Student Seminar on Mathematical Physics

Prof. Vojkan Jaksic, McGill University

Course notes by Dana Mendelson and Mireille Prévost

Winter 2010

Contents

1 Differentiation on Banach Spaces

We consider here Banach spaces over $\mathbb{C}$ or $\mathbb{R}$. Let E and F be two Banach spaces and $\mathcal{U}$ an open set in E, $x \in \mathcal{U}$. Let $f : \mathcal{U} \to F$ be a function.

Definition. *f is differentiable at x if there exist $A \in \mathcal{L}(E, F)$, the Banach space of all continuous linear maps from E to F, such that for h small enough,*

$$f(x+h) = f(x) + A(h) + \varphi(h),$$

where

$$\lim_{h \to 0} \frac{\|\varphi(h)\|_F}{\|h\|_E} = 0 \qquad h \neq 0 \tag{1}$$

*φ is called the error term and we call A the **Fréchet derivative** of f at x. In the sequel, we denote $A = f'(x) = f'_x$.*

If f is differentiable at x then f is continuous there. Indeed,

$$\|f(x+h) - f(x)\| = \|f'(x) \cdot h + \varphi(h)\| \leq \|f'(x) \cdot h\| + \|\varphi(h)\| \longrightarrow 0$$

as $h \to 0$.

Remark. *Regarding the definition, one needs to show that A is well-defined. Let $A, A' \in \mathcal{L}(E, F)$ be such that*

$$f(x+h) = f(x) + A(h) + \varphi_1(h)$$

and

$$f(x+h) = f(x) + A'(h) + \varphi_2(h),$$

where φ_1 and φ_2 are error terms satisfying (1). Then

$$A(h) - A'(h) = \varphi_2(h) - \varphi_1(h).$$

Let $y \in E$ be fixed, $y \neq 0$ and take $h = ty$ for sufficiently small t. Then

$$A(ty) - A'(ty) = \varphi_2(ty) - \varphi_1(ty)$$

and, using linearity of A, and dividing by $t \neq 0$ we can rewrite this as

$$\|Ay - A'y\| = \frac{1}{t}\|\varphi_2(ty) - \varphi_1(ty)\| = \frac{\|y\|}{\|ty\|}\|\varphi_2(ty) - \varphi_1(ty)\|.$$

Taking $t \to 0$ yields $\|Ay - A'y\| = 0$, hence $Ay = A'y$ for all $y \neq 0$. Since $A \cdot 0 = 0 = A' \cdot 0$ $A = A'$

Definition. *If $f : \mathcal{U} \to F$ is differentiable at every point of $\mathcal{U}$ we say that f is differentiable on $\mathcal{U}$. Then we have*

$$f' : \mathcal{U} \longrightarrow \mathcal{L}(E, F).$$

If further this map is continuous, then we say that f is C^1. If this map is differentiable at x then

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

is the second derivative of f at x. We say that f is C^p on $\mathcal{U}$ if the derivatives up to p exist and are continuous on $\mathcal{U}$.

Example $E = \mathbb{R}$, F arbitrary. Then $f : \mathcal{U} \to F$, $f'_x(1) = v$. So for any $\alpha \in \mathbb{R}$, $f'_x(\alpha) = f'_x(\alpha \cdot 1) = \alpha \cdot v$. So the derivative is identified with a vector $f'_x = v \in F$ and the corresponding linear map is just scalar multiplication.

In general, one identifies, $\mathcal{L}(\mathbb{R}, F)$ and F and so f' can be again considered as a map from a subset of $\mathbb{R}$ to F.

Recall in MATH 248, $E = \mathbb{R}^n$, $F = \mathbb{R}$ and $f = (f_1, \dots, f_n)$ is differentiable at x then $\partial f / \partial x_j$ exist and

$$f'_x = \nabla f(x) = \left[\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n}\right]$$

and

$$f'_x(h) = \nabla f(x) \cdot h$$

where $\cdot$ denotes the scalar product. If $\mathcal{U} \subseteq \mathbb{R}^n$, $f : \mathcal{U} \longrightarrow \mathbb{R}^m$, $f = (f_1, \dots, f_n)$ and if f is differentiable at $x \in \mathcal{U}$, then all the f_j are differentiable at x and

$$f'_x = \begin{bmatrix} \nabla f_1(x) \\ \vdots \\ \nabla f_n(x) \end{bmatrix}.$$

1.1 Properties of the derivative

Proposition 1. *If f and g are differentiable at x then so are $f + g$ and cf, and*

$$(f+g)'(x) = f'(x) + g'(x)$$
$$(cf)'(x) = cf'(x).$$

Proof. We write

$$f(x+h) = f(x) = f'_x h + \varphi_1(h)$$

and

$$g(x+h) = g(x) + g'_x h + \varphi_2(h)$$

where the φ_i, $i = 1, 2$ satisfy (1). Then

$$(f+g)(x+h) = f(x+h) + g(x+h) = (f+g)(x) + (f'_x + g'_x)(h) + \varphi_1(h) + \varphi_2(h).$$

Since

$$\lim_{h \to 0} \frac{\|\varphi_1(h) + \varphi_2(h)\|}{\|h\|} \leq \lim_{h \to 0} \frac{\|\varphi_1(h)\| + \|\varphi_2(h)\|}{\|h\|} = 0,$$

by uniqueness of the derivative, $(f+g)'(x) = f'(x) + g'(x)$. □

Proposition 2 (Product rule)**.** *Let F_1, F_2, F be three Banach spaces and $\cdot : F_1 \times F_2 \to F$ a continuous bilinear map. Let $\mathcal{U} \subset E_1$ and $f : \mathcal{U} \to F_1$, $g : \mathcal{U} \to F_2$, two maps that are differentaible at $x \in \mathcal{U}$. Then $f \cdot g : \mathcal{U} \to F$ is differentiable at x and*

$$(f \cdot g)'(x) = f'(x)g(x) + f(x)g'(x)$$

where the meaning of the right hand side above is

$$(f \cdot g)'(x)(v) = (f'(x)v)g(x) + f(x)(g'(x)v).$$

Proof.

$$f(x+h) = f(x) + f'_x h + \varphi_1(h)$$

and

$$g(x+h) = g(x) + g'_x h + \varphi_2(h)$$

where the φ_i, $i = 1,2$ satisfy (1). Then using the bilinearity of $\cdot$, we obtain

$$(f \cdot g)(x+h) = (f \cdot g)(x) + (f'(x)h) \cdot g(x) + (g'(x)h) \cdot f(x) + R,$$

where

$$R = \varphi_1(h) \cdot (g(x) + g'_x h) + (f(x) + f'_x h) \cdot \varphi_2(h).$$

Using bilinearity and continuity of $\cdot$ and that g'_x is bounded

$$\lim_{h \to 0} \frac{\|\varphi_1(h) \cdot (g(x) + g'_x h)\|}{\|h\|} = \lim_{h \to 0} \left\| \frac{\varphi_1(h)}{\|h\|} \cdot (g(x) + g'_x h) \right\| = 0$$

and similarly for the other error term. □

The product rule can be used to derive the formula for integration by parts of Banach space valued functions: Let f and g be piecewise continuous, then by continuity of $\cdot$, $f \cdot g$ is piecewise continuous and we can consider the Banach valued Riemann integral of this function. Supposing $a + tb$, $t \in [0,1]$ is in $\mathcal{U}$, we may use the product rule and a version of the fundamental theorem of calculus (which shall be proven later) to write

$$(f \cdot g)(b) - (f \cdot g)(a) = \int_a^b f'(s)g(s)ds + \int_a^b f(s)g'(s)ds$$

where this is understood upon evaluation at $v \in \mathcal{U}$. Rearranging this relation yields

$$\int_a^b f'(s)g(s)ds = (f \cdot g)(b) - (f \cdot g)(a) + \int_a^b f(s)g'(s)ds.$$

Proposition 3 (Chain Rule)**.** *Let E, F and G be three Banach spaces, $\mathcal{U} \subset E$, $V \subset F$ and $\mathcal{U}, V$ open. Let $f : \mathcal{U} \to V$, $g : V \to G$ be two functions such that f is differentiable at $x \in U$ and g is differentiable at $f(x) \in V$. Then*

$$\ell(x) = (g \circ f)(x), \quad \ell : \mathcal{U} \to G$$

is differentiable at x and

$$\ell'(x) = g'(f(x)) \circ f'(x).$$

Remark. *Just to make sure we have our bearings about us, $f'(x) \in \mathcal{L}(E,F)$, $g'(f(x)) \in \mathcal{L}(F,G)$ so $g'(f(x)) \circ f'(x) \in \mathcal{L}(E,G)$.*

Proof. For h small enough,

$$\begin{aligned} \ell(x+h) - \ell(x) = g(f(x+h)) - g(f(x)) &= g(\underbrace{f(x+h) - f(x)}_{:=K(h)} + f(x)) - g(f(x)) \\ &= g(f(x) + K(h)) - g(f(x) \\ &= g'(f(x))K(h) + \varphi_1(K(h)). \end{aligned}$$

where the error term satisfies (1). If $K(h) = 0$ then $\varphi_1(K(h)) = 0$. For $K \neq 0$, recall that $K(h) := f(x+h) - f(x) = f'(x)h + \varphi_2(h)$, the error term as usual. Thus

$$\begin{aligned} g'(f(x))K(h) + \varphi_1(K(h)) &= g'(f(x))\left[f'(x)h + \varphi_2(h)\right] + \varphi_1(K(h)) \\ &= g'(f(x))f'(x)h + g'(f(x))\varphi_2(h) + \varphi_1(K(h)). \end{aligned}$$

It remains to check that the error terms satisfies our requirement. Since $g'(f(x))$ is bounded

$$\lim_{h\to 0} \frac{\|g'(f(x))\varphi_2(h)\|}{\|h\|} \leq \lim_{h\to 0} \frac{\|g'(f(x))\|\|\varphi_2(h)\|}{\|h\|} = 0.$$

For the second term, choose a sequence $h_n \to 0$, $h_n \neq 0$. If along some subsequence $K(h_n) = 0$ then along that subsequence, $\varphi_1(K(h_n)) = 0$. Without loss of generality, we may assume that $K(h_n) \neq 0$ for all n. Then

$$\frac{\|\varphi_1(K(h_n))\|}{\|h_n\|} = \frac{\|\varphi_1(K(h_n))\|}{\|K(h_n)\|}\frac{\|K(h_n)\|}{\|h_n\|},$$

and

$$\frac{\|K(h_n)\|}{\|h_n\|} = \frac{\|f(x+h_n) - f(x)\|}{\|h_n\|} = \frac{\|f'(x)h_n + \varphi_2(h_n)\|}{\|h_n\|} \leq \|f'(x)\| + \frac{\|\varphi_2(h_n)\|}{\|h_n\|},$$

so

$$\sup_n \frac{\|K(h_n)\|}{\|h_n\|} < \infty.$$

Since $K(h_n) \to 0$ as $h_n \to 0$,

$$\lim_{n\to\infty} \frac{\|\varphi_1(K(h_n))\|}{\|K(h_n)\|} = 0,$$

and the error term vanishes. The result follows by invoking uniqueness. □

Once again, we return to MATH 248 and recall that for $f : \mathbb{R}^n \to \mathbb{R}^p$ and $g : \mathbb{R}^p \to \mathbb{R}^m$, then $(g \circ f)'(x) = g'(f(x)) \circ f'(x)$, where the latter is a composition of linear maps.

Proposition 4. *Let $F_1, \ldots, F_n$ be Banach spaces and let $F := F_1 \times \ldots F_n$, be their product with the maximum norm. Let $\mathcal{U} \subset E$, $f : \mathcal{U} \to F$ be such that $f = (f_1, \ldots, f_n)$ with $f_k : \mathcal{U} \to F_k$. Then f is differentiable at $x \in \mathcal{U}$ if and only if each f_k is a differentiable at $x \in \mathcal{U}$ in which case*

$$f'(x) = (f_1'(x), \ldots, f_n'(x)).$$

Proof. Suppose first that f is differentiable at $x \in \mathcal{U}$ and let $f'(x) \in \mathcal{L}(E, F)$ be the derivative. Let $\pi_k : F \to F_k$ be the projection map. Then π_k is a bounded linear map and

$$A_k = \pi_k \circ f'(x)$$

is a bounded linear map from E to F_k. One writes $f'(x) = (A_1, \ldots, A_n)$. Since f is differentiable,

$$f(x+h) - f(x) = f_x'h + \varphi(h)$$

with φ the usual error term. Then

$$\begin{aligned} f(x+h) - f(x) &= (f_1(x+h) - f_1(x), \dots, f_n(x+h) - f_n(x)) \\ &= (A_1(x)h, \dots, A_n(x)) + (\varphi_1(h), \dots, \varphi_n(h)). \end{aligned}$$

Note that for any $1 \leq k \leq n$,

$$\frac{\|\varphi_k(h)\|}{\|h\|} \leq \max_{1\leq k\leq n} \frac{\|\varphi_k(h)\|}{\|h\|} = \frac{\|\varphi(h)\|}{\|h\|} \longrightarrow 0$$

as $h \to 0$, $h \neq 0$. Then

$$f_k(x+h) - f_k(x) = A_k h + \varphi_k(h)$$

where

$$\lim_{h\to 0} \frac{\|\varphi_k(h)\|}{\|h\|} = 0,$$

so we have that $f_k'(x) = A_k$.

Conversely, if all the f_k are differentiable, with derivative A_k, say, then

$$f_k(x+h) - f_k(x) = A_k h + \varphi_k(h),$$

with $\varphi_k(h)$ the usual error term. We set $f'(x) := (A_1, \dots, A_n)$. This is a map $\mathcal{L}(E, F)$ and we can write

$$f(x+h) - f(x) = (A_1(x)h, \dots, A_n(x)) + (\varphi_1(h), \dots, \varphi_n(h)).$$

Here, the error term vanishes since

$$\lim_{h\to 0} \frac{\|(\varphi_1(h), \dots, \varphi_n(h))\|}{\|h\|} = \lim_{h\to 0} \max_{1\leq k\leq n} \frac{\|\varphi_k(h)\|}{\|h\|} = 0.$$

□

Proposition 5. *Let $A \in \mathcal{L}(E, F)$. Then for any x, $A'(x) = A$.*

Proof.

$$A(x+h) - A(x) = Ah.$$

□

Proposition 6. *Let E, F, G be three Banach spaces, $\mathcal{U} \subset E$ open, $x \in \mathcal{U}$, $f : \mathcal{U} \to F$ differentiable at x and $A \in \mathcal{L}(F, G)$. Then*

$$A \circ f : \mathcal{U} \longrightarrow G$$

is differentiable at x and $(A \circ f)'(x) = A \circ f'(x) \in \mathcal{L}(E, G)$.

Proof. Apply the previous result and the chain rule. □

Proposition 7. *Let $f : [a, b] \to F$ be a map, differentiable on (a, b) and continuous on $[a, b]$. Supposed that $f'(x) = 0$ for all $x \in (a, b)$. Then $f(x) = const$, $f(x) \equiv u$, $u \in F$, for all $x \in [a, b]$.*

Proof. The proof uses the following

Theorem 8 (Hahn-Banach). *Consider $F^* = \mathcal{L}(F, \mathbb{R})$, the Banach space of all continuous linear functionals $x^* : F \to \mathbb{R}$. If $v_1 \neq v_2 \in F$ then there exists $x^* \in F^*$ such that*

$$x^*(v_1) \neq x^*(v_2).$$

Consider now for a given x^* the map $x^* \circ f : [a,b] \to \mathbb{R}$. This function is continuous on $[a,b]$ and differentiable on (a,b) and, by the previous proposition,

$$(x^* \circ f)'(x) = x^* \circ f'(x) = 0$$

so $x^* \circ f$ satisfies the conditions from the mean value theorem of calculus. Hence $x^* \circ f \equiv c$, c a constant, for $x^* \in F^*$. Suppose now that f is not a constant map, that is, there exist $x, y \in [a,b]$ such that $v_1 = f(x) \neq f(y) = v_2$. Choose x^* such that $x^*(v_1) \neq x^*(v_2)$. Then $x^* \circ f$ is not constant. This is a contradiction, so f is constant. □

1.2 Integration on Banach Spaces

Recall from Assignment 1, MATH 355:

Let $f : [a,b] \to F$ be a continuous function. A partition of $[a,b]$ of size n is a sequence of numbers $a = t_0 < \ldots < t_n = b$. We define the diameter of the partition,

$$\operatorname{diam}(P) = \max_{1 \le k \le n} |t_k - t_{k-1}|.$$

Given a partition, one forms

$$S(f,P) = \sum_{k=1}^{n} f(c_k)(t_k - t_{k-1}), \qquad f(c_k) \in F,\ c_k \in [t_{k-1}, t_k].$$

Then

$$\|S(f,P)\| \le \max_{1 \le k \le n} \|f(c_k)\|(b-a) \le \left(\max_{t \in [a,b]} \|f(t)\| \right)(b-a)$$

and

$$\lim_{\substack{n \to \infty \\ \operatorname{diam}(P) \to 0}} S(f,P)$$

exists and is independent of the choice of the points c_k and of the choise of partition, P, in $S(f,P)$. The value of this limit is called the ***Riemann Integral*** of f. It is denoted by

$$\int_a^b f(t)dt \in F.$$

We have seen in the first assignment that

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

is continuous on $[a,b]$ and for $t \in (a,b)$,

$$\lim_{h \to 0} \left\| \frac{F(t+h) - F(t)}{h} - f(t) \right\| = 0.$$

Note that this is equivalent to the definition of derivative we have introduced here. So, $F'(t) = f(t)$.

