

# Notes for Mathematics for Physics: A Guided Tour for Graduate Students

Gregory Feldmann

January 21, 2020

## Contents

<b>1</b>	<b>Calculus of variations</b>	<b>3</b>
1.1	Functionals of $y$ and $y^{(1)}$ . . . . .	3
1.2	Functionals of higher order derivatives . . . . .	4
1.3	Functionals of multiple functions . . . . .	4
1.4	Functionals of multiple functions and higher order derivatives . .	4
1.5	Functionals of multiple functions, multiple independent variables and higher order derivatives . . . . .	5
1.6	Constraints . . . . .	5
1.6.1	Integral constraints . . . . .	5
1.6.2	Holonomic constraints . . . . .	6
1.6.3	Non-holonomic constraints . . . . .	6
1.7	The first integral . . . . .	6
1.8	Free boundary value problems . . . . .	7
1.9	Examples . . . . .	7
1.9.1	Example 1: Brachistochrone . . . . .	7
1.9.2	Interlude: Lagrangian mechanics . . . . .	8
1.9.3	Example 2: Constrained optimisation . . . . .	8
<b>2</b>	<b>Function spaces</b>	<b>10</b>
2.1	Vector spaces . . . . .	10
2.2	Functions as vectors . . . . .	10
2.3	Inner products . . . . .	10
2.4	Norms . . . . .	11
2.5	Convergence of functions . . . . .	11
2.5.1	Pointwise convergence . . . . .	11
2.5.2	Uniform convergence . . . . .	11
2.5.3	Convergence in the mean . . . . .	12
2.6	Banach spaces . . . . .	12
2.7	Hilbert spaces . . . . .	12

<b>3</b>	<b>Linear ordinary differential equations</b>	<b>13</b>
<b>4</b>	<b>Linear differential operators</b>	<b>14</b>
<b>5</b>	<b>Green functions</b>	<b>15</b>

# 1 Calculus of variations

A functional  $J$  is a map  $J : C^\infty(\mathcal{R}^n) \rightarrow \mathcal{R}$ . We restrict ourselves to functionals of the form

$$J[y] = \int_{x_1}^{x_2} F(x, y, y^{(1)}, y^{(2)}, \dots) dx \quad (1)$$

where  $y^{(n)}$  denotes the  $n^{th}$  derivative of  $y$  with respect to  $x$ .

## 1.1 Functionals of $y$ and $y^{(1)}$

For  $J = \int F dx$ , where  $f$  depends only on  $x$ ,  $y$  and  $y'$ , we show how to derive the Euler-Lagrange equation and find the  $y$  that optimises  $J$ .

Let  $\epsilon$  be a small, arbitrary real number,  $\eta$  an arbitrary function of  $x$  and  $y$  an arbitrary function of  $x$  in  $C^\infty$ . A small perturbation in  $y$  is given by  $y = y^* + \epsilon\eta$ , where  $y^*$  is the unperturbed  $y$ . We require that  $y^*(x_1) = y(x_1)$  and  $y^*(x_2) = y(x_2)$ , ie  $\eta(x_1) = \eta(x_2) = 0$ . Let  $\delta J = J[y^* + \epsilon\eta] - J[y^*]$  and  $\delta F = \{F(x, y, y^{(1)}) - F(x, y^*, y^{*(1)})\}$ , where  $\delta$  is the variation operator. The change in  $J$  associated with going from  $y$  to  $y^* + \epsilon\eta$  is given by

$$\delta J = \int_{x_1}^{x_2} \delta F dx \quad (2)$$

A necessary condition for the minimisation of  $J$  is  $\delta J = 0$ . We then use a Taylor expansion of  $\delta F$  around  $\epsilon = 0$ , discarding all terms second order and above.

$$\delta F \approx \epsilon \left( \frac{\partial F}{\partial y} \eta + \frac{\partial F}{\partial y^{(1)}} \eta^{(1)} \right) \quad (3)$$

