

# Calculus of variations

Typeset on March 4, 2021

## Contents

<b>Introduction</b>	<b>3</b>
<b>Integral functionals</b>	<b>3</b>
<b>Fréchet derivative</b>	<b>6</b>
<b>Extrema</b>	<b>7</b>
<b>Spaces we work with</b>	<b>8</b>
<b>Smoothness of integral functionals</b>	<b>10</b>
$C^0$ integrand implies $C^0$ functional . . . . .	10
$C^1$ integrand implies $C^1$ functional . . . . .	11
<b>A few lemmas</b>	<b>12</b>
Lagrange's lemma . . . . .	12
Lemma of Dubois and Raymond . . . . .	13
A generalisation . . . . .	14
<b>Optimisation with fixed endpoints</b>	<b>16</b>
Smoothness of solutions . . . . .	17

<b>Hypersurfaces</b>	<b>19</b>
<b>Conditional extrema in infinite-dimensional spaces</b>	<b>20</b>
<b>Lagrange multipliers</b>	<b>21</b>
An electrostatic example . . . . .	21

## Introduction

Functions are real by default.  $X$  is always a normed vector space over  $\mathbb{R}$ ,  $J: X \rightarrow \mathbb{R}$  a function. The norm of  $x \in X$  may be denoted  $\|x\|$  as well as  $|x|$ .

**Definition** (reminder). Let  $K \in \{\mathbb{R}, \mathbb{C}\}$ . A *norm* on a vector space  $X$  over  $K$  is a function  $f: X \rightarrow \mathbb{R}$  that satisfies the following requirements:

1.  $f(0) = 0$ .
2. If  $x \in X \setminus \{0\}$ , then  $f(x) > 0$ .
3. If  $k \in K$  and  $x \in X$ , then

$$f(kx) = |k|f(x).$$

4. If  $x_1, x_2 \in X$ , then

$$f(x_1 + x_2) \leq f(x_1) + f(x_2).$$

## Integral functionals

**Definition.** Let  $x, h \in X$ . Consider

$$k: \alpha \mapsto J(x + \alpha h).$$

The *variation* or *Gateaux derivative* of  $J$  at  $x$  in the direction  $h$  is the real

$$dJ(x; h) = \left. \frac{dk}{d\alpha} \right|_{\alpha=0}.$$

It is also (misleadingly, since it is not linear in  $h$ ) denoted as

$$J'_G(x)h.$$

If  $J$  is linear, the variation is linear in  $h$ . But this is not generally the case.

For ‘good’ functionals like the following one, it is linear: put

$$J(f) = \int g \circ f,$$

where  $X = C(\mathbb{R}^m, \mathbb{R}^n)$  and  $g \in C^1(\mathbb{R}^n)$ . Then

$$\begin{aligned} J(x + ah) &= \int g \circ (x + ah), \\ J(x) &= \int g \circ x, \\ \int_t \frac{g(x(t) + ah(t)) - g(x(t))}{a} &\xrightarrow{a \rightarrow 0} \int_t dg(x(t); h(t)). \end{aligned}$$

**Remark.** Variation is homogenous in  $h$ , but not necessarily additive.

*Proof.* Homogeneity:

$$\begin{aligned} dJ(x; th) &= \lim_{\alpha \rightarrow 0} \frac{J(x + \alpha th) - J(x)}{\alpha} \\ &= t \lim_{\beta \rightarrow 0} \frac{J(x + \beta h) - J(x)}{\beta}. \end{aligned}$$

■

**G A P**

**Definition.** The *dual*  $X^*$  of  $X$  is the set of linear and **bounded** real functionals on  $X$ . The *norm* of  $\varphi \in X^*$  is defined as

$$\|\varphi\| = \sup_{x \in X \setminus 0} \frac{|\varphi(x)|}{\|x\|}.$$

**Exercise 1.** This is indeed a norm.

*Solution.* Obviously, it is positively homogenous, non-negative, and  $\|0\| = 0$ . We want to show the triangle inequality:

$$\sup_{x \in X \setminus 0} \frac{|(a+b)(x)|}{\|x\|} \leq \sup_{x \in X \setminus 0} \frac{|a(x)|}{\|x\|} + \sup_{x \in X \setminus 0} \frac{|b(x)|}{\|x\|}.$$