Theorem 9 (Newton-Leibniz formula)**.** *If $f : [a, b] \to F$ is differentiable on (a, b), continuous on $[a, b]$ and $f'(x)$ is also continuous on $[a, b]$ then*

$$\int_a^b f'(s)ds = f(b) - f(a)$$

Proof. Let

$$F(t) = \int_a^t f'(s)ds$$

and look at $(F(t) - f(t))' = 0$ for all $t \in (a, b)$. Since $F(t) - f(t)$ is continuous on $[a, b]$ by a previous proposition, $F(t) - f(t) \equiv c$, a constant. Since $F(a) = 0$, $c = -f(a)$. Hence

$$F(b) = \int_a^b f'(s)ds = f(b) - f(a),$$

as required. □

Theorem 10 (Mean-Value Theorem)**.** *Let $\mathcal{U}$ be an open set in E, $x, y \in \mathcal{U}$ such that $x+ty, 0 \leq t \leq 1$ is also in $\mathcal{U}$. Let $f : \mathcal{U} \to F$ be a C^1 function. Then*

$$f(x + y) - f(y) = \int_0^1 f'(x + ty)y\, dt = \underbrace{\left[\int_0^1 f'(x + ty)\, dt\right]}_{\in \mathcal{L}(E,F)} y$$

Remark. *The second equality is just a statement of how we may regard the Riemann Integral of a Banach valued function. Indeed, by our $S(f, P)$ formulation,*

$$\begin{aligned}\int_0^1 f'(x + ty)y\, dt &= \lim_{n\to\infty} \sum_{k=1}^n f'(x + \frac{(k-1)}{n}y)y \cdot \frac{1}{n} \\ &= \lim_{n\to\infty} \left(\sum_{k=1}^n f'(x + \frac{(k-1)}{n}y) \cdot \frac{1}{n} \right) \cdot y \\ &= \left[\int_0^1 f'(x + ty)\, dt\right] \cdot y.\end{aligned}$$

Proof. Set $g(t) = f(x+ty)$. Then g is differentiable on $[-\epsilon, 1+\epsilon]$ and $g'(t)$ is continuous. By chain rule $g'(t) = f'(x + ty)y$, so

$$\int_0^1 g'(t)\, dt = \int_0^1 f'(x + ty)y\, dt = g(1) - g(0) = f(x + y) - f(x)$$

□

Corollary 11. *Under the conditions of the Mean Value theorem*

$$\|f(x + y) - f(x)\| \leq \max_{0\leq t\leq 1} \|f'(x + ty)\|.$$

Proof.

$$\begin{aligned}\|f(x+y)-f(x)\| &= \left\|\left[\int_0^1 f'(x+ty)\,dt\right] y\right\| \\ &\le \left\|\left[\int_0^1 f'(x+ty)\,dt\right]\right\| \|y\| \\ &\le \max_{0\le t\le 1} \|f'(x+ty\| \cdot \|y\|\end{aligned}$$

□

1.3 Higher order derivatives

We now tackle the question of defining higher order derivatives so as to avoid working with $\mathcal{L}(E, \mathcal{L}(E, \mathcal{L}(E, \mathcal{L}\ldots)$.

1.3.1 Background

Let E, F, G be three Banach spaces and let $\mathcal{L}(E, F; G)$ be the collection of all continuous bilinear maps, $\ell : E \times F \longrightarrow G$. In general, the map $\ell : E \times F \to G$ is called bilinear if for each fixed y, $x \mapsto \ell(x, y)$ is linear from $E \to G$ and for each fixed x, $y \mapsto \ell(x, y)$ is linear from $F \to G$.

A bilinear map is called continuous if and only if there exists some $c > 0$ such that for all $x \in E$, $y \in F$, $\|\ell(x, y)\| \le c\|x\|\|y\|$. The smallest of such constants is called the norm of ℓ and is denote by $\|\ell\|$.

The collection of all continuous bilinear maps is a vector space and together with the $\|\cdot\|$ norm, $\mathcal{L}(E, F; G)$ is a Banach space. The proof is identical to the case of linear maps.

Proposition 12. *There is a natural isometry between $\mathcal{L}(E, \mathcal{L}(F, G))$ and $\mathcal{L}(E, F; G)$.*

Proof. Let $\lambda \in \mathcal{L}(E, \mathcal{L}(F, G))$. Associate to this λ the map

$$\ell_\lambda(x, y) = \lambda(x)(y) \qquad x \in E,\ y \in F.$$

For a fixed x, $\ell_\lambda(x, y)$ is linear since $\lambda(x) \in \mathcal{L}(F, G)$. For a fixed y it is linear as well since

$$\lambda \longmapsto \lambda(x) \in \mathcal{L}(F, G)$$

is linear on E. Hence ℓ_λ is bilinear and further

$$\|\ell_\lambda(x, y)\|_G = \|\lambda(x)y\|_G \le \|\lambda(x)\|_{\mathcal{L}(F,G)}\|y\|_f \le \|\lambda\|_{\mathcal{L}(E,\mathcal{L}(F,G))}\|x\|_E\|y\|_F.$$

Thus ℓ_λ is continuous and $\|\ell_\lambda\| \le \|\lambda\|$ and obviously $\lambda \mapsto \ell_\lambda$ is a linear map.

Take now $\ell \in \mathcal{L}(E, F; G)$. Associate to this ℓ a linear map

$$\lambda_\ell(x)y = \ell(x, y).$$

Then $\lambda_\ell(x)$ is a linear map from F to G. For a fixed y we have linearity in x so $x \mapsto \lambda_\ell(x)$ is a linear map from E to the vector space of all linear maps from $F \to G$, $\mathcal{L}(F, G)$. Moreover

$$\|\lambda_\ell(x)y\|_G = \|\ell(x, y)\|_G \le \|\ell\|\|x\|_E\|y\|_F.$$

Hence, $\lambda_\ell(x)$ is a bounded operator and $\|\lambda_\ell(x)\| \leq \|\ell\|\|x\|$. But then $x \mapsto \lambda_\ell(x)$ is bounded and $\|\lambda_\ell\| \leq \|\ell\|$.

The maps $\lambda \mapsto \ell_\lambda$ and $\ell \mapsto \lambda_\ell$ are obviously inverses of each other, both are continuous and

$$\|\lambda\| \leq \|\ell_\lambda\| \leq \|\lambda\|,$$

so $\|\lambda\| = \|\ell_\lambda\|$. Thus we have found a linear bijection, which preserves the norm, hence this is our isometry. □

1.3.2 Higher Order Multilinear Maps

Definition. *Let E_1, ..., E_n and F be Banach spaces. A map $\ell : E_1 \times ... \times E_n \longrightarrow F$ is called multilinear if it is linear with respect to each variable separately. ℓ is continuous if and only if there exist $c > 0$ such that for all $x_1 \in E_1, ..., x_n \in E_n$, $\|\ell(x_1, ..., x_n)\|_F \leq c\|x_1\|_{E_1} ... \|x_n\|_{E_n}$. The smallest of those constants c is denoted $\|\ell\|$ and*

$$\|\ell\| = \sup_{\|x_1\| \leq 1, ..., \|x_n\| \leq 1} \|\ell(x_1, ..., x_n)\|$$

The vector space of all continuous multilinear maps, denoted by $\mathcal{L}(E_1, ..., E_n; F)$, equipped with the above norm, is again a Banach space.

Proposition 13. *There is a natural isometry between $\mathcal{L}(E_1, \mathcal{L}(E_2, ..., \mathcal{L}(E_n, F)...))$ and $\mathcal{L}(E_1, ..., E_n; F)$.*

Proof. We have already done the case $n = 2$. By induction, assume that

$$\mathcal{L}(E_1, \mathcal{L}(E_2, ..., \mathcal{L}(E_{n-1}, F)...))) \cong \mathcal{L}(E_1, \mathcal{L}(E_2, ..., E_{n-1}; F))$$

Then given $\lambda \in \mathcal{L}(E_1, \mathcal{L}(E_2, ..., \mathcal{L}(E_n, F)...))$, let $\ell_\lambda(x_1, ..., x_n) = \lambda(x_1)(x_2, ..., x_n)$. Given $\ell \in \mathcal{L}(E_1, \mathcal{L}(E_2, ..., E_n; F))$, let $\lambda_\ell(x_1)(x_2, ..., x_n) = \lambda(x_1, ..., x_n)$. In the same way as the $n = 2$ case, the maps $\lambda \mapsto \ell_\lambda$ and $\ell \mapsto \lambda_\ell$ are inverses of each other and are norm preserving, hence they are the isometry. □

1.3.3 Second Derivative

Recall that if E and F are Banach spaces, $\mathcal{U} \in E$ is open and $f : \mathcal{U} \to F$ is C^2, then $f' : \mathcal{U} \to \mathcal{L}(E, F)$ and $f'' : \mathcal{U} \to \mathcal{L}(E, \mathcal{L}(E, F)) \cong \mathcal{L}(E, E; F)$.

Proposition 14. *Under the above assumptions and identification, the bilinear form $f''(x)$ is symmetric for all $x \in \mathcal{U}$, ie. $f''(x)(u, v) = f''(x)(v, u)$ $\forall u, v \in E$.*

Remark. *If $E = \mathbb{R}^n$, $F = \mathbb{R}$, then $f' \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}) \cong \mathbb{R}^n$ and $f'' \in \mathcal{L}(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n, \mathbb{R})) \cong \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$. This linear map is given by the usual Hessian matrix:*

$$f''(x) = \begin{bmatrix} \frac{\partial^2 f}{\partial^2 x_1} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial^2 x_2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial^2 x_n} \end{bmatrix}$$

The corresponding bilinear form is given by $f''(x)(u,v) = \langle u, f''(x)v\rangle$, with the usual inner product on $\mathbb{R}^n$. The above statement about the symmetry of the bilinear form is then equivalent to the equality of the mixed partial derivatives.

Proof. Let $x \in \mathcal{U}$ and let the open ball of radius r centered at x, be such that $D(x,r) \in \mathcal{U}$. Let $u, v \in \mathcal{U}$ be such that $\|u\| < \frac{r}{2}$ and $\|v\| < \frac{r}{2}$. Consider $S := f(x+u+v) - f(x+u) - f(x+v) + f(x)$. Let $g(x) = f(x+v) - f(x)$. Then by applying the Mean Value Theorem twice:

$$\begin{aligned} S = g(x+u) - g(x) &= \int_0^1 g'(x+tu)udt \\ &= \int_0^1 [f'(x+v+tu) - f'(x+tu)]udt \\ &= \int_0^1 \left[\int_0^1 f''(x+sv+tu)vds\right] udt \\ &= \int_0^1 \left[\int_0^1 f''(x+sv+tu)v - f''(x)vds\right] udt + \int_0^1 \left[\int_0^1 f''(x)vds\right] udt, \end{aligned}$$

where the last term integrates to $(f''(x)v)u = f''(x)(v,u)$.

Next, let $g_1(x) = f(x+u) - f(x)$. Then $S = g_1(x+v) - g_1(x)$, and by noting that now u and v are interchanged in the above argument, we obtain:

$$S = \int_0^1 \left[\int_0^1 f''(x+sv+tu)v - f''(x)vds\right] udt + f''(x)(u,v)$$

Let now $u, v \in E$ be arbitrary. Choose s' and t' non zero but small enough so that $\|s'u\| \leq \frac{r}{2}$ and $\|t'v\| \leq \frac{r}{2}$. Substituting these vectors into the above formulae:

$$\begin{aligned} &\underbrace{s't' \int_0^1 \left[\int_0^1 [f''(x+ss'u+tt'v) - f''(x)]uds\right] vdt}_{I} + s't'f''(x)(u,v) \\ &= \underbrace{s't' \int_0^1 \left[\int_0^1 [f''(x+st'v+ts'u) - f''(x)]vds\right] udt}_{II} + s't'f''(x)(v,u), \end{aligned}$$

where we made use of the biliearity of $f''(x)$. Now

$$\|I\| \leq \|u\|\|v\| \max_{0\leq s,t\leq 1} \|f''(x+ss'u+tt'v) - f''(x)\|,$$

and similarly for $\|II\|$. Since $f''(x)$ is continuous on $\mathcal{U}$, both $\|I\|$, $\|II\| \longrightarrow 0$ as $s' \longrightarrow 0$ and $t' \longrightarrow 0$. This yields our result. □

1.3.4 Second Order Taylor Formula

Proposition 15. *Let E and F be Banach spaces, $\mathcal{U} \in E$ and $f : \mathcal{U} \longrightarrow F$ a C^2 map. Let also $x, y \in \mathcal{U}$ be such that $x + ty \in \mathcal{U}$ for $0 \leq t \leq 1$. Then,*

$$f(x+y) = f(x) + f'(x)y + \frac{1}{2}f''(x)(y,y) + Q(y),$$

where $Q(y)$ is an error term satisfying: $\lim_{y\to 0} \frac{\|Q(y)\|}{\|y\|^2} = 0$

Proof.

$$\begin{aligned}
f(x+y) &= f(x) + \int_0^1 f'(x+ty)ydt \\
&= f(x) + \int_0^1 f'(x)ydt + \int_0^1 [f'(x+ty) - f'(x)]ydt \\
&= f(x) + f'(x)y + \int_0^1 \left[\int_0^1 f''(x+tsy)tyds\right] ydt \\
&= f(x) + f'(x)y + \int_0^1 \left[\int_0^1 [f''(x+tsy) - f''(x)]tyds\right] ydt + \int_0^1 \left[\int_0^1 f''(x)tyds\right] ydt
\end{aligned}$$

The last term integrates to $\int_0^1 (f''(x)ty)ydt = \int_0^1 tf''(x)(y,y)dt = \frac{1}{2}f''(x)(y,y)$, which gives us the desired equation with error term:

$$Q(y) = \int_0^1 \left[\int_0^1 [f''(x+tsy) - f''(x)]tyds\right] ydt$$

Since

$$\|Q(y)\| \leq \max_{0\leq s,t\leq 1} \|f''(x+tsy) - f''(x)\|\|y\|^2,$$

we get, by continuity of f'' at x, that

$$\lim_{y\to 0} \frac{\|Q(y)\|}{\|y\|^2} \leq \lim_{y\to 0} \max_{0\leq s,t\leq 1} \|f''(x+tsy) - f''(x)\| = 0,$$

□

1.3.5 General Taylor Formula

Let E and F be Banach spaces, $\mathcal{U} \in E$ and $f : \mathcal{U} \longrightarrow F$ a C^n map. Let also $x, y \in \mathcal{U}$ be such that $x + ty \in \mathcal{U}$ for $0 \leq t \leq 1$. Then,

$$f(x+y) = f(x) + f'(x)y + \frac{1}{2}f''(x)(y,y) + ... + \frac{1}{n!}f^{(n)}\underbrace{(y,...,y)}_{n-times} + Q(y),$$

where Q(y) is an error term satisfying: $\lim_{y\to 0} \frac{\|Q(y)\|}{\|y\|^n} = 0$

Exercise. Let $g : (-\epsilon, 1+\epsilon) \longrightarrow F$ be defined by $g(t) = f(x+ty)$, . Then g is C^n and $g'(t) = f'(x+ty)y$ with $g'(t) \in \mathcal{L}(\mathbb{R}, F)$. As discussed earlier, we make the usual identification $\mathcal{L}(\mathbb{R}, F) \cong F$. Then $g' : (-\epsilon, 1+\epsilon) \longrightarrow F$. For higher order derivatives, this identification also yields $g^{(k)} : (-\epsilon, 1+\epsilon) \longrightarrow F$, $g^{(k)}(t) = f^{(k)}(x+ty)\underbrace{(y, ..., y)}_{k-times}$. Then,

$$\begin{aligned}
&\frac{d}{dt}\left(g(t) + (1-t)g'(t) + \frac{1}{2}(1-t)^2 g''(t) ... + \frac{1}{(n-1)!}(1-t)^{n-1}g^{(n-1)}(t)\right) \\
&= g'(t) + [-g'(t) + (1-t)g''(t)] + \left[-(1-t)g''(t) + \frac{1}{2}(1-t)^2 g'''(t)\right] + ... \\
&\quad + \left[-\frac{1}{(n-2)!}(1-t)^{n-2}g^{(n-1)}(t) + \frac{1}{(n-1)!}(1-t)^{n-1}g^{(n)}(t)\right] \\
&= \frac{1}{(n-1)!}(1-t)^{n-1}g^{(n)}(t)
\end{aligned}$$

Integrating over t, from 0 to 1, and using the Newton-Leibniz formula, we obtain

$$g(1) - g(0) - g'(0) - \frac{1}{2}g''(0) - ... - \frac{1}{(n-1)!}g^{(n-1)}(0) = \int_0^1 \frac{1}{(n-1)!}(1-t)^{n-1}g^{(n)}(t)dt \quad (2)$$

Rewriting g as a function of f with our identification,

$$\begin{aligned}
f(x+y) &= f(x) + f'(x)y + \frac{1}{2}f''(x)(y,y) + ... + \frac{1}{(n-1)!}f^{(n-1)}(x)y^{n-1} \\
&\quad + \frac{1}{(n-1)!}\int_0^1 (1-t)^{n-1}f^{(n)}(x)y^n dt + \frac{1}{(n-1)!}\left[\int_0^1 (1-t)^{n-1}\left[f^{(n)}(x+ty) - f^{(n)}(x)\right]y^n dt\right] \\
&= f(x) + f'(x)y + \frac{1}{2}f''(x)(y,y) + ... + \frac{1}{n!}f^{(n)}(x)y^n + \frac{1}{(n-1)!}\left[\int_0^1 (1-t)^{n-1}\left[f^{(n)}(x+ty) - f^{(n)}(x)\right] y
\end{aligned}$$

which is the required equation with

$$Q(y) = \frac{1}{(n-1)!}\left[\int_0^1 (1-t)^{n-1}\left[f^{(n)}(x+ty) - f^{(n)}(x)\right]y^n dt\right]$$

satisfying

$$\frac{\|Q(y)\|}{\|y\|^n} \leq \frac{1}{(n-1)!}\max_{0\leq t\leq 1}\left\|(1-t)^{n-1}\left[f^{(n)}(x+ty) - f^{(n)}(x)\right]\right\|$$

Note that if f is $n+1$ times differentiable, we may integrate by parts equation (2) to obtain

$$Q(y) = \frac{1}{n!}\int_0^1 (1-t)^n f^{(n+1)}(x+ty)y^{n+1}dt$$

In both cases, $\lim_{y\to 0}\frac{\|Q(y)\|}{\|y\|^n} = 0$ □

1.4 Isomorphisms

Our next goal will be to derive the Inverse and Implicit Function Theorems. To do so, we first define the notion of isomorphism between Banach spaces. As ususal, let E and F be a Banach spaces. Define

$$G\mathcal{L}(E,F) = \{A \in \mathcal{L}(E,F) : A \text{ is a bijection}\}.$$

Note that if $A \in G\mathcal{L}(E,F)$ then A^{-1} is linear, so by the open mapping theorem, $A^{-1} \in G\mathcal{L}(E,F)$. The elements of $G\mathcal{L}(E,F)$ are the isomorphisms. Also, if $E = F$, then $G\mathcal{L}(E,E)$ forms a group.

