

# Functional Derivative

Andreas Niedens

May 30, 2020

## 1 Introduction

We want to introduce the notion of functionals and functional derivatives. They can be used to calculate stationary functions of functionals, and help us understand how much does functional value change when we vary given function at certain point.

## 2 Definitions

### 2.1 Functionals

Let's begin with definition of functional. A functional's goal, like function's, is to give us a number. The difference is, functionals take functions and not numbers as input. For example, given trajectory of particle as a function, we could ask, what distance did this particle travel, or we could calculate average kinetic energy. These values depend on path taken (and speed), so functional definition comes to us naturally this way. The term 'functional' is just used so we know that it acts on functions.

We can write functionals like this :  $F[f]$ . An example of functional can be:

- $F[f] = \int_0^1 f(x) dx$

If we provide  $f(x) = x^2$  as input, we will get  $\frac{1}{3}$ .

- $F[f] = \int_{-a}^a 5[f(x)]^2 dx$

For input  $f(x) = x^2$  functional gives us  $2a^5$ .

- $F_{x_0}[f] = \int_{-\infty}^{\infty} f(x)\delta(x - x_0) dx = f(x_0)$

We can see from this example that functionals can be used as functions, so they are in

some sense generalized functions. This functional gives us value of function  $f$  at point  $x_0$ .

## 2.2 Functional Derivatives

If functionals can be thought of as generalized functions, we can try to think of generalized derivatives too. For usual functions, definition of derivative is

$$f'(x) = \lim_{\epsilon \rightarrow 0} \frac{f(x + \epsilon) - f(x)}{\epsilon}$$

It tells us, how fast does a value of function change at point  $x$  by we evaluating it at point  $x + \epsilon$  for arbitrary small  $\epsilon$  and taking the difference. But functionals do not act on points, they act on functions. It would have been convenient to have one parameter to find 'nearest function', but we can change input function at any one point by increasing its value near that point and leaving all other points the same. This way, we have infinite number of closest functions, and instead of having one  $\epsilon$ , we now have to decide two things:

1. Where to increase (or change) the function.
2. How much do we need to change it (Something like choosing  $\epsilon$ )

Also, we need to understand what will the result of differentiation give us. For functions, derivative gives us a number for each input number. For functionals, we would get a function for each input function. This derivative function's value at each point would tell us, how much does functional value change if we change input function's value at that exact point. Another problem is, how do we actually change a function's value at some point? We still would like it to be continuous, so just adding one number would not suffice. But if we change it at many places at once, we would not be able to understand how much each coordinate affects functional value, but just see the global effect. This can be solved with Dirac delta function ( $\delta$  function), which can be thought of limit of functions that are continuous, narrow at one point so that they change value of the function only at that point.

So we arrive at our definition of functional derivative that gives value of function at point  $x_0$ :

$$\frac{\delta F}{\delta f(x)} = \lim_{\epsilon \rightarrow 0} \frac{F[f(x) + \epsilon \delta(x - x_0)] - F[f(x)]}{\epsilon} \quad (1)$$

### 3 Examples

Let's try evaluating several examples.

- $I[f] = \int_{-1}^1 f(x) dx$

$$\begin{aligned} \frac{\delta I[f]}{\delta f(x_0)} &= \lim_{\epsilon \rightarrow 0} \frac{\int_{-1}^1 f(x) + \epsilon \delta(x - x_0) dx - \int_{-1}^1 f(x) dx}{\epsilon} = \\ &= \lim_{\epsilon \rightarrow 0} \frac{\int_{-1}^1 \epsilon \delta(x - x_0) dx}{\epsilon} = \int_{-1}^1 \delta(x - x_0) dx = \begin{cases} 1 & \text{if } -1 \leq x_0 \leq 1 \\ 0 & \text{else} \end{cases} \end{aligned}$$

We see, that derivative is 0 outside integral. We could expect that, since functional does not depend on outside values. It is one inside because we add delta function, whose integral is  $\epsilon * 1$  to some integral and divide difference by  $\epsilon$

- $J[f] = \int [f(y)]^p \phi(y) dy$

$$\frac{\delta J[f]}{\delta f(x)} = \lim_{\epsilon \rightarrow 0} \frac{\int [f(y) + \epsilon \delta(y - x)]^p \phi(y) dy - \int [f(y)]^p \phi(y) dy}{\epsilon}$$

Writing  $[f(y) + \epsilon \delta(y - x)]^p = [f(y)]^p + p[f(y)]^{p-1} \epsilon \delta(y - x)$ , first term cancels and we are left with  $\int p[f(y)]^{p-1} \delta(y - x) \phi(y) dy$  after reducing fraction, and after using delta's property gives us

$$\frac{\delta J[f]}{\delta f(x)} = p[f(x)]^{p-1} \phi(x) \quad (2)$$

In this case functional derivative behaves like usual differentiation, when we are differentiating powers.

