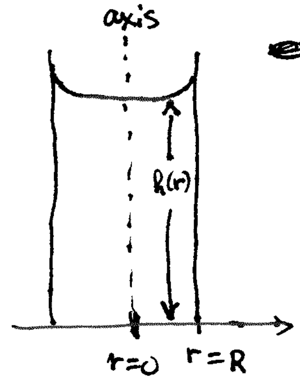
Some notes about functionals and functional derivatives.

A <u>functional</u> is a function whose independent variable is another function. Functionals are very common in physics, and some examples will illustrate the concept.

The <u>potential energy functional</u> for a column of liquid in a capillary tube was discussed in class. Let the tube have a circular cross section with inner radius R, and let $h(r)$ be the height of the liquid.



Then, as shown in class, the gravitational potential energy V is a functional of the height function,

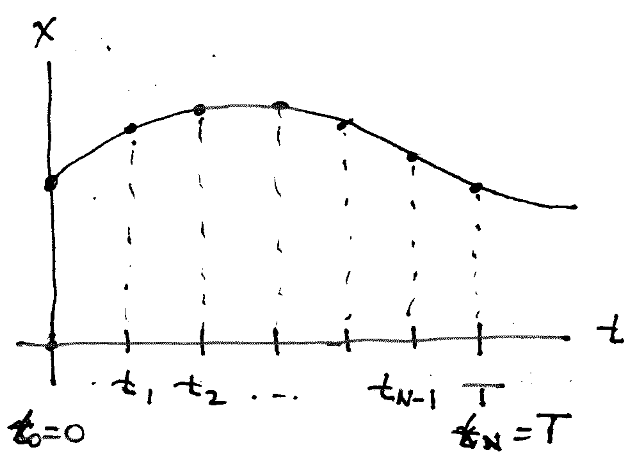
$$V[h(r)] = \pi \rho g \int_0^R r\, dr\; h(r)^2,$$

where ρ is the density of the fluid and g is the acceleration of gravity.

The use of square brackets in $V[h(r)]$ is just a convention that indicates that this is a functional instead of an ordinary function. Notice that V does not depend on r, in fact, r is a dummy variable of integration on the right hand side, and could be replaced by any symbol. To emphasize this, we may prefer to write $V[h(\cdot)]$ instead.

In physics, the word "function" is usually used to refer to a real-valued function of one or more real variables, or something similar. In mathematics, the word "function" has a much broader meaning, and can refer to a map between spaces of any kind of objects. So, from the mathematical standpoint, a functional is just a function or map whose domain is a set of functions.

A function of a continuous variable can be approximated by the values of the function at discrete points. For example, instead of $x(t)$ on the interval $0 \le t \le T$, we can use the collection of numbers, $x_n = x(n\Delta t)$, where $\Delta t = T/N$ and $n = 0, 1, 2, \ldots, N-1, N$



In the limit $N \to \infty$ we can think of the discrete set of numbers $(x_0, x_1, \ldots, x_n)$ as going over to the function $x(t)$.

Suppose we have an ordinary (real-valued) function of a discrete set of variables, $f(x_1, \ldots, x_n)$. Then the partial derivative of f with respect to one of the x's, say, x_i, is defined by

$$\frac{\partial f}{\partial x_i}(x_1, \ldots, x_n) = \lim_{\varepsilon \to 0} \frac{f(x_1, \ldots, x_i + \varepsilon, \ldots x_n) - f(x_1, \ldots, x_i, \ldots x_n)}{\varepsilon}.$$

In other words, we vary one of the x's while holding all the others fixed. Notice that the LHS still depends on the $(x_1, \dots, x_n)$, but it also depends on the index i.

We would like to generalize this to a functional. To do this we rewrite the ~~partial~~ definition of the partial derivative as follows. Let

$$\hat{e}_i = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \begin{matrix} \\ \\ \\ \leftarrow i\text{-th position} \\ \\ \\ \end{matrix}$$

be the i-th "unit vector" in the n-dimensional space with coordinates $\vec{x} = (x_1, \dots, x_n)$. Then we can define

$$\frac{\partial f}{\partial x_i}(\vec{x}) = \frac{d}{d\varepsilon} f(\vec{x} + \varepsilon \hat{e}_i)\Big|_{\varepsilon = 0}.$$

This is equivalent to the earlier definition. We take the derivative of the function f in the direction i.

Here is another definition. Let δx_i, $i = 1, \dots, n$ be a set of n numbers, representing increments in the x_i. Thus, we think of 2 points in $\vec{x}$-space, with coordinates x_i and $x_i + \delta x_i$. Then $\frac{\partial f}{\partial x_i}(\vec{x})$, $i = 1, \dots, n$ are the unique set of numbers such that

$$\delta f = f(x_1 + \delta x_1, \dots, x_n + \delta x_n) - f(x_1, \dots, x_n)$$

$$= \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x_1, \dots, x_n)\, \delta x_i, \qquad \text{for all } \delta x_i,$$

when the δx_i are small.

The final qualification means that for any finite δx_i, the right hand side is an approximation, with (usually) terms of order δx_i^2 as corrections. These corrections can be developed into a Taylor series.

We can make the final formulation of the partial derivative rigorous by letting $(\xi_1, \ldots, \xi_n)$ be a set of n numbers, defining

$$\delta x_i = \varepsilon\, \xi_i ,$$

where ε is a scale factor, and defining $\frac{\partial f}{\partial x_i}(\vec{x})$ as the unique set of n numbers such that

$$\frac{d}{d\varepsilon} f(\vec{x} + \varepsilon \vec{\xi}) \Big|_{\varepsilon=0} = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\vec{x})\, \xi_i , \quad \text{for all } \vec{\xi}.$$

To generalize this to functionals we replace the discrete index i on x_i by a continuous variable t, so $x_i \rightarrow x(t)$. Let's also replace f by F, so $f(x_1, \ldots, x_n) \longrightarrow F[x(t)]$ or $F[x(\cdot)]$. Instead of $\hat{e}_i$ (a vector with zero everywhere except in slot i, where it is 1) we can use $\delta(t-t')$, where t' corresponds to i. This is a function of t that is zero everywhere except at $t = t'$. Then the functional derivative can be defined by

$$\frac{\delta F[x(\cdot)]}{\delta x(t')} = \lim_{\varepsilon \to 0} \frac{F[x(t) + \varepsilon \delta(t-t')] - F[x(t)]}{\varepsilon},$$

or, equivalently, by

$$\frac{\delta F[x(\cdot)]}{\delta x(t')} = \frac{d}{d\varepsilon} F[x(t) + \varepsilon \delta(t-t')]\Big|_{\varepsilon=0}.$$

Another approach is to let $\xi(t)$ be a function, and define

$$\delta x(t) = \varepsilon \xi(t),$$

where ε is a scale factor. Then $\frac{\delta F[x(\cdot)]}{\delta x(t')}$ is the unique functional of $x(\cdot)$ and ordinary function of t' such that

$$\frac{d}{d\varepsilon} F[x(\cdot) + \varepsilon \xi(\cdot)]\Big|_{\varepsilon=0} = \int dt' \, \frac{\delta F[x(\cdot)]}{\delta x(t')} \xi(t'),$$

for all choices of the function $\xi(t')$. Or, we can write,

$$\delta F = F[x(\cdot) + \delta x(\cdot)] - F[x(\cdot)]$$

$$= \int dt' \, \frac{\delta F[x(\cdot)]}{\delta x(t')} \cdot \delta x(t'), \qquad \text{when } \delta x \text{ is small.}$$