16 Delta function (and other fantastic beasts)

> §450. The following story is true. There was a little boy, and his father said, "Do try to be like other people. Don't frown." And he tried and tried, but could not. So his father beat him with a strap; and then he was eaten up by lions.
>
> Reader, if young, take warning by his sad life and death. For though it may be an honour to be different from other people, if Carlyle's dictum[1] about the 30 million be still true, yet other people do not like it. So, if you are different, you had better hide it, and pretend to be solemn and wooden-headed. Until you make your fortune. For most wooden-headed people worship money; and, really, I do not see what else they can do. In particular, if you are going to write a book, remember the wooden-headed. So be rigorous; that will cover a multitude of sins. And do not frown.
>
> Oliver Heaviside (1850–1925)
> Electromagnetic Theory, Vol. 3

16.1 Motivation

My next goal is to introduce the so-called *Green's functions* for solving stationary boundary value problems. There are different ways to define these functions, I am going to pick one that makes a heavy use of the so-called *delta function*, which is, strictly speaking, not a function. (Delta function is often somewhat incorrectly called Dirac delta function, there are strong reasons to believe that Paul Dirac (1902–1984, one of the greatest theoretical physicists of the 20th century) picked delta function from Heaviside's work.) There exists a rigorous theory of *generalized functions* or *distributions*, of which delta function is just one example, but since my use of this theory will be quite limited I will frequently appeal to intuition and natural properties, instead of providing mathematical proofs that the manipulations I perform are legitimate.

To motivate the appearance of delta function consider the following imaginary physical experiment[2]. Assume that a mass m is moving along the x-axis with constant speed v. At the time $t = 0$ an elastic collision with a wall occurs (a collision is called *elastic* if the total kinetic energy of the two bodies after the encounter is equal to their total kinetic energy before the encounter). After the collision the mass moves in the opposite direction with the same speed. If v_1, v_2 denote the speeds at times t_1, t_2 then by the law of mechanics

$$m(v_2 - v_1) = \int_{t_1}^{t_2} F(t)\,\mathrm{d}t,$$

where F denotes the intensity of the force acting on mass m. If $t_1 < t_2 < 0$ or $0 < t_1 < t_2$ then no problem, $v_1 = v_2$ and $F = 0$ since no force is acting. If, however, $t_1 < 0 < t_2$ then the left hand side of my equality is $2mv$, but the right hand side, as we all know from the calculus, must be zero, since

Math 483/683: Partial Differential Equations by Artem Novozhilov
e-mail: artem.novozhilov@ndsu.edu. Spring 2023

[1]To explain the meaning: Thomas Carlyle, when asked the population of England, said "Thirty millions, mostly fools."

[2]I am copying this example from Salsa, Sandro. Partial differential equations in action: from modelling to theory. Vol. 86. Springer, 2015, where also the theory of generalized functions can be found. Another very good source for a deeper journey to the theory of generalized functions is A.S. Demidov, Generalized functions in mathematical physics: Main ideas and concepts, Nova Science Publishers, Inc, 2001.

F is zero everywhere except $t = 0$, and therefore I get a contradiction. To get rid of this contradiction I introduce the delta function by means of a *physical* definition

$$\delta(t) = 0, \quad t \neq 0,$$

and

$$\int_{\mathbf{R}} \delta(t)\, \mathrm{d}t = 1.$$

Then if I put

$$F(t) = 2mv\delta(t),$$

then the contradiction in my reasonings evaporates. The price I pay is that now I need to be very careful when dealing with $\delta(t)$ since *no usual function has these properties* and therefore I am dealing with an unknown (yet) object. Mathematically, I still need a rigorous definition of $\delta(t)$, whereas physically delta function is a mathematical model of something, concentrated at a point (a point mass, a unit charge, a unit intensity of the force, etc).

16.2 Delta-like sequences

To give a mathematical justification to the facts above I choose an intuitively appealing (but not the most general and abstract) way of approximating the newly introduced delta function by a sequence of legitimated functions that depend on a parameter, which I eventually send to zero.

