Supplement 2-A

The Fourier Integral and Delta Functions

Consider a function $f(x)$ that is periodic, with period $2L$, so that

$$f(x) = f(x + 2L) \tag{2A-1}$$

Such a function can be expanded in a Fourier series in the interval $(-L, L)$, and the series has the form

$$f(x) = \sum_{n=0}^{\infty} A_n \cos \frac{n\pi x}{L} + \sum_{n=1}^{\infty} B_n \sin \frac{n\pi x}{L} \tag{2A-2}$$

We can rewrite the series in the form

$$f(x) = \sum_{n=-\infty}^{\infty} a_n e^{in\pi x/L} \tag{2A-3}$$

which is certainly possible, since

$$\cos \frac{n\pi x}{L} = \frac{1}{2}(e^{in\pi x/L} + e^{-in\pi x/L})$$

$$\sin \frac{n\pi x}{L} = \frac{1}{2i}(e^{in\pi x/L} - e^{-in\pi x/L})$$

The coefficients can be determined with the help of the orthonormality relation

$$\frac{1}{2L}\int_{-L}^{L} dx\, e^{in\pi x/L} e^{-im\pi x/L} = \delta_{mn} = \begin{cases} 1 & m = n \\ 0 & m \neq n \end{cases} \tag{2A-4}$$

Thus

$$a_n = \frac{1}{2L}\int_{-L}^{L} dx\, f(x) e^{-in\pi x/L} \tag{2A-5}$$

Let us now rewrite (2A-3) by introducing Δn, the difference between two successive integers. Since this is unity, we have

$$\begin{aligned} f(x) &= \sum_n a_n e^{in\pi x/L}\, \Delta n \\ &= \frac{L}{\pi} \sum_n a_n e^{in\pi x/L} \frac{\pi\, \Delta n}{L} \end{aligned} \tag{2A-6}$$

Let us change the notation by writing

$$\frac{\pi n}{L} = k \tag{2A-7}$$

and

$$\frac{\pi\, \Delta n}{L} = \Delta k \tag{2A-8}$$

We also write

$$\frac{La_n}{\pi} = \frac{A(k)}{\sqrt{2\pi}} \tag{2A-9}$$

Hence (2A-6) becomes

$$f(x) = \sum \frac{A(k)}{\sqrt{2\pi}} e^{ikx}\, \Delta k \tag{2A-10}$$

If we now let $L \to \infty$, then k approaches a continuous variable, since Δk becomes infinitesimally small. If we recall the Riemann definition of an integral, we see that in the limit (2A-10) can be written in the form

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dk\, A(k) e^{ikx} \tag{2A-11}$$

The coefficient $A(k)$ is given by

$$\begin{aligned} A(k) &= \sqrt{2\pi}\,\frac{L}{\pi} \cdot \frac{1}{2L} \int_{-L}^{L} dx\, f(x) e^{-in\pi x/L} \\ &\to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dx\, f(x) e^{-ikx} \end{aligned} \tag{2A-12}$$

Equations (2A-11) and (2A-12) define the Fourier integral transformations. If we insert the second equation into the first we get

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dk\, e^{ikx} \int_{-\infty}^{\infty} dy\, f(y) e^{-iky} \tag{2A-13}$$

Suppose now that we interchange, without question, the order of integrations. We then get

$$f(x) = \int_{-\infty}^{\infty} dy\, f(y) \left[\frac{1}{2\pi} \int_{-\infty}^{\infty} dk\, e^{ik(x-y)} \right] \tag{2A-14}$$

For this to be true, the quantity $\delta(x - y)$ defined by

$$\delta(x - y) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dk\, e^{ik(x-y)} \tag{2A-15}$$

and called the *Dirac delta function* must be a very peculiar kind of function; it must vanish when $x \neq y$, and it must tend to infinity in an appropriate way when $x - y = 0$, since the range of integration is infinitesimally small. It is therefore not a function of the usual mathematical sense, but it is rather a "generalized function" or a "distribution."[1] It does not have any meaning by itself, but it can be defined provided it always appears in the form

$$\int dx\, f(x)\, \delta(x - a)$$

with the function $f(x)$ sufficiently smooth in the range of values that the argument of the delta function takes. We will take that for granted and manipulate the delta function by itself, with the understanding that at the end all the relations that we write down only occur under the integral sign.

