C^1 Dependence on Initial Conditions for ODE

First, two comments about section 3.9:

1. In section 3.9, one can use real Banach spaces in place of Banach spaces.
2. In the statement of Theorem 3.9.2, we should have written

$$f'(a) \text{ is INVERTIBLE, with a bounded inverse.}$$

(In fact, as we shall see in Chapter 12, any invertible map between Banach spaces, or real Banach spaces, has a bounded inverse.)

We will need the following simple lemma.

Lemma 0.1 *Let (X, d_1) and (Y, d_2) be metric spaces. Form the metric space $(X \times Y, d)$ where $d((x_1, y_1), (x_2, y_2)) = d_1(x_1, x_2) + d_2(y_1, y_2)$. Suppose that for each $x \in X$ we have a function $u_x \in C_b(Y)$, the space of bounded continuous functions from Y to $\mathbf{R}^l$. (Here l is fixed.) Suppose moreover that for each $x_0 \in X$,*

$$\lim_{x \to x_0} \|u_x - u_{x_0}\| = 0.$$

(Here $\|\,\|$ denotes sup norm on Y.) For $x \in X$, $y \in Y$, put $u(x, y) = u_x(y)$. Then $h : X \times Y \to \mathbf{R}^l$ is continuous.

Proof Say $\epsilon > 0$. If $(x_0, y_0) \in X \times Y$, then, since $u_{x_0} \in C_b(y)$, we may find $\delta_2 > 0$ such that if $d_2(y, y_0) < \delta_2$, then

$$|u(x_0, y) - u(x_0, y_0)| < \epsilon/2.$$

By hypothesis we can also select $\delta_1 > 0$ such that, if $d_1(x, x_0) < \delta_1$, then

$$\|u_x - u_{x_0}\| < \epsilon/2.$$

It then follows that if $d_1(x, x_0) < \delta_1$ and $d_2(y, y_0) < \delta_2$, then

$$|u(x, y) - u(x_0, y_0)| < \epsilon,$$

as desired.

We next observe that the implicit function theorem has a direct extension for Banach spaces (or real Banach spaces).

If X and Y are Banach spaces (or real Banach spaces), with norms $\| \|_X$ and $\| \|_Y$ respectively, we may define a norm $\| \|$ on $X \times Y$ by setting

$$\|(x, y)\| = \|x\|_X + \|y\|_Y.$$

It is very easy to check that $X \times Y$ is a Banach space (or a real Banach space) with this norm. Note that the metric obtained from this norm on $X \times Y$ is the same as the metric one obtains through the procedure described in Lemma 0.1.

The linear maps $T : X \times Y \to Y$ are precisely the maps of the form

$$T(x, y) = Ax + By \tag{1}$$

where $A : X \to Y$ and $B : Y \to Y$ are linear. In fact, given such A and B, T as defined in (1) is clearly linear. Conversely, given $T : X \times Y \to Y$ linear, we may define linear maps $A : X \to Y$ and $B : Y \to Y$ by $Ax = T(x, 0)$, $By = T(0, y)$, and then (1) clearly holds. In fact, we had no choice here in the definition of A, B; it is forced by (1). Let us write $T = (A \quad B)$.

We then have the following extension of the Implicit Function Theorem for Banach spaces:

Theorem 0.2 Implicit Function Theorem for Banach spaces *Let X, Y be Banach spaces, or real Banach spaces. Suppose $U \subseteq X \times Y$ is open, $F : U \to Y$, $F \in C^1(U)$. Say $(x_0, y_0) \in U$,*

$$F(x_0, y_0) = 0$$

and the derivative of F at (x_0, y_0) is

$$(A \quad B)$$

(here $A : X \to Y$, $B : Y \to Y$). Finally suppose

B is INVERTIBLE, with a bounded inverse.

Then x_0 has an open neighborhood V, and there exists $h_0 \in C^1(V)$, $h_0 : V \to Y$, such that

$$h_0(x_0) = y_0$$

and

$$F(x, h_0(x)) = 0$$

for all $x \in V$.

Further, (x_0, y_0) has an open neighborhood $U_1 \subseteq U$ so that

$$[(x, y) \in U_1, \ F(x, y) = 0] \Longleftrightarrow [x \in V, \ y = h_0(x)].$$

To prove this one needs to make only minor modifications in the proof of Theorem 3.8.2. One defines $G : X \times Y \to X \times Y$ by $G(x, y) = (x, F(x, y))$, and notes that the derivative of G at (x_0, y_0) equals T, where $T(x, y) = (x, Ax + By)$. T is invertible, with a bounded inverse (in fact its inverse is T^{-1}, where $T^{-1}(x, y) = (x, -B^1 Ax + B^{-1} y)$.) One then proceeds to prove the theorem just as in the proof of Theorem 3.8.2.

