac{x}{\sqrt[n-1]{1-(n-1)sx^{n-1}}}}{\sqrt[n-1]{1-(n-1)t\left[\frac{x}{\sqrt[n-1]{1-(n-1)sx^{n-1}}}\right]^{n-1}}} \\ &= \frac{\frac{x}{\sqrt[n-1]{1-(n-1)sx^{n-1}}}}{\sqrt[n-1]{1-(n-1)t\frac{x^{n-1}}{1-(n-1)sx^{n-1}}}} \\ &= \frac{x}{\sqrt[n-1]{1-(n-1)sx^{n-1}-(n-1)tx^{n-1}}} \\ &= \frac{x}{\sqrt[n-1]{1-(n-1)(s+t)x^{n-1}}} = y(t+s,x). \end{aligned}$$

Now suppose $Z(x) = |x|^\alpha$ with $0 < \alpha < 1$ and we now consider the ordinary differential equation

$$\dot{y}(t) = Z(y(t)) = |y(t)|^\alpha \text{ with } y(0) = x \in \mathbb{R}. \tag{15.6}$$

Working as above we find, if $x \neq 0$ that

$$t = \int_0^t \frac{\dot{y}(\tau)}{|y(t)|^\alpha} d\tau = \int_0^{y(t)} |u|^{-\alpha} du = \frac{[y(t)]^{1-\alpha} - x^{1-\alpha}}{1-\alpha},$$

where $u^{1-\alpha} := |u|^{1-\alpha} \operatorname{sgn}(u)$. Since $\operatorname{sgn}(y(t)) = \operatorname{sgn}(x)$ the previous equation implies

$$\begin{aligned} \operatorname{sgn}(x)(1-\alpha)t &= \operatorname{sgn}(x)\left[\operatorname{sgn}(y(t))\,|y(t)|^{1-\alpha} - \operatorname{sgn}(x)\,|x|^{1-\alpha}\right] \\ &= |y(t)|^{1-\alpha} - |x|^{1-\alpha} \end{aligned}$$

and therefore,

$$y(t,x) = \operatorname{sgn}(x)\left(|x|^{1-\alpha} + \operatorname{sgn}(x)(1-\alpha)t\right)^{\frac{1}{1-\alpha}} \tag{15.7}$$

is uniquely determined by this formula until the first time t where $|x|^{1-\alpha} + \operatorname{sgn}(x)(1-\alpha)t = 0$. As before $y(t) = 0$ is a solution to Eq. (15.6), however it is far from being the unique solution. For example letting $x \downarrow 0$ in Eq. (15.7) gives a function

$$y(t, 0+) = ((1-\alpha)t)^{\frac{1}{1-\alpha}}$$

which solves Eq. (15.6) for $t > 0$. Moreover if we define

$$y(t) := \begin{cases} ((1-\alpha)t)^{\frac{1}{1-\alpha}} & \text{if } t > 0 \\ 0 & \text{if } t \le 0 \end{cases},$$

(for example if $\alpha = 1/2$ then $y(t) = \frac{1}{4}t^2 1_{t \ge 0}$) then the reader may easily check y also solve Eq. (15.6). Furthermore, $y_a(t) := y(t-a)$ also solves Eq. (15.6) for all $a \ge 0$, see Figure 15.1 below.

With these examples in mind, let us now go to the general theory. The case of linear ODE's has already been studied in Section 8.3 above.

15.2 Uniqueness Theorem and Continuous Dependence on Initial Data

Lemma 15.2. *Gronwall's Lemma.* *Suppose that f, ε, and k are non-negative functions of a real variable t such that*

$$f(t) \le \varepsilon(t) + \left| \int_0^t k(\tau) f(\tau) d\tau \right|. \tag{15.8}$$

Then

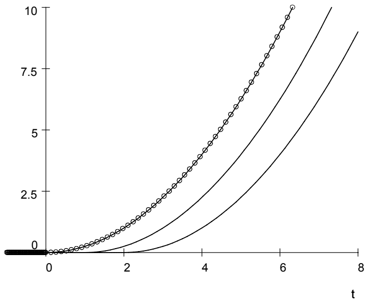

Fig. 15.1. Three different solutions to the ODE $\dot{y}(t) = |y(t)|^{1/2}$ with $y(0) = 0$.

$$f(t) \le \varepsilon(t) + \left| \int_0^t k(\tau)\varepsilon(\tau) e^{\left|\int_\tau^t k(s)ds\right|} d\tau \right|, \tag{15.9}$$

and in particular if ε and k are constants we find that

$$f(t) \le \varepsilon e^{k|t|}. \tag{15.10}$$

Proof. I will only prove the case $t \ge 0$. The case $t \le 0$ can be derived by applying the $t \ge 0$ to $\tilde{f}(t) = f(-t)$, $\tilde{k}(t) = k(-t)$ and $\varepsilon(t) = \varepsilon(-t)$. Set $F(t) = \int_0^t k(\tau) f(\tau) d\tau$. Then by (15.8),

$$\dot{F} = kf \le k\varepsilon + kF.$$

Hence,

$$\frac{d}{dt}(e^{-\int_0^t k(s)ds} F) = e^{-\int_0^t k(s)ds}(\dot{F} - kF) \le k\varepsilon e^{-\int_0^t k(s)ds}.$$

Integrating this last inequality from 0 to t and then solving for F yields:

$$F(t) \le e^{\int_0^t k(s)ds} \cdot \int_0^t d\tau k(\tau)\varepsilon(\tau) e^{-\int_0^\tau k(s)ds} = \int_0^t d\tau k(\tau)\varepsilon(\tau) e^{\int_\tau^t k(s)ds}.$$

But by the definition of F we have that

$$f \le \varepsilon + F,$$

and hence the last two displayed equations imply (15.9). Equation (15.10) follows from (15.9) by a simple integration. ■

Corollary 15.3 (Continuous Dependence on Initial Data). *Let $U \subset_o X$, $0 \in (a, b)$ and $Z : (a, b) \times U \to X$ be a continuous function which is $K-$Lipschitz function on U, i.e. $\|Z(t, x) - Z(t, x')\| \le K\|x - x'\|$ for all x and x' in U. Suppose $y_1, y_2 : (a, b) \to U$ solve*

$$\frac{dy_i(t)}{dt} = Z(t, y_i(t)) \quad \text{with } y_i(0) = x_i \quad \text{for } i = 1, 2. \tag{15.11}$$

Then

$$\|y_2(t) - y_1(t)\| \le \|x_2 - x_1\| e^{K|t|} \text{ for } t \in (a, b) \tag{15.12}$$

and in particular, there is at most one solution to Eq. (15.1) under the above Lipschitz assumption on Z.

Proof. Let $f(t) := \|y_2(t) - y_1(t)\|$. Then by the fundamental theorem of calculus,

$$\begin{aligned} f(t) &= \|y_2(0) - y_1(0) + \int_0^t (\dot{y}_2(\tau) - \dot{y}_1(\tau))\, d\tau\| \\ &\le f(0) + \left| \int_0^t \|Z(\tau, y_2(\tau)) - Z(\tau, y_1(\tau))\|\, d\tau \right| \\ &= \|x_2 - x_1\| + K \left| \int_0^t f(\tau)\, d\tau \right|. \end{aligned}$$

Therefore by Gronwall's inequality we have,

$$\|y_2(t) - y_1(t)\| = f(t) \le \|x_2 - x_1\| e^{K|t|}.$$

■

15.3 Local Existence (Non-Linear ODE)

We now show that Eq. (15.1) under a Lipschitz condition on Z. Another existence theorem was given in Exercise 11.16.

Theorem 15.4 (Local Existence). *Let $T > 0$, $J = (-T, T)$, $x_0 \in X$, $r > 0$ and*

$$C(x_0, r) := \{x \in X : \|x - x_0\| \le r\}$$

be the closed $r-$ball centered at $x_0 \in X$. Assume

$$M = \sup\{\|Z(t, x)\| : (t, x) \in J \times C(x_0, r)\} < \infty \tag{15.13}$$

and there exists $K < \infty$ such that

$$\|Z(t, x) - Z(t, y)\| \le K\|x - y\| \text{ for all } x, y \in C(x_0, r) \text{ and } t \in J. \tag{15.14}$$

Let $T_0 < \min\{r/M, T\}$ and $J_0 := (-T_0, T_0)$, then for each $x \in B(x_0, r - MT_0)$ there exists a unique solution $y(t) = y(t, x)$ to Eq. (15.2) in $C(J_0, C(x_0, r))$. Moreover $y(t, x)$ is jointly continuous in (t, x), $y(t, x)$ is differentiable in t, $\dot{y}(t, x)$ is jointly continuous for all $(t, x) \in J_0 \times B(x_0, r - MT_0)$ and satisfies Eq. (15.1).

Proof. The uniqueness assertion has already been proved in Corollary 15.3. To prove existence, let $C_r := C(x_0, r)$, $Y := C\left(J_0, C(x_0, r)\right)$ and

$$S_x(y)(t) := x + \int_0^t Z(\tau, y(\tau))d\tau. \tag{15.15}$$

With this notation, Eq. (15.2) becomes $y = S_x(y)$, i.e. we are looking for a fixed point of S_x. If $y \in Y$, then

$$\begin{aligned}\|S_x(y)(t) - x_0\| &\le \|x - x_0\| + \left|\int_0^t \|Z(\tau, y(\tau))\| \, d\tau\right| \le \|x - x_0\| + M\,|t| \\ &\le \|x - x_0\| + MT_0 \le r - MT_0 + MT_0 = r,\end{aligned}$$

showing $S_x\left(Y\right) \subset Y$ for all $x \in B(x_0, r - MT_0)$. Moreover if $y, z \in Y$,

$$\begin{aligned}\|S_x(y)(t) - S_x(z)(t)\| &= \left\|\int_0^t \left[Z(\tau, y(\tau)) - Z(\tau, z(\tau))\right] d\tau\right\| \\ &\le \left|\int_0^t \|Z(\tau, y(\tau)) - Z(\tau, z(\tau))\| \, d\tau\right| \\ &\le K\left|\int_0^t \|y(\tau) - z(\tau)\| \, d\tau\right|. \end{aligned} \tag{15.16}$$

Let $y_0(t, x) = x$ and $y_n(\cdot, x) \in Y$ defined inductively by

$$y_n(\cdot, x) := S_x(y_{n-1}(\cdot, x)) = x + \int_0^t Z(\tau, y_{n-1}(\tau, x))d\tau. \tag{15.17}$$

Using the estimate in Eq. (15.16) repeatedly we find

$$\begin{aligned}&||\ y_{n+1}(t) - y_n(t)\ || \\ &\quad\le K\left|\int_0^t \|y_n(\tau) - y_{n-1}(\tau)\| \, d\tau\right| \\ &\quad\le K^2\left|\int_0^t dt_1 \left|\int_0^{t_1} dt_2 \, \|y_{n-1}(t_2) - y_{n-2}(t_2)\|\right|\right| \\ &\quad\vdots \\ &\quad\le K^n\left|\int_0^t dt_1 \left|\int_0^{t_1} dt_2 \dots \left|\int_0^{t_{n-1}} dt_n \, \|y_1(t_n) - y_0(t_n)\|\right| \dots\right|\right| \\ &\quad\le K^n \|y_1(\cdot, x) - y_0(\cdot, x)\|_\infty \int_{\Delta_n(t)} d\tau \\ &\quad= \frac{K^n\,|t|^n}{n!} \|y_1(\cdot, x) - y_0(\cdot, x)\|_\infty \le 2r\frac{K^n\,|t|^n}{n!} \end{aligned} \tag{15.18}$$

wherein we have also made use of Lemma 8.19. Combining this estimate with

$$\|y_1(t, x) - y_0(t, x)\| = \left\|\int_0^t Z(\tau, x)d\tau\right\| \le \left|\int_0^t \|Z(\tau, x)\| \, d\tau\right| \le M_0,$$

where

$$M_0 = T_0 \max\left\{\int_0^{T_0} \|Z(\tau, x)\| \, d\tau, \int_{-T_0}^0 \|Z(\tau, x)\| \, d\tau\right\} \le MT_0,$$

shows

$$\|y_{n+1}(t, x) - y_n(t, x)\| \le M_0\frac{K^n\,|t|^n}{n!} \le M_0\frac{K^n T_0^n}{n!}$$

and this implies

$$\begin{aligned}&\sum_{n=0}^\infty \sup\{\,\|y_{n+1}(\cdot, x) - y_n(\cdot, x)\|_{\infty, J_0} : t \in J_0\} \\ &\qquad\le \sum_{n=0}^\infty M_0\frac{K^n T_0^n}{n!} = M_0 e^{KT_0} < \infty\end{aligned}$$

where

$$\|y_{n+1}(\cdot, x) - y_n(\cdot, x)\|_{\infty, J_0} := \sup\left\{\|y_{n+1}(t, x) - y_n(t, x)\| : t \in J_0\right\}.$$

So $y(t, x) := \lim_{n\to\infty} y_n(t, x)$ exists uniformly for $t \in J$ and using Eq. (15.14) we also have

$$\begin{aligned}&\sup\{\,\|Z(t, y(t)) - Z(t, y_{n-1}(t))\| : t \in J_0\} \\ &\qquad\le K\,\|y(\cdot, x) - y_{n-1}(\cdot, x)\|_{\infty, J_0} \to 0 \text{ as } n \to \infty.\end{aligned}$$

Now passing to the limit in Eq. (15.17) shows y solves Eq. (15.2). From this equation it follows that $y(t, x)$ is differentiable in t and y satisfies Eq. (15.1). The continuity of $y(t, x)$ follows from Corollary 15.3 and mean value inequality (Corollary 8.14):

$$\begin{aligned}\|y(t, x) - y(t', x')\| &\le \|y(t, x) - y(t, x')\| + \|y(t, x') - y(t', x')\| \\ &= \|y(t, x) - y(t, x')\| + \left\|\int_{t'}^t Z(\tau, y(\tau, x'))d\tau\right\| \\ &\le \|y(t, x) - y(t, x')\| + \left|\int_{t'}^t \|Z(\tau, y(\tau, x'))\| \, d\tau\right| \\ &\le \|x - x'\| e^{KT} + \left|\int_{t'}^t \|Z(\tau, y(\tau, x'))\| \, d\tau\right| \\ &\le \|x - x'\| e^{KT} + M\,|t - t'|\,. \end{aligned} \tag{15.19}$$

The continuity of $\dot{y}(t, x)$ is now a consequence Eq. (15.1) and the continuity of y and Z. ∎

Corollary 15.5. *Let $J = (a, b) \ni 0$ and suppose $Z \in C(J \times X, X)$ satisfies*

$$\|Z(t,x) - Z(t,y)\| \le K\|x-y\| \text{ for all } x, y \in X \text{ and } t \in J. \tag{15.20}$$

Then for all $x \in X$, there is a unique solution $y(t,x)$ (for $t \in J$) to Eq. (15.1). Moreover $y(t,x)$ and $\dot{y}(t,x)$ are jointly continuous in (t,x).

Proof. Let $J_0 = (a_0, b_0) \ni 0$ be a precompact subinterval of J and $Y := BC(J_0, X)$. By compactness, $M := \sup_{t\in\bar{J}_0} \|Z(t,0)\| < \infty$ which combined with Eq. (15.20) implies

$$\sup_{t\in\bar{J}_0} \|Z(t,x)\| \le M + K\|x\| \text{ for all } x \in X.$$

Using this estimate and Lemma 8.7 one easily shows $S_x(Y) \subset Y$ for all $x \in X$. The proof of Theorem 15.4 now goes through without any further change. ■

15.4 Global Properties

Definition 15.6 (Local Lipschitz Functions). *Let $U \subset_o X$, J be an open interval and $Z \in C(J\times U, X)$. The function Z is said to be locally Lipschitz in x if for all $x \in U$ and all compact intervals $I \subset J$ there exists $K = K(x, I) < \infty$ and $\varepsilon = \varepsilon(x, I) > 0$ such that $B(x, \varepsilon(x, I)) \subset U$ and*

$$\|Z(t,x_1) - Z(t,x_0)\| \le K(x,I)\|x_1 - x_0\| \ \forall\, x_0, x_1 \in B(x,\varepsilon(x,I)) \ \& \ t \in I. \tag{15.21}$$

For the rest of this section, we will assume J is an open interval containing 0, U is an open subset of X and $Z \in C(J\times U, X)$ is a locally Lipschitz function.

