A quick introduction to the delta function and distributions

Contents

1 Introduction

Physics and engineering textbooks often give a nonrigorous presentation of the following important technique for solving differential equations of the form

$$Lu = g \tag{1}$$

where L is a linear differential operator and g is a continuous function. The technique can be summarized as follows:

1. First we find a function f which satisfies (or supposedly satisfies)

$$Lf = \delta \tag{2}$$

   where δ is the "delta function". The delta function is often described as having the following properties:

$$\delta(x) = \begin{cases} \infty & \text{if } x = 0 \\ 0 & \text{if } x \neq 0 \end{cases} \qquad \text{and} \qquad \int_{\mathbb{R}} \delta(x)\,dx = 1.$$

   Of course, no such function actually exists!

2. Then the function $u = f * g$, the convolution of f and g, is a solution to (1).

Let's illustrate the use of this method by solving the very simple differential equation $u' = g$. We first seek a function f which satisfies

$$f' = \delta. \tag{3}$$

Proceeding nonrigorously, we invoke the fundamental theorem of calculus and assert that the function

$$f(x) = \int_{-\infty}^{x} \delta(y)\,dy = \begin{cases} 1 & \text{if } x > 0, \\ 0 & \text{if } x < 0 \end{cases}$$

satisfies equation (3). This function f is called the Heaviside step function. The next step is to note that if $u = f * g$ then

$$u(x) = \int_{\mathbb{R}} f(y)g(x-y)\,dy = \int_{0}^{\infty} g(x-y)\,dy = \int_{-\infty}^{x} g(z)\,dz$$

(In the final step, we made a substitution $z = x - y$.) Despite our nonrigorous claim that $f' = \delta$, the function u that we have obtained is indeed a genuine solution to $u' = g$. Whatever objections might have been raised against this method, it actually works.

The above method might seem at first to be complete nonsense, given the lack of a coherent statement of what δ is. But, the method is in fact quite intuitive if we interpret δ to be merely an "approximate delta function", by which I mean a smooth function that has the following properties:

- δ has a spike near the origin.
- δ is zero elsewhere.
- $\int_{\mathbb{R}} \delta(x)\,dx = 1$.

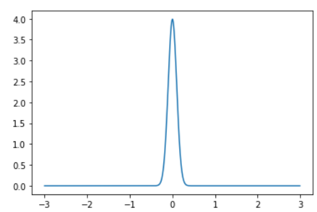

Figure 1: An approximate delta function.

(See figure 1.) The intuitive idea is that, since L is shift invariant, we can shift f to obtain a solution to (1) in the special case where the function on the right is a shifted version of δ (a shifted spike). Then, because L is linear, we can solve (1) whenever the function on the right is a sum of shifted spikes. The punch line is that we can think of *any* continuous function g as a sum of shifted spikes:

$$g(x) \approx \int_{\mathbb{R}} g(y)\delta(x-y)\,dy. \tag{4}$$

So, invoking the linearity and shift invariance of L, we recognize that if

$$u(x) = \int_{\mathbb{R}} g(y) f(x-y)\,dy$$

then $Lu \approx g$.

Despite its intuitive appeal, the above argument remains nonrigorous. If δ is merely an approximate delta function, then (4) is only an approximation. Moreover, physicists and engineers do not hesitate to do other nonrigorous things with the delta function, such as taking the Fourier transform of both sides of equation (2) in order to solve for f. The theory of distributions gives us a way to precisely define the delta function and to make such arguments rigorous.

2 Definitions and basic theory about distributions

Let $V = C_c^\infty(\mathbb{R})$, the vector space of all functions $\phi : \mathbb{R} \to \mathbb{R}$ which are infinitely differentiable and have compact support. The elements of V are called "test functions".

Definition 1. *A* ***distribution*** *is a linear function $F : V \to \mathbb{R}$ which is continuous in the following sense: if $\{\phi_k\}$ is a sequence of functions in V and for each nonnegative integer n we have*

$$\phi_k^{(n)} \to \phi^{(n)} \textit{ uniformly as } k \to \infty$$

then $F(\phi_k) \to F(\phi)$ as $k \to \infty$. Here $\phi^{(n)}$ is the nth-order derivative of ϕ. I'll denote the set of all distributions as V^*.

The strange-sounding definition of continuity that we are using is designed, perhaps by trial and error, so that the "derivative" of F that we will define below is guaranteed to also be a distribution. For more discussion of how this definition can be motivated, see here.

If F is a distribution and ϕ is a test function, we'll use the notation $\langle F, \phi \rangle$ as an alternative way of writing $F(\phi)$. So by definition

$$\langle F, \phi \rangle = F(\phi).$$

Example 2.1. If $f : \mathbb{R} \to \mathbb{R}$ is locally integrable, then the function $F : V \to \mathbb{R}$ defined by

$$F(\phi) = \int_{\mathbb{R}} f(x)\phi(x)\,dx \qquad \text{for all } \phi \in V$$

is a distribution.

