BANACH CALCULUS

CHIN-LUNG WANG

ABSTRACT. This is part of my lecture notes for "Honored Advanced Calculus" at National Taiwan University in 2011-2012. We prove the inverse function theorem for Banach spaces and use it to prove the smooth dependence on initial data for solutions of ordinary differential equations.

The smooth dependence is an essential ingredient in the proofs of many fundamental theorems in modern mathematics. For example in the smoothness of flows generated by vector fields and in the smoothness of exponential map in differential geometry. The inverse function theorem also plays a fundamental role in non-linear problems in analysis as well as in geometry.

Table of Contents

1. Basics in Banach Spaces

1.1. **The category of Banach spaces.** A Banach space E over $\mathbb{F}$ is a complete normed vector space where $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. This means that

Date: April 26, 2012.

there is a norm function $|\cdot| : E \to \mathbb{R}^+$ satisfying

$$|cx| = |c||x|, \quad c \in \mathbb{F}, \qquad |x+y| \le |x| + |y|,$$

and every Cauchy sequence in the metric space $(E, |\cdot|)$ converges. Standard examples are the Euclidean spaces $\mathbb{R}^n$ and $\mathbb{C}^n$, the space of continuous functions over a compact space $C(\Omega)$ under the sup norm, Hölder spaces $C^k(\Omega)$, $C^{k,\alpha}(\Omega)$, L^p spaces, $1 \le p < \infty$, and the more general Sobolev spaces $W^{k,p}$. What we do here are general results true for all Banach spaces. But the reader is reminded that in real applications an adequate knowledge on each involved space is indispensable.

For normed (vector) spaces E, F, we denote by $L(E, F)$ the space of continuous linear maps from E to F. It is easy to see that a linear map $f : E \to F$ is continuous if and only if that it is bounded, in the sense that there exists $C \in \mathbb{R}^+$ such that $|f(x)| \le C|x|$ for all $x \in E$. We then define f to be the inf of such bounds C's. Equivalently

$$|f| = \sup_{x \in E} \frac{|f(x)|}{|x|} \equiv \sup_{|x|=1} |f(x)|.$$

In this sup norm, $L(E.F)$ is also a normed space.

Exercise 1. *If F is a Banach space, show that $L(E, F)$ and $C(\Omega, F)$ are also Banach where Ω is a compact topological space.*

Exercise 2. *The unit sphere $S = \{x \in E \,|\, |x| = 1\}$ of a normed space is compact if and only if that E is finite dimensional.*

We define an $f \in L(E, F)$ to be an isomorphism of normed spaces if it is both a linear and topological isomorphism. For E, F Banach spaces, it turns out that we only need f to be bijective to ensure the continuity of f^{-1}. This follows from the open mapping theorem which we recall later.

For $f \in L(E, F)$, $g \in L(F, G)$, we check easily that $gf \in L(E, G)$. In fact $|gf| \le |g||f|$. Thus when $E = F$, we may consider the Banach algebra $\mathrm{End}(E) = L(E, E)$. In general a Banach space B is called a Banach algebra if B is also an algebra and its multiplication law satisfies $|ab| \le |a||b|$.

Exercise 3. *Let B be a Banach algebra with identity 1. Show that $1 + a$ is invertible in B if $|a| < 1$ and the set of invertible elements $B^\times$ is open.*

1.2. **Multi-linear maps.** Given normed spaces $E_1, \dots, E_k$ and G. As before, a k-linear map $f : E_1 \times \cdots \times E_k \to G$ (i.e. linear in each factor E_i) is continuous if and only if there is a $C \in \mathbb{R}^+$ such that

$$f(x_1, \dots, x_k) \le C|x_1| \cdots |x_k|, \quad \text{for all } x_i \in E_i.$$

Again the norm $|f|$ is defined similarly. We denote by $L(E_1, \dots, E_k; G)$ the normed space of all continuous k-linear maps.

There is a natural bijective linear map

$$L(E_1, L(E_2, \dots, E_k; G)) \to L(E_1, \dots, E_k; G)$$

by sending ϕ to f_ϕ defined by $f_\phi(x_1, \dots, x_k) = \phi(x_1)(x_2, \dots, x_k)$.

Exercise 4. *Show that this map is a norm preserving isomorphism.*

When $E_i = E$ for all i, we use $L^k(E, F)$ to denote the space continuous k-linear maps. When G is Banach, all these spaces are also Banach.

For normed spaces E_i, $1 \le i \le k$, we give $E = \prod_{i=1}^k E_i$ an Euclidean-style norm $|(x_1, \dots, x_k)| = \sqrt{|x_1|^2 + \cdots + |x_k|^2}$. E is Banach if all E_i's are.

1.3. **Two fundamental theorems.** In order to *conceptually* realize that Banach spaces are truly similar to finite dimensional Euclidean spaces, we need suitable structure theorems to replace some finite dimensional operations. For normed space, coordinates are replaced by linear functionals.