Since  $(a+b)(x) = a(x) + b(x)$ , it follows from the triangle inequality for  $\mathbb{R}$  by passing to the supremum. ■

**Definition.** A sequence  $\{x_n\} \subseteq X$  is called *bounded*, iff  $\{\|x_n\|\}$  is bounded.

**Lemma.** In a normed vector space, every Cauchy sequence is bounded.

*Proof.* Let  $\{x_n\}$  be a Cauchy sequence. There is  $N$  such  $\|x_n - x_N\| \leq 1$  for all  $n \geq N$ . Then  $\|x_n\| \leq \|x_N\| + 1$  for all these  $n$ . Hence, for any  $m \in \mathbb{N}$ ,

$$\|x_m\| \leq \max\{\|x_N\| + 1, \|x_{N-1}\|, \dots, \|x_1\|\}.$$

■

**Exercise 2.**  $X^*$  is complete.

*Solution.* Let  $\{f_n\} \subseteq X^*$  be a Cauchy sequence. It is bounded, so we may define

$$a_n = \inf_{k \geq n} \|f_k\|, \quad b_n = \sup_{k \geq n} \|f_k\|.$$

Functions  $a_n$  and  $b_n$  are non-decreasing and non-increasing, respectively. Moreover, since  $f_n$  is Cauchy,  $|b_n - a_n| \rightarrow 0$ .

Let  $\epsilon > 0$ . Note that

$$\left| \|f_j\| - \|f_i\| \right| \leq \|f_j - f_i\| < \epsilon$$

for sufficiently large  $\min\{i, j\}$ , so  $\{\|f_n\|\}$  is a real Cauchy sequence. It converges as such. ■

## Fréchet derivative

**Definition.** The functional  $J: X \rightarrow \mathbb{R}$  is *differentiable in the sense of Fréchet* at  $x$ , iff there exists  $\varphi \in X^*$  such that

$$J(x+h) = J(x) + \varphi(h) + o(\|h\|)$$

with  $h \rightarrow 0$ . In this case we write

$$J'_F(x) = \varphi.$$

**Lemma.** There exists at most 1 Fréchet derivative.

*Proof.* Suppose there is another,  $\varphi_2$ . Then

$$\varphi(h) - \varphi_2(h) = o(\|h\|).$$

Observe that both sides are positively homogenous in  $h$ . Their relation is then constant, but it also tends to zero. ■

**Lemma.** If Fréchet derivative exists, then the Gateaux derivative for every direction does, and it coincides with the Fréchet one.

*Proof.* Let us find the Gateaux derivative. By definition,

$$\begin{aligned} dJ(x; v) &= \lim_{t \rightarrow 0} \frac{J(x + tv) - J(x)}{t} \\ &= \lim_{t \rightarrow 0} \frac{\varphi(tv) + o(\|tv\|)}{t} \\ &= \lim_{t \rightarrow 0} \frac{t\varphi(v) + |t| \cdot o(\|v\|)}{t} \\ &= \lim_{t \rightarrow 0} (\varphi(v) + \text{sign}(t) \cdot o(1)) \\ &= \varphi(v). \end{aligned}$$

■

**Lemma.** The existence of Gateaux derivative in every direction does not imply existence of Fréchet derivative.

*Proof.* This is for the same reason the existence of directional derivative in every direction does not imply existence of a continuous differential. Consider, for example,

$$f(x, y) = \begin{cases} \frac{2xy}{x^2+y^2}, & (x, y) \neq 0, \\ 0, & \text{otherwise.} \end{cases}$$

In spherical coordinates, we can rewrite it as

$$\widehat{f}(r, \alpha) = \begin{cases} \sin(2\alpha), & r \neq 0, \\ 0, & \text{otherwise.} \end{cases}$$

It is clear from this that the limit of  $f$  at 0 depends on the angle, under which we approach the 0. Hence it cannot be continuous, not to say differentiable. On the other hand, the derivative  $\partial_r \widehat{f}(0, \alpha)$  in every direction  $\alpha$  exists and is equal to zero, as  $\sin(2(\alpha + \pi)) = \sin(2\alpha)$ . ■

**Definition.** We write  $J \in C^1(X)$  iff for every  $x \in X$  there is  $J'_F(x)$  and the map

$$x \mapsto J'_F(x)$$

is continuous.