Proposition 16. *$G\mathcal{L}(E,F)$ is an open subset of $\mathcal{L}(E,F)$.*

Lemma 17. *Let $x \in G\mathcal{L}(E,E)$ and suppose that $\|X\| < 1$. Then $I + X$ is invertible and $\|X + I\| \leq \frac{1}{1-\|X\|}$.*

Proof. Let

$$S_n = \sum_{k=0}^{n} (-1)^k X^k = I - X + X^2 - ... + (-1)^n X^n$$

For $n > m$,

$$\begin{aligned} \|S_n - S_m\| &\leq \|X\|^{m+1} + ... + \|X\|^n \\ &\leq \|X\|^{m+1}(1 + \|X\| + ... + \|X\|^{n-m-1}) \\ &\leq \frac{\|X\|^{m+1}}{1 - \|X\|} \end{aligned}$$

since $\|X\| < 1$. Hence $\{S_n\}$ is a Cauchy sequence in the Banach space $\mathcal{L}(E,E)$, so converges to $\lim_{n\to\infty} S_n = S$. Summing, we have

$$(I + X)S_n = S_n(I + X) = I + (-1)^n X^{n+1}$$

Since $\|(I + X)S_n - (I + X)S\| \leq \|I + X\|\|S_n - S\| \to 0$ as $n \to \infty$ and $\|X\|^n \to 0$ as $n \to \infty$, we can take the limit to get the inverse:

$$(I + X)S = S(I + X) = I$$

□

Proof. (Of Proposition) Let $A \in G\mathcal{L}(E,F)$ and let $\epsilon = \frac{1}{2\|A^{-1}\|}$. Consider $B \in \mathcal{L}(E,F)$ such that $B \in D(A,\epsilon)$. Write

$$B = A + (B - A) = A(I + A^{-1}(B - A))$$

Then

$$\|A^{-1}(B - A)\| \leq \|A^{-1}\|\|B - A\| < \frac{1}{2}$$

So $A^{-1}(B - A) \in \mathcal{L}(E,E)$ has norm < 1 and by the lemma, $I + A^{-1}(B - A) \in G\mathcal{L}(E,E)$. B being the composition of $A \in G\mathcal{L}(E,F)$ with $I + A^{-1}(B - A)$, it follows that $B \in G\mathcal{L}(E,F)$. □

Proposition 18. *The map $G\mathcal{L}(E,F) \to G\mathcal{L}(F,E)$ given by $A \mapsto A^{-1}$ is a homeomorphism.*

Proof. $A \to A^{-1}$ is obviously a bijection with $(A^{-1})^{-1} = A$ so it is enough to show only that $A \to A^{-1}$ is continuous. Let $\{A_n\}$ be a sequence converging to A in $G\mathcal{L}(E, F)$. Take n large enough so that

$$\|A_n - A\| \leq \frac{1}{2\|A^{-1}\|}$$

Then

$$\|A^{-1}(A_n - A)\| \leq \|A^{-1}\|\|A_n - A\| < \frac{1}{2}$$

The identity $A_n = A + (A_n - A) = A(I + A^{-1}(A_n - A))$ shows that $A_n^{-1} = (I + A^{-1}(A_n - A))^{-1}A^{-1}$, so

$$\|A_n^{-1}\| \leq \|A^{-1}\|\frac{1}{1 - \|A^{-1}(A_n - A)\|} \leq 2\|A^{-1}\|$$

It follows that

$$\begin{aligned}\|A_n^{-1} - A^{-1}\| &= \|A_n^{-1}(A_n - A)A^{-1}\| \\ &\leq \|A_n^{-1}\|\|(A_n - A)\|\|A^{-1}\| \\ &\leq 2\|A^{-1}\|^2\|A_n - A\|,\end{aligned}$$

as required. □

1.5 Derivatives of Inverse Functions

Theorem 19. *Let E and F be Banach spaces, $\mathcal{U} \in E$ and $\mathcal{V} \in F$ open sets and $f : \mathcal{U} \to \mathcal{V}$ a homeomorphism. Suppose that f is differentiable at $x \in \mathcal{U}$. If the derivative $f'(x)$ is an isomorphism, then $f^{-1} : \mathcal{V} \to \mathcal{U}$ is differentiable at $y = f(x)$ and*

$$(f^{-1})'(y) = (f'(f^{-1}(y)))^{-1} = (f'(x))^{-1}$$

Proof. Fix $x \in \mathcal{U}$ and let $y = f(x)$. Take $k \in F$ small enough so that $y + k \in \mathcal{V}$. Then there exist $h \in E$ such that $x + h \in \mathcal{U}$ and $f(x + h) = y + k$. Note also that $h = 0$ if and only if $k = 0$. This gives $k = f(x + h) - f(x)$ so $k = f'(x)h + \phi(h)$, where $\lim_{h\to 0} \frac{\|\phi(h)\|}{\|h\|} = 0$, $h \neq 0$. Then $f'(x)h = k - \phi(h)$ and using that $f'(x)$ is an isomorphism, $h = (f'(x))^{-1}(k - \phi(h)) = (f'(x))^{-1}k - (f'(x))^{-1}\phi(h)$. The relation $h = f^{-1}(y + k) - f^{-1}(y)$ then gives

$$f^{-1}(y + k) - f^{-1}(y) = (f'(x))^{-1}k - (f'(x))^{-1}\phi(h)$$

By definition of the Fréchet derivative, the statement will follow if we show that

$$\lim_{k\to 0} \frac{\|(f'(x))^{-1}\phi(h)\|}{\|k\|} = 0, k \neq 0$$

Since

$$\frac{\|(f'(x))^{-1}\phi(h)\|}{\|k\|} \leq \|(f'(x))^{-1}\|\frac{\|\phi(h)\|}{\|h\|}\frac{\|h\|}{\|k\|}$$

it is enough to show that $\frac{\|h\|}{\|k\|}$ is bounded.

The relation $h = (f'(x))^{-1}k - (f'(x))^{-1}\phi(h)$ implies that $\|(f'(x))^{-1}k\| \geq \|h\| - \|(f'(x))^{-1}\phi(h)\|$. Since also $\|(f'(x))^{-1}k\| \leq \|(f'(x))^{-1}\|\|k\|$ and $\|(f'(x))^{-1}\phi(h)\| \leq \|(f'(x))^{-1}\|\|\phi(h)\|$,

$$\|(f'(x))^{-1}\|\|k\| \geq \|h\| - \|(f'(x))^{-1}\|\|\phi(h)\| = \|h\| \left(1 - \|(f'(x))^{-1}\| \frac{\|\phi(h)\|}{\|h\|}\right)$$

Since $\frac{\|\phi(h)\|}{\|h\|} \to 0$ as $h \to 0$, for h small enough we have

$$\|(f'(x))^{-1}\| \frac{\|\phi(h)\|}{\|h\|} < \frac{1}{2}$$

So $\frac{\|h\|}{\|k\|}$ is bounded for h sufficiently small:

$$\frac{\|h\|}{\|k\|} < 2\|(f'(x))^{-1}\|$$

Since $h \to 0$ as $k \to 0$, this completes the proof. □

1.6 Diffeomorphisms

Let E and F be Banach spaces, $\mathcal{U} \subset E$ and $\mathcal{V} \subset F$ open sets and $f : \mathcal{U} \to \mathcal{V}$ a homeomorphism. f is called C^n diffeomorphism if f and f^{-1} are C^n.

Proposition 20. *A C^1 homeomorphism $f : \mathcal{U} \to \mathcal{V}$ is a C^1 diffeomorphism if and only if $f'(x)$ is an isomorphism for all $x \in \mathcal{U}$.*

Proof. Assume first that f is a C^1 diffeomorphism. Then $f \circ f^{-1} = id_F$ and $f^{-1} \circ f = id_E$. By the chain rule, $f'(x) \circ (f^{-1})'(f(x)) = id_F$ and $(f^{-1})'(f(x)) \circ f'(x) = id_E$, hence $f'(x)$ has both a left and a right inverse, so is an isomorphism for any $x \in \mathcal{U}$. Conversely, assume that f is a C^1 homeomorphism and that $f'(x)$ is an isomorphism for all $x \in \mathcal{U}$. Then $f^{-1} : \mathcal{V} \to \mathcal{U}$ is differentiable and by the previous theorem, $(f^{-1})'(y) = (f'(f^{-1}(y))^{-1})$. So we only need to prove that $y \mapsto (f^{-1})'(y)$ is continuous. But this map is the composition of the three following continuous functions: $y \mapsto f^{-1}(y)$ from $\mathcal{V} \to \mathcal{U}$, $x \mapsto f'(x)$ from $\mathcal{U} \to G\mathcal{L}(E, F)$ and $A \mapsto A^{-1}$ from $G\mathcal{L}(E, F) \to D\mathcal{L}(F, E)$, so is also continuous. □

We can generalize the previous proposition to the following:

Proposition 21. *A C^n homeomorphism $f : \mathcal{U} \to \mathcal{V}$ is C^n diffeomorphism if and only if $f'(x)$ is an isomorphism for all $x \in \mathcal{U}$.*

Exercise. The proof is the same as above, except we need to show that the map $f^{-1} : \mathcal{V} \to \mathcal{U}$ is C^n. To do so, we will show that $A \mapsto A^{-1}$ going from $G\mathcal{L}(E, F) \to G\mathcal{L}(F, E)$ is actually C^∞. Then, $y \mapsto (f^{-1})'(y)$ is again the composition of the three above functions, with $x \mapsto f'(x)$ a C^{n-1} function by assumption and $y \mapsto f^{-1}(y)$ also C^{n-1} by induction. Let us now prove that $A \mapsto A^{-1}$ is C^∞:

Let $F : G\mathcal{L}(E, F) \to G\mathcal{L}(F, E)$ be defined by $F(A) = A^{-1}$. Then

$$\begin{aligned} F(A + H) - F(A) &= (A + H)^{-1} - A^{-1} \\ &= -(A + H)^{-1} H A^{-1} \\ &= -[(A + H)^{-1} - A^{-1}] H A^{-1} - A^{-1} H A^{-1} \end{aligned}$$

Since $\|[(A+H)^{-1} - A^{-1}]HA^{-1}\| \leq \|H\|\|A^{-1}\|\|(A+H)^{-1} - A^{-1}\|$, the continuity of F then gives

$$\lim_{H \to 0} \frac{\|[(A+H)^{-1} - A^{-1}]HA^{-1}\|}{\|H\|} = 0$$

so $F'_A(H) = -A^{-1}HA^{-1}$. For the second derivative:

$$\begin{aligned}
F'_{A+H_2}(H_1) - F'_A(H_1) &= -\left[(A+H_2)^{-1}H_1(A+H_2)^{-1} - A^{-1}H_1A^{-1}\right] \\
&= -\left[\left((A+H_2)^{-1} - A^{-1}\right)H_1(A+H_2)^{-1} + A^{-1}H_1(A+H_2)^{-1} + \right. \\
&\quad \left. A^{-1}H_1\left((A+H_2)^{-1} - A^{-1}\right) - A^{-1}H_1(A+H_2)^{-1}\right] \\
&= -\left[\left((A+H_2)^{-1} - A^{-1}\right)H_1(A+H_2)^{-1} + A^{-1}H_1\left((A+H_2)^{-1} - A^{-1}\right)\right] \\
&= \left((A+H_2)^{-1}H_2A^{-1}\right)H_1(A+H_2)^{-1} + A^{-1}H_1\left((A+H_2)^{-1}H_2A^{-1}\right) \\
&= \left((A+H_2)^{-1} - A^{-1}\right)H_2A^{-1}H_1(A+H_2)^{-1} + \left[A^{-1}H_2A^{-1}H_1(A+H_2)^{-1}\right] \\
&\quad + A^{-1}H_1\left((A+H_2)^{-1} - A^{-1}\right)H_2A^{-1} + A^{-1}H_1A^{-1}H_2A^{-1} \\
&= \left((A+H_2)^{-1} - A^{-1}\right)H_2A^{-1}H_1(A+H_2)^{-1} + \left[A^{-1}H_2A^{-1}H_1\left((A+H_2)^{-1} - A^{-1}\right) + \right. \\
&\quad \left. A^{-1}H_2A^{-1}H_1A^{-1}\right] + A^{-1}H_1\left((A+H_2)^{-1} - A^{-1}\right)H_2A^{-1} + A^{-1}H_1A^{-1}H_2A^{-1} \\
&= A^{-1}H_1A^{-1}H_2A^{-1} + A^{-1}H_2A^{-1}H_1A^{-1} + \phi(H_2)
\end{aligned}$$

and $\lim_{H_2 \to 0} \frac{\phi(H_2)}{\|H_2\|} = 0$ by continuity of F and boundedness of $\|H_1\|$.

Inductively, let $F^{(k)} : \mathcal{L}(E,F) \to \mathcal{L}(\underbrace{G\mathcal{L}(E,F), ..., G\mathcal{L}(E,F)}_{k-times}; G\mathcal{L}(F,E))$ be given by

$$F_A^{(k)}(H_1, ..., H_k) = (-1)^k \sum_{\sigma \in S_k} A^{-1}H_{\sigma(x_1)}A^{-1}H_{\sigma(x_2)} \cdots A^{-1}H_{\sigma(x_k)}A^{-1}$$

Then we can complete the induction step by first adding and subtracting terms of the form $B_1H_1B_2H_2 \cdots B_kH_kB_{k+1}$ where B_i is either $(A+H_{k+1})^{-1}$ or A^{-1}, in the formula for

$$F^{(k)}_{A+H_{k+1}}(H_1, ..., H_k) - F^{(k)}_A(H_1, ..., H_k).$$

Then, by regrouping the terms by pairs of $((A+H_{k+1})^{-1} - A^{-1})$, using the identity

$$(A+H_{k+1})^{-1} - A^{-1} = -(A+H_{k+1})^{-1}H_{k+1}A^{-1}$$

and finally, regrouping again by pairs the error terms with addition and subtraction of appropriate terms. The limit of the error terms is then seen to be 0 using the boundedness of the function F and of each $\|H_i\|$. □

1.7 Inverse and Implicit Function Theorems

Proposition 22. *Let E be a Banach space, $x_0 \in E$, $D(x_0, R) \subset E$ and $f : D(x_0, R) \to E$ a map such that for all x, $y \in D(x_0, R)$ and some $0 < k < 1$,*

$$\|f(x) - f(y)\| \leq k \cdot \|x - y\|$$

Then

$$\varphi(x) = x - f(x)$$

is a homeomorphism of an open set $V \subset D(x_0, R)$ *onto* $D(x_0 - f(x_0), (1-k)R)$. *Moreover,*

$$\varphi^{-1} : D(x_0 - f(x_0), (1-k)R) \longrightarrow V$$

satisfies

$$\|\varphi^{-1}(x) - \varphi^{-1}(y)\| \leq \frac{1}{1-k}\|x - y\|.$$

Proof. Let $y \in D(x_0 - f(x_0), (1-k)R)$. Then there exists at most one $x \in D(x_0, R)$ such that $\varphi(x) = y$. Indeed, if $x \neq x'$ then

$$\begin{aligned}\|\varphi(x) - \varphi(x')\| = \|x - f(x) - (x' - f(x'))\| &= \|(x - x') - (f(x) - f(x'))\| \\ &\geq \|x - x'\| - \|f(x) - f(x')\| \\ &\geq (1-k)\|x - x'\|\end{aligned}$$

by the contraction property of f. Thus φ is one-one on $D(x_0, R)$, Let now $y \in D(x_0 - f(x_0), (1-k)R)$. Define a sequence $x_n = 0, 1, \dots$ by

$$x_1 = f(x_0) + y, \qquad x_n = f(x_{n-1}) + y.$$

Then

Claim. $x_n \in D(x_0, R)$ *for all* n *and the sequence is well defined.*

Indeed,

$$\|x_1 - x_0\| = \|f(x_0) + y - x_0\| \leq (1-k)R < R$$

since $y \in D(x_0 - f(x_0), (1-k)R)$. Thus $x_1 \in D(x_0, R)$. Suppose that $x_n \in D(x_0, R)$, then

$$\begin{aligned}\|x_{n+1} - x_n\| = \|f(x_n) + y - f(x_{n-1}) - y\| = \|f(x_n) - f(x_{n-1})\| & \\ &\leq k \cdot \|x_n - x_{n-1}\| \\ &\leq \dots \leq k^n\|x_1 - x_0\| < k^n(1-k)R\end{aligned}$$

so

$$\begin{aligned}\|x_{n+1} - x_0\| &\leq \|x_{n+1} - x_n\| + \|x_n - x_{n-1}\| + \dots \|x_1 - x_0\| \\ &\leq (1 + k + \dots k^n)(1-k)R < R\end{aligned}$$

By induction, $x_n \in D(x_0, R)$ for all n and the sequence x_n is well defined. Moreover, the sequence is Cauchy. If $n > m$ then

$$\begin{aligned}\|x_n - x_m\| &\leq \|x_n - x_{n-1}\| + \dots \|x_{m+1} - x_m\| \\ &\leq (k^{n-1} + \dots + k^m)\|x_1 - x_0\| \\ &\leq \frac{k^m}{1-k}\|x_1 - x_0\|.\end{aligned}$$

Thus $\lim_{n\to\infty} x_n := x$ exists. We need to show that $x \in D(x_0, R)$. Recall that

$$\|x_{n+1} - x_0\| \leq (1 + k + \dots + k^n)\|x_1 - x_0\|.$$

Taking $n \to \infty$ yields

$$\|x - x_0\| \le \frac{1}{1-k}\|x_1 - x_0\| < \frac{1}{1-k}(1-k)R = R,$$

so for every $y \in D(x_0 - f(x_0), (1-k)R)$, there exists a unique $x \in D(x_0, R)$ such that $\varphi(x) = x - f(x) = y$. So φ is invertible on $D(x_0, R)$. The initial estimate, when we showed that φ was an injection,

$$\|\varphi(x) - \varphi(x')\| \ge (1-k)\|x - x'\|$$

yields that for $y \ne y' \in D(x_0 - f(x_0), (1-k)R)$

$$\|\varphi^{-1}(y) - \varphi^{-1}(y')\| \le \frac{1}{1-k}\|y - y'\|$$

so φ^{-1} is Lipshitz continuous and if we set

$$V = \varphi^{-1}(D(x_0 - f(x_0), (1-k)R))$$

then V is open ad $\varphi : V \to D(x_0 - f(x_0), (1-k)R)$ is a homeomorphism. □

1.7.1 Inverse Function Theorem

Theorem 23 (Inverse Function Theorem)**.** *Let E and F be Banach spaces and $U \subset E$ open , $f : U \to F$ a C^1 map. Let $x_0 \in U$ and suppose that f'_{x_0} is an isomorphism. Then there exists open sets $U_0 \supset x_0$, and $V \supset f(x_0)$ such that*

$$f : U_0 \to V$$

is a C^1 diffeomorphism.