Let us now look at some important examples of differentiable maps between real Banach spaces.

Example 1 Let $(a, b) \subseteq \mathbf{R}$ be an interval. Let (X, d) be a metric space. Let $C_b(X)$ denote, as usual, the real Banach space of bounded continuous functions from X to $\mathbf{R}^l$ (for some fixed l). What would it mean for a map $G : (a, b) \to C_b(X)$ to be differentiable at a point $s_0 \in (a, b)$? The only linear maps from $\mathbf{R}$ to a real Banach space Y are of the form $T(h) = yh$ for some $y \in Y$ (indeed, $y = T(1)$). We would need, then, that for some $y \in C_b(X)$,

$$\|G(s_0 + h) - G(s_0) - yh\| = o(h) \tag{2}$$

Here $\| \ \|$ denotes sup norm on X. If $s \in (a, b)$, then $G(s) \in C_b(X)$; if $x \in X$, write $G(s, x) = [G(s)](x)$. We see that for all $x \in X$,

$$|G(s_0 + h, x) - G(s_0, x) - y(x)h| = o(h),$$

so that y could only be $\partial G/\partial s$, evaluated at $s = s_0$. Thus this partial exists and is in $C_b(X)$, and we have the uniformity

$$\|G(s_0 + h) - G(s_0) - \frac{\partial G}{\partial s}|_{s=s_0} h\| = o(h) \tag{3}$$

Suppose now that G is differentiable at each point $s_0 \in (a, b)$. Since $(\partial G/\partial s)|_{s=s_0}$ is bounded, (3) implies that

$$\|G(s_0 + h) - G(s_0)\| = O(h).$$

As a consequence of this and Lemma 0.1, we see that $G(s, x)$ *is a continuous function of* $(s, x) \in (a, b) \times X$.

What then would it mean for $G : (a, b) \to C_b(X)$ to be C^1? As we said above, the only linear maps from $\mathbf{R}$ to a real Banach space Y are of the form $T(h) = yh$ for some $y \in Y$; clearly $\|T\| = \|y\|$. The derivative of G at h is the linear transformation $h \to (\partial G/\partial s)t$. So to say that G is C^1 is to say that, for all $s_0 \in (a, b)$, $(\partial G/\partial s)|_{s=s_0}$ is in $C_b(X)$, (3) holds, and

$$\|\frac{\partial G}{\partial s}|_{s=r} - \frac{\partial G}{\partial s}|_{s=s_0}\| \to 0 \tag{4}$$

as $r \to s_0$. Lemma 0.1 and (4) now show that $\partial G/\partial s$ is also continuous as a function of $(s, x) \in (a, b) \times X$.

***Exercise 0.3** *Let $U \subseteq \mathbf{R}^n$ be open, and again let (X, d) be a metric space. Suppose that $G : U \to C_b(X)$.*
(a) What exactly does it mean for G to be differentiable at a point $s_0 \in U$? Show that, if it is, then all partials $\partial G/\partial s_i$ exist at s_0 and are in $C_b(X)$.
(b) If G is differentiable at every $s \in U$, show that G is continuous as a function of $(s, x) \in U \times X$.
(c) What exactly does it mean for G to be C^1 on U? Show that, if it is, then all partials $\partial G/\partial s_i$ are continuous as functions of $(s, x) \in U \times X$.

Solution (a) Every linear map T from $\mathbf{R}^n$ to a real Banach space Y is of the form $T(h) = \sum y_i h_i$ for certain $y_1, \ldots, y_n \in Y$. We would need, then, that for some $y_1, \ldots, y_n \in C_b(X)$,

$$\|G(s_0 + h) - G(s_0) - \sum_{i=1}^{n} y_i h_i\| = o(h). \tag{5}$$

Here $\| \ \|$ denotes sup norm on X. If $s \in U$, then $G(s) \in C_b(X)$; if $x \in X$, write $G(s,x) = [G(s)](x)$. We see that for all $x \in X$, if $T_x h = \sum y_i(x) h_i$, then

$$|G(s_0 + h, x) - G(s_0, x) - T_x h| = o(h).$$

Thus T_x must be the derivative of G at s_0, and consequently y_i must be $\partial G/\partial s_i$, evaluated at $s = s_0$. Thus this partial exists and is in $C_b(X)$, and we have the uniformity

$$\|G(s_0 + h) - G(s_0) - \sum_i \frac{\partial G}{\partial s_i}|_{s=s_0} h_i\| = o(h) \tag{6}$$