Example 2.2. The function $\delta : V \to \mathbb{R}$ defined by

$$\delta(\phi) = \phi(0) \qquad \text{for all } \phi \in V$$

is a distribution. It is called the "delta function" or, more properly, the delta distribution.

The integration by parts formula

$$\int_{\mathbb{R}} f'(x)\phi(x)\,dx = -\int_{\mathbb{R}} f(x)\phi'(x)\,dx \qquad \text{for all } \phi \in V$$

suggests a way to define the derivative of any distribution F.

Definition 2. *Let F be a distribution. The function $DF : V \to \mathbb{R}$ defined by*

$$DF(\phi) = -\langle F, \phi' \rangle \qquad \text{for all } \phi \in V$$

is called the "derivative" of F.

Our strange-sounding definition of continuity for a distribution was designed in order to make the following theorem true:

Theorem 1. *If F is a distribution, then its derivative DF is also a distribution.*

Proof. (Left to reader, for now.) □

The operator $D : V^* \to V^*$ which takes a distribution F as input and returns DF as output is called the "derivative operator" on the space V^* of distributions. We can compute higher derivatives of a distribution F by applying the operator D repeatedly. For example, $D^2 F = D(DF)$. Notice that if $\phi \in V$ then

$$D^2 F(\phi) = -\langle DF, \phi' \rangle = \langle F, \phi'' \rangle.$$

More generally, we have

$$D^n F(\phi) = (-1)^n \langle F, \phi^{(n)} \rangle$$

for any positive integer n. Here $\phi^{(n)}$ denotes the nth derivative of ϕ.

Example 2.3. Let $h : \mathbb{R} \to \mathbb{R}$ be the Heaviside step function defined by

$$h(x) = \begin{cases} 1 & \text{if } x > 0, \\ 0 & \text{if } x < 0. \end{cases}$$

(I haven't specified the value of h when $x = 0$ because it doesn't matter here.) Let H be the distribution defined by

$$H(\phi) = \int_{\mathbb{R}} h(x)\phi(x)\,dx = \int_0^\infty \phi(x)\,dx.$$

Then for any $\phi \in V$ we have

$$\begin{aligned}
DH(\phi) &= -\langle H, \phi' \rangle \\
&= -\int_0^\infty \phi'(x)\, dx \\
&= \phi(0) \\
&= \delta(\phi).
\end{aligned}$$

Thus, $DH = \delta$. This gives a precise meaning to the claim that "the derivative of the Heaviside function is the delta function", which is a statement often heard in nonrigorous treatments of this subject.

If $f : \mathbb{R} \to \mathbb{R}$ is locally integrable, and ϕ is a test function, then the convolution of f and ϕ is the function $f * \phi$ defined by

$$(f * \phi)(x) = \int_{\mathbb{R}} f(y)\phi(x-y)\, dy.$$

The above formula suggests a way to define the convolution of ϕ with a distribution.

Definition 3. *Suppose that F is a distribution and ϕ is a test function. The convolution of F and ϕ is the function $F * \phi : \mathbb{R} \to \mathbb{R}$ defined by*

$$(F * \phi)(x) = \langle F, \phi(x - \cdot) \rangle \quad \text{for all } x \in \mathbb{R}.$$

Here $\phi(x - \cdot)$ is a shorthand notation for the function $y \mapsto \phi(x-y)$.

Example 2.4. Let's compute the convolution of the delta distribution δ with a test function ϕ. If $x \in \mathbb{R}$ then

$$\begin{aligned}
(\delta * \phi)(x) &= \langle \delta, \phi(x - \cdot) \rangle \\
&= \phi(x - 0) \\
&= \phi(x).
\end{aligned}$$

So we see that

$$\delta * \phi = \phi.$$

It's interesting to note that when we convolve a distribution with a test function, the result is a function rather than a distribution. That explains how finding a distribution which satisfies the equation (2) can lead to finding an actual function that satisfies the equation (1). To be specific, we will see that if g is a test function and a distribution F satisfies (2), then the function $F * g$ satisfies (1). In preparation for that, we first state and prove the following key theorem.