## Extrema

**Definition.**  $x$  is a *local maximum* of  $J$ , iff there exists such  $\delta > 0$  that  $J(x) \geq J(x_2)$  for every  $x_2 \neq x$  with  $\|x - x_2\| < \delta$ . We call it *strict*, iff  $J(x) > J(x_2)$ .

**Lemma.** Let  $x$  be a local extremum of  $J$ ,  $h \in X$ . If  $dJ(x; h)$  exists, then  $dJ(x; h) = 0$ .

*Proof.* If  $\psi(\alpha) = J(f + \alpha h)$ , then 0 is the local extremum of  $\psi$ . Then  $\psi'(0) = 0$ . ■

**Definition.**  $x \in X$  is *stationary* for  $J$ , iff the limit  $dJ(x; h)$  exists and is zero for every  $h \in X$ .

Clearly, stationary points are not necessarily extrema.

## Spaces we work with

**Lemma.** Put  $X = C^1([a, b], \mathbb{R}^n)$  and

$$\|f\| = \max_{x \in [a, b]} |f(x)| + \max_{x \in [a, b]} |f'(x)|.$$

The function  $\|\cdot\|: X \rightarrow \mathbb{R}$  is a norm, and the space  $X$  is complete.

The proof repeats that of the theorem on the completeness of the space of bounded operators (from functional analysis) almost precisely.

*Proof.* All the properties of a norm follow trivially from the fact that  $|\cdot|$  is a norm.

We check completeness. Let  $\{f_n\}$  be a Cauchy sequence. Define

$$f: x \mapsto \lim_{n \rightarrow \infty} f_n(x).$$

This definition is correct, since  $\mathbb{R}$  is complete, and  $\{f_n(x)\}$  is a Cauchy sequence for any  $x \in [a, b]$ .

We assert  $f_n$  converges uniformly to  $f$ . Suppose otherwise:

$$\exists \epsilon > 0 \ \forall n_0 \in \mathbb{N} \ \exists n > n_0 \ \exists x \in [a, b]: |f_n(x) - f(x)| > \epsilon.$$

Since  $\{f_n\}$  is Cauchy,

$$\forall \delta > 0 \ \exists n_1 \in \mathbb{N} \ \forall k, l > n_1 \ \forall x \in [a, b]: |f_l(x) - f_k(x)| < \delta.$$

Fix  $\delta > 0$  and take the corresponding  $n_1$ . There exist  $x \in [a, b]$  and  $n > n_1$  such that

$$|f_n(x) - f(x)| > \epsilon. \tag{1}$$

Nevertheless, for all  $m > n_1$  we have

$$\begin{aligned} |f_n(x) - f(x)| &\leq |f_n(x) - f_m(x)| + |f_m(x) - f(x)| \\ &\leq \delta + |f_m(x) - f(x)|. \end{aligned}$$



Here, taking  $n_1$  large enough, we can make the difference  $|f_m(x) - f(x)|$  arbitrarily small, since  $f$  is the pointwise limit of  $\{f_\square\}$ . But with  $n_1 \rightarrow \infty$  we have  $\delta \rightarrow \infty$  — a contradiction to (1).

Having established that, we are, in fact, done, since the same reasoning can be applied to the pointwise limit of  $\{f'_\square\}$ . Therefore,  $X$  is complete in the specified topology. ■

**Definition.** The norm from the previous exercise we'll call the *standard  $C^1$  norm*.

**Lemma.** If a Cauchy sequence in a normed space has a convergent subsequence, it converges.

*Proof.* Let  $a_i$  be the Cauchy sequence,  $a_{n_i}$  its subsequence that converges to some  $a$ . Fix  $\epsilon > 0$ . Choose  $N$  such that  $\|a_j - a_i\| < \epsilon$  and  $\|a_{n_i} - a\| < \epsilon$  for all  $i, j > N$ . Then

$$\|a_i - a\| \leq \|a_i - a_{n_i}\| + \|a_{n_i} - a\| < 2\epsilon.$$

■

**Lemma.** Every absolutely convergent series in  $X$  converges iff  $X$  is complete.

*Proof of  $\Leftarrow$ .* Suppose the series  $\sum \|a_i\|$  converges. It is a Cauchy sequence, so

$$\left\| \sum_{n=i}^j a_j \right\| \leq \sum_{n=i}^j \|a_j\| < \epsilon$$

for sufficiently large  $\min\{i, j\}$ . Then the series  $\sum a_i$  is Cauchy. Since the space  $X$  is Banach, it converges. ■

The converse proof has some thin ice: we need to be careful about how to choose a convergent subsequence.