Hanh-Banach Theorem: *Let F be a subspace of a normed space E. If $\lambda \in L(F, \mathbb{R})$ then λ can be extended to a $\lambda^* \in L(E, \mathbb{R})$ with $|\lambda^*| = |\lambda|$.*

It is enough to consider the case $|\lambda| = 1$. We need to find such an extension λ' on a subspace with one more dimension $F' = F + \mathbb{R}e$ where $e \in E \setminus F$. If we assign $\lambda'(e) = a \in \mathbb{R}$ then $\lambda'(u + te) = \lambda(u) + ta$. By dividing out t, the norm constraint $|\lambda'| = 1$ on F' is equivalent to

$$-|u + e| \le \lambda(u) + a \le |u + e|$$

for all $u \in F$. That is, $a \in [-\lambda(u) - |u + e|, -\lambda(u) + |u + e|]$. To ensure the existence of a, we need to show that when u varies all such intervals have non-empty intersection. Namely we need

$$-\lambda(u) - |u + e| \le -\lambda(v) + |v + e|, \quad \forall\, u, v \in F,$$

which clearly follows from $\lambda(v - u) \le |v - u| \le |v + e| + |u + e|$.

Now we consider the collection of all norm-preserving extensions $\Phi = \{(F_\alpha, \lambda_\alpha)\}$. It is partially ordered by $(F_\alpha, \lambda_\alpha) \le (F_\beta, \lambda_\beta)$ if $F_\alpha \subset F_\beta$ and $\lambda_\beta|_{F_\alpha} = \lambda_\alpha$. If $(F_i, \lambda_i) \in \Phi$ is an increasing chain, then there is an upper bound $(F', \lambda') \in \Phi$ given by $F' := \bigcup_i F_i$ and $\lambda'(x) := \lambda_i(x)$ if $x \in F_i$ (which is clearly independent of the choice of i). By Zorn's lemma, a maximal element $(F_\alpha, \lambda_\alpha) \in \Phi$ exists. We must have $F_\alpha = E$, for otherwise we may find $e \in E \setminus F_\alpha$ and further extend the functional λ_α to $F_\alpha + \mathbb{R}e$, which contradicts to the maximality of $(F_\alpha, \lambda_\alpha)$.

Exercise 5. *Prove the Hahn-Banach theorem for* $\mathbb{F} = \mathbb{C}$.

We recall the open mapping theorem without proof since in this notes we make no essential use of it.

Open Mapping Theorem: *Let* E*,* F *be Banach. If* $f \in L(E, F)$ *is surjective then* f *is an open mapping. That is,* U *open implies* $f(U)$ *open. In particular, if* f *also injective then it is an isomorphism.*

2. Calculus on Banach Spaces

Unless specified, we assume all vector spaces encountered are Banach.

2.1. **Derivative of a map.**

Let $U \subset E$ be an open set. Let $f : U \to F$ be a map. We say that f is differentiable at x if there is a map $T \in L(E, F)$, called the derivative of f at x, such that for $|h|$ small,

$$f(x + h) - f(x) = Th + o(h)$$

where $o(h) = |h|\psi(h)$ is a map with the property that $\lim_{h\to 0} \psi(h) = 0$.

Exercise 6. *Show that the derivative, if exists, is necessarily unique.*

When T exists, it is noted by $f'(x)$ or $Df(x)$. Various notations like Df_p, $df(p)$, df_p or f_{*p} are also commonly used. It is clear that then f is continuous at x.

Take a simple but important example, let $\lambda \in L(E, F)$ be a linear map. Then $\lambda(x + h) - \lambda(x) = \lambda h$, hence $\lambda'(x) = \lambda$ for all $x \in E$.

Exercise 7. *Let* $\lambda \in L(E_1, \ldots, E_k; F)$*. Compute* $\lambda'(x_1, \ldots, x_k)$*.*

From the definition, we see that $f'(x)$ exists implies that

$$f'(x)h = \lim_{t\to 0} \frac{f(x + th) - f(x)}{t} = \frac{d}{dt} f(x + th)\Big|_{t=0}$$

is precisely the directional derivative $D_h f(x)$. But as is well known the existence of all $D_h(x)$ does not imply the existence of $f'(x)$.

The above definition trivially generalizes the one in calculus. Indeed almost all theorems hold true in the Banach category once we formulate them in the *right way*. Namely we should avoid usage of coordinates and inner products.

We now discuss the basic rules of differential calculus: sum rule, product rule (Lebnitz rule), quotient rule and the chain rule.

2.1.1. *Sum rule.*

Let $f, g : U \to F$ with $f'(x)$, $g'(x)$ exists ($x \in U$), then it is obvious that $(cf)'(x) = cf'(x)$ for $c \in \mathbb{R}$ and $(f + g)'(x) = f'(x) + g'(x)$.

2.1.2. *Lebnitz rule.* Let $f : U \to E_1$, $g : U \to E_2$ with $f'(x) \in L(E, E_1)$, $g'(x) \in L(E, E_2)$ exists and let $(,) \in L(E_1, E_2; F)$. Then $(f, g)'(x) \in L(E, F)$ exists and

$$(f, g)'(x)h = (f'(x)h, g(x)) + (f(x), g'(x)h), \quad \forall\, h \in E.$$

Exercise 8. *Verify the Lebnitz rule.*

2.1.3. *Quotient rule.* Let B be a Banach algebra with identity 1. Let $B^\times \subset B$ be the group of invertible elements. Let $\mathcal{I}(x) = x^{-1}$. Then $\mathcal{I} : B^\times \to B$ is differentiable at every $x \in B^\times$. Moreover $\mathcal{I}'(x) \in \mathrm{End}(B)$ with

$$\mathcal{I}'(x)h = -x^{-1}hx^{-1}, \quad \forall\, h \in B.$$

To see this, let $|h|$ be small then

$$\begin{aligned}(x+h)^{-1} - x^{-1} &= (1 + x^{-1}h)^{-1}x^{-1} - x^{-1} \\ &= (1 - x^{-1}h + (x^{-1}h)^2 - (x^{-1}h)^3 + \cdots)x^{-1} - x^{-1} \\ &= -x^{-1}hx^{-1} + o(h).\end{aligned}$$

