1.4 Appendix: The Delta Function

Strictly speaking, the delta function is not a function at all, but a **distribution**. The difference between a distribution and a function is small but important, particularly for the physicist who wishes to avoid unpleasant comments from mathematicians who grumble about "improper functions" and so on. Distributions arise quite naturally in a number of situations. Most important to us are those circumstances where it is desirable to talk about the probability for a continuously distributed random variable to take on a particular value. For example, if a given random variable x is equally likely to be found having any value in the interval $(0, \alpha)$, then it can be characterized by the uniform probability distribution

$$\rho(x) = \begin{cases} 1/\alpha & \text{if } x \in (0, \alpha) \\ 0 & \text{if } x \notin (0, \alpha) \end{cases} . \tag{1.88}$$

This distribution has the property that the probability that x has *some* value is equal to one, i.e.,

$$\int_{-\infty}^{\infty} \rho(x) dx = 1, \tag{1.89}$$

while the probability to find x in any other interval is simply the integral of $\rho(x)$ over that interval. This idea is readily extended to include any sufficiently well-behaved integrable function (where well-behaved is here rather loosely defined, since it obviously includes discontinuous functions). Other common distributions include the Gaussian

$$\rho(x) = \frac{\alpha}{\sqrt{\pi}} \exp[-\alpha(x - x')^2] \tag{1.90}$$

and the Lorentzian

$$\rho(x) = \frac{1}{\pi} \frac{\alpha}{\alpha^2 + (x - x')^2} \tag{1.91}$$

both of which we have written in a form which is centered at an arbitrary point $x = x'$ and in which the width of the distribution is controlled by a parameter α. The distribution "function" $\rho(x)$ is also referred to as the *probability density* to find the random variable taking a value in the neighborhood of x, and we refer to $dP = \rho(x)dx$ as the probability to find the random variable in the interval between x and $x + dx$.

The Dirac delta function $\delta(x - x')$ is to be viewed as the limiting case of a distribution which is entirely concentrated at single point (in the same way that a point charge is to be viewed as the limiting case of a continuous charge distribution in which all of the charge is entirely concentrated at a single point). In other words, if it is impossible that the variable x take on any value *other* than x', we *define* its probability distribution to be the Dirac distribution

$$\rho(x) = \delta(x - x') \tag{1.92}$$

centered at that point. This implies, among other things, that the integral over the Dirac distribution

$$\int_{-\infty}^{\infty} \delta(x - x')\, dx = 1 \tag{1.93}$$

is equal to unity. On the other hand, the integral of this distribution over any interval *not* containing the point $x = x'$ must be zero, since the probability of finding the variable in such an interval vanishes by assumption. Hence we conclude that

$$\int_{x_1}^{x_2} \delta(x - x')\, dx = \begin{cases} 1 & \text{if } x' \in (x_1, x_2) \\ 0 & \text{if } x' \notin (x_1, x_2) \end{cases} . \tag{1.94}$$

It should be evident that this is equivalent to the relations

$$\int_{-\infty}^{x} \delta(x-x')\,dx = \theta(x-x'), \tag{1.95}$$

and

$$\frac{d}{dx}\theta(x-x') = \delta(x-x') \tag{1.96}$$

where $\theta(x-x')$ is the Heaviside step function which increases discontinuously at $x = x'$ from zero to one.

In fact, the delta distribution can be considered the limiting case of any sufficiently narrow distribution (such as the uniform distribution of our original example) in the limit that its spread goes to zero. Thus, the distribution

$$\rho_\alpha(x) = \begin{cases} 1/\alpha & \text{if } x \in (x' - \frac{\alpha}{2}, x' + \frac{\alpha}{2}) \\ \\ 0 & x \notin (x' - \frac{\alpha}{2}, x' + \frac{\alpha}{2}) \end{cases} \tag{1.97}$$

becomes entirely concentrated at the point $x = x'$ in the limit that $\alpha \to 0$. Clearly in order for the integral of the distribution to remain constant, its height must diverge as its width shrinks to zero. This turns out to be a characteristic feature of any distribution which tends to a $\delta-$function in some well defined limit. You should convince yourself that the Lorentzian distribution given above tends to a delta function as $\alpha \to 0$; the Gaussian does so in the opposite limit $\alpha \to \infty$.