*Proof of  $\Rightarrow$ .* Suppose  $a_i$  is a Cauchy sequence. It is sufficient to find a convergent subsequence to show that  $a_i$  converges. We construct one,  $a_{k_i}$ , iteratively. For every  $n \in \mathbb{N}$ , there exists  $m_n$  such that  $\|a_i - a_j\| < 1/2^n$  for all  $i, j > m_n$ . Put  $a_{k_1}, a_{k_2}$  to be such that  $k_1, k_2 > m_1$ , so

$$\|a_{k_2} - a_{k_1}\| < 1/2.$$

Suppose  $a_{k_i}$  has been build up to  $i = 2t$ , and for all  $j \in \{2, \dots, t\}$  we have

$$\|a_{k_{2j}} - a_{k_{2j-1}}\| \leq 1/2^j, \quad \|a_{k_{2j-1}} - a_{k_{2j-2}}\| \leq 1/2^j.$$

We append another two members. Select  $k_{2t+2}, k_{2t+1} > m_{t+1}$ , so

$$\|a_{k_{2t+2}} - a_{k_{2t+1}}\| < 1/2^{t+1} < 1/2^t, \quad \|a_{k_{2t+1}} - a_{k_{2t}}\| < 1/2^t.$$

If we sum this, we get less than

$$\frac{1}{2} + \sum_{j \in \mathbb{N}} \frac{2}{2^j} = 2.5.$$

Then the series  $\sum_i (a_{k_{i+1}} - a_{k_i})$  converges absolutely, and so converges (by hypothesis). But this means  $a_i$  has a convergent subsequence. Since it is Cauchy,  $a_i$  converges as well. ■

**Exercise 3.**  $X$  together with this norm is complete.

*Proof.* Let  $\{f_n\} \subseteq X$  be an absolutely convergent series; that is,

$$\sum_{i=1}^n \|f_i\| \xrightarrow{n \rightarrow \infty} f.$$

The original  $\{f_n\}$  is a Cauchy sequence, since

$$\|f_n - f_m\| \leq \|f_n\| + \|f_m\| \rightarrow 0.$$

Therefore, it is sufficient to find a convergent subsequence. ■

## Smoothness of integral functionals

### $C^0$ integrand implies $C^0$ functional

**Theorem.** Let  $L \in C^0([a, b] \times \mathbb{R}^{2n}) = X$ , where the norm is standard  $C^1$ ;  $J: X \rightarrow \mathbb{R}$  is defined as

$$J: y \mapsto \int_{x=a}^b L(x, y(x), y'(x)).$$

Then  $J$  is continuous.

*Proof.* We will check continuity at  $y_0 \in X$ . By compactness of  $[a, b]$ , for some  $R_0, R_1 \in \mathbb{R}$  hold inequalities  $|y_0(x)| \leq R_0, |y'_0(x)| \leq R_1$  on  $[a, b]$ . Let  $B \subseteq \mathbb{R}^n$  be the closed ball of radius  $R_0 + 1$  with centre at 0. The  $L$  is continuous on  $[a, b] \times B^2$ , so uniformly continuous:  $\forall \epsilon > 0 \exists \delta \in (0, 1)$  such that, if

$$|x_1 - x_2| < \delta, |y_1 - y_2| < \delta, |v_1 - v_2| < \delta,$$

then

$$|L(x_1, y_1, v_1) - L(x_2, y_2, v_2)| < \frac{\epsilon}{b - a}.$$

If

$$\|y - y_0\| < \delta,$$

then

$$|J(y) - J(y_0)| \leq \int_{x=a}^b |L(x, y(x), y'(x)) - L(x, y_0(x), y'_0(x))| < \epsilon.$$

That is,  $J$  is continuous. ■

## **$C^1$ integrand implies $C^1$ functional**

**Theorem.** Let  $L \in C^1([a, b] \times \mathbb{R}^{2n})$  and  $J(y) = \int_{x=a}^b L(x, y(x), y'(x))$ . Then  $J \in C^1(X)$ , and the variation can be found by the formula

$$dJ(y; h) = \int_{x=a}^b \langle \nabla_y L(x, y(x), y'(x)), h(x) \rangle + \langle \nabla_v L(x, y(x), y'(x)), h'(x) \rangle.$$

We start with proving the most immediate conclusion:

**Lemma.** In the conditions of the theorem, if the formula for  $dJ(y; h)$  is true, then the map

$$y \mapsto dJ(y)$$

is continuous in the topology of the standard  $C^1$  norm.