Example 16.1. For this first example I take

$$\delta_{\varepsilon}(t) = \begin{cases} \frac{1}{2\varepsilon}, & -\varepsilon < t < \varepsilon, \\ 0, & \text{otherwise.} \end{cases}$$

For each fixed ε I have

$$\int_{\mathbf{R}} \delta_{\varepsilon}(t)\, \mathrm{d}t = 1,$$

and therefore

$$\lim_{\varepsilon \to 0} \int_{\mathbf{R}} \delta_{\varepsilon}(t)\, \mathrm{d}t = \lim_{\varepsilon \to 0} 1 = 1,$$

which is my mathematical *justification* that $\int_{\mathbf{R}} \delta(t)\, \mathrm{d}t = 1$, and moreover, for any $t \neq 0$ there is $\varepsilon > 0$ such that

$$\delta_{\varepsilon}(t) = 0,$$

which justifies the physical definition that $\delta(t) = 0$ for all $t \neq 0$.

In short, when I will use the notion of delta function, mathematically it means that I consider a *delta*-like sequence of regular functions (as in this example for instance), do the legitimate calculations, and take the limit $\varepsilon \to 0$ (I do not write it explicitly, but usually I take the limit from the right, i.e., $\varepsilon \to 0^+$ or $\varepsilon \downarrow 0$, whatever notation you prefer). The most important part of this approach is that at some point the intermediate (and most tedious) step can be skipped, and I will have enough experience and intuition to jump to the correct conclusion remembering that a proper mathematical justification can always be supplied if needed.

The two properties above (that delta function is zero everywhere except zero and that the integral of delta function is one) are not the most important ones and certainly cannot be used for a definition of a delta-like sequence. Here is actually the key property. Let me make sense of the expression

$$\int_{\mathbf{R}} \delta(t) f(t)\,\mathrm{d}t$$

for a continuous f.

Using the introduced paradigm, I calculate

$$\begin{aligned}\int_{\mathbf{R}} \delta(t) f(t)\,\mathrm{d}t &= \lim_{\varepsilon\to 0}\int_{\mathbf{R}} \delta_\varepsilon(t) f(t)\,\mathrm{d}t = \\ &= \lim_{\varepsilon\to 0}\frac{1}{2\epsilon}\int_{-\varepsilon}^{\varepsilon} f(t)\,\mathrm{d}t = \\ &= \lim_{\varepsilon\to 0}\frac{1}{2\epsilon}\cdot 2\varepsilon f(\xi), \quad -\varepsilon\leq\xi\leq\varepsilon, \\ &= f(0),\end{aligned}$$

where I used the integral mean value theorem for continuous f. By using this calculation I now *gave meaning* to the meaningless before expression, and *proved* that

$$\int_{\mathbf{R}} \delta(t) f(t)\,\mathrm{d}t = f(0). \tag{16.1}$$

I would like to note that I picked probably the simplest example for a delta-like sequence, but not the nicest one to work with, since none of $\delta_\varepsilon(t)$ is continuous, let alone differentiable. I can certainly be more general. Moreover, to make things somewhat more applicable in the future, I now would like to introduce a slight modification, which will allow me to consider a *shifted* delta-function $\delta(t-\xi)$ (in other words, delta function applied at the point ξ).

Example 16.2. Consider function φ defined on the real line such that $\varphi\geq 0$, $\int_{\mathbf{R}}\varphi(t)\,\mathrm{d}t = 1$, and $\varphi(0) = 0$ outside of the interval $(-1,1)$. Now I define a family of functions by

$$\delta_\varepsilon(t) = \frac{1}{\varepsilon}\varphi\left(\frac{t}{\varepsilon}\right).$$

Using the properties above I have, for any interval $(a,b)\subseteq\mathbf{R}$ and point $\xi\in(a,b)$ that

$$\int_{\mathbf{R}}\delta_\varepsilon(t)\,\mathrm{d}t = \int_a^b \delta_\varepsilon(t-\xi)\,\mathrm{d}t = 1, \quad \xi\in(a,b).$$

Now I would like to generalize the property of δ-like sequence from the previous example to my new family.