2.1.4. *Chain rule.* Let $f : U \to V$, $g : V \to G$ with U open in E, V open in F and $f'(x) \in L(E, F)$, $g'(f(x)) \in L(F, G)$ exists. Then gf is differentiable at x and

$$(gf)'(x) = g'(f(x))f'(x) \quad \text{in } L(E, G).$$

The proof is again standard. Let

$$k(h) := f(x + h) - f(x) = f'(x)h + |h|\phi(h).$$

Then

$$\begin{aligned}&g(f(x + h)) - g(f(x)) \\ &\quad = g'(f(x))k(h) + |k(h)|\psi(k(h)) \\ &\quad = g'(f(x))f'(x)h + g'(f(x))|h|\phi(h) + |k(h)|\psi(k(h)).\end{aligned}$$

We need to show that the last two terms are in $o(h)$. For the last term,

$$|k(h)| \le |f'(x)||h| + |h||\phi(h)| \quad \text{and} \quad \lim_{h\to 0} \psi(k(h)) = 0,$$

so

$$\lim_{h\to 0} \frac{|k(h)||\psi(k(h))|}{|h|} \le \lim_{h\to 0}(|f'(x)| + |\phi(h)|)|\psi(k(h))| = 0.$$

The other term is obviously in $o(h)$.

We say $f \in C^1(U, F)$ if $f' : U \to L(E, F)$, $x \mapsto f'(x)$ is continuous.

2.2. **Integration over the real line.** We modify the Riemann integral for $f \in C([a,b],\mathbb{R})$ to functions $f \in C([a,b],F)$ where F is a Banach space. Namely for a partition $P = \{a = t_0 < t_1 < \cdots < t_n = b\}$ we have the Riemann sum

$$I_P(f) = \sum_{i=1}^{n} f(t_i^*)\Delta t_i$$

for arbitrary $t_i^* \in [t_{i-1}, t_i]$. Denote by $|P|$ the maximal of Δt_i. Let P_k be any sequence of partitions such that $\lim_{k\to\infty} |P_k| = 0$. Then we define

$$\int_a^b f(t)\,dt := \lim_{k\to\infty} I_{P_k}(f).$$

By the uniform continuity of f on $[a,b]$ we may show that $I_{P_k}(f)$ is a Cauchy sequence in F hence the limit exists.

Exercise 9. *Give the details of the above uniformity argument and show that the Riemann integral is independent of the choices of P_k's.*

The integral enjoys all the familiar properties.

2.2.1. *Cauchy inequality.* $|\int_a^b f| \le \int_a^b |f|$ implies that $\int_a^b$ is a continuous linear functional: $\int_a^b \in L(C([a,b],F),\mathbb{R})$. If $a > b$ then we set $\int_a^b = -\int_b^a$. The integral is then additive on domains: $\int_a^b + \int_b^c = \int_a^c$ for any a, b, c in the domain interval of f.

2.2.2. *Fundamental theorem of Calculus.* Let $f \in C([a,b],F)$. The integral function

$$F(t) := \int_a^t f(s)\,ds$$

is in $C^1([a,b],F)$ and $F'(t) = f(t)$.

For the proof, we compute

$$F(t+h) - F(t) - f(t)h = \int_t^{t+h} (f(s) - f(t))\,ds.$$

The last term is clearly bounded by $|h| \max_{s\in[t,t+h]} |f(s) - f(t)| = o(h)$.

The one variable function $f \in C([a,b],F)$ are called curves in F. We say $f \in C^1([a,b],F)$ if f can be extended to a C^1 map over some larger open interval. Since $f'(x) \in L(\mathbb{R},F) \cong F$, in this case we usually identify $f'(x)$ as an element in F. There is an equivalent form of the fundamental theorem

$$f(t) - f(a) = \int_a^t f'(s)\,ds.$$

2.2.3. *The mean value theorem.* Let $f : U \to F$ and $\overline{xy} \subset U$. Assume that $f'(c)$ exists for all $c \in \overline{xy}$. Then $|f(x) - f(y)| \leq |f'(c)||x - y|$ for some $c \in \overline{xy}$.

There are two standard proofs. The first one requires that $f \in C^1(U, F)$:

$$\begin{aligned} f(x) - f(y) &= \int_0^1 \frac{d}{dt} f(y + t(x - y))\, dt \\ &= \int_0^1 f'(y + t(x - y))(x - y)\, dt. \end{aligned}$$

Hence that $|f(x) - f(y)| \leq \max_{c \in \overline{xy}} |f'(c)||x - y|$.

The second proof is to apply a linear functional $\lambda \in L(F, \mathbb{R})$ to f and then apply the mean value theorem in Calculus to

$$g(t) := \lambda f(y + t(x - y)).$$

Then

$$\lambda(f(x) - f(y)) = g(1) - g(0) = g'(\xi) = \lambda f'(y + \xi(x - y))(x - y).$$

When $F = \mathbb{R}^n$, we take $\lambda = \langle *, v\rangle$ with v the unit vector in $f(x) - f(y)$ to conclude. For normed spaces we apply the Hahn-Banach theorem to get a linear functional λ with $\lambda(f(x) - f(y)) = |f(x) - f(y)|$ and $|\lambda| = 1$. This proof does not require f to be C^1 along $\overline{xy}$.