*Proof.* In the given formula for  $dJ(\square)$ , the integrand is continuous in  $x, y, y'$ , since  $L$  is continuously differentiable. Therefore, we may apply the theorem on page 10, which says that  $dJ(\square)$  must be

continuous in this case. ■

Now the formula.

*Proof of the theorem.* Since the integrand  $L$  is  $C^1$ , we may use the Taylor's formula:

$$L(x, y + \delta_y, v + \delta_v) = L(x, y, v) + \langle \nabla_y L, \delta_y \rangle + \langle \nabla_v L, \delta_v \rangle + o(|\delta_y| + |\delta_v|).$$

Then

$$\begin{aligned} J(y + \delta_y) &= \int_{x=a}^b L(x, y + \delta_y, y' + \delta'_y) \\ &= J(y) + \underbrace{\int_{x=a}^b \left( \langle \nabla_y L, \delta_y \rangle + \langle \nabla_v L, \delta'_y \rangle \right)}_{\text{Denote this } \varphi(\delta_y)} + o(\|\delta_y\|). \end{aligned}$$

Here, the

$$\int_{x=a}^b o(|\delta_y| + |\delta'_y|)$$

turns

$$o(\|\delta_y\|)$$

after we use the principal estimate for integrals. Observe that the function  $\varphi$ , defined right underneath the expression for  $J(y + \delta_y)$ , is linear. It is also bounded with

$$O(|\delta_y| + |\delta'_y|)$$

by CBS, compactness of  $[a, b]$ , and the fact that  $L \in C^1$ . Therefore,  $\varphi \in X^*$ . This means  $\varphi$  is the Fréchet differential of  $J$ . Continuity of  $dJ(\square)$  has been shown in the preceding lemma. ■

## A few lemmas

### Lagrange's lemma

**Lemma (Lagrange).** Let  $\Omega \subseteq \mathbb{R}^n$  be open,  $f \in L^1(\Omega)$ . Equivalent are:

1.  $f \equiv 0$ .
2. For all  $h \in C_{cs}^\infty(\Omega)$ ,  $\int_{\Omega} fh = 0$ .

*Proof.* Suppose first that  $f$  is continuous.  $f$  is nonzero on an open ball  $B$  of radius  $r$  with centre at  $x_0$ . Then the function

$$h(x) = \begin{cases} \exp \frac{1}{|x-x_0|^2 - r^2}, & x \in B, \\ 0, & x \notin B, \end{cases}$$

contradicts the second condition: when  $x$  tends to the boundary of the ball from the inside, the argument of exp is a large negative number; on the boundary it is zero.

In the general case, recall that  $C(\Omega)$  is dense in  $L^1(\Omega)$ , so there is sequence  $\{f_n\}$  of continuous functions like in the previous paragraph that converges to  $f$  in the  $L^1$  norm. Then

$$\begin{aligned} \left| \int (fh - f_n h) \right| &\leq \int |f - f_n| |h| \\ &\leq \sup_{\Omega} |h| \cdot \int |f - f_n| \\ &\xrightarrow{\|f - f_n\|_1 \rightarrow 0} 0. \end{aligned}$$