Lemma 16.3. *Let f be continuous on (a,b), then*

$$f(\xi) = \lim_{\varepsilon\to 0}\int_a^b f(t)\delta_\varepsilon(t-\xi)\,\mathrm{d}t, \quad \xi\in(a,b). \tag{16.2}$$

Proof. Since f is continuous, I can always find $\varepsilon > 0$ such that for any $\eta > 0$ $|f(t) - f(\xi)| \leq \eta$ if $|t - \xi| \leq \varepsilon$. Using the standard properties of the integral, I have

$$\begin{aligned}\left|\int_a^b f(t)\delta_\varepsilon(t-\xi)\,\mathrm{d}t - f(\xi)\right| &= \left|\int_a^b \big(f(t) - f(\xi)\big)\delta_\varepsilon(t-\xi)\,\mathrm{d}t\right| \\ &\leq \int_{\xi-\varepsilon}^{\xi+\varepsilon} |f(t) - f(\xi)|\delta_\varepsilon(x-\xi)\,\mathrm{d}t \\ &\leq \eta \int_{\xi-\varepsilon}^{\xi+\varepsilon} \delta_\varepsilon(t-\xi)\,\mathrm{d}t = \eta,\end{aligned}$$

as required. ■

Exercise 1. Can you come up with some explicit examples for function φ?

Finally now I am in a position to define the terms I already used.

Definition 16.4. *A sequence of functions $\delta_\varepsilon \colon (a, b) \longrightarrow \mathbf{R}$ such that equality (16.2) holds for any continuous on (a, b) function f is called delta-sequence concentrated near the point ξ. When I use in the computations delta function $\delta(t - \xi)$ concentrated at the point ξ, it precisely means that I first replace the delta function with a delta-like sequence, perform the computations, and then take the limit $\varepsilon \to 0$.*

According to the given deginition the equality (16.2) is a mathematical explanation what the expression

$$\int_{\mathbf{R}} f(t)\delta(t-\xi)\,\mathrm{d}t = f(\xi) \tag{16.3}$$

means. The equality (16.3) is the most important statement so far, since, in words, it says, that any continuous function f can be represented as the *linear combination* of the values of this function at the points t weighted by shifted delta functions $\delta(t - \xi)$ (the latter is often written as δ_ξ).

Certainly the list of delta-like sequences can be extended. For the following I will need a delta-like sequence that is differentiable, here is an example.

Exercise 2. Show that the sequence of functions

$$\delta_\varepsilon(t) = \frac{1}{\pi}\frac{\varepsilon}{t^2 + \varepsilon^2} \tag{16.4}$$

is a delta-like sequence for all bounded and continuous on $\mathbf{R}$ functions (I must add *bounded* here to make sure that all the corresponding integrals converge). Clearly all δ_ε here are infinitely differentiable.

Sometimes it is helpful to get a geometric idea of how the delta-like sequence looks like. In Fig. 1 I plot the delta-like sequences from Example 16.1 and Exercise 2.

I remark that the functions in delta-like sequences do not have to be nonnegative. As an example, I can consider the sequence

$$\delta_\varepsilon(t) = \frac{\sin \varepsilon t}{\pi t}\,, \tag{16.5}$$

which can be shown to satisfy Definition 16.4 (not an easy exercise). This sequence is shown in Fig. 2.

Now I can start practicing with properties of delta function.