In practice the integral representation in the first proof is more powerful and precise in doing estimates. We also call it the mean value theorem.

2.3. **Higher derivatives.** Given a C^1 map $f \in C^1(U, F)$, if the function

$$f' : U \to L(E, F); \quad x \mapsto f'(x)$$

is differentiable at $x \in U$ then we denote the derivative by

$$D(f')(x) \equiv D^2 f(x) \equiv f''(x) \equiv f^{(2)}(x) \in L(E, L(E, F)) = L(E, E; F)$$

and call it the second derivative. $f \in C^2(U, F)$ if $f' \in C^1(U, L(E, F))$.

2.3.1. *Symmetry of $f''(x)$.* The bilinear map $f''(x)$ is not necessarily symmetric. But if $f \in C^2(U, F)$ then it is: $f''(x)(h, k) = f''(x)(k, h)$ for all $h, h \in E$ and $x \in U$. The proof follows form the beautiful formula

$$f''(x)(h.k) = \lim_{t,s \to 0} \frac{f(x + th + sk) - f(x + th) - f(x + sk) + f(x)}{ts}.$$

This is easily seen by repeated applications of the mean value theorem:

$$\begin{aligned}
&f(x+th+sk) - f(x+th) - (f(x+sk) - f(x)) \\
&= \int_0^1 f'(x+th+vsk)(sk)\,dv - \int_0^1 f'(x+vsk)(sk)\,dv \\
&= s\int_0^1 (f'(x+th+vsk) - f'(x+vsk))(k)\,dv \\
&= ts\int_0^1\int_0^1 f''(x+uth+vsk)(h,k)\,du\,dv,
\end{aligned}$$

and notice that the limit of any symmetric expression is necessarily symmetry. We will denote the space of F-valued symmetric continuous k-linear maps by $S^k(E,F)$.

Notice that according to our identification of multilinear maps, the above h corresponds to the second derivative and k to the first.

2.3.2. *Symmetry of $f^{(p)}(x)$.* f is called C^p if $f^{(p-1)}$ is C^1 and then we denote by $f^{(p)}(x) = Df^{(p-1)}(x)$. $f^{(p)}(x)$ is a continuous p-multilinear map from E^p to F.

Exercise 10. *If $f \in C^p(U,F)$ show that $f^{(p)}(x) \in S^p(E,F)$.*

2.3.3. *Taylor expansion.* It should now be clear that the usual Taylor expansion holds for C^{p+1} maps. Let $h^{(k)}$ be the k-tuple $(h,h,\ldots,h) \in E^k$ then

$$f(x+h) = \sum_{k=0}^{p} \frac{1}{k!} f^{(k)}(x)h^{(k)} + \int_0^1 \frac{(1-t)^p}{p!} f^{(p+1)}(x+th)h^{(p+1)}\,dt.$$

Exercise 11. *Prove the Taylor expansion and show that remaining term R_{p+1} satisfies $|R_{p+1}| \le \max_{t\in[0,1]} |f^{(p+1)}(x+th)||h|^{p+1}/(p+1)!$.*

2.4. **Partial derivatives.** Let $E = E_1 \times \cdots \times E_n \to F$ be a product of Banach spaces, $U = U_1 \times \cdots \times U_n \subset E$ with U_i open in E_i. Given a function $f : U \to F$, $x = (x_1,\ldots,x_n) \in U$ we may define partial derivatives $D_i f(x) \equiv f_i(x) \in L(E_i,F)$ as the derivative at x of the restriction map keeping all other variables fixed.

2.4.1. *Total differential.* The derivative $f'(x)$ exists implies that $D_i f(x)$ exists and $D_i f(x)h_i = f'(x)(0,\ldots,0,h_i,0,\ldots,0)$ for all i. We abbreviate the notation by writing it as $f'(x)[h_i]_i$ if no confusion will arise.

As a partial converse, if $D_i f \in C(U, L(E_i,F))$ for $1 \le i \le n-1$ and $D_n f(x)$ exists, then $f'(x)$ exists and

$$f'(x)(h_1,\ldots,h_n) = D_1 f(x)h_1 + \cdots + D_n f(x)h_n.$$

Let us demonstrate the case $n = 2$, the general case is similar. Let $h = (h_1, h_2)$.

$$
\begin{aligned}
&f(x+h) - f(x) = f(x_1 + h_1, x_1 + h_2) - f(x_1, x_2) \\
&= f(x_1 + h_1, x_2 + h_2) - f(x_1, x_2 + h_2) + f(x_1, x_2 + h_2) - f(x_1, x_2) \\
&= \int_0^1 D_1 f(x_1 + th_1, x_2 + h_2) h_1 \, dt + D_2 f(x) h_2 + o(h_2) \\
&= D_1(x) h_1 + D_2 f(x) h_2 + o(h_2) + g(h) h_1.
\end{aligned}
$$

Since $\lim_{h\to 0} g(h_2) = 0$, the last term is also in $o(h)$.

Also, $f \in C^p(U, F)$ if and only if $D_i f \in C^{p-1}(U, L(E_i, F))$ for all i.