Hence the limit integral must be nonzero as well. ■

## Lemma of Dubois and Raymond

**Lemma (Dubois-Raymond).** Let  $g \in C([a, b], \mathbb{R}^n)$ . Equivalent are

1.  $g$  is constant.
2. For every  $h \in C^1([a, b], \mathbb{R}^n)$  such that  $h(a) = h(b) = 0$ , we have

$$\int_a^b \langle g, h' \rangle = 0.$$

*Proof of 1  $\Rightarrow$  2.* By the Newton-Leibniz formula. ■

*Proof of  $2 \Rightarrow 1$ .* Put

$$c = \frac{1}{b-a} \int_a^b g.$$

We assert  $g \equiv c$ . To see this, first contrive a function

$$h(x) = \int_a^x (g - c).$$

Then  $h \in C^1$  and  $h(a) = h(b) = 0$ , so we may apply the hypothesis:

$$\int_a^b \langle g, h' \rangle = 0.$$

Observe that

$$\int_a^b \langle c, h' \rangle = 0.$$

Subtracting the last two equations, we get

$$\begin{aligned} 0 &= \int_a^b \langle g - c, h' \rangle \\ &= \int_a^b \langle g - c, g - c \rangle \\ &= \int_a^b \langle g, g \rangle. \end{aligned}$$

But then the function  $g$  must be zero by the previous lemma. ■

**Lemma.** The previous lemma is also true when  $g \in L^1[a, b]$ .

*Proof.* In the general case ( $g$  might be discontinuous, but  $L^1$ ), recall again that continuous functions are dense in  $L^1$ . Let  $g_n \xrightarrow{n \rightarrow \infty} g$  be a sequence of continuous functions. Estimating in the same manner as in the previous theorem (this time using the CBS), we obtain the desired. ■

## A generalisation

**Lemma** (Dubois-Raymond extended). Let  $g \in L^1[a, b]$ ,  $k \in \mathbb{N}$ . Equivalent are

1.  $g$  is a polynomial of degree  $k - 1$ .
2. For every  $h \in C^k[a, b]$  such that

$$h(a) = h(b) = \dots = h^{(k-1)}(a) = h^{(k-1)}(b) = 0,$$

we have

$$\int_a^b g h^{(k)} = 0.$$

*Proof of  $1 \Rightarrow 2$ .* By induction on  $k$ . The base  $k = 1$  is the Dubois-Raymond.

$$\begin{aligned} \int_a^b g h^{(k)} &= g' h^{(k-1)} \Big|_a^b - \int_a^b g' h^{(k-1)} \\ &= 0. \end{aligned}$$

■

*Proof of  $2 \Rightarrow 1$ .* By induction on  $k$ . The base  $k = 1$  is the Dubois-Raymond.

Suppose  $g \in C^1[a, b]$ . Then

$$\begin{aligned} \int_a^b g h^{(k)} &= g h^{(k-1)} \Big|_a^b - \int_a^b g' h^{(k-1)} \\ &= - \int_a^b g' h^{(k-1)}. \end{aligned}$$

$h' \in C^{k-1}[a, b]$  is, in fact, any; so the result for this case follows by induction. In the general case, observe that we again can approximate with  $C^\infty[a, b]$  functions  $g_n \xrightarrow{L^1} g$ :

$$\begin{aligned} \left| \int_a^b (g - g_n) h^{(k)} \right| &\leq \int_a^b |g - g_n| |h^{(k)}| \\ &\leq \max_{[a, b]} |h^{(k)}| \cdot \int_a^b |g - g_n| \\ &\xrightarrow{n \rightarrow \infty} 0. \end{aligned}$$

■

**Lemma.** The previous lemma is still true if we take  $g \in C^k([a, b], \mathbb{R}^n)$  and interpret the product of vectors in  $\mathbb{R}^n$  as the standard inner product.

*Proof idea.* The proof is the same, but we'll have to use the CBS. ■

## Optimisation with fixed endpoints

**Definition** (reminder). Let  $G \subseteq \mathbb{R}^m, H \subseteq \mathbb{R}^n$ .  $C^k(G, H)$  is the set of functions  $f: G \rightarrow \mathbb{R}$  such that there exists a function  $\widehat{f} \in C^k(\mathbb{R}^m, \mathbb{R}^n)$  with  $\widehat{f}|_G = f$ .

**Theorem** (the Euler-Lagrange equation). Let  $L \in C^1([a, b] \times \mathbb{R}^{2n})$ ,  $J(y) = \int_{x=a}^b L(x, y(x), y'(x))$ . Let  $y_0$  be a local extremum of  $J$  on the set

$$Y_1 = \left\{ y \in C^1([a, b], \mathbb{R}^n) \mid y(a) = A, y(b) = B \right\},$$

where  $A, B \in \mathbb{R}^n$ . Then

$$\left. \frac{\partial L}{\partial y} \right|_{y=y_0} = \frac{d}{dx} \left. \frac{\partial L}{\partial v} \right|_{y=y_0} \quad (2)$$

(in particular, the derivative on the left exists).