- $H[f] = \int_a^b g[f(x)] dx$

$$\frac{\delta H[f]}{\delta f(x_0)} = \lim_{\epsilon \rightarrow 0} \frac{\int_a^b g[f(x) + \epsilon \delta(x_0 - x)] dy - \int_a^b g[f(x)] dy}{\epsilon}$$

Expanding  $g[f(x) + \epsilon\delta(x_0 - x)]$  as  $g[f(x)] + \epsilon\delta(x_0 - x)g'[f(x)]$ , first term cancels and we are left with  $\int \delta(x_0 - x)g'[f(x)]$  that gives us

$$\frac{\delta H[f]}{\delta f(x_0)} = g'[f(x_0)] \quad (3)$$

- $J[f] = \int g(f'(y)) dy$ , where  $f'(y) = \frac{df(y)}{dy}$

$$\frac{\delta J[f]}{\delta f(x)} = \lim_{\epsilon \rightarrow 0} \frac{\int g\left(\frac{\partial}{\partial y}[f(y) + \epsilon\delta(x - y)]\right) dy - \int g(f'(y)) dy}{\epsilon}$$

$g\left(\frac{\partial}{\partial y}[f(y) + \epsilon\delta(x - y)]\right) = g([f'(y) + \epsilon\delta'(x - y)]) = g(f'(y)) + \epsilon\delta'(x - y)g'(f'(y))$ , keeping only first order terms in  $\epsilon$ . We are left with following integral :  $\int \delta'(x - y) \frac{dg(f')}{df'} dy$

We use integration by parts to get  $[\delta(x - y) \frac{dg(f')}{df'}] - \int \delta(x - y) \frac{d}{dy} \frac{dg(f')}{df'} dy$ . First term is 0, if  $x$  is inside integral limits (we get  $0 - 0$ , as  $\delta$  function is 0 on the boundaries) We are left with

$$\frac{\delta J[f]}{\delta f(x)} = - \frac{d}{dx} \frac{dg(f'(x))}{df'(x)} \quad (4)$$

For example, if  $F[\phi] = \int \left(\frac{\partial \phi}{\partial y}\right)^2 dy$ , then  $\frac{\delta F[\phi]}{\delta \phi(x)} = - \frac{d}{dx} 2 \left(\frac{\partial \phi}{\partial x}\right) = -2 \frac{\partial^2 \phi}{\partial x^2}$

## 4 Lagrangians and stationary action

We will try to apply a calculus of variations on a particle moving in potential field. Let's look for some properties that it has along its path, like average potential energy :

$$\bar{V} = \frac{1}{\tau} \int_0^\tau V(x) dt \quad (5)$$

Average potential energy is a functoinal, because we have a number for each trajectory and potential. Using (3), we find its functional derivative :

$$\frac{\delta \bar{V}[x]}{\delta x(t)} = \frac{1}{\tau} V'[x(t)] \quad (6)$$

Of course, we can do the same thing for kinetic energy :

$$\bar{T} = \frac{1}{\tau} \int_0^\tau \frac{1}{2} m \dot{x}^2 dt \quad (7)$$

Using (4), we can write it as :

$$\frac{\delta \bar{T}[x]}{\delta x(t)} = -\frac{1}{\tau} \frac{d}{dx} (m\dot{x}) = -\frac{1}{\tau} m\ddot{x} \quad (8)$$

We also need a formula for forces in potential field :

$$F = -\frac{dV}{dx} = m\ddot{x} \quad (9)$$

Combining this equation with (8) and (6), we see that derivative for average kinetic energy equals to derivative of average potential energy :

$$\boxed{\frac{\delta \bar{T}[x]}{\delta x(t)} = \frac{\delta \bar{V}[x]}{\delta x(t)}} \quad (10)$$

It is indeed an interesting result. On a classical trajectory we can see that deviation from trajectory increases average potential and kinetic energies by *the same amount*. If we try to combine derivative operations together, we can rewrite it as

$$\frac{\delta(\bar{T}[x] - \bar{V}[x])}{\delta x(t)} = 0 \quad (11)$$

We found some quantity that is stationary along classical trajectory, which means that its derivative is zero. We call that quantity **Lagrangian** :

$$\boxed{L := T - V} \quad (12)$$

And integral of Lagrangian over time is called **Action** :

$$\boxed{S := \int_0^\tau L dt} \quad (13)$$

Action is measured in [energy \* time]. We can rewrite equation (11) by rewriting definition of averages and changing them with Lagrangian Lagrangian. The factor of  $\frac{1}{\tau}$  will cancel out and we are left with :

$$\boxed{\frac{\delta S}{\delta x(t)} = 0} \quad (14)$$

Which is known as **Hamilton's principle of least action**