2.4.2. *Mixed partials and symmetries.* We may define the second order partial derivatives $D_{ij} f(x) := D_i D_j f(x) = f_{ji}(x) \in L(E_i, L(E_j, F)) = L(E_i, E_j; F)$ and higher order partial derivatives $D_{ijk\cdots z} f(x) = f_{z\cdots kji}(x)$ in the obvious manner.

The good thing is, when f is C^p, all its partial derivatives up to order p are insensitive to the order of differentiations. This is easy, we simply repeatedly apply the formula $D_i g(x) h_i = Dg(x)[h_i]_i$ to various partial derivatives g to achieve that

$$D_{i_1\cdots i_p} f(x)(h_1, \ldots, h_p) = D^p f(x)([h_1]_{i_1}, \ldots, [h_p]_{i_p})$$

and then use the symmetry property of $D^p f(x)$ to conclude.

Exercise 12. *Let $f \in C([a, b] \times U, F)$ with $D_2 f$ continuous. Let $g(x) = \int_a^b f(t, x)\, dt$. Show that g is C^1 and $Dg(x) = \int_a^b D_2 f(t, x)\, dt$.*

2.4.3. *Matrix representation.* By viewing h as a column vector with components h_i, the total differential equation shows that $f'(x)$ is represented by a row vector $(D_1 f(x), \ldots, D_n f(x))$. If $F = F_1 \times \cdots \times F_m$ with $f(x) = (f^1(x), \cdots, f^m(x))^t$ as a column vector, then $f'(x) \in L(\prod_{j=1}^n E_j, \prod_{i=1}^n F_i) = \prod_{i,j} L(E_j, F_i)$ is represented by an $m \times n$ matrix $T = (T_{ij})$ with $T_{ij} = D_j f^i(x) \in L(E_j, F_i)$.

3. Inverse Function Theorem

3.1. **Contraction mapping principle.** *Let X be a complete metric space and $f : X \to X$ a mapping such that for all $x, y \in X$,*

$$d(f(x), f(y)) \le \rho\, d(x, y) \quad \textit{for some } \rho < 1.$$

Then there is a unique point x with $f(x) = x$.

To prove this, start with any x_0, define a sequence of points by $x_{n+1} = f(x_n)$. Then $d(x_{n+1}, x_n) < \rho^n d(x_1, x_0)$. By the triangle inequality we have

$$d(x_{n+m}, x_n) \le (\rho^{n+m-1} + \cdots + \rho^n) d(x_1, x_0) < \frac{\rho^n}{1-\rho} d(x_1, x_0).$$

This implies that $\{x_n\}$ is a Cauchy sequence hence it converges to a limit x. Since a contraction map is continuous, we take limits to the equation $x_{n+1} = f(x_n)$ to get $x = f(x)$. For uniqueness, if $y = f(y)$ then $d(x, y) = d(f(x), f(y)) < \rho d(x, y)$ implies that $d(x, y) = 0$ and hence $x = y$.

The fixed point depends continuously on the map f: *Let X be a complete metric space and $f : X \times T \to X$ be a continuous map with*

$$d(f(x,t), f(y,t)) \le \rho\, d(x,y)$$

for some $\rho < 1$. Let $x(t)$ be the fixed point of $f_t = f|_{X\times\{t\}}$, i.e. $f_t(a) = f(a,t)$, then $x : T \to X$ is continuous.

Triangle inequality implies that

$$d(f_t^n(x(s)), x(s)) \le \sum_{i=1}^{n} d(f_t^i(x(s)), f_t^{i-1}(x(s))) \le \frac{1}{1-\rho} d(f_t(x(s)), x(s)).$$

Since $x(t) = \lim_{n\to\infty} f_t^n(a)$ for any $a \in X$, by taking limits we get

$$d(x(t), x(s)) \le C d(f_t(x(s)), x(s)).$$

For s fixed, $d(f_t(x(s)), x(s)) = d(f(x(s), t), x(s)) \to 0$ when $t \to s$, so the continuity of x at s follows.

3.2. **Inverse function theorem.** *Let $f : (U \subset E) \to F$ be a C^p map ($p \ge 1$) between Banach spaces and $f'(x_0) : E \to F$ be an isomorphism, then f is locally C^p invertible at x_0.*

We may assume that $x_0 = 0$ and $f(0) = 0$ By identifying F with E through $f'(0)$ we may assume that $E = F$ and $f'(0) = I_E$.

The idea is, given $y \in E$, if there is an x such that $y = f(x)$ then

$$x = y + x - f(x) =: y + g(x) =: g_y(x).$$

That is, x is a fixed point of g_y. Since $g_y' = g'$ and $g'(0) = 0$, we may pick $r > 0$ such that $|g_y'(x)| = |g'(x)| < \frac{1}{2}$ for $x \in \bar{B}_r(0)$. Then $|g_y(a) - g_y(b)| \le \frac{1}{2}|a - b|$. Thus if g_y maps $\bar{B}_r(0)$ into $\bar{B}_r(0)$ we will get a unique fixed point.