Lemma 15.7. *Let $Z \in C(J \times U, X)$ be a locally Lipschitz function in X and E be a compact subset of U and I be a compact subset of J. Then there exists $\varepsilon > 0$ such that $Z(t,x)$ is bounded for $(t,x) \in I \times E_\varepsilon$ and and $Z(t,x)$ is K – Lipschitz on E_ε for all $t \in I$, where*

$$E_\varepsilon := \{x \in U : \operatorname{dist}(x,E) < \varepsilon\}.$$

Proof. Let $\varepsilon(x, I)$ and $K(x, I)$ be as in Definition 15.6 above. Since E is compact, there exists a finite subset $\Lambda \subset E$ such that $E \subset V := \cup_{x\in\Lambda} B(x, \varepsilon(x, I)/2)$. If $y \in V$, there exists $x \in \Lambda$ such that $\|y - x\| < \varepsilon(x, I)/2$ and therefore

$$\begin{aligned}\|Z(t,y)\| &\le \|Z(t,x)\| + K(x,I)\|y-x\| \le \|Z(t,x)\| + K(x,I)\varepsilon(x,I)/2 \\ &\le \sup_{x\in\Lambda, t\in I} \{\|Z(t,x)\| + K(x,I)\varepsilon(x,I)/2\} =: M < \infty.\end{aligned}$$

This shows Z is bounded on $I \times V$. Let

$$\varepsilon := d(E, V^c) \le \frac{1}{2}\min_{x\in\Lambda} \varepsilon(x, I)$$

and notice that $\varepsilon > 0$ since E is compact, V^c is closed and $E \cap V^c = \emptyset$. If $y, z \in E_\varepsilon$ and $\|y - z\| < \varepsilon$, then as before there exists $x \in \Lambda$ such that $\|y - x\| < \varepsilon(x, I)/2$. Therefore

$$\|z-x\| \le \|z-y\| + \|y-x\| < \varepsilon + \varepsilon(x,I)/2 \le \varepsilon(x,I)$$

and since $y, z \in B(x, \varepsilon(x, I))$, it follows that

$$\|Z(t,y) - Z(t,z)\| \le K(x,I)\|y-z\| \le K_0\|y-z\|$$

where $K_0 := \max_{x\in\Lambda} K(x, I) < \infty$. On the other hand if $y, z \in E_\varepsilon$ and $\|y - z\| \ge \varepsilon$, then

$$\|Z(t,y) - Z(t,z)\| \le 2M \le \frac{2M}{\varepsilon}\|y-z\|.$$

Thus if we let $K := \max\{2M/\varepsilon, K_0\}$, we have shown

$$\|Z(t,y) - Z(t,z)\| \le K\|y-z\| \text{ for all } y, z \in E_\varepsilon \text{ and } t \in I.$$

■

Proposition 15.8 (Maximal Solutions). *Let $Z \in C(J \times U, X)$ be a locally Lipschitz function in x and let $x \in U$ be fixed. Then there is an interval $J_x = (a(x), b(x))$ with $a \in [-\infty, 0)$ and $b \in (0, \infty]$ and a C^1–function $y : J \to U$ with the following properties:*

1. *y solves ODE in Eq. (15.1).*
2. *If $\tilde{y} : \tilde{J} = (\tilde{a}, \tilde{b}) \to U$ is another solution of Eq. (15.1) (we assume that $0 \in \tilde{J}$) then $\tilde{J} \subset J$ and $\tilde{y} = y|_{\tilde{J}}$.*

The function $y : J \to U$ is called the maximal solution to Eq. (15.1).

Proof. Suppose that $y_i : J_i = (a_i, b_i) \to U$, $i = 1, 2$, are two solutions to Eq. (15.1). We will start by showing the $y_1 = y_2$ on $J_1 \cap J_2$. To do this[1] let

[1] Here is an alternate proof of the uniqueness. Let

$$T \equiv \sup\{t \in [0, \min\{b_1, b_2\}) : y_1 = y_2 \text{ on } [0,t]\}.$$

(T is the first positive time after which y_1 and y_2 disagree.

Suppose, for sake of contradiction, that $T < \min\{b_1, b_2\}$. Notice that $y_1(T) = y_2(T) =: x'$. Applying the local uniqueness theorem to $y_1(\cdot - T)$ and $y_2(\cdot - T)$ thought as function from $(-\delta, \delta) \to B(x', \epsilon(x'))$ for some δ sufficiently small, we learn that $y_1(\cdot - T) = y_2(\cdot - T)$ on $(-\delta, \delta)$. But this shows that $y_1 = y_2$ on $[0, T + \delta)$ which contradicts the definition of T. Hence we must have the $T = \min\{b_1, b_2\}$, i.e. $y_1 = y_2$ on $J_1 \cap J_2 \cap [0, \infty)$. A similar argument shows that $y_1 = y_2$ on $J_1 \cap J_2 \cap (-\infty, 0]$ as well.

$J_0 = (a_0, b_0)$ be chosen so that $0 \in J_0 \subset J_1 \cap J_2$, and let $E := y_1(J_0) \cup y_2(J_0)$ – a compact subset of X. Choose $\varepsilon > 0$ as in Lemma 15.7 so that Z is Lipschitz on E_ε. Then $y_1|_{J_0}, y_2|_{J_0} : J_0 \to E_\varepsilon$ both solve Eq. (15.1) and therefore are equal by Corollary 15.3. Since $J_0 = (a_0, b_0)$ was chosen arbitrarily so that $[a,b] \subset J_1 \cap J_2$, we may conclude that $y_1 = y_2$ on $J_1 \cap J_2$. Let $(y_\alpha, J_\alpha = (a_\alpha, b_\alpha))_{\alpha \in A}$ denote the possible solutions to (15.1) such that $0 \in J_\alpha$. Define $J_x = \cup J_\alpha$ and set $y = y_\alpha$ on J_α. We have just checked that y is well defined and the reader may easily check that this function $y : J_x \to U$ satisfies all the conclusions of the theorem. ■

Notation 15.9 *For each $x \in U$, let $J_x = (a(x), b(x))$ be the maximal interval on which Eq. (15.1) may be solved, see Proposition 15.8. Set $\mathcal{D}(Z) := \cup_{x \in U} (J_x \times \{x\}) \subset J \times U$ and let $\phi : \mathcal{D}(Z) \to U$ be defined by $\phi(t,x) = y(t)$ where y is the maximal solution to Eq. (15.1). (So for each $x \in U$, $\phi(\cdot, x)$ is the maximal solution to Eq. (15.1).)*

Proposition 15.10. *Let $Z \in C(J \times U, X)$ be a locally Lipschitz function in x and $y : J_x = (a(x), b(x)) \to U$ be the maximal solution to Eq. (15.1). If $b(x) < b$, then either $\limsup_{t \uparrow b(x)} \|Z(t, y(t))\| = \infty$ or $y(b(x)-) := \lim_{t \uparrow b(x)} y(t)$ exists and $y(b(x)-) \notin U$. Similarly, if $a > a(x)$, then either $\limsup_{t \downarrow a(x)} \|y(t)\| = \infty$ or $y(a(x)+) := \lim_{t \downarrow a} y(t)$ exists and $y(a(x)+) \notin U$.*

Proof. Suppose that $b < b(x)$ and $M := \limsup_{t \uparrow b(x)} \|Z(t, y(t))\| < \infty$. Then there is a $b_0 \in (0, b(x))$ such that $\|Z(t, y(t))\| \le 2M$ for all $t \in (b_0, b(x))$. Thus, by the usual fundamental theorem of calculus argument,

$$\|y(t) - y(t')\| \le \left| \int_t^{t'} \|Z(t, y(\tau))\| \, d\tau \right| \le 2M|t - t'|$$

for all $t, t' \in (b_0, b(x))$. From this it is easy to conclude that $y(b(x)-) = \lim_{t \uparrow b(x)} y(t)$ exists. If $y(b(x)-) \in U$, by the local existence Theorem 15.4, there exists $\delta > 0$ and $w \in C^1((b(x) - \delta, b(x) + \delta), U)$ such that

$$\dot{w}(t) = Z(t, w(t)) \text{ and } w(b(x)) = y(b(x)-).$$

Now define $\tilde{y} : (a, b(x) + \delta) \to U$ by

$$\tilde{y}(t) = \begin{cases} y(t) & \text{if } t \in J_x \\ w(t) & \text{if } t \in [b(x), b(x) + \delta) \end{cases}.$$

The reader may now easily show $\tilde{y}$ solves the integral Eq. (15.2) and hence also solves Eq. 15.1 for $t \in (a(x), b(x) + \delta)$.[2] But this violates the maximality of y and hence we must have that $y(b(x)-) \notin U$. The assertions for t near $a(x)$ are proved similarly. ■

[2] See the argument in Proposition 15.13 for a slightly different method of extending y which avoids the use of the integral equation (15.2).

Example 15.11. Let $X = \mathbb{R}^2$, $J = \mathbb{R}$, $U = \left\{(x,y) \in \mathbb{R}^2 : 0 < r < 1\right\}$ where $r^2 = x^2 + y^2$ and

$$Z(x,y) = \frac{1}{r}(x,y) + \frac{1}{1-r^2}(-y,x).$$

The the unique solution $(x(t), y(t))$ to

$$\frac{d}{dt}(x(t), y(t)) = Z(x(t), y(t)) \text{ with } (x(0), y(0)) = (\frac{1}{2}, 0)$$

is given by

$$(x(t), y(t)) = \left(t + \frac{1}{2}\right)\left(\cos\left(\frac{1}{1/2 - t}\right), \sin\left(\frac{1}{1/2 - t}\right)\right)$$

for $t \in J_{(1/2,0)} = (-\infty, 1/2)$. Notice that $\|Z(x(t), y(t))\| \to \infty$ as $t \uparrow 1/2$ and $\mathrm{dist}((x(t), y(t)), U^c) \to 0$ as $t \uparrow 1/2$.

Example 15.12. (Not worked out completely.) Let $X = U = \ell^2$, $\psi \in C^\infty(\mathbb{R}^2)$ be a smooth function such that $\psi = 1$ in a neighborhood of the line segment joining $(1,0)$ to $(0,1)$ and being supported within the $1/10$ – neighborhood of this segment. Choose $a_n \uparrow \infty$ and $b_n \uparrow \infty$ and define

$$Z(x) = \sum_{n=1}^{\infty} a_n \psi(b_n(x_n, x_{n+1}))(e_{n+1} - e_n). \tag{15.22}$$

For any $x \in \ell^2$, only a finite number of terms are non-zero in the above some in a neighborhood of x. Therefor $Z : \ell^2 \to \ell^2$ is a smooth and hence locally Lipshcitz vector field. Let $(y(t), J = (a,b))$ denote the maximal solution to

$$\dot{y}(t) = Z(y(t)) \text{ with } y(0) = e_1.$$

Then if the a_n and b_n are chosen appropriately, then $b < \infty$ and there will exist $t_n \uparrow b$ such that $y(t_n)$ is approximately e_n for all n. So again $y(t_n)$ does not have a limit yet $\sup_{t \in [0,b)} \|y(t)\| < \infty$. The idea is that Z is constructed to blow the particle form e_1 to e_2 to e_3 to e_4 etc. etc. with the time it takes to travel from e_n to e_{n+1} being on order $1/2^n$. The vector field in Eq. (15.22) is a first approximation at such a vector field, it may have to be adjusted a little more to provide an honest example. In this example, we are having problems because $y(t)$ is "going off in dimensions."

Here is another version of Proposition 15.10 which is more useful when $\dim(X) < \infty$.

Proposition 15.13. *Let $Z \in C(J \times U, X)$ be a locally Lipschitz function in x and $y : J_x = (a(x), b(x)) \to U$ be the maximal solution to Eq. (15.1).*

1. *If $b(x) < b$, then for every compact subset $K \subset U$ there exists $T_K < b(x)$ such that $y(t) \notin K$ for all $t \in [T_K, b(x))$.*

2. *When* $\dim(X) < \infty$, *we may write this condition as: if* $b(x) < b$, *then either*
$$\limsup_{t\uparrow b(x)} \|y(t)\| = \infty \text{ or } \liminf_{t\uparrow b(x)} dist(y(t), U^c) = 0.$$

Proof. 1) Suppose that $b(x) < b$ and, for sake of contradiction, there exists a compact set $K \subset U$ and $t_n \uparrow b(x)$ such that $y(t_n) \in K$ for all n. Since K is compact, by passing to a subsequence if necessary, we may assume $y_\infty := \lim_{n\to\infty} y(t_n)$ exists in $K \subset U$. By the local existence Theorem 15.4, there exists $T_0 > 0$ and $\delta > 0$ such that for each $x' \in B(y_\infty, \delta)$ there exists a unique solution $w(\cdot, x') \in C^1((-T_0, T_0), U)$ solving
$$w(t, x') = Z(t, w(t, x')) \text{ and } w(0, x') = x'.$$
Now choose n sufficiently large so that $t_n \in (b(x) - T_0/2, b(x))$ and $y(t_n) \in B(y_\infty, \delta)$. Define $\tilde{y} : (a(x), b(x) + T_0/2) \to U$ by
$$\tilde{y}(t) = \begin{cases} y(t) & \text{if } t \in J_x \\ w(t - t_n, y(t_n)) & \text{if } t \in (t_n - T_0, b(x) + T_0/2). \end{cases}$$
wherein we have used $(t_n - T_0, b(x) + T_0/2) \subset (t_n - T_0, t_n + T_0)$. By uniqueness of solutions to ODE's $\tilde{y}$ is well defined, $\tilde{y} \in C^1((a(x), b(x) + T_0/2), X)$ and $\tilde{y}$ solves the ODE in Eq. 15.1. But this violates the maximality of y. 2) For each $n \in \mathbb{N}$ let
$$K_n := \{x \in U : \|x\| \le n \text{ and } \operatorname{dist}(x, U^c) \ge 1/n\}.$$
Then $K_n \uparrow U$ and each K_n is a closed bounded set and hence compact if $\dim(X) < \infty$. Therefore if $b(x) < b$, by item 1., there exists $T_n \in [0, b(x))$ such that $y(t) \notin K_n$ for all $t \in [T_n, b(x))$ or equivalently $\|y(t)\| > n$ or $\operatorname{dist}(y(t), U^c) < 1/n$ for all $t \in [T_n, b(x))$. ∎

Remark 15.14. In general it is **not** true that the functions a and b are continuous. For example, let U be the region in $\mathbb{R}^2$ described in polar coordinates by $r > 0$ and $0 < \theta < 3\pi/4$ and $Z(x, y) = (0, -1)$ as in Figure 15.2 below. Then $b(x, y) = y$ for all $x, y > 0$ while $b(x, y) = \infty$ for all $x < 0$ and $y \in \mathbb{R}$ which shows b is discontinuous. On the other hand notice that
$$\{b > t\} = \{x < 0\} \cup \{(x, y) : x \ge 0, y > t\}$$
is an open set for all $t > 0$. An example of a vector field for which $b(x)$ is discontinuous is given in the top left hand corner of Figure 15.2. The map ψ would allow the reader to find an example on $\mathbb{R}^2$ if so desired. Some calculations shows that Z transferred to $\mathbb{R}^2$ by the map ψ is given by the new vector
$$\tilde{Z}(x, y) = -e^{-x}\left(\sin\left(\frac{3\pi}{8} + \frac{3}{4}\tan^{-1}(y)\right), \cos\left(\frac{3\pi}{8} + \frac{3}{4}\tan^{-1}(y)\right)\right).$$

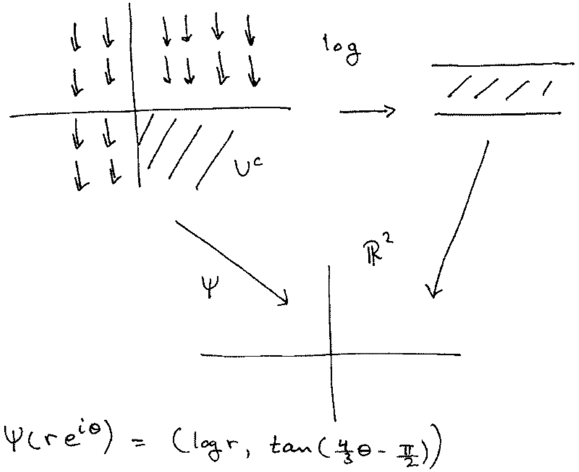

Fig. 15.2. Manufacturing vector fields where $b(x)$ is discontinuous.

Theorem 15.15 (Global Continuity). *Let* $Z \in C(J \times U, X)$ *be a locally Lipschitz function in* x. *Then* $\mathcal{D}(Z)$ *is an open subset of* $J \times U$ *and the functions* $\phi : \mathcal{D}(Z) \to U$ *and* $\dot{\phi} : \mathcal{D}(Z) \to U$ *are continuous. More precisely, for all* $x_0 \in U$ *and all open intervals* J_0 *such that* $0 \in J_0 \sqsubset\sqsubset J_{x_0}$ *there exists* $\delta = \delta(x_0, J_0, Z) > 0$ *and* $C = C(x_0, J_0, Z) < \infty$ *such that for all* $x \in B(x_0, \delta)$, $J_0 \subset J_x$ *and*
$$\|\phi(\cdot, x) - \phi(\cdot, x_0)\|_{BC(J_0, U)} \le C\|x - x_0\|. \tag{15.23}$$

Proof. Let $|J_0| = b_0 - a_0$, $I = \bar{J}_0$ and $E := y(\bar{J}_0)$ – a compact subset of U and let $\varepsilon > 0$ and $K < \infty$ be given as in Lemma 15.7, i.e. K is the Lipschitz constant for Z on E_ε. Also recall the notation: $\Delta_1(t) = [0, t]$ if $t > 0$ and $\Delta_1(t) = [t, 0]$ if $t < 0$. Suppose that $x \in E_\varepsilon$, then by Corollary 15.3,
$$\|\phi(t, x) - \phi(t, x_0)\| \le \|x - x_0\| e^{K|t|} \le \|x - x_0\| e^{K|J_0|} \tag{15.24}$$
for all $t \in J_0 \cap J_x$ such that such that $\phi(\Delta_1(t), x) \subset E_\varepsilon$. Letting $\delta := \varepsilon e^{-K|J_0|}/2$, and assuming $x \in B(x_0, \delta)$, the previous equation implies
$$\|\phi(t, x) - \phi(t, x_0)\| \le \varepsilon/2 < \varepsilon\ \forall\ t \in J_0 \cap J_x \ni \phi(\Delta_1(t), x) \subset E_\varepsilon.$$
This estimate further shows that $\phi(t, x)$ remains bounded and strictly away from the boundary of U for all such t. Therefore, it follows from Proposition 15.8 and "continuous induction[3]" that $J_0 \subset J_x$ and Eq. (15.24) is valid for all

[3] See the argument in the proof of Proposition 8.11.