Example. *Let $U \subset \mathbb{R}^n$, $F = \mathbb{R}^n$ and let $f(x) = (f_1(x), \dots, f_n(x))$. Then*

$$f'(x_0) = \begin{bmatrix} \nabla f_1(x) \\ \vdots \\ \nabla f_n(x) \end{bmatrix}.$$

If $\det f'_{x_0} \ne 0$ then the theorem applies.

Proof. Set

$$\varphi(x) = x - (f'_{x_0})^{-1} \circ f(x)$$

then φ is a C^1 map on U and $\varphi'(x_0) = I - (f'_{x_0})^{-1} \circ f'_{x_0} = 0$. Since φ' is continuous, given $0 < k < 1$, there is $R > 0$ such that $\|\varphi'(x)\| \le k$ for $x \in D(x_0, R)$. Let now $x, y \in D(x_0, R)$ and note too that the points $tx + (1-t)y$ also lie in $D(x_0, R)$. Set

$$g(t) = \varphi(tx + (1-t)y)$$

then $g'(t) = \varphi'(tx + (1-t)y)(x - y)$, and

$$g(1) - g(0) = \int_0^1 g'(t)dt = \int_0^1 \varphi'(tx + (1-t)y)(x - y)dt.$$

So

$$\begin{aligned}\|\varphi(1)-\varphi(0)\| &\le \max_{t\in[0,1]} \|\varphi'(tx+(1-t)y\|\cdot\|x-y\| \\ &\le k\|x-y\|\end{aligned}$$

so φ is a contraction on $D(x_0, R)$. Then, by the previous proposition, there is an open set $\tilde{U} \subset D(x_0, R)$ with $x_0 \in \tilde{U}$ such that

$$I - \varphi = I - (I - (f'_{x_0})^{-1} \circ f) = (f'_{x_0})^{-1} \circ f$$

is a homeomorphism between $\tilde{U}$ and $D(f'^{-1}_{x_0}(f(x_0)), (1-k)R)$. We now compose $(f'_{x_0})^{-1} \circ f$ with f'_{x_0}. Then this implies that f is a homeomorphism between $\tilde{U}$ and $\tilde{V} := f'_{x_0}(D(f'_{x_0}(f(x_0)), (1-k)R))$. So

$$f : \tilde{U} \longrightarrow \tilde{V}$$

is a C^1 homeomorphism, $x_0 \in \tilde{U}$, $f(x_0) \in \tilde{V}$ and both $\tilde{U}$ and $\tilde{V}$ open. Now f'_{x_0} is an isomorphism, $x \mapsto f'_x$ is continuous and $GL(E, F)$ is open. So there is some $U_0 \subset \tilde{U}$ with $x_0 \in U_0$ such that U_0 is open and f'_x is an isomorphism for all $x \in U_0$. So, if $V = f(U_0)$, then V is open, $f(x_0) \in V$ and

$$f : U_0 \longrightarrow V$$

is a C^1 homeomorphism with f'_x an isomorphism for all $x \in U_0$. Hence f is a C^1 diffeomorphism. □

Remark. *If f is C^p, for $1 \le p < \infty$, then $f : U_0 \longrightarrow V$ is a C^p diffeomorphism.*

Remark. *The above is a local result. The ultimate goal is to get global results.*

Example. *Let now $f : E \longrightarrow F$ be a C^1 map such that f'_x is an isomorphism for all $x \in E$. Then f is locally a diffemorphism. However, $f : E \longrightarrow F$ does not have to be either one-one or onto. Take*

$$E = \mathbb{R}^2, F = \mathbb{R}^2, \text{ and } f(x, y) = (e^x \cos y, e^x \sin y)$$

then

$$f'_{(x,y)} = \begin{bmatrix} e^x \cos y & -e^x \sin y \\ e^x \sin y & e^x \cos y \end{bmatrix}.$$

Then $f'_{(x,y)}$ is an isomorphism, and f is a local diffeomorphism, but f is not onto (as $0 \notin$ Ran $f_{(x,y)}$), and nor is it one-one, since

$$f(x, y + 2\pi) = f(x, y).$$

However, if $f : E \longrightarrow F$ is C^1, f'_x is an isomorphism, and there is some $M > 0$ such that for all x,

$$\|(f'_x)^{-1}\| \le M$$

then $f : E \longrightarrow F$ is a C^1 diffeomorphism.

1.7.2 Partial Derivatives on Banach Spaces

Let $E_1, \ldots, E_n$, and F be Banach spaces, $U \subset E_1 \times \ldots \times E_n$ an open set and $f : U \longrightarrow F$ a map. Let $a = (a_1, \ldots, a_n \in U$. Define a map

$$\pi_k : E_k \longrightarrow E_1 \times \ldots \times E_n$$

by

$$\pi_k(y) = (a_1, \ldots, a_{k-1}, y, a_{k+1}, \ldots, a_n)$$

and note that

$$(\pi_k'(y))(h) = (\underbrace{0, \ldots, 0}_{k-1 \text{ times}}, h, 0, \ldots, 0).$$

Consider the map $f \circ \pi_k$ defined in some neighbourhood of a_k. If this map is differentiable at a_k, we call $(f \circ \pi_k')(a_k)$ the k-th partial derivative of f at a_k and we denote it by

$$\frac{\partial f}{\partial x_k}(a).$$

So $\frac{\partial f}{\partial x_k}(a) \in \mathcal{L}(E_k, F)$. We have by direct computation that

$$(f \circ \pi_k)(a_k + h) - (f \circ \pi_k)(a_k) = f(a_1, \ldots, a_k + h, \ldots, a_n) - f(a_1, \ldots, a_n).$$

Suppose that f is differentiable at a. Then the partial derivatives at a exists and by the chain rule

$$\frac{\partial f}{\partial x_k}(a) = f'(a) \circ \pi_k'(a_k).$$

In general with $h = (h_1, \ldots, h_n)$, and

$$\tilde{h_k} = (\underbrace{0, \ldots, 0}_{k-1}, h_k, 0, \ldots, 0)$$

then

$$\begin{aligned} f'(a)h = \sum_{k=1}^{n} f'(a)h_k &= \sum_{k=1}^{n} f'(a)\pi_k'(a_k)h_k \\ &= \sum_{k=1}^{n} \frac{\partial f}{\partial x_k}(a)h_k. \end{aligned}$$

1.7.3 Implicit Function Theorem

Theorem 24 (Implicit Function Theorem)**.** *Let E, F and G be Banach spaces. U an open set in $E \times F$, and $f : U \longrightarrow G$ a C^1 map .Let $(x_0, y_0) \in U$ be such that $f(x_0, y_0) = 0$ and $f_y'(x_0, y_0)$ is an isomorphism from F to G. Then there exists and open set $W \subset E$ with $x_0 \in W$, an open set $U' \subset U$ with $(x_0, y_0) \in U'$ and a C^1 map $g : W \longrightarrow F$ such that*

$$f(x, y) = 0 \text{ for } (x, y) \in U' \iff y = g(x)$$

Proof. Define a map

$$F : U \longrightarrow E \times G, \qquad F(x,y) = (x, f(x,y)).$$

This map is also C^1 and

$$F'_{(x,y)} \in \mathcal{L}(E \times F, E \times G)$$

is described by

$$\begin{bmatrix} I & 0 \\ f'_x(x,y) & f'_y(x,y) \end{bmatrix},$$

that is

$$F'_{(x,y)} \begin{pmatrix} h_1 \\ h_2 \end{pmatrix} = \begin{pmatrix} h_1 \\ f'_x(x,y)h_1 + f'_y(x,y)h_2 \end{pmatrix}.$$

The assumption that $f'_y(x_0, y_0)$ is an isomorphism implies that $F'_{(x_0,y_0)}$ is as well. Indeed, suppose

$$F'_{(x_0,y_0)} \begin{pmatrix} h_1 \\ h_2 \end{pmatrix} = F'_{(x_0,y_0)} \begin{pmatrix} \tilde{h_1} \\ \tilde{h_2} \end{pmatrix}$$

then $h_1 = \tilde{h_1}$ and

$$\cancel{f'_x(x_0,y_0)h_1} + f'_y(x_0,y_0)h_2 = \cancel{f'_x(x_0,y_0)\tilde{h_1}} + f'_y(x_0,y_0)\tilde{h_2}$$

so $h_2 = \tilde{h_2}$. Similarly, we can show that $F'_{(x_0,y_0)}$ is surjective: if

$$F'_{(x_0,y_0)} \begin{pmatrix} h_1 \\ h_2 \end{pmatrix} = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}$$

then $h_1 = z_1$ and

$$h_2 = (f'_y(x_0,y_0))^{-1}[z_2 - f'_x(x_0,y_0)h_1].$$

Hence $F'_{(x_0,y_0)}$ is a bijection, and hence an isomorphism by the open mapping theorem. By the *Inverse Function Theorem* there is an open set $\tilde{U} \in E \times F$ containing (x_0, y_0) and an open set $V \in E \times G$ containing $F(x_0,y_0) = (x_0, f(x_0,y_0)) = (x_0, 0)$ such that

$$F : \tilde{U} \longrightarrow V$$

is a C^1 diffeomorphism. $\square$

Without loss of generality, we may assume that $\tilde{U} = W \times \tilde{W}$ where $W \subset E$, $\tilde{W} \subset F$ is open, $x_0 \in W$ and $y_0 \in \tilde{W}$. So

$$\begin{aligned}
(x,y) \in \tilde{U} \text{ and } f(x,y) = 0 &\iff x \in W,\ (x,y) \in U,\ F(x,y) = (x, f(x,y)) = (x,0) \\
&\iff x \in W,\ y \in \tilde{W},\ (x,y) = F^{-1}(x,0) \\
&\iff x \in W,\ y = \pi_2(F^{-1}(x,0)) \\
&\iff x \in W,\ y = g(x),\ \text{where } g(x) = \pi_2(F^{-1}(x,0)).
\end{aligned}$$

1.8 Minima & Maxima

Let $f : S \longrightarrow \mathbb{R}$, $S \subset E$, E a Banach space.

Definition. *A point $x_0 \in S$ is called a local minimum (resp. maximum) if $x_0 \in \ intS$ and for $x \in$ some open set containing x_0, $f(x) \geq f(x_0)$ (resp. $f(x) \leq f(x_0)$). A local minimum (maximum) is called strict if the above inequality is strict for $x \neq x_0$. A point $x_0 \in S$ is called a global minimum (maximum) if $f(x) \geq f(x_0)$ ($f(x) \leq f(x_0)$) for all $x \in S$. The definition of a strict global minimum (or maximum) is the same.*

Definition. *An interior point $x_0 \in S$ is called a critical point if $f'_{x_0} = 0$.*

Proposition 25. *Suppose that x_0 is a local minimum (maximum) and that f is differentiable at x_0. Then $f'_{x_0} = 0$.*

Proof. Suppose x_0 is a local minimum. Let $g(t) = f(x_0 + th)$, for some given $h \in E$. Since x_0 is an interior point, $g(t)$ is defined for $|t| < \epsilon$. g is also differentiable at 0 with

$$g'(0) = f'(x_0)h.$$

Since f has a minimum at x_0, g has a minimum at 0, that is $g(t) \geq g(0)$ for all sufficiently small t. So $g'(0) = 0$ and hence $f'_{x_0}h = 0$ for all $h \in E$. Thus $f'_{x_0} \equiv 0$. □

Proposition 26. *Let $f : U \longrightarrow \mathbb{R}$ be a C^2 map and $x_0 \in U$ a critical point. $f^{(2)}_{x_0}$ is a bilinear form from $E \times E \longrightarrow \mathbb{R}$.*

1. *If $f^{(2)}_{x_0}$ is strictly positive definite, that is, there exists $k > 0$ such that*

   $$f^{(2)}_{x_0}(h, h) \geq k\|f\|^2 \qquad \forall h \in E$$

   then x_0 is a strict local minimum.

2. *If $f^{(2)}_{x_0}$ is strictly negative definite, that is, there exists $k > 0$ such that*

   $$f^{(2)}_{x_0}(h, h) \leq -k\|f\|^2 \qquad \forall h \in E$$

   then x_0 is a strict local maximum.

3. *If there exist h_1, h_2 in E such that*

   $$f^{(2)}_{x_0}(h_1, h_1) > 0 \quad \textit{and} \quad f^{(2)}_{x_0}(h_2, h_2) < 0$$

   then x_0 is neither a minimum nor a maximum. In this case we call x_0 a saddle point.

Proof. Let $x_0 + th \in U$ for $0 \leq t \leq 1$. Then, by the Taylor formula,

$$f(x_0 + th) = f(x_0) + f'(x_0)h + \frac{1}{2}f^{(2)}_{x_0}(h, h) + \mathcal{O}(h^2)$$

where

$$\lim_{\|h\| \to 0} \frac{|\mathcal{O}(h^2)|}{\|h\|^2} = 0.$$

Since $f'_{x_0} = 0$,

$$f(x_0 + th) = f(x_0) + \frac{1}{2} f^{(2)}_{x_0}(h,h) + \mathcal{O}(h^2)$$

Suppose that (1) holds. Then

$$f^{(2)}_{x_0}(h,h) \geq k\|f\|^2 \qquad \forall h \in E$$

for some $k > 0$. Choose δ such that $D(x_0, \delta) \subset U$ and that for $\|h\| < \delta$,

$$\frac{|\mathcal{O}(h^2)|}{\|h\|^2} \leq \frac{k}{4}.$$

Then $f(x_0 + th) \geq f(x_0) + \frac{k}{4}\|h\|^2$. If $y \in D(x_0, \delta)$, let $x_0 - y = h$, then $\|h\| < \delta$ and

$$f(y) > f(x_0) \qquad \text{if } y \neq x_0.$$

For case (2) we do exactly the same and for (3) we do the same as the above along the h_1 and h_2 directions to yield that x_0 is neither a minimum nor a maximum. □

In practice, obtaining a bound on $f^{(2)}_{x_0}(h,h)$ is a highly non-trivial endeavour. In the case that E is a Hilbert space, however, we have a bit more to work with. Hilbert space theory yields that since $f^{(2)}_{x_0}$ is a symmetric bilinear form, there exists a bounded symmetric operator $A \in \mathcal{L}(E)$ such that

$$f^{(2)}_{x_0}(h_1, h_2) = \langle h_1, Ah_2 \rangle.$$

A is symmetric and bounded, hence

$$\sigma(A) = \{\lambda \in \mathbb{R} : A - \lambda I \text{ is not an isometry}\}$$

is a compact subset of $\mathbb{R}$. Let

$$a = \inf \sigma(A), \quad b = \sup \sigma(A)$$

then if $a > 0$ then A is strictly positive, if $b < 0$ A is strictly negative. Moreover, if $A > 0$ then $f^{(2)}_{x_0}$ is strictly positive definite and if $A < 0$ then $f^{(2)}_{x_0}$ is strictly negative definite. Finally, if $a < 0$, $b > 0$ then $f^{(2)}_{x_0}$ is indefinite.

1.9 Introduction to Calculus of Variations

The general set up on the level of Banach spaces is the following. Let $I = [a,b]$ be a real interval, E a Banach space and denote by $C^0(I;E)$ the vector space of all continuous maps $\alpha : I \longrightarrow E$ equipped with norm $\|\alpha\|_{C^0(I;E)} = \max_{a \leq t \leq b} \|\alpha(t)\|$. This makes $C^0(I;E)$ a Banach space (Analysis 3).

Let now $\alpha : I \longrightarrow E$ be C^p on (a,b). The derivative $\alpha'(t) \in \mathcal{L}(\mathbb{R}, E) \cong E$. So as usual we identify $\alpha'(t)$ with a vector in E and the map $t \mapsto \alpha'(t)$ becomes a map from (a,b) to E. Similarly, for every $1 \leq k \leq p$, $t \mapsto \alpha^{(k)}(t)$ is a map from (a,b) to E.

Note that $\alpha'(t)$ is uniquely specified by $\lim_{\Delta t \to 0} \left\| \frac{\alpha(t+\Delta t) - \alpha(t)}{\Delta t} - \alpha'(t) \right\| = 0$

Let $C^p(I;E)$ be the collection of all C^p maps $\alpha : (a,b) \longrightarrow E$ ($\alpha^{(0)} = \alpha$), which have continuous extension to $[a,b]$. Then $C^p(I;E)$ is a vector space and $\|\alpha\|_{C^p(I,E)} = \max_{1 \leq k \leq p} \max_{a \leq t \leq b} \|\alpha^{(k)}(t)\|_E$ is a norm on $C^p(I;E)$.

Proposition 27. *$C^p(I;E)$ is a Banach space.*

Proof. We will prove the case $p = 1$. The general case is similar. Let α_n be a Cauchy sequence in $C^1(I;E)$. Then by definition of the norm in $C^1(I;E)$, both α_n and α_n' are Cauchy sequences in $C^0(I;E)$ and so there exists α and β in $C^0(I;E)$ such that $\alpha_n \longrightarrow \alpha$ and $\alpha_n' \longrightarrow \beta$ in $C^0(I;E)$.