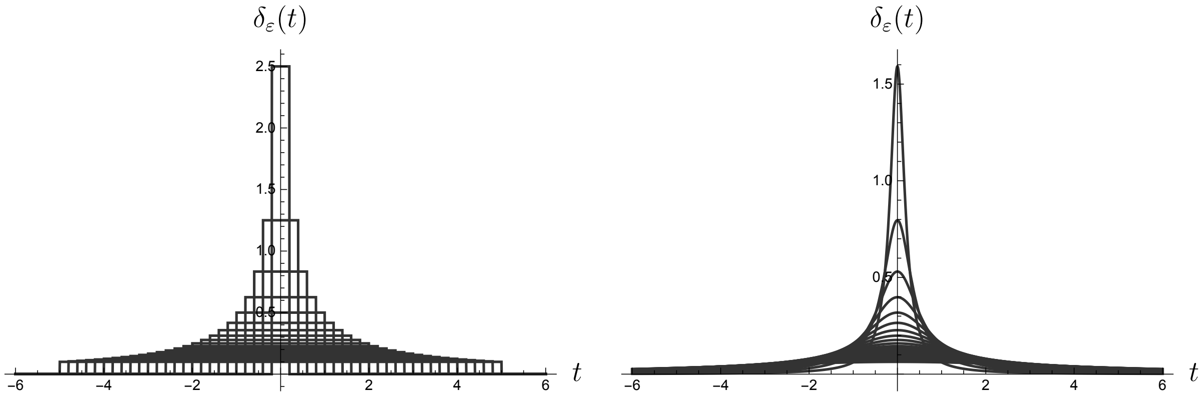

Figure 1: Delta-like sequences from Example 16.1 (left) and Exercise 2 (right).

Example 16.5. For instance, the Wikipedia article on delta function claims that it is an *even* function, i.e.,

$$\delta(t) = \delta(-t).$$

What does it mean if it is not even an ordinary function? I again must look at my delta-like sequences. Clearly, since each function in Example 16.1 and Exercise 2 is even, hence I will conclude that the limit must be even. So let me instead deal with somewhat more general Example 16.2, where I do not assume that φ is an even function.

My goal is to show that

$$\int_{\mathbf{R}} f(t)\delta(t)\,\mathrm{d}t = \int_{\mathbf{R}} f(t)\delta(-t)\,\mathrm{d}t.$$

Ok, using exactly the same steps as in the proof of Lemma 16.3, I find

$$\left|\int_{\mathbf{R}} f(t)\delta_\varepsilon(-t)\,\mathrm{d}t - f(0)\right| \leq \eta \int_{-\varepsilon}^{\varepsilon} \delta_\varepsilon(-t)\,\mathrm{d}t = -\eta \int_{\varepsilon}^{-\varepsilon} \delta_\varepsilon(\tau)\,\mathrm{d}\tau = \eta$$

after the change $\tau = -t$, as required.

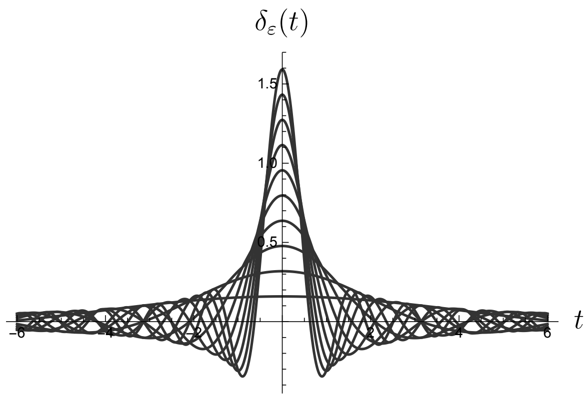

Figure 2: Delta-like sequence (16.5).

The proved property allows me to rewrite (16.3) as more suggestive

$$\int_{\mathbf{R}} f(t)\delta(\xi - t)\,\mathrm{d}t = f(\xi).$$

Now that we have some experience working with the delta function, let me state, leaving the proofs as exercises, that, first, a linear combination of delta functions makes perfect sense, i.e., I can consider

$$\alpha\delta(t-\xi) + \beta\delta(t-\eta),$$

meaning that

$$\int_{\mathbf{R}} f(t)\Big(\alpha\delta(t-\xi) + \beta\delta(t-\eta)\Big)\,\mathrm{d}t = \alpha f(\xi) + \beta f(\eta).$$

Second, for any continuous at ξ function g I have that $g(t)\delta(t-\xi) = g(\xi)\delta(t-\xi)$ (prove it, you have all the required tools).