$t \in J_0$. This proves Eq. (15.23) with $C := e^{K|J_0|}$. Suppose that $(t_0, x_0) \in \mathcal{D}(Z)$ and let $0 \in J_0 \sqsubset\sqsubset J_{x_0}$ such that $t_0 \in J_0$ and δ be as above. Then we have just shown $J_0 \times B(x_0, \delta) \subset \mathcal{D}(Z)$ which proves $\mathcal{D}(Z)$ is open. Furthermore, since the evaluation map

$$(t_0, y) \in J_0 \times BC(J_0, U) \xrightarrow{e} y(t_0) \in X$$

is continuous (as the reader should check) it follows that $\phi = e \circ (x \to \phi(\cdot, x)) : J_0 \times B(x_0, \delta) \to U$ is also continuous; being the composition of continuous maps. The continuity of $\dot{\phi}(t_0, x)$ is a consequences of the continuity of ϕ and the differential equation 15.1 Alternatively using Eq. (15.2),

$$\begin{aligned}\|\phi(t_0, x) - \phi(t, x_0)\| &\leq \|\phi(t_0, x) - \phi(t_0, x_0)\| + \|\phi(t_0, x_0) - \phi(t, x_0)\| \\ &\leq C\|x - x_0\| + \left|\int_t^{t_0} \|Z(\tau, \phi(\tau, x_0))\| \, d\tau\right| \\ &\leq C\|x - x_0\| + M|t_0 - t|\end{aligned}$$

where C is the constant in Eq. (15.23) and $M = \sup_{\tau \in J_0} \|Z(\tau, \phi(\tau, x_0))\| < \infty$. This clearly shows ϕ is continuous. ∎

15.5 Semi-Group Properties of time independent flows

To end this chapter we investigate the semi-group property of the flow associated to the vector-field Z. It will be convenient to introduce the following suggestive notation. For $(t, x) \in \mathcal{D}(Z)$, set $e^{tZ}(x) = \phi(t, x)$. So the path $t \to e^{tZ}(x)$ is the maximal solution to

$$\frac{d}{dt} e^{tZ}(x) = Z(e^{tZ}(x)) \text{ with } e^{0Z}(x) = x.$$

This exponential notation will be justified shortly. It is convenient to have the following conventions.

Notation 15.16 *We write $f : X \to X$ to mean a function defined on some open subset $D(f) \subset X$. The open set $D(f)$ will be called the domain of f. Given two functions $f : X \to X$ and $g : X \to X$ with domains $D(f)$ and $D(g)$ respectively, we define the composite function $f \circ g : X \to X$ to be the function with domain*

$$D(f \circ g) = \{x \in X : x \in D(g) \text{ and } g(x) \in D(f)\} = g^{-1}(D(f))$$

given by the rule $f \circ g(x) = f(g(x))$ for all $x \in D(f \circ g)$. We now write $f = g$ iff $D(f) = D(g)$ and $f(x) = g(x)$ for all $x \in D(f) = D(g)$. We will also write $f \subset g$ iff $D(f) \subset D(g)$ and $g|_{D(f)} = f$.

Theorem 15.17. *For fixed $t \in \mathbb{R}$ we consider e^{tZ} as a function from X to X with domain $D(e^{tZ}) = \{x \in U : (t, x) \in \mathcal{D}(Z)\}$, where $D(\phi) = \mathcal{D}(Z) \subset \mathbb{R} \times U$, $\mathcal{D}(Z)$ and ϕ are defined in Notation 15.9. Conclusions:*

1. *If $t, s \in \mathbb{R}$ and $t \cdot s \geq 0$, then $e^{tZ} \circ e^{sZ} = e^{(t+s)Z}$.*
2. *If $t \in \mathbb{R}$, then $e^{tZ} \circ e^{-tZ} = Id_{D(e^{-tZ})}$.*
3. *For arbitrary $t, s \in \mathbb{R}$, $e^{tZ} \circ e^{sZ} \subset e^{(t+s)Z}$.*

Proof. Item 1. For simplicity assume that $t, s \geq 0$. The case $t, s \leq 0$ is left to the reader. Suppose that $x \in D(e^{tZ} \circ e^{sZ})$. Then by assumption $x \in D(e^{sZ})$ and $e^{sZ}(x) \in D(e^{tZ})$. Define the path $y(\tau)$ via:

$$y(\tau) = \begin{cases} e^{\tau Z}(x) & \text{if } 0 \leq \tau \leq s \\ e^{(\tau - s)Z}(x) & \text{if } s \leq \tau \leq t + s \end{cases}.$$

It is easy to check that y solves $\dot{y}(\tau) = Z(y(\tau))$ with $y(0) = x$. But since, $e^{\tau Z}(x)$ is the maximal solution we must have that $x \in D(e^{(t+s)Z})$ and $y(t + s) = e^{(t+s)Z}(x)$. That is $e^{(t+s)Z}(x) = e^{tZ} \circ e^{sZ}(x)$. Hence we have shown that $e^{tZ} \circ e^{sZ} \subset e^{(t+s)Z}$. To finish the proof of item 1. it suffices to show that $D(e^{(t+s)Z}) \subset D(e^{tZ} \circ e^{sZ})$. Take $x \in D(e^{(t+s)Z})$, then clearly $x \in D(e^{sZ})$. Set $y(\tau) = e^{(\tau+s)Z}(x)$ defined for $0 \leq \tau \leq t$. Then y solves

$$\dot{y}(\tau) = Z(y(\tau)) \text{ with } y(0) = e^{sZ}(x).$$

But since $\tau \to e^{\tau Z}(e^{sZ}(x))$ is the maximal solution to the above initial valued problem we must have that $y(\tau) = e^{\tau Z}(e^{sZ}(x))$, and in particular at $\tau = t$, $e^{(t+s)Z}(x) = e^{tZ}(e^{sZ}(x))$. This shows that $x \in D(e^{tZ} \circ e^{sZ})$ and in fact $e^{(t+s)Z} \subset e^{tZ} \circ e^{sZ}$. Item 2. Let $x \in D(e^{-tZ})$ – again assume for simplicity that $t \geq 0$. Set $y(\tau) = e^{(\tau - t)Z}(x)$ defined for $0 \leq \tau \leq t$. Notice that $y(0) = e^{-tZ}(x)$ and $\dot{y}(\tau) = Z(y(\tau))$. This shows that $y(\tau) = e^{\tau Z}(e^{-tZ}(x))$ and in particular that $x \in D(e^{tZ} \circ e^{-tZ})$ and $e^{tZ} \circ e^{-tZ}(x) = x$. This proves item 2. Item 3. I will only consider the case that $s < 0$ and $t + s \geq 0$, the other cases are handled similarly. Write u for $t + s$, so that $t = -s + u$. We know that $e^{tZ} = e^{uZ} \circ e^{-sZ}$ by item 1. Therefore

$$e^{tZ} \circ e^{sZ} = (e^{uZ} \circ e^{-sZ}) \circ e^{sZ}.$$

Notice in general, one has $(f \circ g) \circ h = f \circ (g \circ h)$ (you prove). Hence, the above displayed equation and item 2. imply that

$$e^{tZ} \circ e^{sZ} = e^{uZ} \circ (e^{-sZ} \circ e^{sZ}) = e^{(t+s)Z} \circ I_{D(e^{sZ})} \subset e^{(t+s)Z}.$$

∎

The following result is trivial but conceptually illuminating partial converse to Theorem 15.17.

Proposition 15.18 (Flows and Complete Vector Fields). *Suppose $U \subset_o X$, $\phi \in C(\mathbb{R} \times U, U)$ and $\phi_t(x) = \phi(t, x)$. Suppose ϕ satisfies:*

1. $\phi_0 = I_U$,
2. $\phi_t \circ \phi_s = \phi_{t+s}$ for all $t, s \in \mathbb{R}$, and
3. $Z(x) := \dot{\phi}(0, x)$ exists for all $x \in U$ and $Z \in C(U, X)$ is locally Lipschitz.

Then $\phi_t = e^{tZ}$.

Proof. Let $x \in U$ and $y(t) := \phi_t(x)$. Then using Item 2.,

$$\dot{y}(t) = \frac{d}{ds}|_0 y(t+s) = \frac{d}{ds}|_0 \phi_{(t+s)}(x) = \frac{d}{ds}|_0 \phi_s \circ \phi_t(x) = Z(y(t)).$$

Since $y(0) = x$ by Item 1. and Z is locally Lipschitz by Item 3., we know by uniqueness of solutions to ODE's (Corollary 15.3) that $\phi_t(x) = y(t) = e^{tZ}(x)$. ■

15.6 Exercises

Exercise 15.1. Find a vector field Z such that $e^{(t+s)Z}$ is not contained in $e^{tZ} \circ e^{sZ}$.

Definition 15.19. *A locally Lipschitz function $Z : U \subset_o X \to X$ is said to be a complete vector field if $\mathcal{D}(Z) = \mathbb{R} \times U$. That is for any $x \in U$, $t \to e^{tZ}(x)$ is defined for all $t \in \mathbb{R}$.*

Exercise 15.2. Suppose that $Z : X \to X$ is a locally Lipschitz function. Assume there is a constant $C > 0$ such that

$$\|Z(x)\| \le C(1 + \|x\|) \text{ for all } x \in X.$$

Then Z is complete. **Hint:** use Gronwall's Lemma 15.2 and Proposition 15.10.

Exercise 15.3. Suppose y is a solution to $\dot{y}(t) = |y(t)|^{1/2}$ with $y(0) = 0$. Show there exists $a, b \in [0, \infty]$ such that

$$y(t) = \begin{cases} \frac{1}{4}(t-b)^2 & \text{if } \quad t \ge b \\ 0 & \text{if } -a < t < b \\ -\frac{1}{4}(t+a)^2 & \text{if } \quad t \le -a. \end{cases}$$

Exercise 15.4. Using the fact that the solutions to Eq. (15.3) are never 0 if $x \neq 0$, show that $y(t) = 0$ is the only solution to Eq. (15.3) with $y(0) = 0$.

Exercise 15.5 (Higher Order ODE). Let X be a Banach space, , $\mathcal{U} \subset_o X^n$ and $f \in C(J \times \mathcal{U}, X)$ be a Locally Lipschitz function in $\mathbf{x} = (x_1, \ldots, x_n)$. Show the n^{th} ordinary differential equation,

$$y^{(n)}(t) = f(t, y(t), \dot{y}(t), \ldots y^{(n-1)}(t)) \text{ with } y^{(k)}(0) = y_0^k \text{ for } k < n \tag{15.25}$$

where $(y_0^0, \ldots, y_0^{n-1})$ is given in $\mathcal{U}$, has a unique solution for small $t \in J$. **Hint**: let $\mathbf{y}(t) = \left(y(t), \dot{y}(t), \ldots y^{(n-1)}(t)\right)$ and rewrite Eq. (15.25) as a first order ODE of the form

$$\dot{\mathbf{y}}(t) = Z(t, \mathbf{y}(t)) \text{ with } \mathbf{y}(0) = (y_0^0, \ldots, y_0^{n-1}).$$

Exercise 15.6. Use the results of Exercises 8.20 and 15.5 to solve

$$\ddot{y}(t) - 2\dot{y}(t) + y(t) = 0 \text{ with } y(0) = a \text{ and } \dot{y}(0) = b.$$

Hint: The 2×2 matrix associated to this system, A, has only one eigenvalue 1 and may be written as $A = I + B$ where $B^2 = 0$.

Exercise 15.7 (Non-Homogeneous ODE). Suppose that $U \subset_o X$ is open and $Z : \mathbb{R} \times U \to X$ is a continuous function. Let $J = (a, b)$ be an interval and $t_0 \in J$. Suppose that $y \in C^1(J, U)$ is a solution to the "non-homogeneous" differential equation:

$$\dot{y}(t) = Z(t, y(t)) \text{ with } y(t_o) = x \in U. \tag{15.26}$$

Define $Y \in C^1(J - t_0, \mathbb{R} \times U)$ by $Y(t) := (t + t_0, y(t + t_0))$. Show that Y solves the "homogeneous" differential equation

$$\dot{Y}(t) = \tilde{Z}(Y(t)) \text{ with } Y(0) = (t_0, y_0), \tag{15.27}$$

where $\tilde{Z}(t, x) := (1, Z(x))$. Conversely, suppose that $Y \in C^1(J - t_0, \mathbb{R} \times U)$ is a solution to Eq. (15.27). Show that $Y(t) = (t + t_0, y(t + t_0))$ for some $y \in C^1(J, U)$ satisfying Eq. (15.26). (In this way the theory of non-homogeneous ode's may be reduced to the theory of homogeneous ode's.)

Exercise 15.8 (Differential Equations with Parameters). Let W be another Banach space, $U \times V \subset_o X \times W$ and $Z \in C(U \times V, X)$ be a locally Lipschitz function on $U \times V$. For each $(x, w) \in U \times V$, let $t \in J_{x,w} \to \phi(t, x, w)$ denote the maximal solution to the ODE

$$\dot{y}(t) = Z(y(t), w) \text{ with } y(0) = x. \tag{15.28}$$

Prove

$$\mathcal{D} := \{(t, x, w) \in \mathbb{R} \times U \times V : t \in J_{x,w}\} \tag{15.29}$$

is open in $\mathbb{R} \times U \times V$ and ϕ and $\dot{\phi}$ are continuous functions on $\mathcal{D}$.

Hint: If $y(t)$ solves the differential equation in (15.28), then $v(t) := (y(t), w)$ solves the differential equation,

$$\dot{v}(t) = \tilde{Z}(v(t)) \text{ with } v(0) = (x, w), \tag{15.30}$$

where $\tilde{Z}(x, w) := (Z(x, w), 0) \in X \times W$ and let $\psi(t, (x, w)) := v(t)$. Now apply the Theorem 15.15 to the differential equation (15.30).

Exercise 15.9 (Abstract Wave Equation). For $A \in L(X)$ and $t \in \mathbb{R}$, let

$$\cos(tA) := \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} t^{2n} A^{2n} \text{ and}$$
$$\frac{\sin(tA)}{A} := \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} t^{2n+1} A^{2n}.$$

Show that the unique solution $y \in C^2(\mathbb{R}, X)$ to

$$\ddot{y}(t) + A^2 y(t) = 0 \text{ with } y(0) = y_0 \text{ and } \dot{y}(0) = \dot{y}_0 \in X \tag{15.31}$$

is given by

$$y(t) = \cos(tA) y_0 + \frac{\sin(tA)}{A}\dot{y}_0.$$

Remark 15.20. Exercise 15.9 can be done by direct verification. Alternatively and more instructively, rewrite Eq. (15.31) as a first order ODE using Exercise 15.5. In doing so you will be lead to compute e^{tB} where $B \in L(X \times X)$ is given by

$$B = \begin{pmatrix} 0 & I \\ -A^2 & 0 \end{pmatrix},$$

where we are writing elements of $X \times X$ as column vectors, $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$. You should then show

$$e^{tB} = \begin{pmatrix} \cos(tA) & \frac{\sin(tA)}{A} \\ -A\sin(tA) & \cos(tA) \end{pmatrix}$$

where

$$A\sin(tA) := \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} t^{2n+1} A^{2(n+1)}.$$

Exercise 15.10 (Duhamel's Principle for the Abstract Wave Equation). Continue the notation in Exercise 15.9, but now consider the ODE,

$$\ddot{y}(t) + A^2 y(t) = f(t) \text{ with } y(0) = y_0 \text{ and } \dot{y}(0) = \dot{y}_0 \in X \tag{15.32}$$

where $f \in C(\mathbb{R}, X)$. Show the unique solution to Eq. (15.32) is given by

$$y(t) = \cos(tA) y_0 + \frac{\sin(tA)}{A}\dot{y}_0 + \int_0^t \frac{\sin((t-\tau)A)}{A} f(\tau) d\tau \tag{15.33}$$

Hint: Again this could be proved by direct calculation. However it is more instructive to deduce Eq. (15.33) from Exercise 8.22 and the comments in Remark 15.20.

16

Banach Space Calculus

In this section, X and Y will be Banach space and U will be an open subset of X.

Notation 16.1 (ε, O, and o notation) *Let $0 \in U \subset_o X$, and $f : U \to Y$ be a function. We will write:*

1. *$f(x) = \varepsilon(x)$ if $\lim_{x\to 0} \|f(x)\| = 0$.*
2. *$f(x) = O(x)$ if there are constants $C < \infty$ and $r > 0$ such that $\|f(x)\| \le C\|x\|$ for all $x \in B(0, r)$. This is equivalent to the condition that $\limsup_{x\to 0} \left(\|x\|^{-1}\|f(x)\|\right) < \infty$, where*

$$\limsup_{x\to 0} \frac{\|f(x)\|}{\|x\|} := \lim_{r\downarrow 0} \sup\{\|f(x)\| : 0 < \|x\| \le r\}.$$

3. *$f(x) = o(x)$ if $f(x) = \varepsilon(x)O(x)$, i.e. $\lim_{x\to 0} \|f(x)\|/\|x\| = 0$.*

Example 16.2. Here are some examples of properties of these symbols.