Let now $a < t < b$. Because of uniform convergence,

$$\lim_{n\to\infty} \int_a^t \alpha_n'(s)ds = \int_a^t \beta(s)ds$$

But also,

$$\int_a^t \alpha_n'(s)ds = \alpha_n(t) - \alpha_n(a)$$

So taking the limit, we obtain

$$\alpha(t) - \alpha(a) = \int_a^t \beta(s)ds$$

Differentiating, we get $\alpha'(t) = \beta(t)$ for $a < t < b$. Hence α is differentiable, has continuous extension to $[a,b]$ and $\alpha_n' \longrightarrow \alpha'$ in $C^0(I;E)$. Since also $\alpha_n \longrightarrow \alpha$ in $C^0(I;E)$, $\alpha_n' \longrightarrow \alpha'$ in $C^1(I;E)$. □

The general setup is as follows. We are given a Banach space E. Consider $E \times E$ with norm $\|(x,y)\| = \max(\|x\|, \|y\|)$. Let $\mathcal{U} \in E \times E$ be an open set and let $L : \mathcal{U} \longrightarrow \mathbb{R}$ be a continuous function (the Lagrangian). We are interested in functionals of the form:

$$S(\alpha) = \int_a^b L(\alpha, \alpha')dt$$

where $\alpha \in C^p(I;E)$ for some $p \geq 1$.

Proposition 28. *The set of all $\alpha \in C^p(I;E)$, $p \geq 1$, such that $(\alpha(t), \alpha') \in \mathcal{U}$ for all $t \in [a,b]$ is an open subset of $C^p(I;E)$, denoted by $\mathcal{V}_p$. Hence for $\alpha \in \mathcal{V}_p$, $L(\alpha(t), \alpha')$ is a well defined continuous function from $[a,b]$ to $\mathbb{R}$ and the map $S(\alpha)$ is also well defined.*

Proof. Let $\alpha_0 \in \mathcal{V}_p$ and let $\Gamma = \{(\alpha_0(t), \alpha_0'(t)) : t \in [a,b]\} \subset E \times E$. Then Γ is a compact set in $\mathcal{U}$. Consider the function $d((x,y), \mathcal{U}^c)$ defined on $E \times E$ with the metric induced by the norm (recall $d(x,A) = \inf_{x' \in A} d(x,x')$). Since Γ is compact, for $(x,y) \in \Gamma$, $d((x,y), \mathcal{U}^c) > 0$ and there exists $\epsilon > 0$ such that

$$\inf_{(x,y)\in\Gamma} d((x,y), \mathcal{U}^c) = \min_{(x,y)\in\Gamma} d((x,y), \mathcal{U}^c) = \epsilon > 0 \tag{3}$$

Consider now the set of all $(x,y) \in E \times E$ such that $d((x,y), \Gamma) < \epsilon$. This set is open by continuity of the metric function and is contained in $\mathcal{U}$: if not, then there exists $(x,y) \in \mathcal{U}^c$ satisfying $d((x,y), \Gamma) < \epsilon$, so there exists $(x',y') \in \Gamma$ such that $d((x,y),(x',y')) < \epsilon$, a contradiction to 3. Hence $\{(x,y) \in E \times E : d((x,y), \Gamma) < \epsilon\} \subset \mathcal{U}$.

Consider now $D(\alpha_0, \epsilon) \in C^p(I;E)$. Then for $p \geq 1$ and for all $\alpha \in D(\alpha_0, \epsilon)$,

$$\max\left(\max_{t\in[a,b]} \|\alpha(t) - \alpha_0(t)\|, \max_{t\in[a,b]} \|\alpha'(t) - \alpha_0'(t)\|\right) < \epsilon$$

Since $(\alpha_0, \alpha_0') \in \Gamma$ and for $\alpha \in D(\alpha_0, \epsilon)$, $(\alpha(t), \alpha'(t))$ is of distance less than ϵ from Γ for all $t \in [a, b]$, we get $D(\alpha_0, \epsilon) \subset \{(x, y) \in E \times E : d((x, y), \Gamma) < \epsilon\} \subset \mathcal{U}$. □

If $L : \mathcal{U} \longrightarrow \mathbb{R}$ is C^1, then we can decompose

$$\begin{aligned} L'(x_0, y_0)(h_1, h_2) &= L'(x_0, y_0)(h_1, 0) + L'(x_0, y_0)(0, h_2) \\ &=: L_x(x_0, y_0)h_1 + L_y(x_0, y_0)h_2 \end{aligned}$$

Similarly, if $L : \mathcal{U} \longrightarrow \mathbb{R}$ is C^2, we can intoduce directly the mixed partial derivatives of $L''(x_0, y_0) \in \mathcal{L}((E \times E) \times (E \times E), \mathbb{R}) \cong \mathcal{L}(E \times E, \mathcal{L}(E \times E, \mathbb{R}))$ as follows. $L''(x_0, y_0)$ is a bilinear form

$$L''(x_0, y_0)\left(\begin{bmatrix} h_1 \\ h_2 \end{bmatrix}, (\begin{bmatrix} \tilde{h}_1 \\ \tilde{h}_2 \end{bmatrix}\right))$$

which can be expressed as a sum of four bilinear forms on $E \times E$ only:

$$L''(x_0, y_0)\left(\begin{bmatrix} h_1 \\ h_2 \end{bmatrix}, \begin{bmatrix} \tilde{h}_1 \\ \tilde{h}_2 \end{bmatrix}\right) = L_{xx}(x_0, y_0)(h_1, \tilde{h}_1) + L_{xy}(x_0, y_0)(h_1, \tilde{h}_2) + L_{yx}(x_0, y_0)(h_2, \tilde{h}_1) + L_{yy}(x_0, y_0)(h_2, \tilde{h}_2)$$

where

$$\begin{aligned} L_{xx}(x_0, y_0)(h_1, \tilde{h}_1) &:= L''(x_0, y_0)\left(\begin{bmatrix} h_1 \\ 0 \end{bmatrix}, \begin{bmatrix} \tilde{h}_1 \\ 0 \end{bmatrix}\right) \\ L_{xy}(x_0, y_0)(h_1, \tilde{h}_2) &:= L''(x_0, y_0)\left(\begin{bmatrix} h_1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ \tilde{h}_2 \end{bmatrix}\right) \end{aligned}$$

and similarly for L_{yx} and L_{yy}. Those are the mixed partial derivatives of order 2. Note also that if $L \in C^2$, we have proved before that L'' is symmetric, which is equivalent to having the three following equations

$$\begin{aligned} L_{xx}(h_1, \tilde{h}_1) &= L_{xx}(\tilde{h}_1, h_1) \\ L_{yy}(h_2, \tilde{h}_2) &= L_{yy}(\tilde{h}_2, h_2) \\ L_{xy}(h_1, \tilde{h}_2) &= L_{yx}(\tilde{h}_2, h_1) \end{aligned}$$

Proposition 29. *If $L : \mathcal{U} \longrightarrow \mathbb{R}$ is C^q for some $q \geq 1$, then the map $S : \mathcal{V}_p \longrightarrow \mathbb{R}$, $p \geq 1$*

$$S(\alpha) = \int_a^b L(\alpha, \alpha')dt$$

is also C^q.

Proof. The only cases of interest for us are the cases $q = 1, 2$. We will prove the case $q = 1$, the proof of $q = 2$ being very similar.

So suppose L is C^1. For $h(t) \in C^p(I; E)$, $p \geq 1$, using the above decomposition and the Mean Value Theorem, write

$$\begin{aligned} L(\alpha + h(t), \alpha' + h'(t)) &= L(\alpha, \alpha') + L'(\alpha, \alpha')(h(t), h'(t)) + \big[L(\alpha + h(t), \alpha' + h'(t)) \\ &\quad - L(\alpha, \alpha') - L'(\alpha, \alpha')(h(t), h'(t))\big] \\ &= L(\alpha, \alpha') + L_x(\alpha, \alpha')h(t) + L_y(\alpha, \alpha')h'(t) + \theta \end{aligned}$$

where

$$\theta = \int_0^1 \left[L'(\alpha + sh(t), \alpha' + sh'(t))(h(t), h'(t)) - L'(\alpha, \alpha')(h(t), h'(t)) \right] ds$$

Integrating over $[a, b]$,

$$S(\alpha + h) = S(\alpha) + \int_a^b L_x(\alpha, \alpha')h(t)dt + \int_a^b L_y(\alpha, \alpha')h'(t)dt + \int_a^b \theta dt \qquad (4)$$

The map $t \mapsto L_x(\alpha, \alpha')h(t)$ is a continuous function since both $t \mapsto L_x(\alpha, \alpha') \in \mathcal{L}(E, \mathbb{R})$ and $t \mapsto h(t) \in E$ are continuous (and similarly for L_y). It follows that

$$(h, h') \longrightarrow \int_a^b L_x(\alpha, \alpha')h(t)dt + \int_a^b L_y(\alpha, \alpha')h'(t)dt$$

is a bounded linear map from $C^p(I; E)$ to $\mathbb{R}$ with norm less than

$$\max_{t \in [a,b]} \|L_x(\alpha(t), \alpha'(t))\| + \max_{t \in [a,b]} \|L_y(\alpha(t), \alpha'(t))\|$$

So from 4

$$S'(\alpha)(h, h') = \int_a^b L_x(\alpha, \alpha')h(t)dt + \int_a^b L_y(\alpha, \alpha')h'(t)dt$$

provided

$$\lim_{h \to 0} \frac{\| \int_a^b \theta dt \|}{\|h\|_{C^p(I;E)}} = 0.$$

$$\begin{aligned}
\| \int_a^b \theta dt\| \leq{}& \|h\|_{C^p(I,E)} \max_{0 \leq s \leq 1} \int_a^b \|L_x(\alpha + sh(t), \alpha' + sh'(t)) - L_x(\alpha, \alpha')\| dt + \\
&+ \|h'\|_{C^p(I,E)} \max_{0 \leq s \leq 1} \int_a^b \|L_y(\alpha + sh(t), \alpha' + sh'(t)) - L_y(\alpha, \alpha')\| dt \\
&\leq (b-a)\|h\|_{C^p(I,E)} \max_{0 \leq s \leq 1} \max_{a \leq t \leq b} \|L_x(\alpha + sh(t), \alpha' + sh'(t)) - L_x(\alpha, \alpha')\| dt + \\
&+ (b-a)\|h\|_{C^p(I,E)} \max_{0 \leq s \leq 1} \max_{a \leq t \leq b} \|L_y(\alpha + sh(t), \alpha' + sh'(t)) - L_y(\alpha, \alpha')\| dt
\end{aligned}$$

Hence it is enough to show (absorbing s in h) that

$$\max_{a \leq t \leq b} \|L_x(\alpha + h(t), \alpha' + h'(t)) - L_x(\alpha, \alpha')\| \to 0$$

$as\|h\|_{C^p(I;E)} \to 0$. Suppose this is not true. Then there exists $\delta > 0$ and a sequence $\{h_n, n \geq 1\}$ such that $\|h_n\|_{C^p(I;E)} \to 0$ and

$$\max_{a \leq t \leq b} \|L_x(\alpha + h_n(t), \alpha' + h_n'(t)) - L_x(\alpha, \alpha')\| \geq \delta$$

for each n. By continuity and compactness, for each n, the maximum is attained, say at t_n. Hence

$$\|L_x(\alpha(t_n) + h_n(t_n), \alpha'(t_n) + h_n'(t_n)) - L_x(\alpha(t_n), \alpha'(t_n))\| \geq \delta$$

for each n. Since $t_n \in [a,b]$, there exist a convergent subsequence $\{t_{n_k}, k \geq 1\}$ with limit, say, $\bar{t}$. Then $\|h_n\|_{C^p(I;E)} \to 0$ for $p \geq 1$ implies that both $h_{n_k}(t_{n_k}) \to 0$ and $h'_{n_k}(t_{n_k}) \to 0$ as $k \to \infty$ and by continuity, $\alpha(t_{n_k}) \to \alpha(\bar{t})$ and $\alpha'(t_{n_k}) \to \alpha'(\bar{t})$. Hence

$$\lim_{k\to\infty} \|L_x(\alpha(t_{n_k}) + h_{n_k}(t_{n_k}), \alpha'(t_{n_k}) + h'_{n_k}(t_{n_k})) - L_x(\alpha(t_{n_k}), \alpha'(t_{n_k}))\| = \|L_x(\alpha(\bar{t}), \alpha'(\bar{t})) - L_x(\alpha(\bar{t}), \alpha'(\bar{t}))\|$$
$$= 0$$

a contradiction. □

Similarly, in the case $q = 2$, the second derivative is

$$S''(\alpha, \alpha')\left(\begin{bmatrix} h \\ h' \end{bmatrix}, \begin{bmatrix} \tilde{h} \\ \tilde{h}' \end{bmatrix}\right) = \int_a^b \Big[L_{xx}(\alpha, \alpha')(h, \tilde{h}) + L_{xy}(\alpha, \alpha')(h, \tilde{h}') +$$
$$+ L_{yx}(\alpha, \alpha')(h', \tilde{h}) + L_{yy}(\alpha, \alpha')(h', \tilde{h}')\Big]\, dt$$

Lemma 30 (Fundamental Lemma of Calculus of Variations)**.** *Let $g : [a,b] \longrightarrow \mathcal{L}(E, \mathbb{R}) \cong E^*$ be a continuous function such that for every $h \in C^p(I; E)$ satisfying $h(a) = h(b) = 0$, we have that the continuous map $t \mapsto g(t)h(t)$ from $[a,b]$ to $\mathbb{R}$ satisfies*

$$\int_a^b g(t)h(t)dt = 0$$

Then $g(t) = 0$ for every $t \in [a,b]$.

Proof. Suppose that $g(t_0) \neq 0$ for some $t_0 \in [a,b]$. Wlog take $t_0 \in (a,b)$. Let $x \in E$ be such that $g(t_0)x \neq 0$. Wlog assume $g(t_0)x > 0$. Choose $\epsilon > 0$ such that $(t_0 - \epsilon, t_0 + \epsilon) \in (a,b)$ and let

$$\chi(t) = \begin{cases} (t - (t_0 - \epsilon))^{2(p+1)}(t - (t_0 + \epsilon))^{2(p+1)}, & t \in (t_0 - \epsilon, t_0 + \epsilon) \\ 0, otherwise \end{cases}$$

Then $\chi : [a,b] \longrightarrow \mathbb{R}$ is p times continuously differentiable and $\chi(a) = \chi(b) = 0$. Let $h(t) = \chi(t)x$. Then $g(t)h(t) = \chi(t)g(t)x$. Then by taking ϵ sufficiently small, we can have $\chi(t)g(t)x$ positive for every $t \in (t_0 - \epsilon, t_0 + \epsilon)$. Hence

$$\int_a^b g(t)h(t)dt = \int_{t_0-\epsilon}^{t_0+\epsilon} \chi(t)g(t)xdt > 0$$

a contradiction. □

Remark (Integration by Parts)**.** *Let $g : [a,b] \longrightarrow \mathcal{L}(E;\mathbb{R}) \cong E^*$ be $C^1(I; \mathcal{L}(E;\mathbb{R}))$ and let $h : [a,b] \longrightarrow E$ be $C^1(I;E)$. Then $g'(x)$ is identified with an element of $\mathcal{L}(E,\mathbb{R})$ and $h'(t)$ with a vector in E. Then*

$$\frac{d}{dt}(g(t)h(t)) = g'(t)h(t) + g(t)h'(t)$$

as can be seen from

$$g(t + \Delta t)h(t + \Delta t) - g(t)h(t) = (g(t + \Delta t) - g(t))h(t + \Delta t) + g(t)(h(t + \Delta t) - h(t)).$$

Integrating on $[a,b]$, we get

$$g(b)h(b) - g(a)h(a) = \int_a^b g'(t)h(t)dt + \int_a^b g(t)h'(t)dt$$

If $h(b) = h(a) = 0$,

$$\int_a^b g'(t)h(t)dt = -\int_a^b g(t)h'(t)dt$$

1.9.1 Euler Lagrange and Jacobi equations

Let as usual $\mathcal{U} \subset E \times E$, $L : \mathcal{U} \longrightarrow \mathbb{R}$, $\mathcal{V}_p \subset C^p(I;E)$ and $S(\alpha) = \int_a^b L(\alpha,\alpha')dt$, $S : \mathcal{V}_p \longrightarrow \mathbb{R}$. Consider $C_0^p(I;E) = \{\alpha \in C^p(I;E) : \alpha(a) = \alpha(b) = 0\}$. $C_0^p(I;E)$ is then a closed subspace of $C^p(I;E)$. Let also $\mathcal{V}_{0,p} = \mathcal{V}_p \bigcap C_0^p(I;E)$. This is an open set in the Banach space $C_0^p(I;E)$.

We now consider S restricted to $\alpha \in \mathcal{V}_{0,p}$. Our formula for the derivative $S'(\alpha)$ is then the same as before (the Fréchet derivative with respect to the subspace is the same as the Fréchet derivative restricted to the subspace):

$$S'(\alpha)h = \int_a^b L_x(\alpha,\alpha')h(t) + L_y(\alpha,\alpha')h'(t)dt$$

for $h \in C_0^p(I;E)$. If $p \geq 2$ and $L : \mathcal{U} \longrightarrow \mathbb{R}$ is C^2, the second term above can be integrated by parts to yield

$$S'(\alpha)h = \int_a^b \left[L_x(\alpha,\alpha') - \frac{d}{dt}L_y(\alpha,\alpha')\right] h(t)dt$$

So $S'(\alpha) = 0$ (zero functional) if and only if

$$L_x(\alpha,\alpha') - \frac{d}{dt}L_y(\alpha,\alpha') = 0.$$

This is the Euler-Lagrange equation and the C^2 paths $\alpha : [a,b] \longrightarrow \mathbb{R}$ which satisfy E-L are called critical paths. Note that E-L has to be coupled to the constraint $\alpha(a) = \alpha(b) = 0$.

Interpreting L_{yx} and L_{yy} as $\mathcal{L}(E,\mathcal{L}(E,\mathbb{R})) \cong \mathcal{L}(E \times E;\mathbb{R})$, E-L becomes

$$L_x(\alpha,\alpha') - L_{yx}(\alpha,\alpha')\alpha' - L_{yy}(\alpha,\alpha')\alpha''(t) = 0$$

for all $t \in [a,b]$.