The following properties of the delta function can be demonstrated:

(i)

$$\delta(ax) = \frac{1}{|a|}\,\delta(x) \tag{2A-16}$$

This can be seen to follow from

$$f(x) = \int dy\, f(y)\, \delta(x - y) \tag{2A-17}$$

If we write $x = a\xi$ and y and $a\eta$, then this reads

$$f(a\xi) = |a| \int d\eta\, f(a\eta)\, \delta[a(\xi - \eta)]$$

On the other hand,

$$f(a\xi) = \int d\eta\, f(a\eta)\, \delta(\xi - \eta)$$

which implies our result.

(ii) A relation that follows from (2A-16) is

$$\delta(x^2 - a^2) = \frac{1}{2|a|}\,[\delta(x - a) + \delta(x + a)] \tag{2A-18}$$

This follows from the fact that the argument of the delta function vanishes at $x = a$ and $x = -a$. Thus there are two contributions:

$$\begin{aligned}
\delta(x^2 - a^2) &= \delta[(x - a)(x + a)] \\
&= \frac{1}{|x + a|}\,\delta(x - a) + \frac{1}{|x - a|}\,\delta(x + a) \\
&= \frac{1}{2|a|}\,[\delta(x - a) + \delta(x + a)]
\end{aligned}$$

More generally, one can show that

$$\delta[f(x)] = \sum_i \frac{\delta(x - x_i)}{|df/dx|_{x=x_i}} \tag{2A-19}$$

where the x_i are the roots of $f(x)$ in the interval of integration.

---

[1]The theory of distributions was developed by the mathematician Laurent Schwartz. An introductory treatment may be found in M. J. Lighthill, *Introduction to Fourier Analysis and Generalized Functions*, Cambridge University Press, Cambridge, England, 1958.

In addition to the representation (2A-15) of the delta function, there are other representations that may prove useful. We discuss several of them.

(a) Consider the form (2A-15), which we write in the form

$$\delta(x) = \frac{1}{2\pi} \operatorname*{Lim}_{L\to\infty} \int_{-L}^{L} dk\, e^{ikx} \tag{2A-20}$$

The integral can be done, and we get

$$\begin{aligned}\delta(x) &= \operatorname*{Lim}_{L\to\infty} \frac{1}{2\pi} \frac{e^{iLx} - e^{-iLx}}{ix} \\ &= \operatorname*{Lim}_{L\to\infty} \frac{\sin\, Lx}{\pi x}\end{aligned} \tag{2A-21}$$

(b) Consider the function $\Delta(x, a)$ defined by

$$\begin{aligned}\Delta(x, a) &= 0 \qquad && x < -a \\ &= \frac{1}{2a} && -a < x < a \\ &= 0 && a < x\end{aligned} \tag{2A-22}$$

Then

$$\delta(x) = \operatorname*{Lim}_{a\to 0} \Delta(x, a) \tag{2A-23}$$

It is clear that an integral of a product of $\Delta(x, a)$ and a function $f(x)$ that is smooth near the origin will pick out the value at the origin

$$\begin{aligned}\operatorname*{Lim}_{a\to 0} \int dx\, f(x)\, \Delta(x, a) &= f(0) \operatorname*{Lim}_{a\to 0} \int dx\, \Delta(x, a) \\ &= f(0)\end{aligned}$$

(c) By the same token, any peaked function, normalized to unit area under it, will approach a delta function in the limit that the width of the peak goes to zero. We will leave it to the reader to show that the following are representations of the delta function:

$$\delta(x) = \operatorname*{Lim}_{a\to 0} \frac{1}{\pi} \frac{a}{x^2 + a^2} \tag{2A-24}$$

and

$$\delta(x) = \operatorname*{Lim}_{\alpha\to\infty} \frac{\alpha}{\sqrt{\pi}}\, e^{-\alpha^2 x^2} \tag{2A-25}$$