(b) Since each $(\partial G/\partial s_i)|_{s=s_0}$ is bounded, (6) implies that

$$\|G(s_0 + h) - G(s_0)\| = O(h).$$

(b) now follows at once from Lemma 0.1.
(c) As we said above, the only linear maps from $\mathbf{R}^n$ to a real Banach space Y are of the form $T(h) = \sum_i y_i h_i$ for certain $y_1, \ldots, y_n \in Y$. We have

$$\max(\|y_1\|, \ldots, \|y_n\|) \le \|T\| \le n \max(\|y_1\|, \ldots, \|y_n\|). \tag{7}$$

Indeed, if $e_i = (0, \ldots, 0, 1, 0, \ldots, 0)$ with the "1" in the ith slot, $T(e_i) = y_i$, so that $\|T\| \ge \|y_i\|$ for all i. But for any h, $\|T(h)\| \le \sum |h_i| \|y_i\| \le (\sum \|y_i\|)|h| \le n \max(\|y_1\|, \ldots, \|y_n\|)|h|$. This proves (7).

The derivative of G at s is the linear transformation $h \to \sum_i (\partial G/\partial s_i) h_i$. So to say that G is C^1 is to say that, for all $s_0 \in U$, and for all i, $(\partial G/\partial s_i)|_{s=s_0}$ is in $C_b(X)$; that (6) holds; and (by (7)), that for any i,

$$\|\frac{\partial G}{\partial s_i}|_{s=r} - \frac{\partial G}{\partial s_i}|_{s=s_0}\| \to 0 \tag{8}$$

as $r \to s_0$. Lemma 0.1 and (4) now show that each $\partial G/\partial s_i$ is also continuous as a function of $(s,x) \in U \times X$.

***Exercise 0.4** *Let $I = [A,B]$ be a closed interval, and $U \subseteq \mathbf{R}^n$ is open. Let $C(I;U)$ denote the set of continuous functions with domain I and range contained in U.*
(a) Show that $C(I;U)$ is an open subset of $C(I)$ (the normed vector space of continuous functions from I to $\mathbf{R}^n$.)
(b) Suppose that $F : I \times U \to \mathbf{R}^n$ is continuous. Suppose that for all $(t, Y) \in I \times U$, all partials $\partial F/\partial Y_i$ exist, and that these partials are also continuous on $I \times U$. Say $t_0 \in I$. For $y \in C(I;U)$, define $Ty \in C(I)$ by

$$Ty(t) = \int_{t_0}^{t} F(s, y(s)) ds.$$

Show that $T : C(I;U) \to C(I)$ is differentiable, and compute its derivative.

Solution (a) This is evident if $U = \mathbf{R}^n$, for then $C(I;U) = C(I)$. So say $U \neq \mathbf{R}^n$, and suppose $y \in C(I;U)$. Let K_1 equal the range of y, a compact subset of U. Let $\delta = \operatorname{dist}(K_1, U^c)$; then $0 < \delta < \infty$. Let $V = \{f \in C(I) : \|f - y\| < \delta\}$; to prove (a), we need only show that V is contained in $C(I;U)$ (since it is an open neighborhood of y, an arbitrary element of $C(I;U)$). But if $f \in V$, then for any $t \in I$, $|f(t) - y(t)| < \delta$. Since $y(t) \in K_1$, we must have $f(t) \in U$. So $f : I \to U$ as desired.
(b) For $i = 1, \ldots, n$, let F^i denote $\partial F/\partial Y_i$. For $s \in I$, define $F_s : U \to \mathbf{R}^n$ by $F_s(Y) = F(s, Y)$. Then, for fixed s, the columns of the matrix of $F_s'(Y)$ are $F^i(s, Y)$.