For $y \in \bar{B}_{r/2}(0)$ and $x \in \bar{B}_r(0)$, we compute (notice that $g(0) = 0$)

$$|g_y(x)| \le |y| + |g(x)| < \frac{r}{2} + |g'(c)||x| < r.$$

Thus we get a well defined inverse $f^{-1} : \bar{B}_{r/2}(0) \to \bar{B}_r(0)$. The continuity of f^{-1} follows from the continuous dependence of fixed points on g_y, hence on y. But here a much precise argument exists: Let $y_i = f(x_i)$, then

$$|x_1 - x_2| = |y_1 - y_2 + g(x_1) - g(x_2)| < |y_1 - y_2| + \frac{1}{2}|x_1 - x_2|,$$

hence $|x_1 - x_2| < 2|y_1 - y_2|$ and f^{-1} is continuous.

From $|x_1| < 2|y_1|$ we also get that $V := f^{-1}(B_{r/2})(0)$ is open in $B_r(0)$. So $f|_V$ is a homeomorphism onto $B_{r/2}(0)$. We claim that $f^{-1} \in C^p(B_{r/2}(0), E)$.

Let $y = f(x)$ and $b = f(a)$, then $f(x) - f(a) - f'(a)(x-a) = o(x-a)$. Thus

$$f(a)^{-1}(y - b) - (f^{-1}(y) - f^{-1}(b)) = f(a)^{-1} o(x - a).$$

Since

$$\lim_{y \to b} \frac{|o(x-a)|}{|y-b|} = \lim_{y \to b} \frac{|o(x-a)|}{|x-a|} \frac{|x-a|}{|y-b|} \le 0 \cdot \frac{1}{2} = 0,$$

we get

$$(f^{-1})'(b) = f'(f^{-1}(b))^{-1}.$$

The mapping of taking inverse is C^∞ and f^{-1} is C^0, this implies that f^{-1} is C^1. Inductively we then conclude that f^{-1} is C^p.

3.3. **Implicit function theorem.** *Let $U \subset E$, $V \subset F$ be open sets in Banach spaces and let $f : U \times V \to G$ be a C^p map. Assume that $D_2 f(a, b) \in L(F, G)$ is an isomorphism. Then there is a unique continuous map $g : U_0 \to V$ on a possibly smaller open set $a \in U_0$ such that $g(a) = b$ and $f(x, g(x)) = f(a, b)$. Moreover g is also C^p.*

By identifying F and G via $D_2 f(a, b)$ we may assume that $G = F$ and $D_2 f(a, b) = I_F$. We may also assume that $f(a, b) = 0$. Consider the map

$$\phi : U \times V \to E \times F \quad \text{via} \quad \phi(x, y) = (x, f(x, y)).$$

Then $\phi'(a, b)$ is represented by $\begin{pmatrix} I_E & 0 \\ D_1 f(a, b) & I_F \end{pmatrix}$, which is clearly an isomorphism on $E \times F$. By the inverse function theorem the C^p local inverse ψ exists and we can write $\psi(x, z) = (x, h(x, z))$ for some C^p map h. Apply ϕ to it we get

$$(x, z) = \phi\psi(x, z) = \phi(x, h(x, z)) = (x, f(x, h(x, z))).$$

Thus $z = f(x, h(x, z))$. Let $g(x) = h(x, 0)$ then we get $f(x, g(x)) = 0$.

Exercise 13. *Verify the uniqueness part of the implicit function theorem.*

4. Ordinary Differential Equations

4.1. **The backgraound.** Let $f : U \to E$ be a C^p map in Banach spaces E with U open in E. We also call f a C^p vector field on U. We want to solve integral curves $\alpha_x : I \to U$ on a maximal interval $I_x = (a_x, b_x)$, $(a_x, b_x) \subset \mathbb{R} \cup \{\pm\infty\}$, such that

$$\alpha_x'(t) = f(\alpha_x(t)) \quad \text{for } t \in I_x.$$

and with given initial data $\alpha_x(0) = x$. As we will see shortly that the existence and uniqueness of this ordinary differentia equation follows readily from the contraction mapping principle and moreover α_x depends continuously on x.

By putting x as a parameter we write $\alpha(t, x) = \alpha_x(t)$ and call it the local flow generated by f. Namely for each t, we get a local homeomorphism ϕ_t on U, namely $\phi_t : U_t \to U_{-t}$ where $U_t = \{x \in U \mid t \in I_x\}$ is the open subset that that can be moved by the integral curve up to time t.

Up to this point all we just says are standard material in any first course in differential equations when $E = \mathbb{R}^n$. And in that case we only need f to be Lipschitz. However, the further smoothness assumption C^p will in fact imply the flow to be C^p too. Unfortunately this is much harder to prove.

There exist at least two proofs of this smooth dependence result. The old one is along the classical methods and deal with non-trivial estimates. A beautifully written proof can be found in the book by Hirsch and Smale.

The second proof is due to Pugh and Robbin in 1968 which uses the Banach implicit function theorem. The issue here is that it is conceptually very simple to get the result and this is why we want to discuss it in this notes. Here we follows the exposition of Lang in his textbook "Real Analysis" closely, with simplification on the proof of continuity of $D_2\alpha$ made by the author.

Far reaching generalization to Fréchet spaces was due to Nash in 1960's and now emerges into the so called Nash-Moser inverse function theorem. The reader may consult Harmilton's beautiful article in Bulletin AMS, 1982, for more details.