Using integration by parts, we can find an alternate expression for  $\frac{\partial F}{\partial y^{(1)}} \eta^{(1)}$

$$\int_{x_1}^{x_2} \frac{\partial F}{\partial y^{(1)}} \eta^{(1)} dx = \left. \frac{\partial F}{\partial y^{(1)}} \eta \right|_{x_1}^{x_2} - \int_{x_1}^{x_2} \frac{d}{dx} \frac{\partial F}{\partial y^{(1)}} \eta dx$$

As we assume  $\epsilon = 0$  at  $x_1$  and  $x_2$ ,  $\left. \frac{\partial F}{\partial y^{(1)}} \eta \right|_{x_1}^{x_2} = 0$ . So we have

$$\frac{\partial F}{\partial y^{(1)}} \eta^{(1)} = - \frac{d}{dx} \frac{\partial F}{\partial y^{(1)}} \eta$$

We substitute this into (3) to obtain

$$\delta F \approx \epsilon \eta \left( \frac{\partial F}{\partial y} - \frac{d}{dx} \frac{\partial F}{\partial y^{(1)}} \right)$$

Substituting this expression into (2) results in

$$\int_{x_1}^{x_2} \delta F dx = \int_{x_1}^{x_2} \epsilon \eta \left( \frac{\partial F}{\partial y} - \frac{d}{dx} \frac{\partial F}{\partial y^{(1)}} \right) dx = 0$$

Supposing that  $\eta$  and  $\epsilon$  are not identically zero, we conclude that

$$\frac{\partial F}{\partial y} - \frac{d}{dx} \frac{\partial F}{\partial y^{(1)}} = 0 \quad (4)$$

(4) is known as an *Euler-Lagrange equation*. The expression on the left hand side is the *functional derivative* of  $J$  with respect to  $y$ ,  $\frac{\delta J}{\delta y}$ .

## 1.2 Functionals of higher order derivatives

If  $F$  is a function of higher order derivatives of  $y$ , e.g.  $y^{(5)}$  or  $y^{(26)}$ , then the Euler-Lagrange equation is extended as follows. Suppose  $F$  is a function of the  $N^{th}$  order derivative of  $y$ . The Euler-Lagrange equation is

$$\frac{\partial F}{\partial y} + \sum_{n=1}^N (-1)^n \frac{d^n}{dx^n} \left( \frac{\partial F}{\partial y^{(n)}} \right) = 0 \quad (5)$$

## 1.3 Functionals of multiple functions

When  $F$  is a function of multiple functions  $y_i$  and their derivatives  $y_{ix}$ , where each  $y_i$  is a function of  $x$ , then we get a separate Euler-Lagrange equation for each  $y_i$

$$\frac{\partial F}{\partial y_i} - \frac{d}{dx} \frac{\partial F}{\partial y_i^{(1)}} = 0 \quad (6)$$

## 1.4 Functionals of multiple functions and higher order derivatives

Combining the previous two sections, the Euler-Lagrange equations for functionals of multiple functions  $y_i$  and higher order derivatives  $y_i^{(n)}$  are given by

$$\frac{\partial F}{\partial y_i} + \sum_{n=1}^N (-1)^n \frac{d^n}{dx^n} \left( \frac{\partial F}{\partial y_i^{(n)}} \right) = 0 \quad (7)$$

for each  $i$  corresponding to a  $y_i$ .