Fix $y \in C(I;U)$. For $h = (h_1, \ldots, h_n) \in C(I)$, define $Qh \in C(I)$ by

$$Qh(t) = \int_{t_0}^{t} F_s'(y(s))h(s)\, ds. \tag{9}$$

(Here, $F_s'(y(s))h(s) = \sum_{i=1}^n F^i(s, y(s))h_i(s)$.) It is enough to show that Q is the derivative of T at y. Choose K_1, δ for y as in the solution of (a). Let K be the compact set defined by

$$K = \{x \in \mathbf{R}^n : \ \operatorname{dist}(x, K_1) \leq \delta/2\}.$$

Then K is a compact subset of U. Moreover, if $Y \in K_1$ then the open ball of radius $\delta/2$, centered at Y, is contained in K. In particular, if $Y \in K_1$, $H \in \mathbf{R}^n$ and $|H| < \delta/2$, then the line segment joining Y to $Y + H$ is completely contained in K. In addition, for any i, F^i is uniformly continuous on $I \times K$. Thus, although K is not necessarily convex, the proof of Exercise 3.6.16, as applied to this K, shows this uniformity: for every $\epsilon > 0$, there exists $\eta > 0$ such that whenever $a \in K_1$, $|H| < \eta$ and $s \in I$, we have $|F_s(Y + H) - F_s(Y) - F_s'(Y)H| \leq \frac{\epsilon}{B-A}|H|$. Accordingly, if $h \in C(I)$ and $\|h\| < \eta$, we have that for any $t \in I$,

$$\begin{aligned}
|[T(y+h) - Ty - Qh](t)| &= \left|\int_{t_0}^{t} [F_s([y+h](s)) - F_s(y(s)) - F_s'(y(s))h(s)] ds\right| \\
&\leq \left|\int_{t_0}^{t} |F_s([y+h](s)) - F_s(y(s)) - F_s'(y(s))h(s)| ds\right. \\
&\leq \left|\int_{t_0}^{t} \frac{\epsilon}{B-A}\|h\| ds\right| \\
&\leq \epsilon\|h\|.
\end{aligned}$$

Thus $\|T(y+h) - Ty - Qh\| \leq \epsilon\|h\|$, as desired.

We may now establish an important result that states that the solution of an ODE depends in a C^1 manner on the initial conditions. We shall work in the situation of the Fundamental Local Existence theorem for ODE (Theorem 2.2.2), but we shall assume the stronger condition that for all i, $\partial F/\partial Y_i$ exists and is continuous on $[t_0 - h, t_0 + h] \times U$, where $U \subseteq \mathbf{R}^n$ is open and contains $B_R(Y_0)$. As we shall review in a moment, this implies that F is Lipschitz in Y on the set S of (2.6), in the sense of (2.7).

Theorem 0.5 **C^1 dependence on initial conditions for solutions of ODE** *Say $t_0 \in \mathbf{R}$, $Y_0 \in \mathbf{R}^n$, $R > 0$. Let*

$$S = [t_0 - h, t_0 + h] \times B_R(Y_0) \tag{10}$$

where $B_R(Y_0)$ is the closed ball of radius $R > 0$ about Y_0.

Suppose:

(i) The domain of $F(t, Y)$ contains a set $S_1 = [t_0 - h, t_0 + h] \times U$, where $U \subseteq \mathbf{R}^n$ is open and contains $B_R(Y_0)$;
(ii) The range of $F(t, Y)$ is contained in $\mathbf{R}^n$;
(iii) F is continuous on S;
(iv) For $i = 1, \ldots, n$, the function $F^i = \partial F/\partial Y_i$ exists and is continuous on S_1.

Then there exists $P_1 > 0$, satisfying the following three properties:

(A) whenever $|Y - Y_0| < R/2$, the equation $dy/dt = F(t, y)$ has a unique solution $y_Y(t)$ in $(t_0 - P_1, t_0 + P_1)$ with $y(t_0) = Y$;

(B) whenever $|Y - Y_0| < R/2$, if $0 < r < P_1$, then $y_Y(t)$ is also the unique solution of $dy/dt = F(t, y)$ in $(t_0 - r, t_0 + r)$ with $y_Y(t_0) = Y$; and

(C) if we set $y(Y, t) = y_Y(t)$ for $|Y - Y_0| < R/2$, $t \in (t_0 - P_1, t_0 + P_1)$, then $y(Y, t)$ is a C^1 function of Y, t.