1. A function $f : U \subset_o X \to Y$ is continuous at $x_0 \in U$ if $f(x_0 + h) = f(x_0) + \varepsilon(h)$.
2. If $f(x) = \varepsilon(x)$ and $g(x) = \varepsilon(x)$ then $f(x) + g(x) = \varepsilon(x)$.
   Now let $g : Y \to Z$ be another function where Z is another Banach space.
3. If $f(x) = O(x)$ and $g(y) = o(y)$ then $g \circ f(x) = o(x)$.
4. If $f(x) = \varepsilon(x)$ and $g(y) = \varepsilon(y)$ then $g \circ f(x) = \varepsilon(x)$.

16.1 The Differential

Definition 16.3. *A function $f : U \subset_o X \to Y$ is **differentiable** at $x_0 \in U$ if there exists a linear transformation $\Lambda \in L(X, Y)$ such that*

$$f(x_0 + h) - f(x_0) - \Lambda h = o(h). \tag{16.1}$$

*We denote Λ by $f'(x_0)$ or $Df(x_0)$ if it exists. As with continuity, f is **differentiable on** U if f is differentiable at all points in U.*

Remark 16.4. The linear transformation Λ in Definition 16.3 is necessarily unique. Indeed if Λ_1 is another linear transformation such that Eq. (16.1) holds with Λ replaced by Λ_1, then

$$(\Lambda - \Lambda_1)h = o(h),$$

i.e.

$$\limsup_{h\to 0} \frac{\|(\Lambda - \Lambda_1)h\|}{\|h\|} = 0.$$

On the other hand, by definition of the operator norm,

$$\limsup_{h\to 0} \frac{\|(\Lambda - \Lambda_1)h\|}{\|h\|} = \|\Lambda - \Lambda_1\|.$$

The last two equations show that $\Lambda = \Lambda_1$.

Exercise 16.1. Show that a function $f : (a, b) \to X$ is a differentiable at $t \in (a, b)$ in the sense of Definition 8.8 iff it is differentiable in the sense of Definition 16.3. Also show $Df(t)v = v\dot{f}(t)$ for all $v \in \mathbb{R}$.

Example 16.5. If $T \in L(X, Y)$ and $x, h \in X$, then

$$T(x + h) - T(x) - Th = 0$$

which shows $T'(x) = T$ for all $x \in X$.

Example 16.6. Assume that $GL(X, Y)$ is non-empty. Then by Corollary 7.20, $GL(X, Y)$ is an open subset of $L(X, Y)$ and the inverse map $f : GL(X, Y) \to GL(Y, X)$, defined by $f(A) := A^{-1}$, is continuous. We will now show that f is differentiable and

$$f'(A)B = -A^{-1}BA^{-1} \text{ for all } B \in L(X, Y).$$

This is a consequence of the identity,

$$f(A + H) - f(A) = (A + H)^{-1}(A - (A + H))A^{-1} = -(A + H)^{-1}HA^{-1}$$

which may be used to find the estimate,

$$\begin{aligned}\left\|f(A + H) - f(A) + A^{-1}HA^{-1}\right\| &= \left\|\left[A^{-1} - (A + H)^{-1}\right]HA^{-1}\right\| \\ &\le \left\|A^{-1} - (A + H)^{-1}\right\|\|H\|\left\|A^{-1}\right\| \\ &\le \frac{\|A^{-1}\|^3\|H\|^2}{1 - \|A^{-1}\|\|H\|} = O\left(\|H\|^2\right)\end{aligned}$$

wherein we have used the bound in Eq. (7.8) of Corollary 7.20 for the last inequality.

16.2 Product and Chain Rules

The following theorem summarizes some basic properties of the differential.

Theorem 16.7. *The differential D has the following properties:*

1. ***Linearity:*** *D is linear, i.e. $D(f + \lambda g) = Df + \lambda Dg$.*
2. ***Product Rule:*** *If $f : U \subset_o X \to Y$ and $A : U \subset_o X \to L(X, Z)$ are differentiable at x_0 then so is $x \to (Af)(x) := A(x)f(x)$ and*
$$D(Af)(x_0)h = (DA(x_0)h)f(x_0) + A(x_0)Df(x_0)h.$$
3. ***Chain Rule:*** *If $f : U \subset_o X \to V \subset_o Y$ is differentiable at $x_0 \in U$, and $g : V \subset_o Y \to Z$ is differentiable at $y_0 := f(x_0)$, then $g \circ f$ is differentiable at x_0 and $(g \circ f)'(x_0) = g'(y_0)f'(x_0)$.*
4. ***Converse Chain Rule:*** *Suppose that $f : U \subset_o X \to V \subset_o Y$ is* ***continuous*** *at $x_0 \in U$, $g : V \subset_o Y \to Z$ is differentiable $y_0 := f(h_o)$, $g'(y_0)$ is invertible, and $g \circ f$ is differentiable at x_0, then f is differentiable at x_0 and*
$$f'(x_0) := [g'(x_0)]^{-1}(g \circ f)'(x_0). \tag{16.2}$$

Proof. Linearity. Let $f, g : U \subset_o X \to Y$ be two functions which are differentiable at $x_0 \in U$ and $\lambda \in \mathbb{R}$, then

$$\begin{aligned}&(f + \lambda g)(x_0 + h) \\ &\quad = f(x_0) + Df(x_0)h + o(h) + \lambda(g(x_0) + Dg(x_0)h + o(h) \\ &\quad = (f + \lambda g)(x_0) + (Df(x_0) + \lambda Dg(x_0))h + o(h),\end{aligned}$$

which implies that $(f + \lambda g)$ is differentiable at x_0 and that

$$D(f + \lambda g)(x_0) = Df(x_0) + \lambda Dg(x_0).$$

Product Rule. The computation,

$$\begin{aligned}&A(x_0 + h)f(x_0 + h) \\ &\quad = (A(x_0) + DA(x_0)h + o(h))(f(x_0) + f'(x_0)h + o(h)) \\ &\quad = A(x_0)f(x_0) + A(x_0)f'(x_0)h + [DA(x_0)h]f(x_0) + o(h),\end{aligned}$$

verifies the product rule holds. This may also be considered as a special case of Proposition 16.9. **Chain Rule.** Using $f(x_0 + h) - f(x_0) = O(h)$ (see Eq. (16.1)) and $o(O(h)) = o(h)$,

$$\begin{aligned}&(g\circ f)(x_0 + h) \\ &\quad = g(f(x_0)) + g'(f(x_0))(f(x_0 + h) - f(x_0)) + o(f(x_0 + h) - f(x_0)) \\ &\quad = g(f(x_0)) + g'(f(x_0))(Df(x_0)x_0 + o(h)) + o(f(x_0 + h) - f(x_0) \\ &\quad = g(f(x_0)) + g'(f(x_0))Df(x_0)h + o(h).\end{aligned}$$

Converse Chain Rule. Since g is differentiable at $y_0 = f(x_0)$ and $g'(y_0)$ is invertible,

$$\begin{aligned} &g(f(x_0+h)) - g(f(x_0)) \\ &\quad = g'(f(x_0))(f(x_0+h) - f(x_0)) + o(f(x_0+h) - f(x_0)) \\ &\quad = g'(f(x_0))\left[f(x_0+h) - f(x_0) + o(f(x_0+h) - f(x_0))\right]. \end{aligned}$$

And since $g \circ f$ is differentiable at x_0,

$$(g \circ f)(x_0+h) - g(f(x_0)) = (g \circ f)'(x_0)h + o(h).$$

Comparing these two equations shows that

$$\begin{aligned} &f(x_0+h) - f(x_0) + o(f(x_0+h) - f(x_0)) \\ &\qquad = g'(f(x_0))^{-1}\left[(g \circ f)'(x_0)h + o(h)\right] \end{aligned}$$

which is equivalent to

$$\begin{aligned} &f(x_0+h) - f(x_0) + o(f(x_0+h) - f(x_0)) \\ &\quad = g'(f(x_0))^{-1}\left[(g \circ f)'(x_0)h + o(h)\right] \\ &\quad = g'(f(x_0))^{-1}\{(g \circ f)'(x_0)h + o(h) - o(f(x_0+h) - f(x_0))\} \\ &\quad = g'(f(x_0))^{-1}(g \circ f)'(x_0)h + o(h) + o(f(x_0+h) - f(x_0)). \quad (16.3) \end{aligned}$$

Using the continuity of f, $f(x_0+h) - f(x_0)$ is close to 0 if h is close to zero, and hence

$$\|o(f(x_0+h) - f(x_0))\| \le \frac{1}{2}\|f(x_0+h) - f(x_0)\| \qquad (16.4)$$

for all h sufficiently close to 0. (We may replace $\frac{1}{2}$ by any number $\alpha > 0$ above.) Taking the norm of both sides of Eq. (16.3) and making use of Eq. (16.4) shows, for h close to 0, that

$$\begin{aligned} &\|f(x_0+h) - f(x_0)\| \\ &\quad \le \|g'(f(x_0))^{-1}(g \circ f)'(x_0)\|\|h\| + o(\|h\|) + \frac{1}{2}\|f(x_0+h) - f(x_0)\|. \end{aligned}$$

Solving for $\|f(x_0+h) - f(x_0)\|$ in this last equation shows that

$$f(x_0+h) - f(x_0) = O(h). \qquad (16.5)$$

(This is an improvement, since the continuity of f only guaranteed that $f(x_0+h) - f(x_0) = \varepsilon(h)$.) Because of Eq. (16.5), we now know that $o(f(x_0+h) - f(x_0)) = o(h)$, which combined with Eq. (16.3) shows that

$$f(x_0+h) - f(x_0) = g'(f(x_0))^{-1}(g \circ f)'(x_0)h + o(h),$$

i.e. f is differentiable at x_0 and $f'(x_0) = g'(f(x_0))^{-1}(g \circ f)'(x_0)$. ■

Corollary 16.8 (Chain Rule). *Suppose that $\sigma : (a,b) \to U \subset_o X$ is differentiable at $t \in (a,b)$ and $f : U \subset_o X \to Y$ is differentiable at $\sigma(t) \in U$. Then $f \circ \sigma$ is differentiable at t and*

$$d(f \circ \sigma)(t)/dt = f'(\sigma(t))\dot{\sigma}(t).$$

Proposition 16.9 (Product Rule II). *Suppose that $X := X_1 \times \cdots \times X_n$ with each X_i being a Banach space and $T : X_1 \times \cdots \times X_n \to Y$ is a multilinear map, i.e.*

$$x_i \in X_i \to T(x_1, \dots, x_{i-1}, x_i, x_{i+1}, \dots, x_n) \in Y$$

is linear when $x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n$ are held fixed. Then the following are equivalent:

1. *T is continuous.*
2. *T is continuous at $0 \in X$.*
3. *There exists a constant $C < \infty$ such that*
$$\|T(x)\|_Y \le C \prod_{i=1}^n \|x_i\|_{X_i} \qquad (16.6)$$
for all $x = (x_1, \dots, x_n) \in X$.
4. *T is differentiable at all $x \in X_1 \times \cdots \times X_n$.*

Moreover if T the differential of T is given by

$$T'(x)h = \sum_{i=1}^n T(x_1, \dots, x_{i-1}, h_i, x_{i+1}, \dots, x_n) \qquad (16.7)$$

where $h = (h_1, \dots, h_n) \in X$.

Proof. Let us equip X with the norm

$$\|x\|_X := \max\left\{\|x_i\|_{X_i}\right\}.$$

If T is continuous then T is continuous at 0. If T is continuous at 0, using $T(0) = 0$, there exists a $\delta > 0$ such that $\|T(x)\|_Y \le 1$ whenever $\|x\|_X \le \delta$. Now if $x \in X$ is arbitrary, let $x' := \delta\left(\|x_1\|_{X_1}^{-1} x_1, \dots, \|x_n\|_{X_n}^{-1} x_n\right)$. Then $\|x'\|_X \le \delta$ and hence

$$\left\|\left(\delta^n \prod_{i=1}^n \|x_i\|_{X_i}^{-1}\right) T(x_1, \dots, x_n)\right\|_Y = \|T(x')\| \le 1$$

from which Eq. (16.6) follows with $C = \delta^{-n}$.

Now suppose that Eq. (16.6) holds. For $x, h \in X$ and $\varepsilon \in \{0,1\}^n$ let $|\varepsilon| = \sum_{i=1}^n \varepsilon_i$ and

$$x^{\varepsilon}(h) := ((1-\varepsilon_1)x_1 + \varepsilon_1 h_1, \dots, (1-\varepsilon_n)x_n + \varepsilon_n h_n) \in X.$$

By the multi-linearity of T,

$$\begin{aligned} T(x+h) = T(x_1+h_1,\dots,x_n+h_n) &= \sum_{\varepsilon\in\{0,1\}^n} T(x^{\varepsilon}(h)) \\ &= T(x) + \sum_{i=1}^{n} T(x_1,\dots,x_{i-1},h_i,x_{i+1},\dots,x_n) \\ &\quad + \sum_{\varepsilon\in\{0,1\}^n : |\varepsilon|\ge 2} T(x^{\varepsilon}(h)). \end{aligned} \tag{16.8}$$

From Eq. (16.6),

$$\left\| \sum_{\varepsilon\in\{0,1\}^n : |\varepsilon|\ge 2} T(x^{\varepsilon}(h)) \right\| = O\left(\|h\|^2\right),$$

and so it follows from Eq. (16.8) that $T'(x)$ exists and is given by Eq. (16.7). This completes the proof since it is trivial to check that T being differentiable at $x \in X$ implies continuity of T at $x \in X$. ■

Exercise 16.2. Let $\det : L(\mathbb{R}^n) \to \mathbb{R}$ be the determinant function on $n \times n$ matrices and for $A \in L(\mathbb{R}^n)$ we will let A_i denote the i^{th} – column of A and write $A = (A_1|A_2|\dots|A_n)$.

1. Show $\det'(A)$ exists for all $A \in L(\mathbb{R}^n)$ and
$$\det{}'(A)H = \sum_{i=1}^{n} \det(A_1|\dots|A_{i-1}|H_i|A_{i+1}|\dots|A_n) \tag{16.9}$$
for all $H \in L(\mathbb{R}^n)$. **Hint**: recall that $\det(A)$ is a multilinear function of its columns.
2. Use Eq. (16.9) along with basic properties of the determinant to show $\det'(I)H = \operatorname{tr}(H)$.
3. Suppose now that $A \in GL(\mathbb{R}^n)$, show
$$\det{}'(A)H = \det(A)\operatorname{tr}(A^{-1}H).$$
Hint: Notice that $\det(A+H) = \det(A)\det\left(I + A^{-1}H\right)$.
4. If $A \in L(\mathbb{R}^n)$, show $\det\left(e^A\right) = e^{\operatorname{tr}(A)}$. **Hint:** use the previous item and Corollary 16.8 to show
$$\frac{d}{dt}\det\left(e^{tA}\right) = \det\left(e^{tA}\right)\operatorname{tr}(A).$$

Definition 16.10. *Let X and Y be Banach spaces and let $\mathcal{L}^1(X,Y) := L(X,Y)$ and for $k \ge 2$ let $\mathcal{L}^k(X,Y)$ be defined inductively by $\mathcal{L}^{k+1}(X,Y) = L(X,\mathcal{L}^k(X,Y))$. For example $\mathcal{L}^2(X,Y) = L(X,L(X,Y))$ and $\mathcal{L}^3(X,Y) = L(X,L(X,L(X,Y)))$.*

Suppose $f : U \subset_o X \to Y$ is a function. If f is differentiable on U, then it makes sense to ask if $f' = Df : U \to L(X,Y) = \mathcal{L}^1(X,Y)$ is differentiable. If Df is differentiable on U then $f'' = D^2 f := DDf : U \to \mathcal{L}^2(X,Y)$. Similarly we define $f^{(n)} = D^n f : U \to \mathcal{L}^n(X,Y)$ inductively.

Definition 16.11. *Given $k \in \mathbb{N}$, let $C^k(U,Y)$ denote those functions $f : U \to Y$ such that $f^{(j)} := D^j f : U \to \mathcal{L}^j(X,Y)$ exists and is continuous for $j = 1, 2, \dots, k$.*

Example 16.12. Let us continue on with Example 16.6 but now let $X = Y$ to simplify the notation. So $f : GL(X) \to GL(X)$ is the map $f(A) = A^{-1}$ and

$$f'(A) = -L_{A^{-1}}R_{A^{-1}}, \text{ i.e. } f' = -L_f R_f.$$

where $L_A B = AB$ and $R_A B = AB$ for all $A, B \in L(X)$. As the reader may easily check, the maps

$$A \in L(X) \to L_A, R_A \in L(L(X))$$

are linear and bounded. So by the chain and the product rule we find $f''(A)$ exists for all $A \in L(X)$ and

$$f''(A)B = -L_{f'(A)B}R_f - L_f R_{f'(A)B}.$$

More explicitly

$$[f''(A)B]\,C = A^{-1}BA^{-1}CA^{-1} + A^{-1}CA^{-1}BA^{-1}. \tag{16.10}$$

Working inductively one shows $f : GL(X) \to GL(X)$ defined by $f(A) := A^{-1}$ is C^{∞}.

16.3 Partial Derivatives

Definition 16.13 (Partial or Directional Derivative). *Let $f : U \subset_o X \to Y$ be a function, $x_0 \in U$, and $v \in X$. We say that f is differentiable at x_0 in the direction v iff $\frac{d}{dt}|_0(f(x_0+tv)) =: (\partial_v f)(x_0)$ exists. We call $(\partial_v f)(x_0)$ the directional or partial derivative of f at x_0 in the direction v.*

Notice that if f is differentiable at x_0, then $\partial_v f(x_0)$ exists and is equal to $f'(x_0)v$, see Corollary 16.8.

