

Distributions – generalized functions

Andras Vasy

March 25, 2004

The problem

One of the main achievements of 19th century mathematics was to carefully analyze concepts such as the continuity and differentiability of functions. Recall that f is differentiable at x , and its derivative is $f'(x) = L$, if the limit

$$\lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h}$$

exists, and is equal to L .

While it was always clear that not every continuous function is differentiable, e.g. the function $f : \mathbb{R} \rightarrow \mathbb{R}$ given by $f(x) = |x|$ is not differentiable at 0, it was not until the work of Bolzano and Weierstrass that the full extent of the problem became clear: there are nowhere differentiable continuous functions.

Let u be the saw-tooth function: $u(0) = 0$, $u(1/2) = 1/2$, u is periodic with period 1, and linear on $[0, 1/2]$ as well as on $[1/2, 1]$. Then let

$$f(x) = \sum_{j=0}^{\infty} c_j u(q^j x),$$

for suitable c_j and q – e.g. $q = 16$, $c_j = 2^{-j}$ work. Then the sum converges to a continuous function f , but the difference quotients do not have limits. In fact, u could be replaced even by $u(x) = \sin(2\pi x)$.

However, one can make sense of f' and even the 27th derivative of f for *any* continuous f if one relaxes the requirement that f' be a function. So, for instance, we cannot expect f' to have values at any point – it will be a distribution, i.e. a ‘generalized function’, introduced by Schwartz and Sobolev.

Why care?

- PDE's: most PDE's are not explicitly solvable. Related techniques play a crucial role in analyzing PDEs.
- Another PDE example: take the wave equation on the line:

$$u_{tt} = c^2 u_{xx},$$

u a function on $\mathbb{R}_x \times \mathbb{R}_t$, $u_{tt} = \frac{\partial^2 u}{\partial t^2}$, etc. The general solution of this PDE, obtained by d'Alembert in the 18th century, is

$$u(x, t) = f(x + ct) + g(x - ct),$$

where f and g are 'arbitrary' functions on \mathbb{R} . Indeed, it is easy to check by the chain rule that u solves the PDE – as long as we can make sense of the differentiation. So, in the 'classical sense', f, g twice continuously differentiable, written as $f, g \in C^2(\mathbb{R})$, suffice.

But shouldn't this also work for rougher f, g ? For instance, what about the step function $f: f(x) = 1$ if $x \geq 0$, $f(x) = 0$ for $x < 0$?

- Limits of familiar objects are often distributions. For example, for $\epsilon > 0$, define $f_\epsilon: \mathbb{R} \rightarrow \mathbb{C}$ by

$$f_\epsilon(x) = \frac{1}{x + i\epsilon}.$$

What is $\lim_{\epsilon \rightarrow 0} f_\epsilon$? For $x \neq 0$, of course the limit makes sense directly – it is $f(x) = \frac{1}{x}$. But what about $x = 0$? For instance, does $\int_{-1}^1 f(x) dx$ make sense, and what is it? Note that this integral does not converge due to the behavior of the integrand at 0!

However, we can take

$$\lim_{\epsilon \rightarrow 0} \int_{-1}^1 f_\epsilon(x) dx = \lim_{\epsilon \rightarrow 0} \log(x + i\epsilon) \Big|_{-1}^1$$

$$= \log(1) - \log(-1) = 0 - (i\pi) = -i\pi.$$

So, the integral of the limit f on $[-1, 1]$ should be $-i\pi$. Can we make sense of this directly?

- Idealization of physical problems often results in distributions. For instance, the sharp front for the wave equation discussed above, or point charges (the electron is supposed to be such!) are good examples.

I will usually talk about functions on \mathbb{R} , but almost everything makes sense on \mathbb{R}^n , $n \geq 1$ arbitrary.

Notation:

- We say that f is C^0 if f is continuous.
- We say that f is C^k , $k \geq 1$ integer, if f is k times continuously differentiable, i.e. if f is C^{k-1} and its $(k-1)$ st derivative, $f^{(k-1)}$, is differentiable, and its derivative, $f^{(k)}$ is continuous.
- We say that f is C^∞ , i.e. f is infinitely differentiable, if f is C^k for every k .

Motivation: to deal with very 'bad' objects, first we need very 'good' ones.

Example of an interesting C^∞ function on \mathbb{R} :
 $f(x) = 0$ for $x \leq 0$, $f(x) = e^{-1/x}$ for $x > 0$.

An even more interesting example: $g(x) = f(1 - x^2)$. Note that g is 0 for $|x| \geq 1$.

Our very good functions then will be the (complex-valued) functions ϕ which are C^∞ and which are 0 outside a bounded set, i.e. there is $R > 0$ such that $\phi(x) = 0$ for $|x| \geq R$. The set of such functions is denoted by $C_c^\infty(\mathbb{R})$, and its elements are called 'compactly supported smooth functions' or simply 'test functions'.