Example 16.6. Next, let me calculate

$$\int_{-\infty}^{t} \delta(t)\,\mathrm{d}t.$$

Again I use the same strategy (I will use the delta-like sequence from Example 16.1 but any of them leads to the same result):

$$\begin{aligned}\int_{-\infty}^{t} \delta(t)\,\mathrm{d}t &= \lim_{\varepsilon\to 0}\int_{-\infty}^{t} \delta_\varepsilon(t)\,\mathrm{d}t = \\ &= \lim_{\varepsilon\to 0}\begin{cases} 0, & t < -\varepsilon \\ \frac{1}{2\varepsilon}(t+\varepsilon), & -\varepsilon \leq t \leq \varepsilon = \\ 1, & t > \varepsilon \end{cases} \\ &= \begin{cases} 0, & t < 0 \\ \frac{1}{2}, & t = 0 =: \chi(t). \\ 1, & t > 0 \end{cases}\end{aligned}$$

That is the result is the *unit step function* or, how it is also frequently called, *Heaviside's* function[3] χ (please note that different textbooks use different notations for the Heaviside function).

Generalizing slightly, I get

$$\int_{-\infty}^{t} \delta(t-\xi)\,\mathrm{d}t = \chi(t-\xi).$$

Now I can integrate the delta function, what about differentiation? We know from Calculus that if

$$\int_{a}^{x} f(t)\,\mathrm{d}t = F(x)$$

then (by the fundamental theorem of Calculus)

$$F'(x) = f(x).$$

[3]In a way this is really mocking to keep the name of Heaviside attached to the most trivial object of the whole theory.

Therefore, I postulate that

$$\frac{\mathrm{d}\chi}{\mathrm{d}t}(t-\xi)=\delta(t-\xi).$$

That is, delta function (which is, one more time, is not a function) is the derivative the Heaviside function.

Let me try to explain a little more here what just has happened. The students should not be confused by the stated fact (and what follows after) although formally it simply contradicts many facts from other math classes (especially from Calculus). Here I am taking derivatives in some *generalized* sense. Following the ideology about the delta function I am looking at my objects not as independent entities but from the *averaging* or *integral* point of view. That is, when I am writing that $\chi'=\delta$ I actually mean that for some sufficiently nice functions f (for all at the same time!) I get

$$\int_{\mathbf{R}}\chi'(t)f(t)\,\mathrm{d}t=\int_{\mathbf{R}}\delta(t)f(t)\,\mathrm{d}t=f(0),$$

where I used (16.1) at the end, since I already justified this fact. Let me prove it, by choosing differentiable functions f that vanish outside of some sufficiently large interval $(-a,a)$ for finite a (such functions are said to have *compact support*). Indeed, using the integration by parts,

$$\begin{aligned}\int_{\mathbf{R}}\chi'(t)f(t)\,\mathrm{d}t&=f(t)\chi(t)|_{-\infty}^{\infty}-\int_{\mathbf{R}}\chi(t)f'(t)\,\mathrm{d}t=-\int_{\mathbf{R}}\chi(t)f'(t)\,\mathrm{d}t\\&=-\int_0^{\infty}f'(t)\,\mathrm{d}t=-f(t)|_0^{\infty}=f(0)\end{aligned}$$

as required.

In the same *generalized* sense should be understood all other derivatives below; it is very important to embrace the idea that we do not care any more about values of our functions at specific points (say, the value of delta function at zero does not make any sense), from the more general prospective, we care only how my functions (or any objects I use) act under the integral signs.

The fact that I now can differentiate the Heaviside function allows me to find derivatives of *any* piecewise continuously differentiable functions at all points including the points of discontinuity. Let me show it by an example.

Example 16.7. As a first example I will find the derivative of

$$\operatorname{sgn}(t)=\begin{cases}-1, & t<0,\\ 1, & t>0.\end{cases}$$

I can represent my function as

$$\operatorname{sgn}(t)=1\cdot\chi(t)+(-1)\cdot\chi(-t).$$

Hence, using the ordinary rules of differentiating, I get

$$\operatorname{sgn}'(t)=\begin{cases}0, & t\neq 0,\\ 2\delta(t), & t=0,\end{cases}$$

since

$$(1)' = (-1)' = 0, \quad \chi'(t) = \delta(t), \quad \chi'(-t) = -\delta(-t) = -\delta(t),$$

using the chain rule and the proven fact that delta function is even.

(If such computations are still worrying, I invite the reader to replace them with a sequence of rigorous steps, involving integrals, as I did for the derivative of the Heaviside function.)