Proposition 16.14. *Let $f : U \subset_o X \to Y$ be a continuous function and $D \subset X$ be a dense subspace of X. Assume $\partial_v f(x)$ exists for all $x \in U$ and $v \in D$, and there exists a continuous function $A : U \to L(X,Y)$ such that $\partial_v f(x) = A(x)v$ for all $v \in D$ and $x \in U \cap D$. Then $f \in C^1(U,Y)$ and $Df = A$.*

Proof. Let $x_0 \in U$, $\varepsilon > 0$ such that $B(x_0, 2\varepsilon) \subset U$ and $M := \sup\{\|A(x)\| : x \in B(x_0, 2\varepsilon)\} < \infty$[1]. For $x \in B(x_0,\varepsilon) \cap D$ and $v \in D \cap B(0,\varepsilon)$, by the fundamental theorem of calculus,

$$\begin{aligned} f(x+v) - f(x) &= \int_0^1 \frac{df(x+tv)}{dt}\, dt \\ &= \int_0^1 (\partial_v f)(x+tv)\, dt = \int_0^1 A(x+tv)\, v\, dt. \end{aligned} \tag{16.11}$$

For general $x \in B(x_0,\varepsilon)$ and $v \in B(0,\varepsilon)$, choose $x_n \in B(x_0,\varepsilon) \cap D$ and $v_n \in D \cap B(0,\varepsilon)$ such that $x_n \to x$ and $v_n \to v$. Then

$$f(x_n + v_n) - f(x_n) = \int_0^1 A(x_n + tv_n)\, v_n\, dt \tag{16.12}$$

holds for all n. The left side of this last equation tends to $f(x+v) - f(x)$ by the continuity of f. For the right side of Eq. (16.12) we have

$$\begin{aligned} &\| \int_0^1 A(x+tv)\, v\, dt - \int_0^1 A(x_n + tv_n)\, v_n\, dt \| \\ &\leq \int_0^1 \|A(x+tv) - A(x_n+tv_n)\| \|v\|\, dt + M\|v - v_n\|. \end{aligned}$$

It now follows by the continuity of A, the fact that $\|A(x+tv) - A(x_n+tv_n)\| \leq M$, and the dominated convergence theorem that right side of Eq. (16.12) converges to $\int_0^1 A(x+tv)\, v\, dt$. Hence Eq. (16.11) is valid for all $x \in B(x_0,\varepsilon)$ and $v \in B(0,\varepsilon)$. We also see that

$$f(x+v) - f(x) - A(x)v = \varepsilon(v)v, \tag{16.13}$$

where $\varepsilon(v) := \int_0^1 [A(x+tv) - A(x)]\, dt$. Now

[1] It should be noted well, unlike in finite dimensions closed and bounded sets <u>need not</u> be compact, so it is not sufficient to choose ϵ sufficiently small so that $\overline{B(x_0, 2\epsilon)} \subset U$. Here is a counter example. Let $X \equiv H$ be a Hilbert space, $\{e_n\}_{n=1}^\infty$ be an orthonormal set. Define $f(x) \equiv \sum_{n=1}^\infty n\phi(\|x - e_n\|)$, where ϕ is any continuous function on $\mathbb{R}$ such that $\phi(0) = 1$ and ϕ is supported in $(-1,1)$. Notice that $\|e_n - e_m\|^2 = 2$ for all $m \neq n$, so that $\|e_n - e_m\| = \sqrt{2}$. Using this fact it is rather easy to check that for any $x_0 \in H$, there is an $\epsilon > 0$ such that for all $x \in B(x_0,\epsilon)$, only one term in the sum defining f is non-zero. Hence, f is continuous. However, $f(e_n) = n \to \infty$ as $n \to \infty$.

$$\begin{aligned} \|\varepsilon(v)\| &\leq \int_0^1 \|A(x+tv) - A(x)\|\, dt \\ &\leq \max_{t\in[0,1]} \|A(x+tv) - A(x)\| \to 0 \text{ as } v \to 0, \end{aligned}$$

by the continuity of A. Thus, we have shown that f is differentiable and that $Df(x) = A(x)$. ∎

Corollary 16.15. *Suppose now that $X = \mathbb{R}^d$, $f : U \subset_o X \to Y$ be a continuous function such that $\partial_i f(x) := \partial_{e_i} f(x)$ exists and is continuous on U for $i = 1, 2, \ldots, d$, where $\{e_i\}_{i=1}^d$ is the standard basis for $\mathbb{R}^d$. Then $f \in C^1(U,Y)$ and $Df(x)\, e_i = \partial_i f(x)$ for all i.*

Proof. For $x \in U$, let $A(x) : \mathbb{R}^d \to Y$ be the unique linear map such that $A(x)\, e_i = \partial_i f(x)$ for $i = 1, 2, \ldots, d$. Then $A : U \to L(\mathbb{R}^d, Y)$ is a continuous map. Now let $v \in \mathbb{R}^d$ and $v^{(i)} := (v_1, v_2, \ldots, v_i, 0, \ldots, 0)$ for $i = 1, 2, \ldots, d$ and $v^{(0)} := 0$. Then for $t \in \mathbb{R}$ near 0, using the fundamental theorem of calculus and the definition of $\partial_i f(x)$,

$$\begin{aligned} f(x+tv) - f(x) &= \sum_{i=1}^d \left[f\left(x + tv^{(i)}\right) - f\left(x + tv^{(i-1)}\right)\right] \\ &= \sum_{i=1}^d \int_0^1 \frac{d}{ds} f\left(x + tv^{(i-1)} + stv_i e_i\right) ds \\ &= \sum_{i=1}^d tv_i \int_0^1 \partial_i f\left(x + tv^{(i-1)} + stv_i e_i\right) ds \\ &= \sum_{i=1}^d tv_i \int_0^1 A\left(x + tv^{(i-1)} + stv_i e_i\right) e_i ds. \end{aligned}$$

Using the continuity of A, it now follows that

$$\begin{aligned} \lim_{t\to 0} \frac{f(x+tv) - f(x)}{t} &= \sum_{i=1}^d v_i \lim_{t\to 0} \int_0^1 A\left(x + tv^{(i-1)} + stv_i e_i\right) e_i ds \\ &= \sum_{i=1}^d v_i \int_0^1 A(x)\, e_i ds = A(x)\, v \end{aligned}$$

which shows $\partial_v f(x)$ exists and $\partial_v f(x) = A(x)\, v$. The result now follows from an application of Proposition 16.14. ∎

16.4 Higher Order Derivatives

It is somewhat inconvenient to work with the Banach spaces $\mathcal{L}^k(X,Y)$ in Definition 16.10. For this reason we will introduce an isomorphic Banach space, $M_k(X,Y)$.

Definition 16.16. *For $k \in \{1,2,3,\dots\}$, let $M_k(X,Y)$ denote the set of functions $f : X^k \to Y$ such that*

1. *For $i \in \{1,2,\dots,k\}$, $v \in X \to f\langle v_1, v_2, \dots, v_{i-1}, v, v_{i+1}, \dots, v_k\rangle \in Y$ is linear* [2] *for all $\{v_i\}_{i=1}^n \subset X$.*
2. *The norm $\|f\|_{M_k(X,Y)}$ should be finite, where*

$$\|f\|_{M_k(X,Y)} := \sup\{\frac{\|f\langle v_1, v_2, \dots, v_k\rangle\|_Y}{\|v_1\|\|v_2\|\cdots\|v_k\|} : \{v_i\}_{i=1}^k \subset X \setminus \{0\}\}.$$

Lemma 16.17. *There are linear operators $j_k : \mathcal{L}^k(X,Y) \to M_k(X,Y)$ defined inductively as follows: $j_1 = Id_{L(X,Y)}$ (notice that $M_1(X,Y) = \mathcal{L}^1(X,Y) = L(X,Y)$) and*

$$(j_{k+1}A)\langle v_0, v_1, \dots, v_k\rangle = (j_k(Av_0))\langle v_1, v_2, \dots, v_k\rangle \quad \forall v_i \in X.$$

(Notice that $Av_0 \in \mathcal{L}^k(X,Y)$.) Moreover, the maps j_k are isometric isomorphisms.

Proof. To get a feeling for what j_k is let us write out j_2 and j_3 explicitly. If $A \in \mathcal{L}^2(X,Y) = L(X, L(X,Y))$, then $(j_2A)\langle v_1, v_2\rangle = (Av_1)v_2$ and if $A \in \mathcal{L}^3(X,Y) = L(X, L(X, L(X,Y)))$, $(j_3A)\langle v_1, v_2, v_3\rangle = ((Av_1)v_2)v_3$ for all $v_i \in X$. It is easily checked that j_k is linear for all k. We will now show by induction that j_k is an isometry and in particular that j_k is injective. Clearly this is true if $k = 1$ since j_1 is the identity map. For $A \in \mathcal{L}^{k+1}(X,Y)$,

$$\begin{aligned}
&\|j_{k+1}A\|_{M_{k+1}(X,Y)} \\
&\quad := \sup\{\frac{\|(j_k(Av_0))\langle v_1, v_2, \dots, v_k\rangle\|_Y}{\|v_0\|\|v_1\|\|v_2\|\cdots\|v_k\|} : \{v_i\}_{i=0}^k \subset X \setminus \{0\}\} \\
&\quad = \sup\{\frac{\|(j_k(Av_0))\|_{M_k(X,Y)}}{\|v_0\|} : v_0 \in X \setminus \{0\}\} \\
&\quad = \sup\{\frac{\|Av_0\|_{\mathcal{L}^k(X,Y)}}{\|v_0\|} : v_0 \in X \setminus \{0\}\} \\
&\quad = \|A\|_{L(X,\mathcal{L}^k(X,Y))} := \|A\|_{\mathcal{L}^{k+1}(X,Y)},
\end{aligned}$$

wherein the second to last inequality we have used the induction hypothesis. This shows that j_{k+1} is an isometry provided j_k is an isometry. To finish the proof it suffices to show that j_k is surjective for all k. Again this is true for $k = 1$. Suppose that j_k is invertible for some $k \geq 1$. Given $f \in M_{k+1}(X,Y)$ we must produce $A \in \mathcal{L}^{k+1}(X,Y) = L(X, \mathcal{L}^k(X,Y))$ such that $j_{k+1}A = f$. If such an equation is to hold, then for $v_0 \in X$, we would have $j_k(Av_0) = f\langle v_0, \cdots\rangle$. That is $Av_0 = j_k^{-1}(f\langle v_0, \cdots\rangle)$. It is easily checked that A so defined is linear, bounded, and $j_{k+1}A = f$. ∎

From now on we will identify $\mathcal{L}^k$ with M_k without further mention. In particular, we will view $D^k f$ as function on U with values in $M_k(X,Y)$.

[2] I will routinely write $f\langle v_1, v_2, \dots, v_k\rangle$ rather than $f(v_1, v_2, \dots, v_k)$ when the function f depends on each of variables linearly, i.e. f is a multi-linear function.

Theorem 16.18 (Differentiability). *Suppose $k \in \{1,2,\dots\}$ and D is a dense subspace of X, $f : U \subset_o X \to Y$ is a function such that $(\partial_{v_1}\partial_{v_2}\cdots\partial_{v_l} f)(x)$ exists for all $x \in D \cap U$, $\{v_i\}_{i=1}^l \subset D$, and $l = 1, 2, \dots k$. Further assume there exists continuous functions $A_l : U \subset_o X \to M_l(X,Y)$ such that such that $(\partial_{v_1}\partial_{v_2}\cdots\partial_{v_l} f)(x) = A_l(x)\langle v_1, v_2, \dots, v_l\rangle$ for all $x \in D \cap U$, $\{v_i\}_{i=1}^l \subset D$, and $l = 1, 2, \dots k$. Then $D^l f(x)$ exists and is equal to $A_l(x)$ for all $x \in U$ and $l = 1, 2, \dots, k$.*

Proof. We will prove the theorem by induction on k. We have already proved the theorem when $k = 1$, see Proposition 16.14. Now suppose that $k > 1$ and that the statement of the theorem holds when k is replaced by $k-1$. Hence we know that $D^l f(x) = A_l(x)$ for all $x \in U$ and $l = 1, 2, \dots, k-1$. We are also given that

$$(\partial_{v_1}\partial_{v_2}\cdots\partial_{v_k} f)(x) = A_k(x)\langle v_1, v_2, \dots, v_k\rangle \quad \forall x \in U \cap D, \{v_i\} \subset D. \tag{16.14}$$

Now we may write $(\partial_{v_2}\cdots\partial_{v_k} f)(x)$ as $(D^{k-1}f)(x)\langle v_2, v_3, \dots, v_k\rangle$ so that Eq. (16.14) may be written as

$$\begin{aligned}
&\partial_{v_1}(D^{k-1}f)(x)\langle v_2, v_3, \dots, v_k\rangle) \\
&\qquad = A_k(x)\langle v_1, v_2, \dots, v_k\rangle \quad \forall x \in U \cap D, \{v_i\} \subset D.
\end{aligned} \tag{16.15}$$

So by the fundamental theorem of calculus, we have that

$$\begin{aligned}
&((D^{k-1}f)(x + v_1) - (D^{k-1}f)(x))\langle v_2, v_3, \dots, v_k\rangle \\
&\qquad = \int_0^1 A_k(x + tv_1)\langle v_1, v_2, \dots, v_k\rangle\, dt
\end{aligned} \tag{16.16}$$

for all $x \in U \cap D$ and $\{v_i\} \subset D$ with v_1 sufficiently small. By the same argument given in the proof of Proposition 16.14, Eq. (16.16) remains valid for all $x \in U$ and $\{v_i\} \subset X$ with v_1 sufficiently small. We may write this last equation alternatively as,

$$(D^{k-1}f)(x + v_1) - (D^{k-1}f)(x) = \int_0^1 A_k(x + tv_1)\langle v_1, \cdots\rangle\, dt. \tag{16.17}$$

Hence

$$\begin{aligned}
&(D^{k-1}f)(x + v_1) - (D^{k-1}f)(x) - A_k(x)\langle v_1, \cdots\rangle \\
&\qquad = \int_0^1 [A_k(x + tv_1) - A_k(x)]\langle v_1, \cdots\rangle\, dt
\end{aligned}$$

from which we get the estimate,

$$\|(D^{k-1}f)(x + v_1) - (D^{k-1}f)(x) - A_k(x)\langle v_1, \cdots\rangle\| \leq \varepsilon(v_1)\|v_1\| \tag{16.18}$$

where $\varepsilon(v_1) := \int_0^1 \|A_k(x + tv_1) - A_k(x)\|\, dt$. Notice by the continuity of A_k that $\varepsilon(v_1) \to 0$ as $v_1 \to 0$. Thus it follow from Eq. (16.18) that $D^{k-1}f$ is differentiable and that $(D^k f)(x) = A_k(x)$. ∎

Example 16.19. Let $f : GL(X,Y) \to GL(Y,X)$ be defined by $f(A) := A^{-1}$. We assume that $GL(X,Y)$ is not empty. Then f is infinitely differentiable and

$$(D^k f)(A)\langle V_1, V_2, \ldots, V_k\rangle = (-1)^k \sum_\sigma \{B^{-1}V_{\sigma(1)}B^{-1}V_{\sigma(2)}B^{-1}\cdots B^{-1}V_{\sigma(k)}B^{-1}\}, \quad (16.19)$$

where sum is over all permutations of σ of $\{1,2,\ldots,k\}$.

Let me check Eq. (16.19) in the case that $k=2$. Notice that we have already shown that $(\partial_{V_1} f)(B) = Df(B)V_1 = -B^{-1}V_1B^{-1}$. Using the product rule we find that

$$(\partial_{V_2}\partial_{V_1} f)(B) = B^{-1}V_2B^{-1}V_1B^{-1} + B^{-1}V_1B^{-1}V_2B^{-1} =: A_2(B)\langle V_1, V_2\rangle.$$

Notice that $\|A_2(B)\langle V_1, V_2\rangle\| \le 2\|B^{-1}\|^3\|V_1\|\cdot\|V_2\|$, so that $\|A_2(B)\| \le 2\|B^{-1}\|^3 < \infty$. Hence $A_2 : GL(X,Y) \to M_2(L(X,Y), L(Y,X))$. Also

$$\begin{aligned}
\|(A_2(B) - A_2(C))\langle V_1, V_2\rangle\| &\le 2\|B^{-1}V_2B^{-1}V_1B^{-1} - C^{-1}V_2C^{-1}V_1C^{-1}\| \\
&\le 2\|B^{-1}V_2B^{-1}V_1B^{-1} - B^{-1}V_2B^{-1}V_1C^{-1}\| \\
&\quad + 2\|B^{-1}V_2B^{-1}V_1C^{-1} - B^{-1}V_2C^{-1}V_1C^{-1}\| \\
&\quad\quad + 2\|B^{-1}V_2C^{-1}V_1C^{-1} - C^{-1}V_2C^{-1}V_1C^{-1}\| \\
&\le 2\|B^{-1}\|^2\|V_2\|\|V_1\|\|B^{-1} - C^{-1}\| \\
&\quad + 2\|B^{-1}\|\|C^{-1}\|\|V_2\|\|V_1\|\|B^{-1} - C^{-1}\| \\
&\quad\quad + 2\|C^{-1}\|^2\|V_2\|\|V_1\|\|B^{-1} - C^{-1}\|.
\end{aligned}$$

This shows that

$$\|A_2(B) - A_2(C)\| \le 2\|B^{-1} - C^{-1}\|\{\|B^{-1}\|^2 + \|B^{-1}\|\|C^{-1}\| + \|C^{-1}\|^2\}.$$

Since $B \to B^{-1}$ is differentiable and hence continuous, it follows that $A_2(B)$ is also continuous in B. Hence by Theorem 16.18 $D^2 f(A)$ exists and is given as in Eq. (16.19)

Example 16.20. Suppose that $f : \mathbb{R} \to \mathbb{R}$ is a C^∞– function and $F(x) := \int_0^1 f(x(t))\,dt$ for $x \in X := C([0,1], \mathbb{R})$ equipped with the norm $\|x\| := \max_{t\in[0,1]} |x(t)|$. Then $F : X \to \mathbb{R}$ is also infinitely differentiable and

$$(D^k F)(x)\langle v_1, v_2, \ldots, v_k\rangle = \int_0^1 f^{(k)}(x(t))v_1(t)\cdots v_k(t)\,dt, \quad (16.20)$$

for all $x \in X$ and $\{v_i\} \subset X$.