## 1.5 Functionals of multiple functions, multiple independent variables and higher order derivatives

The next logical step is to introduce multiple independent variables. Let  $\mu_m$  be indices spanning the independent variables  $x_{\mu_m}$ . Let  $y_i^{(n)\mu_m}$  denote the  $n^{th}$  partial derivative of  $y_i$  with respect to  $x_{\mu_m}$ . Given a total of  $M$  independent variables, functionals of multiple functions  $y_i$  and higher order derivatives  $y_i^{(n)\mu_m}$ , the Euler-Lagrange equation for each  $y_i$  is as follows

$$\frac{\partial F}{\partial y_i} + \sum_{n=1}^N \sum_{\mu_1 \leq \dots \leq \mu_n} (-1)^n \frac{d^n}{dx_{\mu_1} \dots dx_{\mu_n}} \left( \frac{\partial F}{\partial y_i^{(1)\mu_1} \dots \partial y_i^{(1)\mu_n}} \right) = 0 \quad (8)$$

Note that  $\mu_m$  is an arbitrary reference to any  $x_m$ . Hence  $x_{\mu_1}$  may be equal to  $x_{\mu_2}$  even though  $x_1 \neq x_2$ . For 3 independent variables  $x_1, x_2$  and  $x_3$ , and  $n = 2$ , we have that

$$\sum_{\mu_1 \leq \dots \leq \mu_2} dx_{\mu_1} dx_{\mu_2} = dx_1^2 + dx_1 dx_2 + dx_1 dx_3 + dx_2^2 + dx_2 dx_3 + dx_3^2$$

For the same example, but with  $n = 1$ , we have

$$\sum_{\mu_1 \leq \dots \leq \mu_1} dx_{\mu_1} dx_{\mu_1} = dx_1^2 + dx_2^2 + dx_3^2$$

## 1.6 Constraints

We will frequently want to find the stationary point of  $\delta J$  subject to certain constraints.

### 1.6.1 Integral constraints

The simplest constraints are integral constraints

$$K = \int_{x_1}^{x_2} G(x, y, y^{(1)}, y^{(2)}, \dots) dx \quad (9)$$

where  $K$  is a constant and  $G$  is a functional.

To incorporate this into our Euler-Lagrange equations, we define  $F_c$  as follows

$$F_c = F + \lambda G$$

and work with  $F_c$  in the same way we would with  $F$ , subject to the constraints being satisfied. e.g. (8) becomes

$$\frac{\partial F_c}{\partial y_i} + \sum_{n=1}^N \sum_{\mu_1 \leq \dots \leq \mu_n} (-1)^n \frac{d^n}{dx_{\mu_1} \dots dx_{\mu_n}} \left( \frac{\partial F_c}{\partial y_i^{(1)\mu_1} \dots \partial y_i^{(1)\mu_n}} \right) = 0$$

and our solutions must also satisfy (9).

Given multiple functionals  $G_m$  in integral constraints,  $F_c$  is given by

$$F_c = F + \sum_m \lambda_m G_m$$

and solutions to the Euler-Lagrange equations must satisfy the constraints.

### 1.6.2 Holonomic constraints

Equivalence constraints not involving derivatives of functions that can be expressed without an integral are referred to as holonomic constraints. They take the form

$$K = G(x, y) \quad (10)$$

where  $G$  and  $K$  are defined as in the previous section. We now define  $F_c$  as

$$F_c = F + \int_{x_1}^{x_2} \lambda(x) G(x, y) dx \quad (11)$$

where  $\lambda(x)$  is a Lagrange multiplier. Take note that our Lagrange multiplier is now dependant on our independent variables. Given  $F_c$ , we proceed in the same manner as for integral constraints.

### 1.6.3 Non-holonomic constraints

Non-holonomic constraints are equivalence constraints that involve derivatives of functions. They are not covered.