Another important property associated with general distributions is the concept of mean values. If the random variable x is described by the distribution $\rho(x)$, and if $f(x)$ is some function whose value depends upon that of the random variable x, then the **mean value** associated with the function $f(x)$ is given by the expression

$$\langle f \rangle = \int_{-\infty}^{\infty} \rho(x)\,f(x)\,dx. \tag{1.98}$$

On the other hand, if the random variable x *always* takes the value x', as it does when x is characterized by the delta distribution, then the mean value of $f(x)$ must just be the value of f evaluated at this one value. This then motivates the property that

$$\int_{-\infty}^{\infty} \delta(x-x')\,f(x)\,dx = f(x') \tag{1.99}$$

which holds for any function $f(x)$ which is continuous at $x = x'$. Thus, the delta function tends to simply "pick out" the value of any function at the point at which it is centered. This gives us a general rule for evaluating any integral containing a delta function: it is simply equal to the value of the factors in the integrand multiplying the delta function, evaluated at the integration point where the argument of the delta function equals zero. In summary, we list the basic properties of the simple one-dimensional delta function below.

$$\delta(x-x') = \begin{cases} 0 & \text{if } x \neq x' \\ \\ \infty & \text{if } x = x' \end{cases} \tag{1.100}$$

$$\int_{x_1}^{x_2} \delta(x-x')\,dx = \begin{cases} 1 & \text{if } x' \in (x_1, x_2) \\ \\ 0 & \text{if } x' \notin (x_1, x_2) \end{cases} \quad . \tag{1.101}$$

$$\int_{x_1}^{x_2} f(x)\delta(x-x')\,dx = \begin{cases} f(x') & \text{if } x' \in (x_1, x_2) \\ \\ 0 & \text{if } x' \notin (x_1, x_2) \end{cases} \tag{1.102}$$

The purist might well claim that the second of these properties is really a special case of the last. Other properties of the delta function can be proven either from these basic properties or from the definition of the delta function as the limiting form of a suitable family of distributions as discussed above. Such properties include the following, which will be given without proof:

$$\delta(ax) = \frac{1}{|a|}\delta(x). \tag{1.103}$$

$$\delta[f(x)] = \sum_i \frac{1}{|f'(x_i)|}\delta(x-x_i) \tag{1.104}$$

In these expressions, a is a constant, the quantities x_i denote the zeroes of the function $f(x)$, and $f'(x)$ denotes the derivative of f with respect to its argument. Note that these relations imply that the delta function is symmetric with respect to its argument, i.e., $\delta(x-x') = \delta(x'-x)$.

Finally, it is possible to provide a relatively well-defined meaning to the *derivative* of the delta function. The idea of taking a derivative of a "function" which is everywhere zero except at one point (where it is presumably infinite) tends to make mathematicians squirm. In fact, what we mean is that limiting distribution which is the derivative of any distribution whose limit is itself a delta function. The "function" which we shall denote by

$$\delta'(x-x') = \frac{d}{dx}\delta(x-x') \tag{1.105}$$

is even stranger than the delta function, insofar as it appears to be zero everywhere except the point $x = x'$, its integral is zero, but it has the property that (note carefully where the primes go in this expression

$$\int_{x_1}^{x_2} \delta'(x-x')f(x')\,dx' = \begin{cases} f'(x) & \text{if } x \in (x_1, x_2) \\ \\ 0 & \text{if } x \notin (x_1, x_2) \end{cases} \tag{1.106}$$

so the derivative of $\delta(x-a)$ returns the value of the *derivative* of the function at the point where it is centered. This relation can formally be obtained through an integration by parts, i.e., consider