To verify this example, notice that

$$\begin{aligned}
(\partial_v F)(x) &:= \frac{d}{ds}|_0 F(x+sv) = \frac{d}{ds}|_0 \int_0^1 f(x(t)+sv(t))\,dt \\
&= \int_0^1 \frac{d}{ds}|_0 f(x(t)+sv(t))\,dt = \int_0^1 f'(x(t))v(t)\,dt.
\end{aligned}$$

Similar computations show that

$$(\partial_{v_1}\partial_{v_2}\cdots\partial_{v_k} f)(x) = \int_0^1 f^{(k)}(x(t))v_1(t)\cdots v_k(t)\,dt =: A_k(x)\langle v_1, v_2, \ldots, v_k\rangle.$$

Now for $x, y \in X$,

$$\begin{aligned}
|A_k(x)&\langle v_1, v_2, \ldots, v_k\rangle - A_k(y)\langle v_1, v_2, \ldots, v_k\rangle| \\
&\le \int_0^1 |f^{(k)}(x(t)) - f^{(k)}(y(t))|\cdot|v_1(t)\cdots v_k(t)|\,dt \\
&\le \prod_{i=1}^k \|v_i\| \int_0^1 |f^{(k)}(x(t)) - f^{(k)}(y(t))|dt,
\end{aligned}$$

which shows that

$$\|A_k(x) - A_k(y)\| \le \int_0^1 |f^{(k)}(x(t)) - f^{(k)}(y(t))|dt.$$

This last expression is easily seen to go to zero as $y \to x$ in X. Hence A_k is continuous. Thus we may apply Theorem 16.18 to conclude that Eq. (16.20) is valid.

16.5 Inverse and Implicit Function Theorems

In this section, let X be a Banach space, $R > 0$, $U = B = B(0,R) \subset X$ and $\varepsilon : U \to X$ be a continuous function such that $\varepsilon(0) = 0$. Our immediate goal is to give a sufficient condition on ε so that $F(x) := x + \varepsilon(x)$ is a homeomorphism from U to $F(U)$ with $F(U)$ being an open subset of X. Let's start by looking at the one dimensional case first. So for the moment assume that $X = \mathbb{R}$, $U = (-1,1)$, and $\varepsilon : U \to \mathbb{R}$ is C^1. Then F will be injective iff F is either strictly increasing or decreasing. Since we are thinking that F is a "small" perturbation of the identity function we will assume that F is strictly increasing, i.e. $F' = 1 + \varepsilon' > 0$. This positivity condition is not so easily interpreted for operators on a Banach space. However the condition that $|\varepsilon'| \le \alpha < 1$ is easily interpreted in the Banach space setting and it implies $1 + \varepsilon' > 0$.

Lemma 16.21. *Suppose that $U = B = B(0,R)$ $(R > 0)$ is a ball in X and $\varepsilon : B \to X$ is a C^1 function such that $\|D\varepsilon\| \le \alpha < \infty$ on U. Then*

$$\|\varepsilon(x) - \varepsilon(y)\| \le \alpha\|x - y\| \text{ for all } x, y \in U. \quad (16.21)$$

Proof. By the fundamental theorem of calculus and the chain rule:

$$\begin{aligned}\varepsilon(y)-\varepsilon(x) &= \int_0^1 \frac{d}{dt}\varepsilon(x+t(y-x))dt \\ &= \int_0^1 [D\varepsilon(x+t(y-x))](y-x)dt.\end{aligned}$$

Therefore, by the triangle inequality and the assumption that $\|D\varepsilon(x)\| \le \alpha$ on B,

$$\|\varepsilon(y)-\varepsilon(x)\| \le \int_0^1 \|D\varepsilon(x+t(y-x))\|dt \cdot \|(y-x)\| \le \alpha\|(y-x)\|.$$

■

Remark 16.22. It is easily checked that if $\varepsilon : U = B(0,R) \to X$ is C^1 and satisfies (16.21) then $\|D\varepsilon\| \le \alpha$ on U.

Using the above remark and the analogy to the one dimensional example, one is lead to the following proposition.

Proposition 16.23. *Suppose $\alpha \in (0,1)$, $R > 0$, $U = B(0,R) \subset_o X$ and $\varepsilon : U \to X$ is a continuous function such that $\varepsilon(0) = 0$ and*

$$\|\varepsilon(x)-\varepsilon(y)\| \le \alpha\|x-y\| \quad \forall\, x,y \in U. \tag{16.22}$$

Then $F : U \to X$ defined by $F(x) := x + \varepsilon(x)$ for $x \in U$ satisfies:

1. *F is an injective map and $G = F^{-1} : V := F(U) \to U$ is continuous.*
2. *If $x_0 \in U$, $z_0 = F(x_0)$ and $r > 0$ such the $B(x_0,r) \subset U$, then*
$$B(z_0,(1-\alpha)r) \subset F(B(x_0,r)) \subset B(z_0,(1+\alpha)r). \tag{16.23}$$
In particular, for all $r \le R$,
$$B(0,(1-\alpha)r) \subset F(B(0,r)) \subset B(0,(1+\alpha)r), \tag{16.24}$$
see Figure 16.1 below.
3. *$V := F(U)$ is open subset of X and $F : U \to V$ is a homeomorphism.*

Proof.

1. Using the definition of F and the estimate in Eq. (16.22),
$$\begin{aligned}\|x-y\| &= \|(F(x)-F(y))-(\varepsilon(x)-\varepsilon(y))\| \\ &\le \|F(x)-F(y)\| + \|\varepsilon(x)-\varepsilon(y)\| \\ &\le \|F(x)-F(y)\| + \alpha\|(x-y)\|\end{aligned}$$
for all $x,y \in U$. This implies

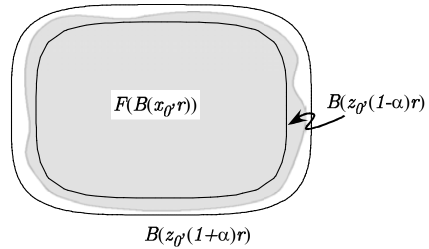

Fig. 16.1. Nesting of $F(B(x_0,r))$ between $B(z_0,(1-\alpha)r)$ and $B(z_0,(1+\alpha)r)$.

$$\|x-y\| \le (1-\alpha)^{-1}\|F(x)-F(y)\| \tag{16.25}$$

which shows F is injective on U and hence shows the inverse function $G = F^{-1} : V := F(U) \to U$ is well defined. Moreover, replacing x,y in Eq. (16.25) by $G(x)$ and $G(y)$ respectively with $x,y \in V$ shows

$$\|G(x)-G(y)\| \le (1-\alpha)^{-1}\|x-y\| \text{ for all } x,y \in V. \tag{16.26}$$

Hence G is Lipschitz on V and hence continuous.

2. Let $x_0 \in U$, $r > 0$ and $z_0 = F(x_0) = x_0 + \varepsilon(x_0)$ be as in item 2. The second inclusion in Eq. (16.23) follows from the simple computation:
$$\begin{aligned}\|F(x_0+h)-z_0\| &= \|h+\varepsilon(x_0+h)-\varepsilon(x_0)\| \\ &\le \|h\| + \|\varepsilon(x_0+h)-\varepsilon(x_0)\| \\ &\le (1+\alpha)\|h\| < (1+\alpha)r\end{aligned}$$
for all $h \in B(0,r)$. To prove the first inclusion in Eq. (16.23) we must find, for every $z \in B(z_0,(1-\alpha)r)$, an $h \in B(0,r)$ such that $z = F(x_0+h)$ or equivalently an $h \in B(0,r)$ solving
$$z - z_0 = F(x_0+h) - F(x_0) = h + \varepsilon(x_0+h) - \varepsilon(x_0).$$
Let $k := z - z_0$ and for $h \in B(0,r)$, let $\delta(h) := \varepsilon(x_0+h) - \varepsilon(x_0)$. With this notation it suffices to show for each $k \in B(z_0,(1-\alpha)r)$ there exists $h \in B(0,r)$ such that $k = h + \delta(h)$. Notice that $\delta(0) = 0$ and
$$\|\delta(h_1)-\delta(h_2)\| = \|\varepsilon(x_0+h_1)-\varepsilon(x_0+h_2)\| \le \alpha\|h_1-h_2\| \tag{16.27}$$
for all $h_1,h_2 \in B(0,r)$. We are now going to solve the equation $k = h + \delta(h)$ for h by the method of successive approximations starting with $h_0 = 0$ and then defining h_n inductively by
$$h_{n+1} = k - \delta(h_n). \tag{16.28}$$
A simple induction argument using Eq. (16.27) shows that

$$\|h_{n+1} - h_n\| \le \alpha^n \|k\| \text{ for all } n \in \mathbb{N}_0$$

and in particular that

$$\begin{aligned}\|h_N\| &= \left\|\sum_{n=0}^{N-1}(h_{n+1}-h_n)\right\| \le \sum_{n=0}^{N-1}\|h_{n+1}-h_n\| \\ &\le \sum_{n=0}^{N-1}\alpha^n \|k\| = \frac{1-\alpha^N}{1-\alpha}\|k\|. \qquad (16.29)\end{aligned}$$

Since $\|k\| < (1-\alpha) r$, this implies that $\|h_N\| < r$ for all N showing the approximation procedure is well defined. Let

$$h := \lim_{N\to\infty} h_n = \sum_{n=0}^{\infty}(h_{n+1}-h_n) \in X$$

which exists since the sum in the previous equation is absolutely convergent. Passing to the limit in Eqs. (16.29) and (16.28) shows that $\|h\| \le (1-\alpha)^{-1}\|k\| < r$ and $h = k - \delta(h)$, i.e. $h \in B(0,r)$ solves $k = h + \delta(h)$ as desired.

3. Given $x_0 \in U$, the first inclusion in Eq. (16.23) shows that $z_0 = F(x_0)$ is in the interior of $F(U)$. Since $z_0 \in F(U)$ was arbitrary, it follows that $V = F(U)$ is open. The continuity of the inverse function has already been proved in item 1.

■

For the remainder of this section let X and Y be two Banach spaces, $U \subset_o X$, $k \ge 1$, and $f \in C^k(U,Y)$.

Lemma 16.24. *Suppose $x_0 \in U$, $R > 0$ is such that $B^X(x_0,R) \subset U$ and $T : B^X(x_0,R) \to Y$ is a C^1 – function such that $T'(x_0)$ is invertible. Let*

$$\alpha(R) := \sup_{x\in B^X(x_0,R)} \left\|T'(x_0)^{-1}T'(x) - I\right\|_{L(X)} \qquad (16.30)$$

and $\varepsilon \in C^1\left(B^X(0,R), X\right)$ be defined by

$$\varepsilon(h) = T'(x_0)^{-1}[T(x_0+h) - T(x_0)] - h \qquad (16.31)$$

so that

$$T(x_0+h) = T(x_0) + T'(x_0)(h + \varepsilon(h)). \qquad (16.32)$$

Then $\varepsilon(h) = o(h)$ as $h \to 0$ and

$$\|\varepsilon(h') - \varepsilon(h)\| \le \alpha(R)\|h'-h\| \text{ for all } h,h' \in B^X(0,R). \qquad (16.33)$$

If $\alpha(R) < 1$ (which may be achieved by shrinking R if necessary), then $T'(x)$ is invertible for all $x \in B^X(x_0,R)$ and

$$\sup_{x\in B^X(x_0,R)} \left\|T'(x)^{-1}\right\|_{L(Y,X)} \le \frac{1}{1-\alpha(R)}\left\|T'(x_0)^{-1}\right\|_{L(Y,X)}. \qquad (16.34)$$

Proof. By definition of $T'(x_0)$ and using $T'(x_0)^{-1}$ exists,

$$T(x_0+h) - T(x_0) = T'(x_0)h + o(h)$$

from which it follows that $\varepsilon(h) = o(h)$. In fact by the fundamental theorem of calculus,

$$\varepsilon(h) = \int_0^1 \left(T'(x_0)^{-1}T'(x_0+th) - I\right) h dt$$

but we will not use this here. Let $h, h' \in B^X(0,R)$ and apply the fundamental theorem of calculus to $t \to T(x_0 + t(h'-h))$ to conclude

$$\begin{aligned}\varepsilon(h') - \varepsilon(h) &= T'(x_0)^{-1}[T(x_0+h') - T(x_0+h)] - (h'-h) \\ &= \left[\int_0^1 \left(T'(x_0)^{-1}T'(x_0+t(h'-h)) - I\right) dt\right](h'-h).\end{aligned}$$

Taking norms of this equation gives

$$\begin{aligned}\|\varepsilon(h') - \varepsilon(h)\| &\le \left[\int_0^1 \left\|T'(x_0)^{-1}T'(x_0+t(h'-h)) - I\right\| dt\right]\|h'-h\| \\ &\le \alpha(R)\|h'-h\|\end{aligned}$$

It only remains to prove Eq. (16.34), so suppose now that $\alpha(R) < 1$. Then by Proposition 7.19, $T'(x_0)^{-1}T'(x) = I - \left(I - T'(x_0)^{-1}T'(x)\right)$ is invertible and

$$\left\|\left[T'(x_0)^{-1}T'(x)\right]^{-1}\right\| \le \frac{1}{1-\alpha(R)} \text{ for all } x \in B^X(x_0,R).$$

Since $T'(x) = T'(x_0)\left[T'(x_0)^{-1}T'(x)\right]$ this implies $T'(x)$ is invertible and

$$\left\|T'(x)^{-1}\right\| = \left\|\left[T'(x_0)^{-1}T'(x)\right]^{-1}T'(x_0)^{-1}\right\| \le \frac{1}{1-\alpha(R)}\left\|T'(x_0)^{-1}\right\|$$

for all $x \in B^X(x_0,R)$. ■

Theorem 16.25 (Inverse Function Theorem). *Suppose $U \subset_o X$, $k \ge 1$ and $T \in C^k(U,Y)$ such that $T'(x)$ is invertible for all $x \in U$. Further assume $x_0 \in U$ and $R > 0$ such that $B^X(x_0,R) \subset U$.*

1. For all $r \le R$,

$$T(B^X(x_0,r)) \subset T(x_0) + T'(x_0)B^X(0,(1+\alpha(r))r). \qquad (16.35)$$

2. If we further assume that

$$\alpha(R) := \sup_{x\in B^X(x_0,R)} \left\|T'(x_0)^{-1}T'(x) - I\right\| < 1,$$

which may always be achieved by taking R sufficiently small, then

$$T(x_0) + T'(x_0)B^X(0,(1-\alpha(r))r) \subset T(B^X(x_0,r)) \qquad (16.36)$$

for all $r \le R$, see Figure 16.2.

3. *$T : U \to Y$ is an open mapping, in particular $V := T(U) \subset_o Y$.*
4. *Again if R is sufficiently small so that $\alpha(R) < 1$, then $T|_{B^X(x_0,R)} : B^X(x_0,R) \to T(B^X(x_0,R))$ is invertible and $T|^{-1}_{B^X(x_0,R)} : T\left(B^X(x_0,R)\right) \to B^X(x_0,R)$ is a C^k – map.*
5. *If T is injective, then $T^{-1} : V \to U$ is also a C^k – map and*

$$\left(T^{-1}\right)'(y) = \left[T'(T^{-1}(y))\right]^{-1} \text{ for all } y \in V.$$

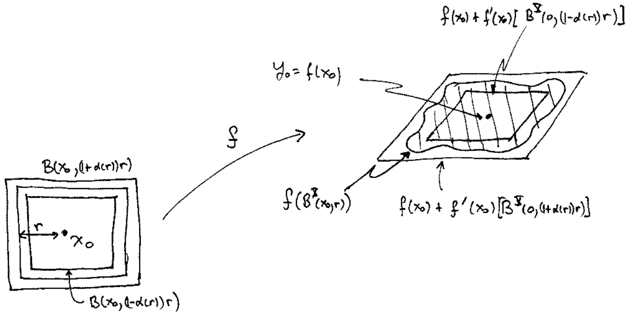

Fig. 16.2. The nesting of $T(B^X(x_0,r))$ between $T(x_0)+T'(x_0)\,B^X(0,(1-\alpha(r))r)$ and$T(x_0)+T'(x_0)\,B^X(0,(1+\alpha(r))r)$.

Proof. Let $\varepsilon \in C^1\left(B^X(0,R),X\right)$ be as defined in Eq. (16.31).