## 1.7 The first integral

The quantity  $I = F - y^{(1)} \frac{\partial F}{\partial y^{(1)}}$  is called the *first integral*. Suppose  $F$  is directly dependent only on  $y$  and  $y^{(1)}$ . We then have the following

$$\begin{aligned} \frac{d}{dx} \left( F - y^{(1)} \frac{\partial F}{\partial y^{(1)}} \right) &= y^{(1)} \frac{\partial F}{\partial y} + y^{(2)} \frac{\partial F}{\partial y^{(1)}} - y^{(2)} \frac{\partial F}{\partial y^{(1)}} - y^{(1)} \frac{d}{dx} \left( \frac{\partial F}{\partial y^{(1)}} \right) \\ &= y^{(1)} \left( \frac{\partial F}{\partial y} - \frac{d}{dx} \left( \frac{\partial F}{\partial y^{(1)}} \right) \right) \end{aligned} \quad (12)$$

The bracketed expression on the second line is the Euler-Lagrange equation and hence is set to 0. We therefore have the result that the first integral  $I$  is equal 0 when  $F$  has no direct dependence on  $x$ .

For multiple functions  $y_i$  the first integral becomes

$$I = F - \sum_i y_i^{(1)} \frac{\partial F}{\partial y_i^{(1)}} \quad (13)$$

## 1.8 Free boundary value problems

When an end-point is free to vary, eg  $x_2$  is not specified, then we must also determine  $x_2$  as part of the optimisation solution.

The general procedure is to proceed as if  $x_2$  is fixed and find the extremal  $y$ .  $x_2$  can then be determined by solving

$$\mathcal{L}(x_2, y(x_2), y'(x_2), \dots) = 0 \quad (14)$$

for  $x_2$ . This is referred to as the *natural boundary condition* for  $x_2$ . A corresponding condition applies to determine  $x_1$ , should it also not be predetermined.

## 1.9 Examples

### 1.9.1 Example 1: Brachistochrone

Suppose a bead is sliding down a wire from a height of  $y_0$  to  $y_1$  along a length of wire with start and end  $x$  coordinates being  $x_0$  and  $x_1$ . We want to know the shape of the wire that results in the particle moving from  $x_0$  to  $x_1$  as quickly as possible under the force of gravity alone.

The time taken by the bead to slide down the wire is given by the distance travelled divided by the velocity of the bead. Using the fact that  $T + V = \text{constant}$  and choosing a reference point of  $V$  such that the *constant* = 0 allows us to derive an expression for the velocity,  $v = \sqrt{2gy}$ . The distance travelled is simply the integral of  $ds$ , hence our functional is

$$\begin{aligned} t_{total} &= \int_{y_0}^{y_1} \frac{ds}{\sqrt{2gy}} \\ &= \int_{y_0}^{y_1} \sqrt{\frac{1 + y'^2}{2gy}} dt \end{aligned} \quad (15)$$

We thus have

$$F(y, y') = \sqrt{\frac{1 + y'^2}{2gy}} \quad (16)$$

Using the Euler-Lagrange equation and solving the resulting differential equation yields a solution of

$$\begin{aligned} x &= at - asin(t) \\ y &= a - acos(t) \end{aligned} \quad (17)$$

where  $a$  is some constant chosen to satisfy the boundary conditions below

$$\begin{aligned} (x_0, y_0) &= (at - asin(0), a - acos(0)) \\ (x_1, y_1) &= (at - asin(t_{total}), a - acos(t_{total})) \end{aligned} \quad (18)$$

### 1.9.2 Interlude: Lagrangian mechanics

Lagrangian mechanics is a formulation of classical mechanics that uses functionals composed of the difference between kinetic and potential energy, as opposed to the traditional approach based on Newton's laws. The Lagrangian  $\mathcal{L}$  is given as follows

$$\mathcal{L} = T - V \quad (19)$$

where  $T$  is the kinetic energy and  $V$  is the potential energy.

We are free to write  $T$  and  $V$  in terms of whatever variables we would like. Typically the independent variable is time and functions dependent on time are denoted by  $q_i$ .

The Principle of Least Action states that the time evolution of a physical system between time  $t_0$  and  $t_1$  minimises the *action* of the system. The action  $\mathcal{S}$  is defined as

$$\mathcal{S} = \int_{t_0}^{t_1} \mathcal{L} dt \quad (20)$$

By minimising the action using the calculus of variations we can therefore find the equations describing the time evolution of a system.