To go beyond E-L, we need $S''(\alpha) : C_0^p(I;E) \times C_0^p(I;E) \longrightarrow \mathbb{R}$. Then from before,

$$S''(\alpha,\alpha')\left(\begin{bmatrix} h \\ h' \end{bmatrix}, \begin{bmatrix} \tilde{h} \\ \tilde{h}' \end{bmatrix}\right) = \int_a^b \Big[L_{xx}(\alpha,\alpha')(h,\tilde{h}) + L_{xy}(\alpha,\alpha')(h,\tilde{h}')+ \\ + L_{yx}(\alpha,\alpha')(h',\tilde{h}) + L_{yy}(\alpha,\alpha')(h',\tilde{h}')\Big] dt$$

Take $h = \tilde{h}$. Then

$$S''(\alpha,\alpha')(h,h) = \int_a^b \big[[L_{xx}(\alpha,\alpha')h(t)]h(t) + 2[L_{xy}(\alpha,\alpha')h(t)]h'(t) - \\ + [L_{yy}(\alpha,\alpha')h'(t)]h'(t)\big] dt$$

If in addition L is C^3, integration by parts yields

$$S''(\alpha,\alpha')(h,h) = \int_a^b \Big[L_{xx}(\alpha,\alpha')h(t) - 2\frac{d}{dt}[L_{xy}(\alpha,\alpha')h(t)] - \\ - \frac{d}{dt}[L_{yy}(\alpha,\alpha')h'(t)]\Big]\, h(t)dt$$

If there exist h satisfying

$$L_{xx}(\alpha,\alpha')h(t) - 2\frac{d}{dt}[L_{xy}(\alpha,\alpha')h(t)] - \frac{d}{dt}[L_{yy}(\alpha,\alpha')h'(t)] = 0$$

for α a fixed critical path, then h is called Jacobi field. In general, one cannot find a solution satisfying $h(a) = h(b) = 0$. If such a solution does exist, this leads to the notion of conjugate points. Note that the above equation, called Jacobi equation, is linear in h.

Given $x_0, y_0 \in E$ let

$$C^p_{x_0,y_0} = \{\alpha \in C^p(I;E) : \alpha(a) = x_0,\ \alpha(b) = y_0\}.$$

$C^p_{x_0,y_0}$ is not a vector space but it is a closed subset of $C_p(I;E)$ and hence a metric space with the induced metric subspace structure. Let now

$$V_{x_0,y_0,p} = C^p_{x_0,y_0} \cap V_p,$$

and assume that this set is non-empty (it is certainly open in the subspace topology if this is the case). Consider our functional

$$S(\alpha) = \int_a^b L(\alpha,\alpha')dt$$

restricted to $V_{x_0,y_0,p}$. A path $\alpha_0 \in V_{x_0,y_0,p}$ is a local minimum of S if there is an open set O_{x_0,y_0} in $V_{x_0,y_0,p}$ such that for all $\alpha \in O_{x_0,y_0}$,

$$S(\alpha) \geq S(\alpha_0).$$

As we have done previously, we may similarly define a local maximum, and strict local minima and maxima. Note that if $\alpha \in O_{x_0,y_0}$ then $h = \alpha - \alpha_0$ is a path satisfying $h(a) = h(b) = 0$ with the same regularity properties as α. If $\alpha_0 \in V_{x_0,y_0,p}$ is fixed, O_{x_0,y_0} is an open set containing α_0, then

$$O_{x_0,y_0} - \alpha_0 = \{\alpha - \alpha_0 : \alpha \in O_{x_0,y_0}\}$$

is an open set in $C^p(I;E)$ and we set

$$S(\alpha) := \tilde{S}(\alpha - \alpha_0)$$

where $\tilde{S}(h)$ is defined on $O_0 = O_{x_0,y_0} - \alpha_0$ and equal to

$$\int_a^b L(\alpha_0 + h, \alpha_0' + h')dt.$$

α_0 is a strict local minimum of $S(\alpha)$ if and only if $h = 0$ is a strict local minimum of $\tilde{S}(h)$. We are thus back to the usual framework except for one thing:

$$\tilde{S}(h) = \int_a^b \tilde{L}(t, h(t), h'(t))dt,$$

where $\tilde{L}(t, x, y) = L(t, \alpha_0 + x, \alpha_0 + y)$, i.e. our Langrangian is dependent on **t**!

However, the entire theory described so far is independent of the fact that L is t independent.

$L(t, x, y)$ is just fine!

The Euler-Lagrange equation, second variation, and Jacobi equation all have exactly the same form. In particular if $\alpha_0(t)$ is a local minimum of S (on $V_{x_0,y_0,p}$) then the Euler-Lagrange equation holds. Indeed if $\alpha_0(t) \in V_{x_0,y_0,p}$ satisfies

$$L_x(\alpha_0, \alpha_0') - \frac{d}{dt}L_y(\alpha_0, \alpha_0') = 0$$

then we say that $\alpha_0(t)$ is a critical (stationary) path. When L is C^2, the nature of the critical path is then determined by a second variation of $\tilde{S}(h)$ given by

$$\begin{aligned}\tilde{S}''(0)(h, h) = \int_a^b \big[&L_{xx}(\alpha_0, \alpha_0')(h, h) + 2L_{xy}(\alpha_0, \alpha_0')(h, h') \\ &+ L_{yy}(\alpha_0, \alpha_0')(h', h')\big]dt.\end{aligned}$$

For such L we can write the Euler-Lagrange equation for the stationary path by

$$\begin{aligned}&L_x(\alpha_0, \alpha_0') - \frac{d}{dt}L_y(\alpha_0, \alpha_0') \\ &= L_x(\alpha_0, \alpha_0') - L_{xx}(\alpha_0, \alpha_0')\alpha_0' - L_{xy}(\alpha_0, \alpha_0')\alpha_0''(t) \\ &= 0.\end{aligned}$$

This is a non-linear ODE! Of the second order! We will always assume the non-degenerate case that $L_{yy} \neq 0$ on $\mathcal{U}$. Then this equation will take the form

$$0 = \alpha_0'' + \alpha_0'\left(\frac{-L_{xy}(\alpha_0, \alpha_0')}{L_{yy}(\alpha_0, \alpha_0')}\right) + \frac{L_x(\alpha_0, \alpha_0')}{L_{yy}(\alpha_0, \alpha_0')}.$$

The coefficients are continuous functions on $[a, b]$. If L is C^3, they are differentiable and hence Lipshitz.

By the Picard method (an application of the Banach fixed point theorem) a solution exists and is uniquely determined by $\alpha(a)$ and $\alpha'(a)$ if $|b - a|$ is sufficiently small. In this case, the second variation takes the form

$$\begin{aligned}\tilde{S}''(0)(h, h) = \int_a^b \big[&L_{xx}(\alpha_0, \alpha_0')h(t)^2 + \underbrace{2L_{xy}(\alpha_0, \alpha_0')h(t)h'(t)}_{I} \\ &+ L_{yy}(\alpha_0, \alpha_0')(h(t))^2\big]dt.\end{aligned}$$

The second variation can be rewritten in two different ways

1. We integrate I by parts and obtain

$$\tilde{S}''(0)(h,h) = \int_a^b [A(t)h(t)^2 + \underbrace{B(t)(h'(t))^2}_{II}]dt$$

with

$$A(t) = L_{xx}(\alpha_0(t), \alpha_0'(t)) - \frac{d}{dt}L_{xy}(\alpha_0(t), \alpha_0'(t))$$

and

$$B(t) = L_{yy}(\alpha_0(t), \alpha_0'(t)).$$

2. Integrate II by parts:

$$(h'(t))^2 = h'(t)\frac{d}{dt}h(t)$$

yielding

$$\tilde{S}''(0)(h,h) = \int_a^b [A(t)h(t) - \frac{d}{dt}B(t)h'(t)]h(t)dt$$

(assuming L is C^3).

Then

$$J(h) = A(t)h(t) - \frac{d}{dt}B(t)h'(t)$$

is a second order linear operator. The equation $J(h) = 0$ is called the Jacobi equation and its solutions are called Jacobi fields. Due to linearity, a solution exists and is uniquely determined by $h(0)$ and $h'(0)$. Any solution is a linear combination of the two fundamental solutions associated to

$$\begin{matrix} h(a) = 0 \\ h'(a) = 1 \end{matrix} \qquad \text{and} \qquad \begin{matrix} h(a) = 1 \\ h'(a) = 0. \end{matrix}$$

If there exists a Jacobi field $h \neq 0$ such that $h(a) = h(b) = 0$ then the points a and b are called conjugate to each other.

Proposition 31 (Necessary condition for a minimum (Lagrange Condition))**.** *If α_0 is a stationary parth which is a local minimum then*

$$B(t) = L_{yy}(\alpha_0(t), \alpha_0'(t)) > 0$$

on $[a,b]$.

Proof. Suppose that α_0 is a local minimum and that $B(t) < 0$ on $[a,b]$. Let $\bar{t} \in (a,b)$ be given. Consider the second variation

$$\tilde{S}''(0)(h,h) = \int_a^b [A(t)h(t)^2 + B(t)(h'(t))^2]dt$$

and consider a sequence h_n such that

$$\int_a^b |h_n(t)|^2 dt \to 0 \qquad \text{and } |h_n'(t)|^2 \to \delta_{\bar{t}}.$$

By the usual methods, we may, without loss of generality, take the h_n to be smooth. Then

$$\tilde{S}''(0)(h_n, h_n) \underset{n}{\longrightarrow} B(\bar{t}) < 0.$$

So for n sufficiently large, $\tilde{S}''(0)(h_n, h_n) < 0$ which contradicts the necessary condition for a minimum. □

Remember the condition for strict local minimum. If for some $c > 0$

$$\tilde{S}''(0)(h, h) \geq c\|h\|_{C_0^2(I;\mathbb{R})}$$

then 0 is a strict local minimum of $\tilde{S}$ and α_0 will be a strict local minimum of S. This never works in our setup for the following reason. We have $B(t) > 0$. Take any $\bar{t} \in (a, b)$ and let h_n be as before. Then

$$\tilde{S}''(0)(h_n, h_n) \underset{n}{\longrightarrow} B(\bar{t}) > 0.$$

But we also have that

$$\|h_n\|_{C_0^2(I;\mathbb{R})} \geq \max_{t\in[a,b]} |h_n'(t)|^2 \to \infty$$

Theorem 32 (Jacobi)**.** *Let L be C^3 and α_0 a stationary path. Then*

1. If α_0 is a local minimum then no point in (a, b) is conjugate to a.
2. If there is no point in $(a, b]$ conjugate to a then α_0 is a local minimum.

Remark. *$c \in (a, b]$ is conjugate to a if there exists $h_n \not\equiv 0$ on $[a, c]$ and $h(a) = h(c) = 0$, and Jacobi equation*

$$A(t)h(t) - \frac{d}{dt}\left[B(t)h'(t)\right] = 0$$

holds on $[a, c]$.

We will not prove this here. The crucial thing given a stationary path α_0 ($L_{yy} > 0$) is to study the Jacobi equation and to understand conjugate points. There are various ways of looking at Jacobi's equation.

1. $\tilde{S}''(0)(h, h) = \int_a^b [A(t)h_n(t)^2 + B(t)(h_n'(t))^2]dt$.

   $$\tilde{L}(t, x, y) = A(t)x^2 + B(t)y^2$$

   yielding

   $$\tilde{S}''(0)(h, h) = \int_a^b \tilde{L}(t, h(t), h'(t))dt.$$

   The Euler-Lagrange equation for a stationary path of $\tilde{S}''(0)(h, h)$ is given by

   $$\tilde{L}_x(t, h_0(t), h_0'(t)) - \frac{d}{dt}\tilde{L}_y(t, h_0(t), h_0'(t)) = 0$$

   or

   $$A(t)h_0(t) - \frac{d}{dt}B(t)h_0'(t) = 0 \qquad \text{Jacobi equation!}$$

   With $h_0(a) = h_0(b) = 0$, h_0 is non-trivial if and only if the second variation has a non-trivial stationary path. Along that stationary path $\tilde{S}''(0)(h, h) = 0$. This observation is the starting point for the sufficiency portion of Jacobi's theorem.

2. Small disturbance equation.

   Given α_0 a C^2 stationary path connecting $\alpha_0(a) = x_0$ to $\alpha_0(b) = y_0$. Now suppose α_0 can be embedded into a family of stationary paths $\alpha_0(t,u)$, with $u \in (-\epsilon, \epsilon)$, which are C^2 and connect $x_0(u)$ to $y_0(u)$.

   Then
$$h(t) = \frac{\partial}{\partial u}\alpha_0(t,u)\Big|_{u=0}$$
   is a solution of the Jacobi equation.

   Proof. $\alpha_0(t,u)$ satisfies the Euler-Lagrange equation,
$$L_x(\alpha_0(t,u), \alpha_0'(t,u)) - \frac{d}{dt}L_y(\alpha_0(t,u), \alpha_0'(t,u)) = 0.$$
   Take the derivative with respect to u to obtain
$$\begin{aligned}&L_{xx}(\alpha_0(t,u), \alpha_0'(t,u))\frac{\partial}{\partial u}\alpha_0(t,u) + L_{xy}(\alpha_0(t,u), \alpha_0'(t,u))\frac{\partial}{\partial u}\alpha_0'(t,u)\\ &- \frac{d}{dt}\left[L_{xy}(\alpha_0(t,u), \alpha_0'(t,u))\frac{\partial}{\partial u}\alpha_0(t,u) - L_{yy}(\alpha_0(t,u), \alpha_0'(t,u))\frac{\partial}{\partial u}\alpha_0'(t,u)\right].\end{aligned}$$
   Setting $u = 0$ we have
$$\begin{aligned}&L_{xx}(\alpha_0, \alpha_0')h_0(t) + L_{xy}(\alpha_0, \alpha_0')h_0'(t)\\ &- \frac{d}{dt}\left[L_{xy}(\alpha_0, \alpha_0')h_0(t) - L_{yy}(\alpha_0, \alpha_0')h_0'(t)\right]\\ &= (L_{xx}(\alpha_0, \alpha_0') - L_{xy}(\alpha_0, \alpha_0'))h_0(t) - \frac{d}{dt}L_{yy}(\alpha_0, \alpha_0')h_0'(t)\\ &= A(t)h(t) - \frac{d}{dt}B(t)h'(t) = 0.\end{aligned}$$
   We generated a solution of the Jacobi equation through a variation of the stationary path. □

3. Given the interval $[a,b]$ and
$$J(h) = A(t)h(t) - \frac{d}{dt}\left[B(t)h'(t)\right],$$
   with $h(a) = h(b) = 0$, h is C^2. Suppose
$$\int_a^b J(h_1)h_2 dt = \int_a^b J(h_2)h_1 dt$$
   for h_1, h_2 both C^2 then $J(h)$ is a symmetric operator. J can then be extended to what is called an unbounded self-adjoint operator on $L^2([a,b], dx)$. J will have a discrete spectrum and there is a sequence $\lambda_n \to \infty$ such that $J(h_n) = \lambda_n h_n$. Further, for each

n there exists an h_n, a unique solution to this equation, with $h_n(a) = h_n(b) = 0$. The $\{h_n\}$ form a complete orthonormal basis for $L^2([a,b], dx)$. So

$$\int_a^b (J(h)h)dt \geq \lambda_1 \geq \int_a^b |h(t)|^2 dt \qquad \text{if } \lambda_1 > 0.$$

If $\lambda_1 < 0$ then

$$\int_a^b J(h_1)h_1 dt \geq \lambda_1 = \int_a^b |h_1(t)|^2 dt < 0$$

and hence for some α_0 this is not an extremum.

1.9.2 First example: Free motion

The Lagrangian in this case is given by $L(x,y) = \frac{1}{2}y^2$, so $S(\alpha) = \frac{1}{2}\int_a^b [\alpha(t)]^2 dt$ with boundary conditions $\alpha(a) = x_0$ and $\alpha(b) = y_0$. Hence $L_x = 0$, $L_{xy} = 0$ and $L_{yy} = 1$. The Euler-Lagrange equation then reduces to $-\alpha''(t) = 0$. The solution is given by $\alpha_0''(t) = At + B$ with A and B satisfying $Aa + B = x_0$ and $Ab + B = y_0$. Hence $A = \frac{x_0 - y_0}{a-b}$ and $B = \frac{a(y_0 - x_0)}{a-b}$.

For the Jacobi equation, we have $A(t) = 0$ and $B(t) = 1$ so the equation reduces to $-h''(t) = 0$. This is solved by $h(t) = \tilde{A}t + \tilde{B}$. Imposing $h(a) = h(b) = 0$, we get $\tilde{A} = \tilde{B} = 0$ and so $h \equiv 0$ is the only solution. There are therefore no conjugate points and the critical path $\alpha(t) = At + B$ is a local minimum by the Jacobi Theorem.

Note that we can get this result directly as follows. Let h be a C^1 path such that $h(a) = h(b) = 0$.

$$\begin{aligned} S(\alpha_0 + h) &= \frac{1}{2}\int_a^b [\alpha_0'(t) + h'(t)]^2 dt \\ &= \frac{1}{2}\int_a^b [\alpha_0'(t)]^2 dt + \int_a^b \alpha_0'(t)h'(t)dt + \frac{1}{2}\int_a^b [h'(t)]^2 dt \end{aligned}$$

For the second term, integrating by parts and using $h(a) = h(b) = 0$,

$$\int_a^b \alpha_0(t)\frac{d}{dt}h(t)dt = -\int_a^b \alpha_0''(t)h(t)dt = 0$$

, where the last equality comes from the E-L equation $\alpha_0''(t) = 0$.

Hence

$$S(\alpha_0 + h) = S(\alpha_0) + \frac{1}{2}\int_a^b [h'(t)]^2 dt > 0$$

unless $h \equiv 0$.

1.9.3 Second example: Harmonic Oscillator

The Lagrangian is given by $L(x,y) = \frac{1}{2}(y^2 - x^2)$ so that $L_x = -x$, $L_{xx} = -1$, $L_{yy} = 1$ and $L_{xy} = 0$. Hence for $S(\alpha) = \frac{1}{2}\int_a^b [\alpha'(t)]^2 - [\alpha(t)]^2 dt$, E-L becomes $-\alpha_0(t) - \alpha_0''(t) = 0$, which is solved by $\alpha_0(t) = A\cos(t) + B\sin(t)$. Applying boundary conditions:

$$\alpha_0(a) = A\cos(a) + B\sin(a) = x_0$$

$$\alpha_0(b) = A\cos(b) + B\sin(b) = y_0$$

The determinant of this system is $AB(\cos(a)\sin(b) - \cos(b)\sin(a) = AB(\sin(a-b))$ which is zero if $b - a = n\pi$. So is $b - a \neq n\pi$, then we have the unique solution for all x_0 and y_0. If $b - a = n\pi$ and n is odd, the two equations of our system are the same so there are infinitely many solutions if $x_0 = y_0$ and none otherwise. Similarly, if $b - a = n\pi$ and n is odd, the equations are negative of each other so there are infinitely many solutions if $x_0 = -y_0$ and none otherwise.