1. Using Eqs. (16.32) and (16.24),
$$\begin{aligned} T\left(B^X(x_0,r)\right) &= T(x_0) + T'(x_0)(I+\varepsilon)\left(B^X(0,r)\right) \\ &\subset T(x_0) + T'(x_0)\,B^X(0,(1+\alpha(r))\,r) \end{aligned} \tag{16.37}$$
which proves Eq. (16.35).
2. Now assume $\alpha(R) < 1$, then by Eqs. (16.37) and (16.24),
$$\begin{aligned} &T(x_0) + T'(x_0)\,B^X(0,(1-\alpha(r))\,r) \\ &\quad \subset T(x_0) + T'(x_0)(I+\varepsilon)\left(B^X(0,r)\right) = T\left(B^X(x_0,r)\right) \end{aligned}$$
which proves Eq. (16.36).
3. Notice that $h \in X \to T(x_0) + T'(x_0)\,h \in Y$ is a homeomorphism. The fact that T is an open map follows easily from Eq. (16.36) which shows that $T(x_0)$ is interior of $T(W)$ for any $W \subset_o X$ with $x_0 \in W$.
4. The fact that $T|_{B^X(x_0,R)} : B^X(x_0,R) \to T(B^X(x_0,R))$ is invertible with a continuous inverse follows from Eq. (16.32) and Proposition 16.23. It now follows from the converse to the chain rule, Theorem 16.7, that $g := T|^{-1}_{B^X(x_0,R)} : T\left(B^X(x_0,R)\right) \to B^X(x_0,R)$ is differentiable and
$$g'(y) = \left[T'(g(y))\right]^{-1} \text{ for all } y \in T\left(B^X(x_0,R)\right).$$
This equation shows g is C^1. Now suppose that $k \geq 2$. Since $T' \in C^{k-1}(B, L(X))$ and $i(A) := A^{-1}$ is a smooth map by Example 16.19, $g' = i \circ T' \circ g$ is C^1, i.e. g is C^2. If $k \geq 2$, we may use the same argument to now show g is C^3. Continuing this way inductively, we learn g is C^k.
5. Since differentiability and smoothness is local, the assertion in item 5. follows directly from what has already been proved.

■

Theorem 16.26 (Implicit Function Theorem). *Suppose that X, Y, and W are three Banach spaces, $k \geq 1$, $A \subset X \times Y$ is an open set, (x_0, y_0) is a point in A, and $f : A \to W$ is a C^k – map such $f(x_0,y_0) = 0$. Assume that $D_2 f(x_0,y_0) := D(f(x_0,\cdot))(y_0) : Y \to W$ is a bounded invertible linear transformation. Then there is an open neighborhood U_0 of x_0 in X such that for all connected open neighborhoods U of x_0 contained in U_0, there is a unique continuous function $u : U \to Y$ such that $u(x_0) = y_o$, $(x,u(x)) \in A$ and $f(x,u(x)) = 0$ for all $x \in U$. Moreover u is necessarily C^k and*

$$Du(x) = -D_2 f(x,u(x))^{-1} D_1 f(x,u(x)) \text{ for all } x \in U. \tag{16.38}$$

Proof. By replacing f by $(x,y) \to D_2 f(x_0,y_0)^{-1} f(x,y)$ if necessary, we may assume with out loss of generality that $W = Y$ and $D_2 f(x_0,y_0) = I_Y$. Define $F : A \to X \times Y$ by $F(x,y) := (x, f(x,y))$ for all $(x,y) \in A$. Notice that

$$DF(x,y) = \begin{bmatrix} I & D_1 f(x,y) \\ 0 & D_2 f(x,y) \end{bmatrix}$$

which is invertible iff $D_2 f(x,y)$ is invertible and if $D_2 f(x,y)$ is invertible then

$$DF(x,y)^{-1} = \begin{bmatrix} I & -D_1 f(x,y) D_2 f(x,y)^{-1} \\ 0 & D_2 f(x,y)^{-1} \end{bmatrix}.$$

Since $D_2 f(x_0,y_0) = I$ is invertible, the inverse function theorem guarantees that there exists a neighborhood U_0 of x_0 and V_0 of y_0 such that $U_0 \times V_0 \subset A$, $F(U_0 \times V_0)$ is open in $X \times Y$, $F|_{(U_0 \times V_0)}$ has a C^k–inverse which we call F^{-1}. Let $\pi_2(x,y) := y$ for all $(x,y) \in X \times Y$ and define C^k – function u_0 on U_0 by $u_0(x) := \pi_2 \circ F^{-1}(x,0)$. Since $F^{-1}(x,0) = (\tilde{x}, u_0(x))$ iff

$$(x,0) = F(\tilde{x}, u_0(x)) = (\tilde{x}, f(\tilde{x}, u_0(x))),$$

it follows that $x = \tilde{x}$ and $f(x,u_0(x)) = 0$. Thus

$$(x, u_0(x)) = F^{-1}(x,0) \in U_0 \times V_0 \subset A$$

and $f(x, u_0(x)) = 0$ for all $x \in U_0$. Moreover, u_0 is C^k being the composition of the C^k– functions, $x \to (x, 0)$, F^{-1}, and π_2. So if $U \subset U_0$ is a connected set containing x_0, we may define $u := u_0|_U$ to show the existence of the functions u as described in the statement of the theorem. The only statement left to prove is the uniqueness of such a function u. Suppose that $u_1 : U \to Y$ is another continuous function such that $u_1(x_0) = y_0$, and $(x, u_1(x)) \in A$ and $f(x, u_1(x)) = 0$ for all $x \in U$. Let

$$O := \{x \in U | u(x) = u_1(x)\} = \{x \in U | u_0(x) = u_1(x)\}.$$

Clearly O is a (relatively) closed subset of U which is not empty since $x_0 \in O$. Because U is connected, if we show that O is also an open set we will have shown that $O = U$ or equivalently that $u_1 = u_0$ on U. So suppose that $x \in O$, i.e. $u_0(x) = u_1(x)$. For $\tilde{x}$ near $x \in U$,

$$0 = 0 - 0 = f(\tilde{x}, u_0(\tilde{x})) - f(\tilde{x}, u_1(\tilde{x})) = R(\tilde{x})(u_1(\tilde{x}) - u_0(\tilde{x})) \qquad (16.39)$$

where

$$R(\tilde{x}) := \int_0^1 D_2 f((\tilde{x}, u_0(\tilde{x}) + t(u_1(\tilde{x}) - u_0(\tilde{x})))dt. \qquad (16.40)$$

From Eq. (16.40) and the continuity of u_0 and u_1, $\lim_{\tilde{x} \to x} R(\tilde{x}) = D_2 f(x, u_0(x))$ which is invertible.[3] Thus $R(\tilde{x})$ is invertible for all $\tilde{x}$ sufficiently close to x which combined with Eq. (16.39) implies that $u_1(\tilde{x}) = u_0(\tilde{x})$ for all $\tilde{x}$ sufficiently close to x. Since $x \in O$ was arbitrary, we have shown that O is open. ∎

16.6 Smooth Dependence of ODE's on Initial Conditions*

In this subsection, let X be a Banach space, $U \subset_o X$ and J be an open interval with $0 \in J$.

Lemma 16.27. *If $Z \in C(J \times U, X)$ such that $D_x Z(t, x)$ exists for all $(t, x) \in J \times U$ and $D_x Z(t, x) \in C(J \times U, X)$ then Z is locally Lipschitz in x, see Definition 15.6.*

Proof. Suppose $I \sqsubset\sqsubset J$ and $x \in U$. By the continuity of DZ, for every $t \in I$ there an open neighborhood N_t of $t \in I$ and $\varepsilon_t > 0$ such that $B(x, \varepsilon_t) \subset U$ and

$$\sup \{\|D_x Z(t', x')\| : (t', x') \in N_t \times B(x, \varepsilon_t)\} < \infty.$$

By the compactness of I, there exists a finite subset $\Lambda \subset I$ such that $I \subset \cup_{t \in I} N_t$. Let $\varepsilon(x, I) := \min \{\varepsilon_t : t \in \Lambda\}$ and

[3] Notice that $DF(x, u_0(x))$ is invertible for all $x \in U_0$ since $F|_{U_0 \times V_0}$ has a C^1 inverse. Therefore $D_2 f(x, u_0(x))$ is also invertible for all $x \in U_0$.

$$K(x, I) := \sup \{\|DZ(t, x')\| (t, x') \in I \times B(x, \varepsilon(x, I))\} < \infty.$$

Then by the fundamental theorem of calculus and the triangle inequality,

$$\begin{aligned}\|Z(t, x_1) - Z(t, x_0)\| &\le \left(\int_0^1 \|D_x Z(t, x_0 + s(x_1 - x_0)\| \, ds\right) \|x_1 - x_0\| \\ &\le K(x, I) \|x_1 - x_0\|\end{aligned}$$

for all $x_0, x_1 \in B(x, \varepsilon(x, I))$ and $t \in I$. ∎

Theorem 16.28 (Smooth Dependence of ODE's on Initial Conditions). *Let X be a Banach space, $U \subset_o X$, $Z \in C(\mathbb{R} \times U, X)$ such that $D_x Z \in C(\mathbb{R} \times U, X)$ and $\phi : \mathcal{D}(Z) \subset \mathbb{R} \times X \to X$ denote the maximal solution operator to the ordinary differential equation*

$$\dot{y}(t) = Z(t, y(t)) \text{ with } y(0) = x \in U, \qquad (16.41)$$

see Notation 15.9 and Theorem 15.15. Then $\phi \in C^1(\mathcal{D}(Z), U)$, $\partial_t D_x \phi(t, x)$ exists and is continuous for $(t, x) \in \mathcal{D}(Z)$ and $D_x \phi(t, x)$ satisfies the linear differential equation,

$$\frac{d}{dt} D_x \phi(t, x) = [(D_x Z)(t, \phi(t, x))] D_x \phi(t, x) \text{ with } D_x \phi(0, x) = I_X \qquad (16.42)$$

for $t \in J_x$.

Proof. Let $x_0 \in U$ and J be an open interval such that $0 \in J \subset \bar{J} \sqsubset\sqsubset J_{x_0}$, $y_0 := y(\cdot, x_0)|_J$ and

$$\mathcal{O}_\varepsilon := \{y \in BC(J, U) : \|y - y_0\|_\infty < \varepsilon\} \subset_o BC(J, X).$$

By Lemma 16.27, Z is locally Lipschitz and therefore Theorem 15.15 is applicable. By Eq. (15.23) of Theorem 15.15, there exists $\varepsilon > 0$ and $\delta > 0$ such that $G : B(x_0, \delta) \to \mathcal{O}_\varepsilon$ defined by $G(x) := \phi(\cdot, x)|_J$ is continuous. By Lemma 16.29 below, for $\varepsilon > 0$ sufficiently small the function $F : \mathcal{O}_\varepsilon \to BC(J, X)$ defined by

$$F(y) := y - \int_0^{\cdot} Z(t, y(t))dt. \qquad (16.43)$$

is C^1 and

$$DF(y)v = v - \int_0^{\cdot} D_y Z(t, y(t))v(t)dt. \qquad (16.44)$$

By the existence and uniqueness Theorem 8.21 for linear ordinary differential equations, $DF(y)$ is invertible for any $y \in BC(J, U)$. By the definition of ϕ, $F(G(x)) = h(x)$ for all $x \in B(x_0, \delta)$ where $h : X \to BC(J, X)$ is defined by $h(x)(t) = x$ for all $t \in J$, i.e. $h(x)$ is the constant path at x. Since h is a bounded linear map, h is smooth and $Dh(x) = h$ for all $x \in X$.

We may now apply the converse to the chain rule in Theorem 16.7 to conclude $G \in C^1(B(x_0,\delta),\mathcal{O})$ and $DG(x) = [DF(G(x))]^{-1}Dh(x)$ or equivalently, $DF(G(x))DG(x) = h$ which in turn is equivalent to

$$D_x\phi(t,x) - \int_0^t [DZ(\phi(\tau,x)]D_x\phi(\tau,x)\,d\tau = I_X.$$

As usual this equation implies $D_x\phi(t,x)$ is differentiable in t, $D_x\phi(t,x)$ is continuous in (t,x) and $D_x\phi(t,x)$ satisfies Eq. (16.42). ∎

Lemma 16.29. *Continuing the notation used in the proof of Theorem 16.28 and further let*

$$f(y) := \int_0^{\cdot} Z(\tau, y(\tau))\,d\tau \text{ for } y \in \mathcal{O}_\varepsilon.$$

Then $f \in C^1(\mathcal{O}_\varepsilon, Y)$ and for all $y \in \mathcal{O}_\varepsilon$,

$$f'(y)h = \int_0^{\cdot} D_xZ(\tau,y(\tau))h(\tau)\,d\tau =: \Lambda_y h.$$

Proof. Let $h \in Y$ be sufficiently small and $\tau \in J$, then by fundamental theorem of calculus,

$$\begin{aligned} Z(\tau,&y(\tau)+h(\tau)) - Z(\tau,y(\tau)) \\ &= \int_0^1 [D_xZ(\tau,y(\tau)+rh(\tau)) - D_xZ(\tau,y(\tau))]dr \end{aligned}$$

and therefore,

$$\begin{aligned} f(y+h) &- f(y) - \Lambda_y h(t) \\ &= \int_0^t [Z(\tau,y(\tau)+h(\tau)) - Z(\tau,y(\tau)) - D_xZ(\tau,y(\tau))h(\tau)]\,d\tau \\ &= \int_0^t d\tau \int_0^1 dr[D_xZ(\tau,y(\tau)+rh(\tau)) - D_xZ(\tau,y(\tau))]h(\tau). \end{aligned}$$

Therefore,

$$\|(f(y+h) - f(y) - \Lambda_y h)\|_\infty \le \|h\|_\infty \delta(h) \tag{16.45}$$

where

$$\delta(h) := \int_J d\tau \int_0^1 dr\,\|D_xZ(\tau,y(\tau)+rh(\tau)) - D_xZ(\tau,y(\tau))\|.$$

With the aide of Lemmas 16.27 and Lemma 15.7,

$$(r,\tau,h) \in [0,1]\times J \times Y \to \|D_xZ(\tau,y(\tau)+rh(\tau))\|$$

is bounded for small h provided $\varepsilon > 0$ is sufficiently small. Thus it follows from the dominated convergence theorem that $\delta(h) \to 0$ as $h \to 0$ and hence Eq. (16.45) implies $f'(y)$ exists and is given by Λ_y. Similarly,

$$\begin{aligned} &\|f'(y+h) - f'(y)\|_{op} \\ &\quad \le \int_J \|D_xZ(\tau,y(\tau)+h(\tau)) - D_xZ(\tau,y(\tau))\|\,d\tau \to 0 \text{ as } h \to 0 \end{aligned}$$

showing f' is continuous. ∎

Remark 16.30. If $Z \in C^k(U,X)$, then an inductive argument shows that $\phi \in C^k(\mathcal{D}(Z),X)$. For example if $Z \in C^2(U,X)$ then $(y(t),u(t)) := (\phi(t,x), D_x\phi(t,x))$ solves the ODE,

$$\frac{d}{dt}(y(t),u(t)) = \tilde{Z}\left((y(t),u(t))\right) \text{ with } (y(0),u(0)) = (x, Id_X)$$

where $\tilde{Z}$ is the C^1 – vector field defined by

$$\tilde{Z}(x,u) = (Z(x), D_xZ(x)u).$$

Therefore Theorem 16.28 may be applied to this equation to deduce: $D_x^2\phi(t,x)$ and $D_x^2\dot{\phi}(t,x)$ exist and are continuous. We may now differentiate Eq. (16.42) to find $D_x^2\phi(t,x)$ satisfies the ODE,

$$\begin{aligned} \frac{d}{dt}D_x^2\phi(t,x) &= [\left(\partial_{D_x\phi(t,x)}D_xZ\right)(t,\phi(t,x))]D_x\phi(t,x) \\ &\quad + [(D_xZ)(t,\phi(t,x))]D_x^2\phi(t,x) \end{aligned}$$

with $D_x^2\phi(0,x) = 0$.

16.7 Existence of Periodic Solutions

A detailed discussion of the inverse function theorem on Banach and Frechét spaces may be found in Richard Hamilton's, "The Inverse Function Theorem of Nash and Moser." The applications in this section are taken from this paper. In what follows we say $f \in C^k_{2\pi}(\mathbb{R},(c,d))$ if $f \in C^k_{2\pi}(\mathbb{R},(c,d))$ and f is 2π – periodic, i.e. $f(x+2\pi) = f(x)$ for all $x \in \mathbb{R}$.

Theorem 16.31 (Taken from Hamilton, p. 110.). *Let $p : U := (a,b) \to V := (c,d)$ be a smooth function with $p' > 0$ on (a,b). For every $g \in C^\infty_{2\pi}(\mathbb{R},(c,d))$ there exists a unique function $y \in C^\infty_{2\pi}(\mathbb{R},(a,b))$ such that*

$$\dot{y}(t) + p(y(t)) = g(t).$$

Proof. Let $\tilde{V} := C^0_{2\pi}(\mathbb{R},(c,d)) \subset_o C^0_{2\pi}(\mathbb{R},\mathbb{R})$ and $\tilde{U} \subset_o C^1_{2\pi}(\mathbb{R},(a,b))$ be given by

$$\tilde{U} := \left\{y \in C^1_{2\pi}(\mathbb{R},\mathbb{R}) : a < y(t) < b \;\&\; c < \dot{y}(t) + p(y(t)) < d \;\forall\; t\right\}.$$

The proof will be completed by showing $P : \tilde{U} \to \tilde{V}$ defined by

$$P(y)(t) = \dot{y}(t) + p(y(t)) \text{ for } y \in \tilde{U} \text{ and } t \in \mathbb{R}$$

is bijective. Note that if $P(y)$ is smooth then so is y.