(d) We will have occasion to deal with *orthonormal polynomials*, which we denote by the general symbol $P_n(x)$. These have the property that

$$\int dx\, P_m(x) P_n(x) w(x) = \delta_{mn} \tag{2A-26}$$

where $w(x)$ may be unity or some simple function, called the weight function. For functions that may be expanded in a series of these orthogonal polynomials, we can write

$$f(x) = \sum_n a_n P_n(x) \tag{2A-27}$$

If we multiply both sides by $w(x)P_m(x)$ and integrate over x, we find that

$$a_m = \int dy\, w(y) f(y) P_m(y) \tag{2A-28}$$

We can insert this into (2A-27) and, prepared to deal with "generalized functions," we freely interchange sum and integral. We get

$$\begin{aligned} f(x) &= \sum_n P_n(x) \int dy\, w(y) f(y) P_n(y) \\ &= \int dy\, f(y) \left(\sum_n P_n(x) w(y) P_n(y) \right) \end{aligned} \tag{2A-29}$$

Thus we get still another representation of the delta function. Examples of the $P_n(x)$ are Legendre polynomials, Hermite polynomials, and Laguerre polynomials, all of which make their appearance in quantum mechanical problems.

Since the delta function always appears multiplied by a smooth function under an integral sign, we can give meaning to its derivatives. For example,

$$\begin{aligned} \int_{-\varepsilon}^{\varepsilon} dx\, f(x) \frac{d}{dx} \delta(x) &= \int_{-\varepsilon}^{\varepsilon} dx \frac{d}{dx} [f(x)\, \delta(x)] - \int_{-\varepsilon}^{\varepsilon} dx \frac{df(x)}{dx} \delta(x) \\ &= -\int_{-\varepsilon}^{\varepsilon} dx \frac{df(x)}{dx} \delta(x) \\ &= -\left(\frac{df}{dx} \right)_{x=0} \end{aligned} \tag{2A-30}$$

and so on. The delta function is an extremely useful tool, and the student will encounter it in every part of mathematical physics.

The integral of a delta function is

$$\begin{aligned} \int_{-\infty}^{x} dy\, \delta(y - a) &= 0 \qquad x < a \\ &= 1 \qquad x > a \\ &\equiv \theta(x - a) \end{aligned} \tag{2A-31}$$

which is the standard notation for this discontinuous function. Conversely, the derivative of the so-called *step function* is the Dirac delta function:

$$\frac{d}{dx} \theta(x - a) = \delta(x - a) \tag{2A-32}$$

Supplement 2-B

A Brief Tutorial on Probability

In this supplement we give a brief discussion of probability. For simplicity we start with discrete events. Consider the toss of a six-faced die. If the die is not perfectly symmetric, the outcome for face n ($n = 1, 2, \ldots, 6$) has a probability p_n. In a large number N of tosses, the face "n" turns up a_n times, and we say that the probability of getting the face n is

$$p_n = \frac{a_n}{N} \tag{2B-1}$$

Since $\Sigma a_n = N$, it follows that

$$\Sigma p_n = 1 \tag{2B-2}$$

Let us now assign to each face a certain "payoff." We assign points in the following manner: 1 point when face 1 turns up, 2 points when face 2 turns up, and so on. In N tosses we get n points a_n times, so that the total number of points is $\Sigma n a_n$. This, of course, grows with N. We thus focus on the average value (points per toss), so that

$$\langle n \rangle = \frac{1}{N} \sum_n n a_n = \sum_n n p_n \tag{2B-3}$$

We may be interested in an average of n^2, say, and in that case we calculate

$$\langle n^2 \rangle = \frac{1}{N} \sum_n n^2 a_n = \sum_n n^2 p_n \tag{2B-4}$$

and so on.

When we do not have a discrete outcome, we must deal with densities. To be specific, consider a quantity that varies continuously, for example the height of a population of students. We can make this discrete by making a *histogram* (Fig. 2B-1) plotting the height by listing people's heights in 10-cm intervals. Nobody will be exactly 180 cm or 190 cm tall, so we just group people, and somehow round things up at the edges. We may want a finer detail, and list the heights in 1-cm intervals or 1-mm intervals, which will continue to make things discrete, but as the intervals become smaller, the histogram resembles more and more a continuous curve. Let us take some interval, dx, and treat it as infinitesimal. The number of people whose height lies between x and $x + dx$ is $n(x)\,dx$. The proportionality to dx is obvious: twice as many people will fall into the interval $2dx$ as fall into dx. It is here that the infinitesimal character of dx comes in: we do not need to decide whether $n(x)\,dx$ or $n(x + dx/2)\,dx$ is to be taken, since we treat $(dx)^2$ as vanishingly small. If the total population has N members, we can speak of the probability of falling into the particular interval as