We now turn to the Jacobi equation to investigate the nature of the critical paths. Assume $a = 0$. Then if $b \neq n\pi$ and $b > 0$, the above solution becomes

$$\alpha_0(t) = x_0\cos(t) + \frac{(y_0 - x_0\cos(b))}{\sin(b)}\sin(t).$$

Case 1: $0 < b < \pi$. Let $\alpha_0(t)$ be a critical path. The solutions to the Jacobi equations $-h(t) - h''(t) = 0$ are given by: $h(t) = \tilde{A}\cos(t) + \tilde{B}\sin(t)$. $h(0) = 0$ implies $\tilde{A} = 0$ and $h(b) = 0$ imples $\tilde{B}\sin(b) = 0$. So we need $b = n\pi$, which cannot happen for our choice of b. Hence there are no conjugate points and by the Jacobi Theorem, $\alpha_0(t)$ is a local minimum.

This can also be seen as follows. Take $h \in C^1[0, b]$ such that $h(0) = h(b) = 0$. Then

$$\begin{aligned} S(\alpha_0 + h) &= \frac{1}{2}\int_0^b [\alpha_0'(t) + h'(t)]^2 dt - \frac{1}{2}\int_0^b [\alpha_0(t) + h(t)]^2 dt \\ &= \frac{1}{2}\int_0^b (\alpha_0'(t))^2 - \alpha_0(t)^2 dt - \int_0^b \alpha_0'(t)h'(t) - \alpha_0(t)h(t)dt + \frac{1}{2}\int_0^b (h'(t))^2 - h(t)^2 dt \end{aligned}$$

As in example (1), we can integrate by parts the second term to see that it must be zero for a critical path $\alpha_0(t)$ by the E-L equation. So

$$S(\alpha_0 + h) = S(\alpha_0) + \frac{1}{2}\int_0^b (h'(t))^2 - h(t)^2 dt.$$

To get the sign of the last term, expand h in a Fourier sine series:

$$h(t) = c_0 \sum_{n=0}^{\infty} c_n \sin\left(\frac{n\pi t}{b}\right)$$

Then $\int_0^b |h(t)|^2 dt = c_0^2 \sum_{n=0}^{\infty} |c_n|^2$ and

$$h'(t) = c_0 \sum_{n=0}^{\infty} \frac{n\pi}{b}\cos\left(\frac{n\pi}{b}\right)$$

So

$$\int_0^b |h'(t)|^2 dt = |c_0|^2 \sum_{n=1}^{\infty} \frac{n^2\pi^2}{b^2}|c_n|^2 \geq \frac{\pi^2}{b^2}|c_0|^2 \sum_{n=1}^{\infty} |c_n|^2 = \frac{\pi^2}{b^2}\int_0^b |h(t)|^2 dt$$

In our case, $0 < b < \pi$, so $\int_0^b |h'(t)|^2 dt \geq \int_0^b |h(t)|^2 dt$ for all such h. α_0 is thus a global maximum.

A complete set of eigenfunction of Jacobi operator $h''(t) - h(t) = 0$ coupled with the condition $h(0) = h(b) = 0$ is the set $\{\sin(\frac{n\pi t}{b}), n \geq 1\}$ with eigenvalues $\{\frac{n^2\pi^2}{b^2} - 1, n \geq 1\}$. Note that the eigenvalues are positive for $b < \pi$.

Case 2: $b = \pi$. For the Jacobi operator, the lowest eigenvalue is zero. So there exists a solution of the Jacobi equation which is nontrivial: $\sin(t)$. Since π is conjugate to 0, we cannot apply the Jacobi Theorem (there is a conjugate point in $(0, \pi]$). For the system to have a solution, we need $\alpha_0(0) = x_0$ and $\alpha_0(\pi) = -x_0$. Then $\alpha_0^A(t) = x_0 \cos(t) + A \sin(t)$ is a stationary path for each A.

As before, write

$$\begin{aligned} S(\alpha_0^A + h) &= \frac{1}{2}\int_0^\pi [(\alpha_0^A + h)']^2 dt - \frac{1}{2}\int_0^\pi [\alpha_0^A + h]^2 dt \\ &= S(\alpha_0^A) + \frac{1}{2}\int_0^\pi [h'(t)]^2 - [h(t)]^2 dt, \end{aligned}$$

where the last term is ≥ 0. So we still have local minima, but not strict local minima. To see this, take $h_{\epsilon(t)} = \epsilon \cos(t)$: the second term above then integrates to zero.

We conclude that for this case, there are infinitely many stationary paths of the form $\alpha_0(t) = x_0 \cos(t) + A \sin(t)$, $A \in \mathbb{R}$. Each of these is a local minimum, but not a strict local minimum.

Case 3: $b > 2\pi$ ($\pi < b < 2\pi$, $b = 2\pi$). Here π is conjugate to 0 and $0 < \pi < b$. So by the Jacobi Theorem, $\alpha_0(t)$ cannot be a local minimum. It must therefore be a saddle point. Once again, write

$$S(\alpha_0 + h) = S(\alpha_0) + \frac{1}{2}\int_0^b [h'(t)]^2 - [h(t)]^2 dt.$$

and take $h_\epsilon(t) = \epsilon \sin(\frac{\pi t}{b})$ so the last term becomes:

$$\frac{\epsilon^2}{2}\int_0^b \frac{\pi^2}{b^2}\cos^2\left(\frac{\pi t}{b}\right) - \sin^2\left(\frac{\pi t}{b}\right) dt = \frac{\epsilon^2(\pi^2 - b^2)}{4b} < 0.$$

We then have $S(\alpha_0) > S(\alpha_0 + h_\epsilon)$ which implies that α_0 cannot be a local minimum. It is a saddle point.

Note that as b crosses π, some eigenvalues of the Jacobi equation become negative. Also, if $b = 2\pi$, there are infinitely many solutions, each of which is a saddle point.

1.9.4 Third example: Soap Film

The setup consists in two rings of radius r and R whose boundaries are connected by a soap film. The distance l between the two rings is then increased while the film naturally settles in the position of minimal surface area. Center both ring about the t axis and let $\alpha(t)$ be the function whose revolution about the t axis generates the surface of the soap film.

In this case, our Lagrangian is:

$$L(x, y) = 2\pi x\sqrt{1 + y^2}$$

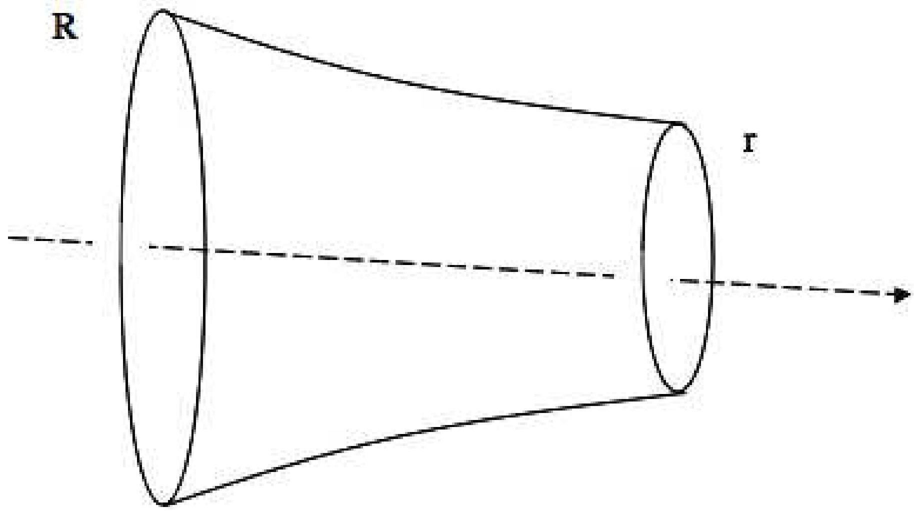

Figure 1: The system.

defined on $\mathcal{U} = \{(x, y) \in \mathbb{R}^2 \, x > 0\}$. Correspondingly, $\alpha : [0, l] \to \mathbb{R}$ and $\alpha(t)$ is assumed to be positive and C^1. The area of the surface of revolution generated by $\alpha(t)$ is then:

$$S(\alpha) = \int_0^l 2\pi\alpha(t)\sqrt{1 + [\alpha'(t)]^2}dt.$$

We will denote by $\alpha_0(t)$ a path solving the E-L equation and satisfying $\alpha_0(0) = r$ and $\alpha_0(l) = R$. To express E-L, we first calculate the derivatives $L_x = 2\pi x\sqrt{1 + y^2}$, $L_{yx} = \frac{2\pi y}{\sqrt{1+y^2}}$ and $\frac{2\pi x}{(1+y^2)^{\frac{3}{2}}}$. So E-L becomes:

$$\begin{aligned}
&2\pi\sqrt{1 + [\alpha_0'(t)]^2} - 2\pi\frac{[\alpha_0'(t)]^2}{\sqrt{1 + [\alpha_0'(t)]^2}} - 2\pi\frac{\alpha_0(t)\alpha_0''(t)}{(1 + [\alpha_0)]^2)^{\frac{3}{2}}} = 0 \\
&\iff \frac{\alpha_0''(t)\alpha_0(t)}{(1 + [\alpha_0'(t)]^2)^{\frac{5}{2}}} = \frac{1}{\sqrt{1 + [\alpha_0'(t)]^2}} \\
&\iff \alpha_0''(t)\alpha_0(t) = 1 + [\alpha_0'(t)]^2
\end{aligned}$$

We now turn to the following general fact:

$$\begin{aligned}
\frac{d}{dt}[L(\alpha_0, \alpha_0') - \alpha_0' L_y(\alpha_0, \alpha_0')] &= L_x(\alpha_0, \alpha_0')\alpha_0' + L_y(\alpha_0, \alpha_0')\alpha_0'' - \alpha_0'' L_y(\alpha_0, \alpha_0') - \\
&\quad - \alpha_0' L_{yx}(\alpha_0, \alpha_0')\alpha_0' - \alpha_0' L_{yy}(\alpha_0, \alpha_0')\alpha_0'' \\
&= L_x(\alpha_0, \alpha_0')\alpha_0' - \alpha_0' L_{yx}(\alpha_0, \alpha_0')\alpha_0' - \alpha_0' L_{yy}(\alpha_0, \alpha_0')\alpha_0'' \\
&= \alpha_0'[L_x(\alpha_0, \alpha_0') - L_{yx}(\alpha_0, \alpha_0')\alpha_0' - L_{yy}(\alpha_0, \alpha_0')\alpha_0''] \\
&= 0
\end{aligned}$$

because the term in bracket in the last equation is E-L. Hence

$$\frac{\alpha_0(t)}{\sqrt{1 + [\alpha_0'(t)]^2}} = A$$

and we have

$$\alpha_0(t) = A\sqrt{1 + [\alpha_0'(t)]^2}.$$

We square the second equation and divide to find

$$\alpha_0''(t) = \frac{1}{A^2}\alpha_0(t),$$

yielding the solution

$$\alpha_0(t) = C_1 e^{t/A} + C_2 e^{-t/A}.$$

We must ensure that C_1 and C_2 are compatible with

$$[\alpha_0(t)]^2 = A^2(1 + [\alpha_0'(t)]^2),$$

so substituting our solution in this equation, we have

$$C_1^2 e^{2t/A} + C_2^2 e^{-2t/A} + 2C_1C_2 = A^2\left[1 + \frac{C_1^2}{A^2}e^{2t/A} + \frac{C_2^2}{A^2}e^{-2t/A} - 2\frac{C_1C_2}{A^2}\right]$$

that is,

$$C_1 = \frac{A^2}{4C_1}.$$

We can now rewrite our solution in the form

$$\alpha_0(t) = \frac{A}{2}\left[\frac{A}{2C_2}e^{t/A} + \frac{2C_2}{A}e^{-t/A}\right], \qquad C_1, C_2 > 0, A > 0.$$

Let $\frac{A}{2C_2} = e^{-D}$ then, with $AD = B$ we obtain

$$\alpha_0(t) = A\cosh\left(\frac{t-B}{A}\right).$$

One verifies that this family of functions is indeed the solution of the E-L equations. Moreover, we now find

$$\begin{aligned}
S(\alpha_0(t)) &= 2\pi\int_0^l \alpha_0(t)\sqrt{1 + [\alpha_0(t)]^2}dt \\
&= 2\pi\int_0^l \cosh\left(\frac{t-B}{A}\right)\sqrt{1 + \sinh^2\left(\frac{t-B}{A}\right)}dt \\
&= 2\pi\int_0^l \cosh^2\left(\frac{t-B}{A}\right)dt \\
&= \frac{2\pi A}{2}\int_0^l \left[1 + \cosh\left(\frac{2(t-B)}{A}\right)\right]dt \\
&= \pi Al + \frac{\pi A^2}{2}\left[\sinh\left(\frac{2(t-B)}{A}\right)\right]_0^l \\
&= \pi Al + \frac{\pi A^2}{2}\sinh\left(\frac{2(l-B)}{A}\right) + \sinh\left(\frac{2b}{A}\right)
\end{aligned}$$

as the formula for the area of $S(\alpha_0)$. We now work with the boundary conditions

$$\alpha_0(0) = r, \qquad \alpha_0(0) = R.$$

Here we do *only* the case $r = R$ (in fact $r = R = 1$, without loss of generality). The boundary conditions become

$$A \cosh\left(\frac{B}{A}\right) = 1$$
$$A \cosh\left(\frac{l - B}{A}\right) = 1.$$

Since

$$\cosh x = \cosh y \iff x = \pm y$$

we must have here that

$$\frac{B}{A} = \pm\left(\frac{l-B}{A}\right), \qquad \text{i.e.} B = \pm(l - b).$$

(-) Case: $l = 0$.
(+) Case: $2B = l \Rightarrow B = l/2$.

We are still looking for A such that $A\cosh(l/2A) = 1$. Let $f(A) = A\cosh(l/2A)$, a function of $A > 0$. Then

$$\lim_{A\to 0} f(A) = \infty, \qquad \lim_{A\to\infty} f(A) = \infty$$

and with

$$f'(A) = \cosh\left(\frac{l}{2A}\right) - \frac{l}{2A}\sinh\left(\frac{l}{2A}\right)$$

then

$$f'(A) = 0 \iff \frac{2A}{l} = \frac{\sinh\left(\frac{l}{2A}\right)}{\cosh\left(\frac{l}{2A}\right)}.$$

Consider now the additional function

$$t(x) = \frac{\sinh(x)}{\cosh x} - \frac{1}{x}$$

then

$$\lim_{t\to 0} t(x) = -\infty, \qquad \lim_{x\to\infty} t(x) = 1$$

and

$$\begin{aligned} t'(x) &= \frac{\cosh^2 x - \sinh^2 x}{\cosh^2 x} + \frac{1}{x^2} \\ &= \frac{1}{\cosh^2 x} + \frac{1}{x^2} > 0. \end{aligned}$$

So there exists a unique x_0 such that $t(x_0) = 0$, i.e. for which

$$\frac{\sinh x_0}{\cosh x_0} = \frac{1}{x_0}$$

yielding numerically that $x_0 \simeq 1.21$. Further $t(x) < 0$ for $x < x_0$ and $t(x) > 0$ for $x > x_0$. Thus our function is decreasing for $A \in (0, A_l)$, increasing on (A_l, ∞) and has a global minimum at

$$A_l = \frac{l}{2x_0}.$$

What is

$$f(A_l) = \min_A f = \frac{l}{2x_0} \cosh x_0 = \frac{l}{2} \sinh x_0.$$

There are three cases:

1. $\frac{l}{2} \sinh x_0 < 1$. The equation $A \cosh(l/2A) = 1$ has two solutions, $A_l^{(1)} < l/2x_0 < A_l^{(2)}$.

2. $\frac{l}{2} \sinh x_0 = 1$. Then we have a unique solution at $A_l = l/2x_0$. Finally,

3. $\frac{l}{2} \sinh x_0 > 1$. Here there are no solutions.

The critical length is given by $l_0 = 2/\sinh x_0$.

Conclusion.

1. For $0 < l < l_0$ there are two solutiont, $\alpha_0^{(1)}$, $\alpha_0^{(2)}$ of the E-L equations satisfying the required boundary conditions

$$\alpha_0^{(k)}(0) = \alpha_0^{(k)}(l) = 1$$

given by

$$\alpha_0^{(k)}(t) = A_l^{(k)} \cosh\left(\frac{t - l/2}{A_l^{(k)}}\right).$$

2. For $l = l_0$, there exists a unique solution

$$\alpha_0(t) = A_{l_0} \cosh\left(\frac{t - l/2}{A_{l_0}}\right).$$

with

$$A_{l_0} = \frac{l_0}{2x_0} = \frac{1}{x_0 \sinh x_0} = \frac{1}{\cosh x_0}.$$

3. When $l > l_0$ there are no solutions.

1.9.5 The Nature of the Critical Path

We have $A_l^{(k)}$, for $k = 1, 2$, $0 < l < l_0$, with

$$A_l^{(k)} \cosh\left(\frac{l}{2A_l^{(k)}}\right) = 1,$$

and $A_l^{(1)} < l/2x_0 < A_l^{(2)}$. By the implicit function theorem, the $A_l^{(k)}$ are differentiable for $0 < l < l_0$ (in fact infinitely differentiable, in fact real analytic).