### 1.9.3 Example 2: Constrained optimisation

Suppose a chain is fixed to two walls and we would like to determine the shape in which it will hang. We can do this by minimising the potential energy using a functional. The potential energy at each point in the chain is given by  $PE =$



$\mu g y(x) ds$ , where  $\mu$  is the density of the chain per unit length,  $g$  is the force of gravity and  $y(x)$  is the height. Integrating over the length of the chain gives us the total potential energy  $V$ , and is a functional that we can minimise with respect to  $y$ .

$$V = g\mu \int_a^b y \sqrt{1 - y'^2} dx \quad (21)$$

We then have

$$F(y, y') = y \sqrt{1 - y'^2} \quad (22)$$

The associated Euler-Lagrange equation is as follows

$$\frac{\partial F}{\partial y} - \frac{d}{dx} \frac{\partial F}{\partial y'} = \sqrt{1 - y'^2} - \quad (23)$$

## References

- [1] R.E. Hunt, *Chapter 1 variational methods*, 2002.
- [2] R.S. Schechter, *The variational method in engineering*, McGraw-Hill chemical engineering series, McGraw-Hill, 1967.
- [3] M. Stone and P. Goldbart, *Mathematics for physics: A guided tour for graduate students*, Cambridge University Press, United States, 2009.

## 2 Function spaces

### 2.1 Vector spaces

A vector space consists of a set  $V$  (whose elements are referred to as vectors), a field  $\mathbb{F}$  (whose elements are called scalars) and two operators:

1. A vector addition operation  $+$  such that for  $\mathbf{u}, \mathbf{v} \in V$ ,  $\mathbf{u} + \mathbf{v} \in V$ .
2. A scalar multiplication operation such that for  $\lambda \in \mathbb{F}$  and  $\mathbf{u} \in V$ ,  $\lambda \mathbf{u} \in V$ .

which satisfy the following conditions

1. For  $\forall \mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ ,

$$(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w}) \quad (24)$$

2. There exists some  $\mathbf{0} \in V$  such that for  $\forall \mathbf{v} \in V$ ,

$$\mathbf{v} + \mathbf{0} = \mathbf{v} \quad (25)$$

3. For  $\forall \mathbf{v} \in V$ , there exists some  $-\mathbf{v} \in V$  such that

$$\mathbf{v} + (-\mathbf{v}) = \mathbf{v} - \mathbf{v} = \mathbf{0} \quad (26)$$

4. For  $a, b \in \mathbb{F}$  and  $\mathbf{v} \in V$ ,  $(ab)\mathbf{v} = a(b\mathbf{v})$ .

5. For  $\forall \mathbf{v} \in V$ ,  $1\mathbf{v} = \mathbf{v}$ .

6. For  $a, b \in \mathbb{F}$  and  $\mathbf{u}, \mathbf{v} \in V$ ,

$$(a + b)\mathbf{u} = a\mathbf{u} + b\mathbf{u} \quad (27)$$

and

$$a(\mathbf{u} + \mathbf{v}) = a\mathbf{u} + a\mathbf{v} \quad (28)$$

### 2.2 Functions as vectors

Taking functions to be infinite-dimensional vectors, we can take the vectors in a vector space to be functions. We use this to build up Hilbert spaces. An example of a vector space of functions is  $C^n[a, b]$  - the set of all functions in  $[a, b]$  with continuous  $n^{th}$  order derivatives.