Proof Let M be the maximum of $|F(t, Y)|$ on S. We may assume $M > 0$. For $t \in [t_0 - h, t_0 + h]$, define $F_t : U \to \mathbf{R}^n$ by $F_t(Y) = F(t, Y)$. Let K be the maximum of $\|F_t'(Y)\|$ for $(t, Y) \in S$ (unless this maximum is zero, in which case take $K = 1$.) We claim that we may take $P_1 = \min(\frac{R}{2M}, \frac{1}{K}, h)$.

First note that F is Lipschitz in Y on S. Indeed, if $(t, Y^1), (t, Y^2) \in S$, then

$$|F(t, Y^1) - F(t, Y^2)| = |F_t(Y^1) - F_t(Y^2)| \leq K|Y^1 - Y^2|,$$

by Proposition 3.6.11 (b).

Now say $|Y - Y_0| < R/2$, and note $[t_0 - h, t_0 + h] \times B_{R/2}(Y) \subseteq S$. Then, by Theorem 2.2.2, Corollary 2.2.4 and their proofs, the differential equation $dy/dt = F(t, y)$, $y(t_0) = Y$, has a unique solution y_Y on $(t_0 - P_1, t_0 + P_1)$. Moreover, if $0 < r < P_1$, then $y_Y(t)$ is also the unique solution of $dy/dt = F(t, y)$ in $(t_0 - r, t_0 + r)$ with $y_Y(t_0) = Y$. Moreover, on the interval $(t_0 - P_1, t_0 + P_1)$, the range of y_Y is contained in $B_{R/2}(Y)$.

Now, fix r with $0 < r < P_1$, and set $J = [t_0 - r, t_0 + r]$. Define $\mathcal{T} : \mathbf{R}^n \times C(J; U) \to C(J)$ by

$$(\mathcal{T}(Y, f))(t) = Y + \int_{t_0}^{t} F(s, f(s))ds - f(t)$$

for $(Y, f) \in \mathbf{R}^n \times C(J; U)$. The proofs of Theorem 2.2.2 and Corollary 2.2.4 further show that $\mathcal{T}(Y, y_Y) = 0$ whenever $|Y - Y_0| < R/2$.

Pick any Y_0' with $|Y_0' - Y_0| < R/2$. To complete the proof, we must establish conclusion (C), and for this, it suffices to show that, for some open neighborhood V of Y_0' in $\mathbf{R}^n$, $y(Y, t) = y_Y(t)$ is a C^1 function of (Y, t) for $(Y, t) \in V \times J$. Put $y_0 = y_{Y_0'}$. Since the derivative of a sum is the sum of the derivatives of the summands, by Exercise 0.4 we see that $\mathcal{T}$ is differentiable at y_0, with derivative $(I \quad Q - I)$, where Q is as in (9) (with y there replaced by y_0).

We now want to use the implicit function theorem for Banach spaces, and for this we will need to know that $Q - I : C(J) \to C(J)$ is invertible (with a bounded inverse). For this it suffices to show that $\|Q\| < 1$. But if $h \in C(J)$, then

$$\begin{aligned} |Qh(t)| &= \left|\int_{t_0}^{t} F_s'(y_0(s))h(s)\,ds\right| \\ &\leq \left|\int_{t_0}^{t} |F_s'(y_0(s))h(s)|\,ds\right| \\ &\leq \left|\int_{t_0}^{t} K\|h\|\,ds\right| \\ &\leq Kr\|h\| \end{aligned}$$

Accordingly, $\|Qh\| \leq Kr\|h\|$. Since $Kr < 1$, it follows that $\|Q\| < 1$, as claimed.

Surely $\mathcal{T}(Y_0', y_0) = 0$, so the implicit function theorem for Banach spaces tells us that for some open neighborhood V of Y_0' in $\mathbf{R}^n$, there is a C^1 mapping $z : V \to C(J; U)$ such that $\mathcal{T}(Y, z(Y)) = 0$ for all $Y \in V$. Writing $z_Y = z(Y)$, we surely have $z_Y(t_0) = Y$ and $dz_Y/dt = F(t, z_Y)$ on J; so $z_Y = y_Y$. What we have gained by using the implicit function theorem is the knowledge that the map z is C^1. This implies, by the results of Exercise 0.3, that $y(Y, t)$ is continuous as a function of $(Y, t) \in V \times J$, and that all $(\partial y/\partial Y_i)(Y, t)$ exist and are continuous for $(Y, t) \in V \times J$. Finally $(\partial y/\partial t)(Y, t) = F(t, y(Y, t))$ is also continuous for $(Y, t) \in V \times J$. This shows that y is C^1 on $V \times J$, as desired.