Theorem 2. *Suppose that F is a distribution and ϕ is a test function. Then $F * \phi$ is differentiable and*

$$(F * \phi)' = F * \phi' = (DF) * \phi.$$

Proof. Let $x \in \mathbb{R}$. If $t \in \mathbb{R}$ and $t \neq 0$ then

$$(F * \phi)(x+t) = \langle F, \phi(x + t - \cdot) \rangle \quad \text{and} \quad (F * \phi)(x) = \langle F, \phi(x - \cdot) \rangle$$

so

$$\frac{(F * \phi)(x+t) - (F * \phi)(x)}{t} = \left\langle F, \underbrace{\frac{\phi(x+t-\cdot) - \phi(x - \cdot)}{t}}_{\psi_t} \right\rangle = \langle F, \psi_t \rangle,$$

where ψ_t is the function defined by $\psi_t(y) = \frac{\phi(x+t-y)-\phi(x-y)}{t}$. It can be shown that ψ_t converges uniformly to $\phi'(x - \cdot)$ as $t \to 0$, and likewise for each positive integer n the nth derivative of ψ_t converges uniformly to the nth derivative of $\phi'(x - \cdot)$ as $t \to 0$. By the continuity property of distributions, we see that

$$\lim_{t \to 0} \frac{(F * \phi)(t+h) - (F * \phi)(x)}{t} = \langle F, \phi'(x - \cdot) \rangle = (F * \phi')(x).$$

Thus, $F * \phi$ is differentiable and $(F * \phi)' = F * \phi'$.

Next, notice that

$$\begin{aligned}
(DF * \phi)(x) &= \langle DF, \phi(x - \cdot) \rangle \\
&= \langle F, -[\phi(x - \cdot)]' \rangle \\
&= \langle F, \phi'(x - \cdot) \rangle \qquad \text{(by the chain rule)} \\
&= (F * \phi')(x).
\end{aligned}$$

This shows that $DF * \phi = F * \phi'$. ☐

Repeatedly applying theorem 2 yields the following corollary.

Corollary 1. *If F is a distribution and ϕ is a test function, then $F * \phi$ is infinitely differentiable, and for any positive integer n we have*

$$(F * \phi)^{(n)} = F * \phi^{(n)} = (D^n F) * \phi. \tag{5}$$

Proof. By theorem 2, we know that $F * \phi$ is differentiable and that $(F * \phi)' = F * \phi' = (DF) * \phi$. Invoking theorem (2) again, we see that $(DF) * \phi$ is differentiable and its derivative is $D^2 F * \phi$. Because ϕ' is a test function, theorem (2) also tells us that the derivative of $F * \phi'$ is $F * \phi''$. Thus, $(F * \phi)'$ is differentiable and

$$(F * \phi)'' = F * \phi'' = D^2 F * \phi.$$

Continuing like this, we can see that $F * \phi$ is in fact infinitely differentiable, and that equation (5) holds for any positive integer n. ☐

Corollary 2. *Let $L : V^* \to V^*$ be a differential operator, which means that*

$$L = \sum_{n=1}^{N} c_n D^n$$

for some real numbers $c_1, \ldots, c_N$. Suppose that F is a distribution which satisfies $L(F) = \delta$. If g is a test function, then the function $u = F * g$ satisfies

$$\sum_{n=1}^{N} c_n u^{(n)} = g.$$

(Here $u^{(n)}$ denotes the nth derivative of u.)

Proof. Using corollary 1, we see that

$$\begin{aligned}
\sum_{n=1}^{N} c_n (F * g)^{(n)} &= \sum_{n=1}^{N} c_n (D^n F) * g \\
&= \left(\sum_{n=1}^{N} c_n D^n \right) * g \\
&= \delta * g \\
&= g.
\end{aligned}$$

☐

3 Example: Solving a simple differential equation

In this section we'll illustrate the use of distributions in solving differential equations by solving the very simple differential equation

$$u' = g$$

where g is a test function. First note that the Heaviside distribution H defined in example 2.3 satisfies

$$DH = \delta.$$

It follows from corollary 2 that the function $u = H * g$ satisfies $u' = g$. All that remains is to find a more explicit formula for $H * g$. If $x \in \mathbb{R}$, then

$$\begin{aligned}(H * g)(x) &= \langle H, g(x - \cdot)\rangle \\ &= \int_0^\infty g(x - y)\, dy.\end{aligned}$$

Making a change of variable $z = x - y$, we see that

$$(H * g)(x) = -\int_x^{-\infty} g(z)\, dz = \int_{-\infty}^x g(z)\, dz.$$

Of course, we could have found this antiderivative directly using the fundamental theorem of calculus. But, the above calculation illustrates a general technique — we first find a fundamental solution (in this case H), and then we convolve the fundamental solution with g to obtain a solution to $Lu = g$.

4 Questions

- When defining the convolution of a distribution F and a function g, can I loosen the restrictions on g? It seems overly restrictive to assume that g is infinitely differentiable and has compact support.
- How can I extend this approach to solve $Lu = g$, where g is a less nice function? For example, can I drop the assumption that g has compact support?