Step 1. The differential of P is given by $P'(y)h = \dot{h} + p'(y)h$, see Exercise 16.8. We will now show that the linear mapping $P'(y)$ is invertible. Indeed let $f = p'(y) > 0$, then the general solution to the Eq. $\dot{h} + fh = k$ is given by

$$h(t) = e^{-\int_0^t f(\tau)d\tau} h_0 + \int_0^t e^{-\int_\tau^t f(s)ds} k(\tau) d\tau$$

where h_0 is a constant. We wish to choose h_0 so that $h(2\pi) = h_0$, i.e. so that

$$h_0 \left(1 - e^{-c(f)}\right) = \int_0^{2\pi} e^{-\int_\tau^t f(s)ds} k(\tau) d\tau$$

where

$$c(f) = \int_0^{2\pi} f(\tau) d\tau = \int_0^{2\pi} p'(y(\tau)) d\tau > 0.$$

The unique solution $h \in C^1_{2\pi}(\mathbb{R}, \mathbb{R})$ to $P'(y)h = k$ is given by

$$\begin{aligned} h(t) &= \left(1 - e^{-c(f)}\right)^{-1} e^{-\int_0^t f(\tau)d\tau} \int_0^{2\pi} e^{-\int_\tau^t f(s)ds} k(\tau) d\tau + \int_0^t e^{-\int_\tau^t f(s)ds} k(\tau) d\tau \\ &= \left(1 - e^{-c(f)}\right)^{-1} e^{-\int_0^t f(s)ds} \int_0^{2\pi} e^{-\int_\tau^t f(s)ds} k(\tau) d\tau + \int_0^t e^{-\int_\tau^t f(s)ds} k(\tau) d\tau. \end{aligned}$$

Therefore $P'(y)$ is invertible for all y. Hence by the inverse function Theorem 16.25, $P : \tilde{U} \to \tilde{V}$ is an open mapping which is locally invertible.

Step 2. Let us now prove $P : \tilde{U} \to \tilde{V}$ is injective. For this suppose $y_1, y_2 \in \tilde{U}$ such that $P(y_1) = g = P(y_2)$ and let $z = y_2 - y_1$. Since

$$\dot{z}(t) + p(y_2(t)) - p(y_1(t)) = g(t) - g(t) = 0,$$

if $t_m \in \mathbb{R}$ is point where $z(t_m)$ takes on its maximum, then $\dot{z}(t_m) = 0$ and hence

$$p(y_2(t_m)) - p(y_1(t_m)) = 0.$$

Since p is increasing this implies $y_2(t_m) = y_1(t_m)$ and hence $z(t_m) = 0$. This shows $z(t) \leq 0$ for all t and a similar argument using a minimizer of z shows $z(t) \geq 0$ for all t. So we conclude $y_1 = y_2$.

Step 3. Let $W := P(\tilde{U})$, we wish to show $W = \tilde{V}$. By step 1., we know W is an open subset of $\tilde{V}$ and since $\tilde{V}$ is connected, to finish the proof it suffices to show W is relatively closed in $\tilde{V}$. So suppose $y_j \in \tilde{U}$ such that $g_j := P(y_j) \to g \in \tilde{V}$. We must now show $g \in W$, i.e. $g = P(y)$ for some $y \in W$. If t_m is a maximizer of y_j, then $\dot{y}_j(t_m) = 0$ and hence $g_j(t_m) = p(y_j(t_m)) < d$ and therefore $y_j(t_m) < b$ because p is increasing. A similar argument works for the minimizers then allows us to conclude $\operatorname{Ran} p \circ y_j) \subset \operatorname{Ran} g_j) \sqsubset\sqsubset (c, d)$

for all j. Since g_j is converging uniformly to g, there exists $c < \gamma < \delta < d$ such that $\operatorname{Ran}(p \circ y_j) \subset \operatorname{Ran}(g_j) \subset [\gamma, \delta]$ for all j. Again since $p' > 0$,

$$\operatorname{Ran}(y_j) \subset p^{-1}([\gamma, \delta]) = [\alpha, \beta] \sqsubset\sqsubset (a, b) \text{ for all } j.$$

In particular $\sup\{|\dot{y}_j(t)| : t \in \mathbb{R} \text{ and } j\} < \infty$ since

$$\dot{y}_j(t) = g_j(t) - p(y_j(t)) \subset [\gamma, \delta] - [\gamma, \delta] \tag{16.46}$$

which is a compact subset of $\mathbb{R}$. The Ascoli-Arzela Theorem 11.29 now allows us to assume, by passing to a subsequence if necessary, that y_j is converging uniformly to $y \in C^0_{2\pi}(\mathbb{R}, [\alpha, \beta])$. It now follows that

$$\dot{y}_j(t) = g_j(t) - p(y_j(t)) \to g - p(y)$$

uniformly in t. Hence we concluded that $y \in C^1_{2\pi}(\mathbb{R}, \mathbb{R}) \cap C^0_{2\pi}(\mathbb{R}, [\alpha, \beta])$, $\dot{y}_j \to y$ and $P(y) = g$. This has proved that $g \in W$ and hence that W is relatively closed in $\tilde{V}$. ■

16.8 Contraction Mapping Principle

Some of the arguments uses in this chapter and in Chapter 15 may be abstracted to a general principle of finding fixed points on a complete metric space. This is the content of this chapter.

Theorem 16.32. *Suppose that (X, ρ) is a complete metric space and $S : X \to X$ is a contraction, i.e. there exists $\alpha \in (0, 1)$ such that $\rho(S(x), S(y)) \leq \alpha\rho(x, y)$ for all $x, y \in X$. Then S has a unique fixed point in X, i.e. there exists a unique point $x \in X$ such that $S(x) = x$.*

Proof. For uniqueness suppose that x and x' are two fixed points of S, then

$$\rho(x, x') = \rho(S(x), S(x')) \leq \alpha\rho(x, x').$$

Therefore $(1 - \alpha)\rho(x, x') \leq 0$ which implies that $\rho(x, x') = 0$ since $1 - \alpha > 0$. Thus $x = x'$. For existence, let $x_0 \in X$ be any point in X and define $x_n \in X$ inductively by $x_{n+1} = S(x_n)$ for $n \geq 0$. We will show that $x := \lim_{n\to\infty} x_n$ exists in X and because S is continuous this will imply,

$$x = \lim_{n\to\infty} x_{n+1} = \lim_{n\to\infty} S(x_n) = S(\lim_{n\to\infty} x_n) = S(x),$$

showing x is a fixed point of S. So to finish the proof, because X is complete, it suffices to show $\{x_n\}_{n=1}^\infty$ is a Cauchy sequence in X. An easy inductive computation shows, for $n \geq 0$, that

$$\rho(x_{n+1}, x_n) = \rho(S(x_n), S(x_{n-1})) \leq \alpha\rho(x_n, x_{n-1}) \leq \cdots \leq \alpha^n \rho(x_1, x_0).$$

Another inductive argument using the triangle inequality shows, for $m > n$, that,

$$\rho(x_m, x_n) \le \rho(x_m, x_{m-1}) + \rho(x_{m-1}, x_n) \le \dots \le \sum_{k=n}^{m-1} \rho(x_{k+1}, x_k).$$

Combining the last two inequalities gives (using again that $\alpha \in (0,1)$),

$$\rho(x_m, x_n) \le \sum_{k=n}^{m-1} \alpha^k \rho(x_1, x_0) \le \rho(x_1, x_0)\alpha^n \sum_{l=0}^{\infty} \alpha^l = \rho(x_1, x_0)\frac{\alpha^n}{1-\alpha}.$$

This last equation shows that $\rho(x_m, x_n) \to 0$ as $m, n \to \infty$, i.e. $\{x_n\}_{n=0}^{\infty}$ is a Cauchy sequence. ■

Corollary 16.33 (Contraction Mapping Principle II). *Suppose that (X, ρ) is a complete metric space and $S : X \to X$ is a continuous map such that $S^{(n)}$ is a contraction for some $n \in \mathbb{N}$. Here*

$$S^{(n)} := \overbrace{S \circ S \circ \ldots \circ S}^{n \text{ times}}$$

and we are assuming there exists $\alpha \in (0,1)$ such that $\rho(S^{(n)}(x), S^{(n)}(y)) \le \alpha\rho(x, y)$ for all $x, y \in X$. Then S has a unique fixed point in X.

Proof. Let $T := S^{(n)}$, then $T : X \to X$ is a contraction and hence T has a unique fixed point $x \in X$. Since any fixed point of S is also a fixed point of T, we see if S has a fixed point then it must be x. Now

$$T(S(x)) = S^{(n)}(S(x)) = S(S^{(n)}(x)) = S(T(x)) = S(x),$$

which shows that $S(x)$ is also a fixed point of T. Since T has only one fixed point, we must have that $S(x) = x$. So we have shown that x is a fixed point of S and this fixed point is unique. ■

Lemma 16.34. *Suppose that (X, ρ) is a complete metric space, $n \in \mathbb{N}$, Z is a topological space, and $\alpha \in (0,1)$. Suppose for each $z \in Z$ there is a map $S_z : X \to X$ with the following properties:*

Contraction property $\rho(S_z^{(n)}(x), S_z^{(n)}(y)) \le \alpha\rho(x, y)$ for all $x, y \in X$ and $z \in Z$.
Continuity in z For each $x \in X$ the map $z \in Z \to S_z(x) \in X$ is continuous.

By Corollary 16.33 above, for each $z \in Z$ there is a unique fixed point $G(z) \in X$ of S_z.
Conclusion: *The map $G : Z \to X$ is continuous.*

Proof. Let $T_z := S_z^{(n)}$. If $z, w \in Z$, then

$$\begin{aligned}\rho(G(z), G(w)) &= \rho(T_z(G(z)), T_w(G(w))) \\ &\le \rho(T_z(G(z)), T_w(G(z))) + \rho(T_w(G(z)), T_w(G(w))) \\ &\le \rho(T_z(G(z)), T_w(G(z))) + \alpha\rho(G(z), G(w)).\end{aligned}$$

Solving this inequality for $\rho(G(z), G(w))$ gives

$$\rho(G(z), G(w)) \le \frac{1}{1-\alpha}\rho(T_z(G(z)), T_w(G(z))).$$

Since $w \to T_w(G(z))$ is continuous it follows from the above equation that $G(w) \to G(z)$ as $w \to z$, i.e. G is continuous. ■

16.9 Exercises

Exercise 16.3. Suppose that $A : \mathbb{R} \to L(X)$ is a continuous function and $V : \mathbb{R} \to L(X)$ is the unique solution to the linear differential equation

$$\dot{V}(t) = A(t)V(t) \text{ with } V(0) = I. \tag{16.47}$$

Assuming that $V(t)$ is invertible for all $t \in \mathbb{R}$, show that $V^{-1}(t) := [V(t)]^{-1}$ must solve the differential equation

$$\frac{d}{dt}V^{-1}(t) = -V^{-1}(t)A(t) \text{ with } V^{-1}(0) = I. \tag{16.48}$$

See Exercise 8.13 as well.

Exercise 16.4 (Differential Equations with Parameters). Let W be another Banach space, $U \times V \subset_o X \times W$ and $Z \in C^1(U \times V, X)$. For each $(x, w) \in U \times V$, let $t \in J_{x,w} \to \phi(t, x, w)$ denote the maximal solution to the ODE

$$\dot{y}(t) = Z(y(t), w) \text{ with } y(0) = x \tag{16.49}$$

and

$$\mathcal{D} := \{(t, x, w) \in \mathbb{R} \times U \times V : t \in J_{x,w}\}$$

as in Exercise 15.8.

1. Prove that ϕ is C^1 and that $D_w\phi(t, x, w)$ solves the differential equation:

$$\frac{d}{dt}D_w\phi(t, x, w) = (D_x Z)(\phi(t, x, w), w)D_w\phi(t, x, w) + (D_w Z)(\phi(t, x, w), w)$$

with $D_w\phi(0, x, w) = 0 \in L(W, X)$. **Hint:** See the hint for Exercise 15.8 with the reference to Theorem 15.15 being replace by Theorem 16.28.

2. Also show with the aid of Duhamel's principle (Exercise 8.23) and Theorem 16.28 that

$$D_w\phi(t,x,w) = D_x\phi(t,x,w)\int_0^t D_x\phi(\tau,x,w)^{-1}(D_wZ)(\phi(\tau,x,w),w)d\tau$$

Exercise 16.5. (Differential of e^A) Let $f : L(X) \to GL(X)$ be the exponential function $f(A) = e^A$. Prove that f is differentiable and that

$$Df(A)B = \int_0^1 e^{(1-t)A}Be^{tA}\,dt. \tag{16.50}$$

Hint: Let $B \in L(X)$ and define $w(t,s) = e^{t(A+sB)}$ for all $t,s \in \mathbb{R}$. Notice that

$$dw(t,s)/dt = (A+sB)w(t,s) \text{ with } w(0,s) = I \in L(X). \tag{16.51}$$

Use Exercise 16.4 to conclude that w is C^1 and that $w'(t,0) := dw(t,s)/ds|_{s=0}$ satisfies the differential equation,

$$\frac{d}{dt}w'(t,0) = Aw'(t,0) + Be^{tA} \text{ with } w(0,0) = 0 \in L(X). \tag{16.52}$$

Solve this equation by Duhamel's principle (Exercise 8.23) and then apply Proposition 16.14 to conclude that f is differentiable with differential given by Eq. (16.50).

Exercise 16.6 (Local ODE Existence). Let S_x be defined as in Eq. (15.15) from the proof of Theorem 15.4. Verify that S_x satisfies the hypothesis of Corollary 16.33. In particular we could have used Corollary 16.33 to prove Theorem 15.4.

Exercise 16.7 (Local ODE Existence Again). Let $J = (-1,1)$, $Z \in C^1(X,X)$, $Y := BC(J,X)$ and for $y \in Y$ and $s \in J$ let $y_s \in Y$ be defined by $y_s(t) := y(st)$. Use the following outline to prove the ODE

$$\dot{y}(t) = Z(y(t)) \text{ with } y(0) = x \tag{16.53}$$

has a unique solution for small t and this solution is C^1 in x.

1. If y solves Eq. (16.53) then y_s solves

$$\dot{y}_s(t) = sZ(y_s(t)) \text{ with } y_s(0) = x$$

or equivalently

$$y_s(t) = x + s\int_0^t Z(y_s(\tau))d\tau. \tag{16.54}$$

Notice that when $s = 0$, the unique solution to this equation is $y_0(t) = x$.

2. Let $F : J \times Y \to J \times Y$ be defined by

$$F(s,y) := (s, y(t) - s\int_0^t Z(y(\tau))d\tau).$$

Show the differential of F is given by

$$F'(s,y)(a,v) = \left(a, t \to v(t) - s\int_0^t Z'(y(\tau))v(\tau)d\tau - a\int_0^\cdot Z(y(\tau))d\tau\right).$$

3. Verify $F'(0,y) : \mathbb{R}\times Y \to \mathbb{R}\times Y$ is invertible for all $y \in Y$ and notice that $F(0,y) = (0,y)$.
4. For $x \in X$, let $C_x \in Y$ be the constant path at x, i.e. $C_x(t) = x$ for all $t \in J$. Use the inverse function Theorem 16.25 to conclude there exists $\varepsilon > 0$ and a C^1 map $\phi : (-\varepsilon,\varepsilon)\times B(x_0,\varepsilon) \to Y$ such that

$$F(s,\phi(s,x)) = (s, C_x) \text{ for all } (s,x) \in (-\varepsilon,\varepsilon)\times B(x_0,\varepsilon).$$

5. Show, for $s \le \varepsilon$ that $y_s(t) := \phi(s,x)(t)$ satisfies Eq. (16.54). Now define $y(t,x) = \phi(\varepsilon/2,x)(2t/\varepsilon)$ and show $y(t,x)$ solve Eq. (16.53) for $|t| < \varepsilon/2$ and $x \in B(x_0,\varepsilon)$.

Exercise 16.8. Show P defined in Theorem 16.31 is continuously differentiable and $P'(y)h = \dot{h} + p'(y)h$.

Exercise 16.9. Embedded sub-manifold problems.

Exercise 16.10. Lagrange Multiplier problems.

16.9.1 Alternate construction of g. To be made into an exercise.

Suppose $U \subset_o X$ and $f : U \to Y$ is a C^2 – function. Then we are looking for a function $g(y)$ such that $f(g(y)) = y$. Fix an $x_0 \in U$ and $y_0 = f(x_0) \in Y$. Suppose such a g exists and let $x(t) = g(y_0 + th)$ for some $h \in Y$. Then differentiating $f(x(t)) = y_0 + th$ implies

$$\frac{d}{dt}f(x(t)) = f'(x(t))\dot{x}(t) = h$$

or equivalently that

$$\dot{x}(t) = [f'(x(t))]^{-1}h = Z(h,x(t)) \text{ with } x(0) = x_0 \tag{16.55}$$

where $Z(h,x) = [f'(x(t))]^{-1}h$. Conversely if x solves Eq. (16.55) we have $\frac{d}{dt}f(x(t)) = h$ and hence that

$$f(x(1)) = y_0 + h.$$

Thus if we define

$$g(y_0 + h) := e^{Z(h,\cdot)}(x_0),$$

then $f(g(y_0+h)) = y_0 + h$ for all h sufficiently small. This shows f is an open mapping.