Implicit Differentiation We have

$$\frac{d}{dl} A_l^{(k)} \cosh\left(\frac{l}{2A_l^{(k)}}\right) = 0$$

that is

$$\left(\frac{d}{dl} A_l^{(k)}\right) \cosh\left(\frac{l}{2A_l^{(k)}}\right) + \frac{1}{2} \sinh\left(\frac{l}{2A_l^{(k)}}\right) A_l^{(k)} \left[\frac{A_l^{(k)} - l\frac{d}{dl} A_l^{(k)}}{[A_l^{(k)}]^2}\right] = 0.$$

Hence

$$\frac{d}{dl} A_l^{(k)} \left[\cosh\left(\frac{l}{2A_l^{(k)}}\right) - \frac{l}{2A_l^{(k)}} \sinh\left(\frac{l}{2A_l^{(k)}}\right)\right] = -\frac{1}{2} \sinh\left(\frac{l}{2A_l^{(k)}}\right).$$

and from this we can isolate $\frac{d}{dl} A_l^{(k)}$. We have that

$$\frac{l}{2A_l^{(1)}} > x_0 > \frac{l}{2A_l^{(2)}},$$

and

$$\frac{\sinh x}{\cosh x} - \frac{1}{x} \begin{cases} > 0 & x > 0 \\ < 0 & x < 0 \end{cases}$$

so we see that

$$\frac{d}{dl} A_l^{(1)} > 0, \qquad \frac{d}{dl} A_l^{(2)} < 0$$

for $0 < l < l_0$. Since $A_l^{(1)} < l/2x_0$, $\lim_{l\to 0} A_l^{(1)} = 0$. Since $A_l^{(2)}$ is decreasing, $\lim_{l\to 0} A_l^{(2)}$ exists, equals a, say, and is positive. From

$$A_l^{(2)} \cosh\left(\frac{l}{2A_l^{(2)}}\right) = 1,$$

we take the limit as $l \to 0$ to find

$$a \cosh 0 = 1 \Rightarrow a = 1.$$

Moreover,

$$\lim_{l\to l_0} A_l^{(k)} = \frac{l_0}{2x_0} = A_{l_0}.$$

Let us now compute

$$S(\alpha) = 2\pi \int_0^l \alpha_0 \sqrt{1 + [\alpha_0']^2} dt.$$

For the critical paths $\alpha_0^{(k)}(t)$, $k = 1, 2$. For a general formula, we substitute

$$
\begin{aligned}
S(\alpha_0^{(k)}) &= \pi A_l^{(k)} l + \cancel{2} \sinh\left(\frac{l}{A_l^{(k)}}\right) \frac{\pi (A_l^{(k)})^2}{\cancel{2}} \\
&= \pi A_l^{(k)} l + \pi (A_l^{(k)})^2 \sinh\left(\frac{l}{2A_l^{(k)}}\right) \cosh\left(\frac{l}{2A_l^{(k)}}\right) \\
&= \pi A_l^{(k)} l + 2\pi A_l^{(k)} \sinh\left(\frac{l}{2A_l^{(k)}}\right) \\
&= \pi A_l^{(k)} l + 2\pi A_l^{(k)} \sqrt{\cosh\left(\frac{l}{2A_l^{(k)}}\right) - 1}
\end{aligned}
$$

yielding

$$S(\alpha_0^{(k)}) = \pi A_l^{(k)} l + 2\pi \sqrt{1 - (A_l^{(k)})^2}.$$

We want to study this as a function of l. Noting that

$$\frac{d}{dl} S(\alpha_0^{(k)}) = 2\pi A_l^{(k)}$$

we conclude that $S(\alpha_0^{(k)})$ is an increasing function of l! The derivative is positive and from the expression

$$\frac{d^2}{dl^2} S(\alpha_0^{(k)}) = 2\pi \frac{d}{dl} A_l^{(k)}$$

we have that $S(\alpha_0^{(1)})$ is convex while $S(\alpha_0^{(2)})$ is concave. We also know that

$$\lim_{l \to 0} S(\alpha_0^{(1)}) = 2\pi, \qquad \text{since } \lim_{l \to 0} A_l^{(1)} = 0$$

and

$$\lim_{l \to 0} S(\alpha_0^{(2)}) = 0, \qquad \text{since } \lim_{l \to 0} A_l^{(1)} = 1.$$

We also know that

$$\lim_{l \to l_0} S(\alpha_0^{(k)}) + \pi A_{l_0} l_0 + 2\pi \sqrt{1 - (A_{l_0})^2} > 2\pi,$$

by the relations

$$A_{l_0} = \frac{l_0}{2x_0}, \qquad l_0 = \frac{2}{\sinh x_0}, \qquad \frac{\sinh x_0}{\cosh x_0} = \frac{1}{x_0}.$$

and that $x_0 = 1.21$. Thus

$$\lim_{l \to l_0} \frac{d}{dl} S(\alpha_0^{(k)}) = 2\pi A_{l_0} = \frac{\pi l_0}{x_0} = \frac{2\pi}{\cosh x_0}$$

and

$$\lim_{l \to l_0} \frac{d^2}{dl^2} S(\alpha_0^{(k)}) = \begin{cases} -\infty & k = 2 \\ \infty & k = 1 \end{cases}$$

and we can plot the graph of these two equations (see Figure 2 below).

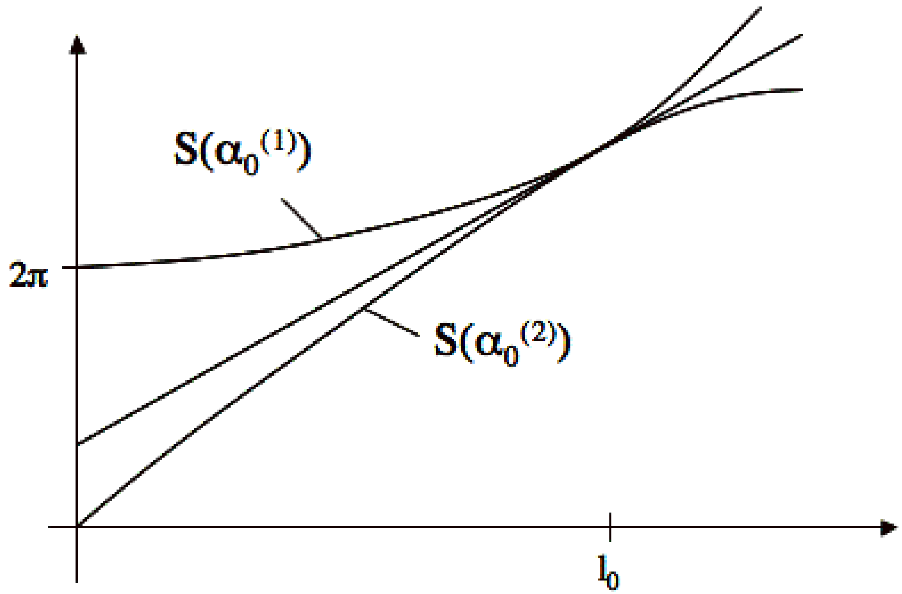

Figure 2: $S(\alpha_0^{(1)})$ is a non-physical path while $S(\alpha_0^{(2)})$ is physical.

1.9.6 The Nature of Stationary Points

At this stage, we must solve the Jacobi equation for conjugate points. We need to look at the Jacobi equations along each critical path.

Recall

$$L(x, y) = 2\pi\sqrt{1+y^2}, \qquad L_{yy} = \frac{2\pi x}{(1+y^2)^{3/2}} > 0, \quad x > 0.$$

Thus the critical path is either a local minimum or a saddle point. We solve the Jacobi equation using variations through the sationary path. Let

$$\alpha_0^{(k)}(t, u_1, u_2) = (A_l^{(k)} + u_1)\cosh\left(\frac{t - l/2 - u_2}{A_l^{(k)} + u_1}\right)$$

then $\alpha_0^{(k)}(t, u_1, u_2)$ solves the E-L equations for all u_1, u_2 and $\alpha_0^{(k)}(t, 0, 0) = \alpha_0^{(k)}(t)$. We compute

$$\left.\frac{\partial}{\partial u_1}\alpha_0^{(k)}(t, u_1, u_2)\right|_{u_1=u_2=0} = \cosh\left(\frac{t-l/2}{A_l^{(k)}}\right) - \left(\frac{t-l/2}{A_l^{(k)}}\right)\sinh\left(\frac{t-l/2}{A_l^{(k)}}\right) \equiv h_1^{(k)}(t)$$

and

$$\left.\frac{\partial}{\partial u_2}\alpha_0^{(k)}(t, u_1, u_2)\right|_{u_1=u_2=0} = \sinh\left(\frac{t-l/2}{A_l^{(k)}}\right) \equiv h_2^{(k)}(t)$$

two linearly independent solutions of the Jacobi equation. Thus any solution of the Jacobi equation associated to $\alpha_0^{(k)}$ is of the form

$$C_1 h_1^{(k)}(t) + C_2 h_2^{(k)}(t).$$

We are now looking for points $a \in (0, l]$ conjugate to 0. Recall that a points a is conjugate to 0 is there exists a non-trivial solution $h^{(k)}$ of the Jacobi equation such that

$$h^{(k)}(a) = h^{(k)}(0) = 0.$$

We set up the sustem of equations

$$0 = C_1 h_1^{(k)}(a) + C_2 h_2^{(k)}(a)$$

$$0 = C_1 h_1^{(k)}(0) + C_2 h_2^{(k)}(0)$$

and note that a is conjugate to 0 if the system has a non-trivial solutions, which is true if and only if

$$\Delta^{(k)}(a) = \begin{vmatrix} h_1^{(k)}(0) & h_1^{(k)}(a) \\ h_2^{(k)}(0) & h_2^{(k)}(a) \end{vmatrix},$$

which we compute as

$$\begin{aligned}\Delta^{(k)}(a) = &\left[\cosh\left(\frac{l}{2A_l^{(k)}}\right) - \frac{l}{2A_l^{(k)}}\sinh\left(\frac{l}{2A_l^{(k)}}\right)\right]\sinh\left(\frac{a-l/2}{A_l^{(k)}}\right) \\ &+ \left[\cosh\left(\frac{a-l/2}{A_l^{(k)}}\right) - \frac{a-l/2}{A_l^{(k)}}\sinh\left(\frac{a-l/2}{A_l^{(k)}}\right)\right]\sinh\left(\frac{l}{2A_l^{(k)}}\right).\end{aligned}$$

We have

$$\Delta^{(k)}(0) = 0$$

$$\Delta^{(k)}\left(\frac{l}{2}\right) = \sinh\left(\frac{l}{2A_l^{(k)}}\right) > 0.$$

$$\begin{aligned}\Delta^{(k)}(l) = &\left[\cosh\left(\frac{l}{2A_l^{(k)}}\right) - \frac{l}{2A_l^{(k)}}\sinh\left(\frac{l}{2A_l^{(k)}}\right)\right]\sinh\left(\frac{l}{2A_l^{(k)}}\right) \\ &+ \left[\cosh\left(\frac{l}{2A_l^{(k)}}\right) - \frac{l}{2A_l^{(k)}}\sinh\left(\frac{l}{2A_l^{(k)}}\right)\right]\sinh\left(\frac{l}{2A_l^{(k)}}\right).\end{aligned}$$

When $l = l_0$, $l_0/2A_{l_0}^{(k)} = x_0$, hence we obtain that $\Delta^{(k)}(l_0) = 0$. The cricitcal length is conjugate to 0! We don't know if the critical path at $l = l_0$ is a local minimum or a saddle point. Yet. We now look at $0 < l < l_0$. The sign of $\Delta^{(k)}(l)$ depends on the factor

$$\cosh\left(\frac{l}{2A_l^{(k)}}\right) - \frac{l}{2A_l^{(k)}}\sinh\left(\frac{l}{2A_l^{(k)}}\right).$$

Therefore

$$A_l^{(1)} < \frac{l}{2x_0} < A_l^{(2)} \Longrightarrow \Delta^{(1)}(l) < 0, \;\; \Delta^{(2)}(l) > 0,$$

hence there exists an $a \in (l/2, l]$ such that $\Delta^{(1)}(a) = 0$. So there is a conjugate point for the Jacobi equation associated to $\alpha_0^{(1)}$ - it is a saddle point!

It thus remains to show that there is no solution for $\Delta^{(2)}(a) = 0$. We differentiate further

$$\frac{d}{da}\Delta^{(2)}(a) = \left[\cosh\left(\frac{l}{2A_l^{(2)}}\right) - \frac{l}{2A_l^{(2)}}\right]\cosh\left(\frac{a-l/2}{A_l^{(2)}}\right)\cdot\frac{1}{A_l^{(2)}}$$
$$-\frac{a-l/2}{A_l^{(2)}}\frac{1}{A_l^{(2)}}\cosh\left(\frac{a-l/2}{A_l^{(2)}}\right)\sinh\left(\frac{l}{2A_l^{(2)}}\right).$$

since $a < l/2$, $\frac{d}{da}\Delta^{(2)}(a) > 0$, that is, $\Delta^{(2)}(a)$ is increasing. By factoring out the cosine terms above (which are strictly > 0), the above is a linear equation with a unique solution. Hence there are no conjugate points and $\alpha_0^{(2)}$ is a local minimum.

1.9.7 The $\mathbb{R}^n$ case

Here, we take the Banach space E to be $\mathbb{R}^n$. Let $\mathcal{U} \subset \mathbb{R}^n \times \mathbb{R}^n$ be an open set and $L : \mathcal{U} \to \mathbb{R}$ be the Lagrangian (assume also that it is C^2 or C^3). Let $A = (a_1, ., a_n)$ and $B = (b_1, \ldots, b_n)$ be points in $\mathbb{R}^n$. Using the same notation as before, denote $V_{A,B,p} = C^p_{A,B}(I, \mathbb{R}^n) \cap \mathcal{V}_p$, where as usual $I = [a, b]$. Take $\alpha \in V_{A,B,p}$. Then $(\alpha(t), \alpha'(t)) \in \mathcal{U}$, $t \in [a, b]$ and $\alpha(a) = A$, $\alpha(b) = B$. We are interested in the local extrema of $S(\alpha) = \int_a^b L(\alpha(t), \alpha'(t))dt$.

Recall that

$$L_x = \nabla_x L(x, y), \qquad L_y = \nabla_y L(x, y), \qquad L_{xx} = [\frac{\partial^2 L}{\partial x_i \partial x_j}]$$

(and similarly for L_{yy}), $L_{xy} = [\frac{\partial^2 L}{\partial x_i \partial y_j}]$ and $L_{yx} = L_{xy}^T$. Hence for a critical path $\alpha_0(t)$ to satisfy the E-L equation

$$L_x(\alpha_0(t), \alpha_0'(t)) - L_{xy}(\alpha_0(t), \alpha_0'(t))\alpha_0'(t) - L_{yy}(\alpha_0(t), \alpha_0'(t))\alpha_0''(t) = 0,$$

there are n nonlinear ODE's to be solved (one per component).

If L_{yy} is invertible in $\mathcal{U}$, then the system can be written as

$$\alpha_0''(t) = [L_{yy}(\alpha_0(t), \alpha_0'(t))]^T[L_x(\alpha_0(t), \alpha_0'(t)) - L_{xy}(\alpha_0(t), \alpha_0'(t))\alpha_0'(t)].$$

By the Banach fixed point theorem, if L is C^3, there exists solutions for short time. Moreover, the solution is unique given $\alpha_0(a)$ and $\alpha_0'(a)$. A solution satisfying the boundary conditions $\alpha_0(a) = A$ and $\alpha_0(b) = B$ however may not exist even for short time and if it does exist, it may not be unique.

If the critical path exists (i.e. satisfying E-L and the boundary conditions), we look at the second variation to determine the nature of the critical path. Specifically, consider $\tilde{(S)}(h) = S(h - \alpha_0)$, where $h \in C_{0,0}$, and look at

$$\tilde{S}''(o)(h, h) = \int_a^b [L_{xx}(\alpha_0(t), \alpha_0'(t))h(t)] \cdot h(t) + 2[L_{xy}(\alpha_0(t), \alpha_0'(t))h(t)] \cdot h'(t)$$
$$+ [L_{yy}(\alpha_0(t), \alpha_0'(t))h'(t)] \cdot h'(t)dt$$

The second variation can be rewritten using integration by parts as:

$$\tilde{S}''(0)(h, h) = \int_a^b [A(t)h(t)] \cdot h(t) + [B(t)h'(t)] \cdot h'(t)$$

where $A(t) = L_{xx}(\alpha_0(t), \alpha_0'(t)) - \frac{d}{dt}L_{xy}(\alpha_0(t), \alpha_0'(t))$ and $B(t) = L_{yy}(\alpha_0(t), \alpha_0'(t))$. Note that the derivative of L_{xy} is to be taken on every entry of the matrix.

Proposition 33. *(Lagrange necessary condition) If $\alpha_0(t)$ is a local minimum, then $B(t) \geq 0$ for all $t \in [a, b]$, that is all eigenvalues of $B(t)$ are non negative.*

Proof. Suppose there is $\bar{t}$ such that $B(\bar{t})$ has a negative eigenvalue $\lambda(\bar{t})$. Let $v(\bar{(t)}) = [v_1, \dots, v_n]$ be the eigenvector and take $\tilde{h}_n(t) = [v_1 h_n(t), \dots, v_n h_n(t)]$, where $h_n(t)$ is the function that was constructed in the same proof for the $E = \mathbb{R}$ case. Then $\lim_{n\to\infty} \tilde{S}(0)(h_n, h_n) = \lambda(\bar{t}) < 0$, so that we cannot have a minimum. □

Integrating by parts the above equation for the second variation, we get:

$$\tilde{S}(0)(h, h) = \int_a^b \left[A(t)h(t) + \frac{d}{dt}(B(t)h'(t)) \right] \cdot h'(t)$$

Let $J(h) = A(t)h(t) + \frac{d}{dt}(B(t)h'(t))$, called the Jacobi operator, which is linear in h and second order. The Jacobi equation is the linear ODE $J(h) = 0$. If L is C^3, solutions exist, form a vector space and are uniquely determined by $h(a)$ and $h'(a)$. The dimension of the vector space of solutions is $2n$. The solutions are called Jacobi fields and can be found as before using variations to stationary paths (and this is why the Jacobi equation is called the small disturbance equation).

Theorem 34. *Let L be C^3 and $\alpha_0(t)$ be a stationary path.*

1. *If α_0 is a local minimum, then there is no point in (a, b) conjugate to a.*

2. *If $L_{yy}(\alpha_0(t), \alpha_0'(t)) > 0$ for all $t \in [a, b]$ and there is no conjugate point to a in $(a, b]$, then α_0 is a local minimum.*