### 2.3 Inner products

An inner product associates two vectors  $\mathbf{u}, \mathbf{v} \in V$  with a real number and is denoted by  $\langle \mathbf{u}, \mathbf{v} \rangle$ . It obeys three axioms:

1. Bilinearity

$$\langle c\mathbf{u} + d\mathbf{v}, \mathbf{w} \rangle = c\langle \mathbf{u}, \mathbf{w} \rangle + d\langle \mathbf{v}, \mathbf{w} \rangle \quad (29)$$

$$\langle \mathbf{u}, c\mathbf{v} + d\mathbf{w} \rangle = c\langle \mathbf{u}, \mathbf{v} \rangle + d\langle \mathbf{u}, \mathbf{w} \rangle \quad (30)$$

2. Symmetry

$$\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle \quad (31)$$

3. Non-negative

$$\langle \mathbf{u}, \mathbf{v} \rangle \geq 0 \quad (32)$$

with  $\langle \mathbf{v}, \mathbf{v} \rangle = 0$  iff  $\mathbf{v} = \mathbf{0}$ .

## 2.4 Norms

The norm of a vector space provides a measure of the 'size' of its elements. The norm  $\|f\|$  of some function  $f$  is a real number satisfying the following conditions:

1.  $\|f\| \geq 0$ , with  $\|f\| = 0$  iff  $f = 0$ ,
2.  $\|f + g\| \leq \|f\| + \|g\|$ ,
3.  $\|\lambda f\| = |\lambda| \|f\|$  for constant  $\lambda$ .

We can obtain a norm from an inner product by taking the  $p^{th}$  root of the inner product of some  $\mathbf{v}$  with itself:  $\langle \mathbf{v}, \mathbf{v} \rangle^{1/p}$ .

The space  $L^p[a, b]$  is defined to be a function space over the interval  $[a, b]$ , function set  $F[a, b]$ , for  $1 \leq p < \infty$  and a norm of

$$\|f\|_p = \left( \int_a^b |f(x)|^p dx \right)^{1/p} \quad (33)$$

$\|f\|_p$  must be finite for  $\forall f \in F$ .

The elements of  $L^p[a, b]$  are equivalence classes of functions, with functions in the same equivalence class differing by a function of norm 0. Note that a function can have a norm of zero despite not being identically 0 if it is only non-zero at a finite number of points. This violates the first condition for being a norm, but forget that for now.

## 2.5 Convergence of functions

Sequence of functions  $f_n$  may converge to  $f$  *pointwise*, *uniformly* or *in the mean*.

### 2.5.1 Pointwise convergence

For all  $x$  and all  $\epsilon > 0$ , there exists some  $N$  such that, if  $n \geq N$ , then  $|f_n - f| < \epsilon$ .

### 2.5.2 Uniform convergence

For all  $\epsilon > 0$  there exists some  $N$  such that, for all  $x$  and  $n > N$ ,  $\sup |f_n(x) - f(x)| < \epsilon$ .

### 2.5.3 Convergence in the mean

For all  $\epsilon > 0$  there exists some  $N$  such that  $\int |f_n(x) - f(x)|dx < \epsilon$  for  $n > N$ .

## 2.6 Banach spaces

A normed vector space is *complete* if every Cauchy sequence in it converges to an element contained in the normed vector space. A complete normed vector space is called a *Banach space*.  $L^p[a, b]$  spaces are complete if the integrals in the  $L^p$  norm are taken to be Lebesgue integrals.

## 2.7 Hilbert spaces

A Banach space whose norm is derived from an inner product is called a *Hilbert space*. e.g. if we define the inner product as

$$\langle f, g \rangle = \int_a^b f^* g dx \quad (34)$$

then the  $L^2[a, b]$  norm can be written as

$$\|f\|_2 = \sqrt{\langle f, f \rangle} \quad (35)$$

$\langle f, g \rangle$  is a continuous functional of  $f$  and  $g$ . A set of functions  $\{u_n\}$  is orthonormal if  $\langle u_n, u_m \rangle = \delta_{nm}$ . An orthonormal set of functions  $\{u_n\}$  within some Hilbert space is *complete* if any element  $f$  inside the Hilbert space can be written as

$$f(x) = \sum_{n=0}^{\infty} a_n u_n(x) \quad (36)$$

### 3 Linear ordinary differential equations

## 4 Linear differential operators

## 5 Green functions