This is a common feature of many variational problems: the optimal function is in some way better than the functions from the consideration domain.

*Proof.* By the theorem on page 11, the functional  $J$  has a Fréchet differential  $dJ$ , and is, in particular, differentiable in the sense of Gateaux in every direction. By the same theorem,

$$dJ(y; h) = \int_{x=a}^b \langle \nabla_y L(x, y, y'), h(x) \rangle + \langle \nabla_v L(x, y, y'), h'(x) \rangle.$$

Fix  $h \in C^1([a, b], \mathbb{R}^n)$  with  $h(a) = h(b) = 0$ . Since the function  $y_0$  is a local extremum,

$$dJ(y_0; h) = 0$$



(we rely on  $h$  having ends in zero). Put

$$G(x) = \int_{t=a}^x \nabla_y L(t, y(t), y'(t)).$$

Using integration by parts, we obtain

$$\begin{aligned} 0 &= \int_{x=a}^b \langle \nabla_y L(x, y, y'), h(x) \rangle + \langle \nabla_v L(x, y, y'), h'(x) \rangle \\ &= \langle G, h \rangle \Big|_a^b + \int_{x=a}^b -\langle G(x), h'(x) \rangle + \langle \nabla_v L(x, y, y'), h'(x) \rangle \\ &= \int_{x=a}^b \langle -G(x) + \nabla_v L(x, y, y'), h'(x) \rangle. \end{aligned}$$

From the lemma on page 16 we obtain that

$$(x \mapsto -G(x) + \nabla_v L(x, y(x), y'(x))) = \text{const}.$$

Since  $G \in C^1$ , the function  $\nabla_v L \in C^1$ , and

$$\left( x \mapsto \partial_3 L(x, y_0(x), y'_0(x)) \right)' = G' = \left( x \mapsto \partial_2 L(x, y_0(x), y'_0(x)) \right).$$

■

## Smoothness of solutions

**Remark.** While the differential equation is of degree 2, the solution  $y_0$  is not necessarily  $C^2$ .

*Proof.* Consider, for example,  $L(x, y, v) = y^2(v - 2x)^2$  on  $[a, b] = [-1, 1]$ . It can be shown that

$$y_0(x) = \begin{cases} 0, & x < 0, \\ x^2, & x \geq 0 \end{cases}$$

is an extremum, satisfies the boundary conditions  $\{y(1) = 1, y(0) = -1\}$ , and  $J(y) = 0$ . Nevertheless,  $y_0$  has no second derivative at 0. ■

Here two theorems were skipped due to the difficulty in typesetting all the formulas. One of them is concerned with conditions, upon which  $y_0 \in C^2$ ; the other shows that the  $\nabla_v L = 0$  at the ends of the interval, when the values of  $y$  on them are not fixed.

Hereon

$$X = [a, b] \times \mathbb{R}^{2n},$$

$$Y = C^1([a, b], \mathbb{R}^n).$$

**Theorem** (extrema with ends fixed are  $C^2$ ). Let  $L \in C^2(X)$ , and

$$\det \left( \frac{\partial^2 L}{\partial v_i \partial v_j} (x, y, v) \right)_{i,j \in [1,n]} \neq 0$$

for all  $(x, y, v) \in X$ . Suppose that  $y_0$  is a local extremum of  $J$  on

$$\{y \in Y \mid y(a) = A, y(b) = B\}.$$

Then  $y_0 \in C^2([a, b], \mathbb{R}^n)$ .

**Theorem** (extrema with ends fixed are  $C^2$ , bis). Let  $L \in C^2(X)$ , and

$$\det \left( \frac{\partial^2 L}{\partial v_i \partial v_j} (x, y, v) \right)_{i,j \in [1,n]} \neq 0$$

for all  $(x, y, v) \in X$ . Suppose that  $y_0$  is a local extremum of  $J$  on  $Y$ . Then  $y_0 \in C^2([a, b], \mathbb{R}^n)$ .