Example 16.8. Generalizing the previous example, let me assume that function f is continuously differentiable everywhere except the point $t = \xi$, where it has the jump discontinuity of magnitude α. Using the Heaviside function I can write that

$$f(t) = g(t) + \alpha\chi(t - \xi),$$

where now g is continuous and differentiable anywhere but $t = \xi$. According to the previous, I get

$$f'(t) = g'(t) + \alpha\delta(t - \xi).$$

I (for the n-th time) emphasize that I still do not know what the value of f' at the point $t = \xi$ (since there is no meaning for the value of delta function concentrated at the point $t = \xi$)! However, what I gained is the fact that now I can consider integrals of the form $\int h(t)f'(t)\,\mathrm{d}t$ and have perfect meaning for them.

Example 16.9. What is the derivative of $|t|$? For three semesters of calculus the students were told that there is no such thing as the derivative of $|t|$ on the interval that includes 0. Now, armed with the previous examples, I get that, in the generalized sense (supply the missing details),

$$|t|' = \begin{cases} +1, & t > 0 \\ -1, & t < 0 \end{cases} = \operatorname{sgn} t,$$

the so-called sign function. (Once again, I do not know the value of the derivative at the point $t = 0$, but I do know that

$$\int_{\mathbf{R}} f(t)|t|'\,\mathrm{d}t = \int_{\mathbf{R}} f(t)\operatorname{sgn} t\,\mathrm{d}t$$

for any reasonable continuous f.)

I already know the derivative of sgn, and hence

$$|t|'' = 2\delta(t).$$

I can even find the derivative of delta function! In the spirit of the examples above, I need to understand how new object (the derivative of delta function) acts on other functions under the integral sign. Here I will assume that my functions are differentiable at 0.

I have, using integration by parts, that

$$\begin{aligned}\int_{\mathbf{R}} \delta'(t)f(t)\,\mathrm{d}t &= \lim_{\varepsilon\to 0}\int_{\mathbf{R}} \delta'_{\varepsilon}(t)f(t)\,\mathrm{d}t = \\ &= \lim_{\varepsilon\to 0}\left(\delta_{\varepsilon}(t)f(t)|_{-\infty}^{\infty} - \int_{\mathbf{R}} \delta_{\varepsilon}(t)f'(t)\,\mathrm{d}t\right) = \\ &= -\lim_{\varepsilon\to 0}\int_{\mathbf{R}} \delta_{\varepsilon}(t)f'(t)\,\mathrm{d}t = \\ &= -f'(0).\end{aligned}$$

That is, the derivative of delta function is a new object that, if pared with a differentiable function under the integral sign, returns the value $-f'(0)$ as the output. At this point it should be clear how to find *any* derivative of delta function. E.g.,

$$\int_{\mathbf{R}} \delta''(t)f(t)\,\mathrm{d}t = f''(0),$$

and so on.

In the following I will need also delta functions that depend on more than one independent variable. In this case it is convenient to set, for, e.g., $\boldsymbol{x} = (x_1, x_2, x_3)$, $\boldsymbol{\xi} = (\xi_1, \xi_2, \xi_3)$,

$$\delta(\boldsymbol{x} - \boldsymbol{\xi}) = \delta(x_1 - \xi_1)\delta(x_2 - \xi_2)\delta(x_3 - \xi_3),$$

and this is a very rare case when I am allowed to multiply delta functions. In general, the operation of multiplication for delta functions is not well defined (for instance, the expression $\delta(x-1)\delta(x-2)$ does not make any sense).

Exercise 3. Recall that when I discussed the differentiation of Fourier series, I mentioned that one must be careful. It is a remarkable fact that if I understand my derivatives in the generalized sense discussed above, no more care with taking derivatives should be exercised. Everything is (correctly) differentiable!

Find the Fourier series for the Heaviside function and the delta function on $[-\pi, \pi]$. Observe that the Fourier series for the delta function can be obtained by formally differentiating the Fourier series for the Heaviside function (keep in mind that what you find is the Fourier series for 2π periodic extensions). Do you recognize any of the results (recall me differentiating the Fourier sine series that I found for the function x)? Can you now reconcile all the results?

16.3 Weak solutions to PDE

16.4 Test yourself

16.5 Solutions to the exercises