4.2. **The Lipschitz case.** We can allow f to be time dependent. Let $0 \in J$ be a fixed interval. $f : J \times U \to E$ a map with $|f(t, x) - f(t, y)| \le K|x - y|$ for all $x, y \in U$ and $t \in J$, where $K \ge 1$ is a fixed constant. Assume that $|f| \le M$ for some $M \ge 1$.

For any $z \in U$, let $B_{3r}(z) \subset U$ with $0 < r < 1$ and let $b < r/MK$, then there is a unique continuous flow $\alpha : J_b \times B_r(z) \to U$. (We set $I_b = [-b, b]$ and $J_b = (-b, b)$.)

Let $x \in \bar{B}_r(z)$ and $X = C(I_b, \bar{B}_{2r}(z))$, which is Banach.

$$(S_x\alpha)(t) := x + \int_0^t f(u, \alpha(u))\, du$$

define a map $S_x : X \to X$ since $|S_x\alpha| < r + bM < 2r$. Also $bK < r/M < 1$ and

$$|S_x\alpha - S_x\beta| \le b \sup |f(u, \alpha(u)) - f(u, \beta(u))| \le bK|\alpha - \beta|,$$

hence there exist a unique fixed point $\alpha_x \in X$ of S_x. The Lipschitz continuity of α_x in x is standard (as was shown before): Since $|S_x\alpha_y - \alpha_y| = |x - y|$,

$$\begin{aligned} |S_x^n\alpha_y - \alpha_y| &\le \sum_{i=i}^{n} |S_x^i\alpha_y - S_x^{i-1}\alpha_y| \\ &\le ((bk)^n + \cdots + (bk) + 1)|x - y| \le C|x - y| \end{aligned}$$

with $C = 1/bK$. Take $n \to \infty$ we get $|\alpha_x - \alpha_y| < C|x - y|$. Finally

$$\begin{aligned} |\alpha(t, x) - \alpha(s, y)| &\le |\alpha(t, x) - \alpha(t, y)| + |\alpha_y(t) - \alpha_y(s)| \\ &\le C|x - y| + |\alpha_y(t) - \alpha_y(s)| \end{aligned}$$

implies the continuity of the flow α.

The uniqueness statement is easy and is left to the reader.

4.3. **Reduction to time/parameter independent case.** Although we prove the result for f depends on time, it is indeed equivalent to the time-independent case. Indeed, given a time-dependent vector field $f : J \times U \to E$, we define a time -independent vector field $\bar{f} : J \times U \to \mathbb{R} \times E$ by $\bar{f}(s, x) = (1, f(s, x))$ over $J \times U$.

Let $\bar{\alpha} = (\beta, \gamma)$ be its flow. Namely

$$\bar{\alpha}'(t, (s, x)) = \bar{f}(\bar{\alpha}(t, (s, x))) \quad \text{and} \quad \bar{\alpha}(0, (s, x)) = (s, x).$$

Then $\beta'(t, (s, x)) = 1$ with $\beta(0, (s, x)) = s$ shows that $\beta(t, (s, x)) = t+s$. Thus

$$\bar{f}(\bar{\alpha}(t, (s, x))) = \bar{f}(t + s, \gamma(t, (s, x))) = (1, f(t + s, \gamma(t, (s, x)))$$

and then γ satisfies

$$\gamma'(t, (s, x)) = f(t + s, \gamma(t, (s, x)) \quad \text{and} \quad \gamma(0, (s, x)) = x.$$

The map $\alpha(t, x) := \gamma(t, (0, x))$ is then the flow of the original f.

Now we consider a time/parameter-dependent vector field

$$\tilde{f} : J \times V \times U \to E.$$

with parameter space V being open in some Banach space B.

Let $f : J \times (V \times U) \to F \times E$ be the time-dependent vector field on $V \times U$ defined by $f(t,(z,x)) = (0, \tilde{f}(t,z,x))$ and $\alpha = (\beta, \gamma)$ be its flow. Then

$$\begin{aligned}(\beta'(t,(z,x)), \gamma'(t,(z,x))) &= (0, F(t,z,x)),\\ (\beta(0,(z,x)), \gamma(0,(z,x)) &= (z,x)\end{aligned}$$

implies that $\tilde{\alpha}(t,z,x) := \gamma(t,(z,x))$ is the flow of the original $\tilde{f}$.

Notice that the above reduction preserve the C^p condition of the vector field as well as the flow.

4.4. **Local smoothness of the flow: The C^p case.** Given a C^p vector field $f : U \to E$ with U open in E. Let $F = C^0(I_b, E)$ be the Banach space under the sup norm and $V = C^0(I_b, U)$ be an open subset. Consider the map $T : U \times V \to F$

$$T(x,\sigma)(t) := x + \int_0^t f(\sigma(u))\,du - \sigma(t).$$

Notice that the solution (x,σ) of the equation $T(x,\sigma) = 0$ is precisely the integral curve $\sigma = \alpha_x$. We already have the continuous implicit function $x \mapsto \alpha_x$ locally. The plan is to use implicit function theorem to show that this map is indeed C^p. For this purpose we need to first show that T is a C^p map.

