APPENDIX D

CALCULUS OF VARIATIONS

At several points in this book we make use of the technique of *functional differentiation,* also known as *calculus of variations.* Here we give a brief introduction to this topic, using an analogy to conventional differentiation. We can regard a function $f(x)$ as a transformation which takes x as input, and which generates f as output. For this function we can define its derivative df/dx by considering the change in $f(x)$ when the value of x is changed by a small amount δx so that

$$\delta f = \frac{df}{dx}\delta x + \mathcal{O}(\delta x^2). \tag{D.1}$$

A function of many variables $f(x_1, \ldots, x_d)$ can be regarded as a transformation which depends on a discrete set of independent variables. For such a function we have

$$\delta f = \sum_{i=1}^{d} \frac{\partial f}{\partial x_i}\delta x_i + \mathcal{O}(\delta x^2). \tag{D.2}$$

In the same way, we can consider a *functional,* written as $E[f]$, which takes a function $f(x)$ as input and returns a scalar value E. As an example of a functional, consider

$$E[f] = \int \left\{ \left(\frac{df}{dx}\right)^2 + f^2 \right\} dx \tag{D.3}$$

so that the value of $E[f]$ depends on the particular choice of the function $f(x)$. The concept of a functional derivative arises when we consider how much $E[f]$ changes when we make a small change $\delta f(x)$ to the function $f(x)$, where $\delta f(x)$ is a function of x which has small magnitude everywhere but which is otherwise arbitrary. We denote the functional derivative of $E[f]$ with respect to $f(x)$ by $\delta E/\delta f(x)$, and define it by the following relation:

$$\delta E = E[f + \delta f] - E[f] = \int \frac{\delta E}{\delta f(x)}\delta f(x)dx + \mathcal{O}(\delta f^2). \tag{D.4}$$

This can be seen as a natural extension of (D.2) where now $E[f]$ depends on a continuous set of variables, namely the values of f at all points x. As an illustration, we can calculate the derivative of the functional given in (D.3):

$$E[f+\delta f] = E[f] + 2\int \left\{ \frac{df}{dx}\frac{d}{dx}\delta f + f\delta f \right\} dx + O(\delta f^2). \tag{D.5}$$

This can be expressed in the form (D.4) if we integrate by parts, and assume that the boundary term vanishes. We then obtain the following result for the functional derivative:

$$\frac{\delta E}{\delta f(x)} = -2\frac{d^2 f}{dx^2} + 2f. \tag{D.6}$$

Note that, from (D.4) we also have the following useful result:

$$\frac{\delta f(x)}{\delta f(x')} = \delta(x - x') \tag{D.7}$$

where $\delta(x)$ is the Dirac delta function. This result is easily verified by taking $E[f] = f(x)$ and then substituting (D.7) into (D.4).

If we require that, to lowest order in $\delta f(x)$, the functional $E[f]$ be stationary then from (D.4) we have

$$\int \frac{\delta E}{\delta f(x)} \delta f(x) dx = 0. \tag{D.8}$$

Since this must hold for an arbitrary choice of $\delta f(x)$ we can choose $\delta f(x) = \delta(x - x')$ where $\delta(x)$ is the Dirac delta function. Hence it follows that

$$\frac{\delta E}{\delta f(x)} = 0 \tag{D.9}$$

so that, requiring the functional to be stationary with respect to arbitrary variations in the function is equivalent to requiring that the functional derivative vanish.

If we define a differential operator $D \equiv d/dx$ then (D.3) can be written as

$$E = \int \left\{ (Df)^2 + f^2 \right\} dx. \tag{D.10}$$

Following the same argument as before we see that the functional derivative becomes

$$\frac{\delta E}{\delta f(x)} = 2\widehat{D}Df(x) + 2f(x) \tag{D.11}$$

where $\widehat{D} \equiv -d/dx$ is the *adjoint operator* to the operator D. Similar forms of adjoint operator arise in the discussion of radial basis functions and regularization in Section 5.4.