$$\frac{1}{N} n(x)\,dx = p(x)\,dx \tag{2B-5}$$

$p(x)$ is called a *probability density*. Since the total number of students is N, we have

$$\int dx\, n(x) = N \tag{2B-6}$$

or

$$\int dx\, p(x) = 1 \tag{2B-7}$$

If we want the probability of finding a student of height between a (say, 150 cm) and b (say, 180 cm), we calculate

$$P(a \le x \le b) = \int_a^b dx\, p(x) \tag{2B-8}$$

The average height is calculated in the same way as for the discrete case

$$\langle x \rangle = \int xp(x)\, dx \tag{2B-9}$$

and we can calculate other quantities such as

$$\langle x^2 \rangle = \int x^2 p(x)\, dx \tag{2B-10}$$

and so on. Instead of calling these quantities averages, we call them *expectation values*, a terminology that goes back to the roots of probability theory in gambling theory.

We are often interested in a quantity that gives a measure of how the heights, say, are distributed about the average. The deviations from the average must add up to zero. Formally it is clear that

$$\langle x - \langle x \rangle \rangle = 0 \tag{2B-11}$$

since the average of any number, including $\langle x \rangle$, is just that number. We can, however, calculate the average value of the square of the deviation, and this quantity will not be zero. In fact,

$$\begin{aligned} \langle (x - \langle x \rangle)^2 &= \langle x^2 - 2x\langle x \rangle + \langle x \rangle^2 \rangle = \langle x^2 \rangle - 2\langle x \rangle\langle x \rangle + \langle x \rangle^2 \\ &= \langle x^2 \rangle - \langle x \rangle^2 = (\Delta x)^2 \end{aligned} \tag{2B-12}$$

This quantity is called the *dispersion*.

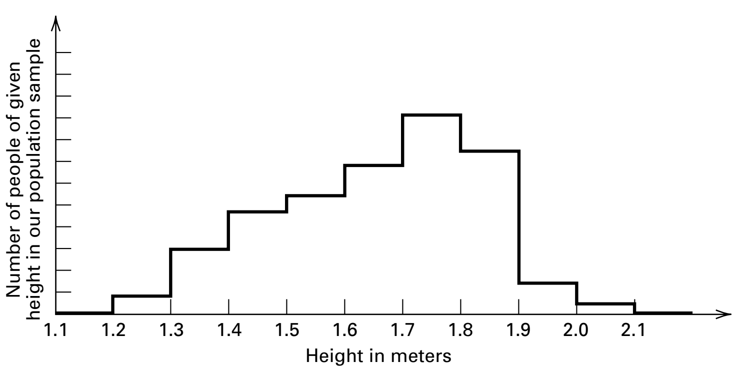

Figure 2B-1 Example of a histogram.

The dispersion is used in quantum mechanics to define the *uncertainty*, so that in the Heisenberg uncertainty relations proper definitions of $(\Delta x)^2$ and $(\Delta p)^2$ are given by

$$(\Delta x)^2 = \langle x^2 \rangle - \langle x \rangle^2 \tag{2B-13}$$

and

$$(\Delta p)^2 = \langle p^2 \rangle - \langle p \rangle^2 \tag{2B-14}$$

There is one more point to be made that is directly relevant to quantum physics. We may illustrate it by going back to the population of students. Let us ask for the probability that a randomly chosen student's height lies between 160 cm and 170 cm, *and* that his or her telephone number ends in an even integer. Assuming that there is nothing perverse about how telephone numbers are distributed in the community, half the students will fall into the even category, and thus the desired probability is $P(160 \le x \le 170) \times (1/2)$. This is just an example of a general rule *that the probability of two (or more) uncorrelated events is the product of the individual probabilities*.

Thus, if the probability that a particle is in some state "n" in our laboratory is $P(n)$, and the probability that a different particle in a laboratory across the country is in a state "m" is $P(m)$, then the joint probability that one finds "n" in the first laboratory *and* "m" in the second laboratory is

$$P(m, n) = P(n)P(m) \tag{2B-15}$$

We will find that a similar result holds for probability amplitudes—that is, for wave functions.