It is clear that $D_1T(x,\sigma) = I_E$. We claim that

$$D_2T(x,\sigma) = \int_0^* f' \circ \sigma - I_F.$$

We only need to handle the main term, namely by the mean value theorem

$$\begin{aligned}&\left|\int f\circ(\sigma+h) - \int f\circ\sigma - \int (f'\circ\sigma)h\right|\\ &\le \int |f(\sigma(u)+h(u)) - f(\sigma(u)) - f'(\sigma(u))h(u)|\,du\\ &\le |h| \int \sup|f'(z_u) - f'(\sigma(u))|\,du,\end{aligned}$$

where z_u is a point inside the line segment between $\sigma(u)$ and $\sigma(u)+h(u)$. Since $\sigma(I_b)$ is compact and f' is continuous, we see that the sup goes to zero when $|h| \to 0$.

We need to show that D_2T is continuous to conclude that T is C^1. We compute

$$D_2T(x,\sigma) - D_2T(y,\tau) = \int (f'(\sigma(u)) - f'(\tau(u)))\, du.$$

This is small when $|\sigma - \tau|$ is small since $\sigma(I_b)$ is compact and f' is continuous.

Exercise 14. *Fill in the details of the above two arguments.*

There is no difficulty to handle the C^p case because the shape of D_2T is essentially the same as T, and in fact even easier since it does not involve x. So by the same argument and induction we get that D_2T is C^{p-1} and hence T is C^p.

Now we focus on a point (x_0, σ_0) with $T(x_0, \sigma_0) = 0$ and let $|f'| \leq C$ on $B_r(x_0)$. Pick $b < 1/C$, then

$$|D_2T(x_0, \sigma_0) + I_F| = \left| \int f' \circ \sigma_0 \right| \leq bC < 1.$$

This implies that $D_2T(x_0, \sigma_0) \in L(F, F)$ is an isomorphism. By the implicit function theorem we thus proved in a neighborhood of x_0 the map $x \mapsto \alpha_x$ is C^p.

It remains to show that the (continuous) flow $\alpha : J_b \times B_r(x_0) \to U$ is C^p with possibly smaller r and b. It suffices to show that $D_1\alpha$ and $D_2\alpha$ are both C^{p-1}.

We do the case $p = 1$ first. Since $D_1\alpha(t, x) = f(\alpha(t, x))$, which is a composition of continuous functions, we get the continuity of $D_1\alpha$.

For $D_2\alpha$, we need to first show that it exists. let ϕ be the derivative of $x \mapsto \alpha_x$. So $\phi : B_r(x_0) \to L(E, C(I_b, E)) = L(E, F)$ is C^{p-1}. Then $\alpha_{x+h} - \alpha_x = \phi(x)h + |h|\psi(h)$ with $\lim_{h\to 0} \psi(h) = 0$. So

$$\alpha(t, x + h) - \alpha(t, x) = (\phi(x)h)(t) + |h|\phi(h)(t).$$

Hence $D_2\alpha$ exists and $D_2\alpha(t, x)h = (\phi(x)h)(t)$. In particular, $D_2\alpha(t, x)h$ is continuous for any h. For any fixed h, we have an equation for functions in (t, s):

$$\alpha(t, x + sh) = (x + sh) + \int_0^t f(\alpha(u, x + sh))\, du.$$

We differentiate s under integral sign to get (why can we do so!)

$$D_2\alpha(t, x + sh)h = h + \int_0^t f'(\alpha(u, x + sh))D_2\alpha(u, x + sh)h\, du.$$

For $s = 0$ this gives

$$D_2\alpha(t,x)h = h + \int_0^t f'(\alpha(u,x))D_2\alpha(u,x)h\,du.$$

Then for any h with $|h| = 1$ we compute

$$\begin{aligned}
&|D_2\alpha(t',x')h - D_2\alpha(t,x)h| \\
&\quad = |(\phi(x')h)(t') - (\phi(x)h)(t)| \\
&\quad \le |((\phi(x') - \phi(x))h)(t')| + |(\phi(x)h)(t') - (\phi(x)h)(t)|.
\end{aligned}$$

The first term is a good term since it is bounded by $|\phi(x') - \phi(x)|$. The second term is

$$\left| \int_t^{t'} f'(\alpha(u,x))(\phi(x)h)(u)\,du \right| \le |t' - t| C |\phi(x)|.$$

These estimates shows that $|D_2\alpha(t',x') - D_2\alpha(t,x)| \to 0$ when $(t',x') \to (t,x)$. This finish the proof that the flow α is C^1.

For the C^p case, we prove by induction on p and we may suppose that the theorem holds for the C^{p-1} case. That is, α is C^{p-1}.

Since $D_2\alpha$ is continuous, we may differentiate the integral equation to get

$$D_2\alpha(t,x) = I_E + \int_0^t f'(\alpha(u,x))D_2\alpha(u,x)\,du.$$

Equivalently $D_2\alpha$ satisfy the linear ordinary differential equation

$$\frac{d}{dt}D_2\alpha(t,x) = f'(\alpha(t,x))D_2\alpha(t,x), \quad D_2(0,x) = I_E.$$

Since $f'(\alpha(t,x))$ is C^{p-1} by the induction hypothesis, we see that $D_2\alpha$ is C^{p-1}. Here one may wonder the coefficient $f'(\alpha(t,x))$ depends on both time and parameter. This is OK since we have shown that such cases can be reduced to the time/parameter independent case.

Exercise 15. *Give a direct proof of the theorem for time/parameter dependent linear equations.*

For $D_1\alpha$, since it satisfies $D_1\alpha(t,x) = f(\alpha(t,x))$. The induction hypotheses shows that $D_1\alpha$ is C^{p-1} too. Putting together we get that α is C^p and the proof is completed.