**Theorem.** Let  $L \in C^1(X)$  and let  $y_0$  be a local extremum of  $J$  on the whole of  $Y$ . Then

$$\nabla_v L|_{x=a} = \nabla_v L|_{x=b} = 0. \quad (3)$$

**Lemma.** The equalities (2) and (3) together are equivalent to the  $y_0$  being a stationary point of  $J$ .

**G A P**

# Hypersurfaces

**Definition.** Let  $X$  be a normed space, and  $f \in C^1(X)$ . A set  $M \subseteq X$  is called a *hypersurface*, iff

$$f(x) = 0, \quad \text{and} \quad df(x) \neq 0$$

for all  $x \in M$ . The *tangent space*  $T_x M$  at  $x \in M$  is the set

$$T_x M = \ker df(x).$$

**Theorem** (tangent space in terms of curves). Suppose  $f$  and  $M$  are as above,  $p \in M$ , and  $h \in \ker df(p)$ . Then exists a neighbourhood  $U \subseteq \mathbb{R}$  of 0 and a  $C^1$  function  $w: U \rightarrow X$  such that

1.  $w(U) \subseteq M$ .
2.  $w(0) = p$ .
3.  $w'(0) = h$ .

*Proof.* Let  $v \notin T_p M$ . Introduce

$$g(s, t) := f(p + sv + th).$$

Hence

$$\frac{\partial g}{\partial t}(0, 0) = 0.$$

By the IFT, there exists  $s: U \rightarrow \mathbb{R}$  such that

$$g(s(t), t) = 0$$

and

$$s'(t) = -\frac{\partial g / \partial t}{\partial g / \partial s}.$$

The last implies

$$s'(0) = 0.$$

Put

$$w(t) := p + s(t)v + th.$$

Then  $w(0) = p$ ,  $f(w(t)) = 0$ ,  $w'(0) = s'(0)v + h = 0$ . ■

## Conditional extrema in infinite-dimensional spaces

Let  $J, G \in C^1(X)$ . Define

$$M = \{x \in X \mid G(x) = 0, \det dG(x) \neq 0\}.$$

**Lemma.** Let  $u$  be a local extremum of  $J$  on  $M$  and  $h \in T_u M$ . Then  $dJ(u; h) = 0$ .

*Proof.* Let again  $v \notin T_u M$ . Put

$$w(t) := u + s(t)v + th,$$

where

$$s(0) = 0, \quad s'(0) = 0.$$

Since  $u$  is a local extremum of  $J$ ,

$$0 = (J \circ w)'(0) = dJ(u; h).$$

■

**Theorem.** Let  $G, J \in C^1(X)$ , and let  $u$  be a local extremum of  $J$  on  $M$ . Then exists  $\lambda \in \mathbb{R}$  such that  $u$  is a stationary point of  $J + \lambda G$ .

*Proof.* Let  $v \notin \ker dG(u)$ , and

$$\lambda = -\frac{dJ(u; v)}{dG(u; v)}.$$

Let  $y \in X$ , and put

$$h = y - \frac{dG(u; y)}{dG(u; v)}.$$

By the previous lemma,

$$\begin{aligned} 0 &= dJ(u; h) \\ &= dJ(u; y) + \lambda dG(u; y), \end{aligned}$$

what was asserted.

■

# Lagrange multipliers

Let  $G \in C^1(X, \mathbb{R}^n)$ , and  $J$  and  $M$  as before.

**Theorem** (on Lagrange multipliers). Let  $u$  be a local extremum of  $J$  on  $M$ . Suppose  $u$  is a regular value. Then exists  $\lambda \in \mathbb{R}^n$  (a column of *Lagrange multipliers*) such that

$$d_u(J + \langle \lambda, G \rangle) = 0.$$

This section is completely dedicated to the proof, and all lemmas are stated within the context of the theorem.

**Lemma.** Let  $h \in \ker dG(u)$ . Then exists a neighbourhood  $U \subseteq \mathbb{R}$  of 0 and a curve  $\gamma \in C^1(U, X)$  such that  $\text{im } \gamma \subseteq M$ ,  $\gamma(0) = u$  and  $\gamma'(0) = h$ .

**Lemma.** For any  $h \in T_u M$ ,

$$dJ(u; h) = 0.$$

**G A P**

## An electrostatic example

Let  $\Omega \subseteq \mathbb{R}^3$ .