$$\begin{aligned} \int_{x_1}^{x_2} \delta(x-x')\frac{df(x')}{dx'}\,dx' &= \delta(x-x')f(x')|_{x_1}^{x_2} - \int_{x_1}^{x_2} \frac{d}{dx'}\delta(x-x')f(x')\,dx' \\ f'(x) &= \int_{x_1}^{x_2} \delta'(x-x')f(x')\,dx' \end{aligned}$$

where we have used the fact that the integrand on the left evaluates to $f'(x)$, dropped the first term since the delta function vanishes at the endpoints of the integration, and used the fact that $d/dx'\,[\delta(x-x')] = -\delta'(x-x')$, with the minus sign coming from an application of the chain rule. Equally odd higher order derivatives of the delta function can also be defined, but we shall not have much use of them, and so leave them to the interested student to pursue.

Extension to Higher Dimensions

The extension of the delta function to higher dimensions is straightforward. Consider two random variables x and y, governed by the *joint* probability distribution function $\rho(x, y)$, which is defined so that $dP = \rho(x, y)dxdy$ gives the probability that the first variable takes a value between x and $x + dx$, while the second takes a value in the interval y to $y + dy$. This implies the normalization

$$\int\int dxdy\, \rho(x, y) = 1, \tag{1.107}$$

where the integration is over the entire xy plane. It is convenient to view this as defining the probability density governing a randomly distributed vector $\vec{r} = x\hat{\imath} + y\hat{\jmath}$. Thus, we write $\rho(\vec{r}) = \rho(x, y)$, so that

$$\int d^2r\, \rho(\vec{r}) = 1. \tag{1.108}$$

The two-dimensional delta function is the limiting form of such a distribution when there is only one possibility for the random vector $\vec{r}$. That is, if $\vec{r} = \vec{r}_0$ with unit probability, then by definition $\rho(\vec{r}) = \delta(\vec{r}-\vec{r}_0)$. It is not hard to see that $\delta(\vec{r}-\vec{r}_0) = \delta(x-x_0)\delta(y-y_0)$, where x_0 and y_0 are the components of the vector $\vec{r}_0$.

More generally, if $\vec{r}\,'$ is a vector in d-dimensions we can define the d-dimensional delta function $\delta(\vec{r} - \vec{r}\,')$ having the following properties:

$$\delta(\vec{r} - \vec{r}\,') = \begin{cases} 0 & \text{if } \vec{r} \neq \vec{r}\,' \\ \infty & \text{if } \vec{r} = \vec{r}\,' \end{cases} \tag{1.109}$$

$$\int dV\, \delta(\vec{r} - \vec{r}\,') = 1 \tag{1.110}$$

$$\int dV\, f(\vec{r})\, \delta(\vec{r} - \vec{r}\,') = f(\vec{r}\,'). \tag{1.111}$$

where dV denotes the infinitesimal volume element in d-dimensions, and it is assumed that the integrations are over any region containing $\vec{r}\,'$ (the integral over any region not containing $\vec{r}_0$ vanishing). The d-dimensional delta function can be written as the d-fold product

$$\delta(\vec{r} - \vec{r}\,') = \delta(x_1 - x_1')\delta(x_2 - x_2') \cdots \delta(x_d - x_d') \tag{1.112}$$

of delta functions associated with each of the cartesian components.

In higher dimensions the *gradient* of the delta function takes the part of the one-dimensional derivative of the simple delta function. Thus we can define the distribution $\vec{\nabla}\delta(\vec{r} - \vec{r}\,')$ which in three dimensions, for example, has the property that

$$\int d^3r'\ \vec{\nabla}\delta(\vec{r} - \vec{r}\,')f(\vec{r}\,') = \vec{\nabla} f(\vec{r}). \tag{1.113}$$

In a similar fashion it is possible to give meaning to the Laplacian of the delta function $\nabla^2\delta(\vec{r} - \vec{r}\,')$, which has the property that

$$\int d^3r'\ \nabla^2\delta(\vec{r} - \vec{r}\,')f(\vec{r}\,') = \nabla^2 f(\vec{r}) \tag{1.114}$$

Note that this allows us to express a differential operator as an integral operator.