There are other sets of very good functions with which analogous conclusions are possible: e.g. C^∞ functions which decrease faster than $C_k|x|^{-k}$ at infinity for all k , and analogous estimates hold for their derivatives. Such functions are called Schwartz functions.

The set $C_c^\infty(\mathbb{R})$ is a vector space with the usual pointwise addition of functions and pointwise multiplication by scalars $c \in \mathbb{C}$. Since this is an infinite dimensional vector space, we need one more notion:

Suppose that $\phi_n, n \in \mathbb{N}$, is a sequence in $C_c^\infty(\mathbb{R})$, and $\phi \in C_c^\infty(\mathbb{R})$. We say that $\phi_n \rightarrow \phi$ in $C_c^\infty(\mathbb{R})$ if

1. there is an $R > 0$ such that $\phi_n(x) = 0$ for all n and for all $|x| \geq R$,
2. and for all k , $\max_{x \in \mathbb{R}} \left| \frac{d^k}{dx^k}(\phi_n - \phi) \right| \rightarrow 0$ as $n \rightarrow \infty$, i.e. for all k and for all $\epsilon > 0$, there is N such that

$$n \geq N, x \in \mathbb{R} \Rightarrow \left| \frac{d^k}{dx^k}(\phi_n - \phi) \right| < \epsilon.$$

Now we 'dualize' $C_c^\infty(\mathbb{R})$ to define distributions:

A distribution $u \in \mathcal{D}'(\mathbb{R})$ is a continuous linear functional $u : C_c^\infty(\mathbb{R}) \rightarrow \mathbb{C}$. That is:

1. u is linear:

$$u(c_1\phi_1 + c_2\phi_2) = c_1u(\phi_1) + c_2u(\phi_2)$$

for all $c_j \in \mathbb{C}$, $\phi_j \in C_c^\infty(\mathbb{R})$, $j = 1, 2$.

2. u is continuous: if $\phi_n \rightarrow \phi$ in $C_c^\infty(\mathbb{R})$ then $u(\phi_n) \rightarrow u(\phi)$, i.e. $\lim_{n \rightarrow \infty} u(\phi_n) = u(\phi)$, in \mathbb{C} .

The simplest example is the delta distribution: for $a \in \mathbb{R}$, δ_a is the distribution given by $\delta_a(\phi) = \phi(a)$ for $\phi \in C_c^\infty(\mathbb{R})$.

Another example: for $\phi \in C_c^\infty(\mathbb{R})$, let $u(\phi) = \phi'(1) - \phi''(-2)$.

Why is this a generalization of functions?

If f is continuous (or indeed just locally integrable), we can associate a distribution $\iota(f) = \iota_f$ to it:

$$\iota_f(\phi) = \int_{\mathbb{R}} f(x)\phi(x) dx.$$

Note that $\iota : C^0(\mathbb{R}) \rightarrow \mathcal{D}'(\mathbb{R})$ is injective, i.e. $\iota_{f_1} = \iota_{f_2}$ implies $f_1 = f_2$, or equivalently $\iota_f = 0$ implies $f = 0$, so we can think of $C^0(\mathbb{R})$ as a subset of $\mathcal{D}'(\mathbb{R})$, identifying f with ι_f .

Here we already used that $\mathcal{D}'(\mathbb{R})$ is a vector space: $u_1 + u_2$ is the distribution given by $(u_1 + u_2)(\phi) = u_1(\phi) + u_2(\phi)$, while cu is the distribution given by $(cu)(\phi) = cu(\phi)$ ($c \in \mathbb{C}$).

Convergence: suppose that u_n is a sequence of distributions and $u \in \mathcal{D}'(\mathbb{R})$. We say that $u_n \rightarrow u$ in $\mathcal{D}'(\mathbb{R})$ if for all $\phi \in C_c^\infty(\mathbb{R})$, $\lim_{n \rightarrow \infty} u_n(\phi) = u(\phi)$.

Example: Suppose that $u_n \geq 0$ are continuous functions (i.e. $u_n = \iota_{f_n}$, f_n continuous), $u_n(x) = 0$ for $|x| \geq \frac{1}{n}$, and $\int_{\mathbb{R}} u_n(x) dx = 1$. Then $\lim_{n \rightarrow \infty} u_n = \delta_0$.

Example: Suppose $u_\epsilon(x) = \frac{1}{x+i\epsilon}$. Then for $\epsilon > 0$, $\phi \in C_c^\infty(\mathbb{R})$,

$$\begin{aligned} \int u_\epsilon(x) \phi(x) dx &= \int \frac{1}{x+i\epsilon} \phi(x) dx \\ &= - \int \log(x+i\epsilon) \phi'(x) dx, \end{aligned}$$

But the last expression has a limit as $\epsilon \rightarrow 0$, for \log is locally integrable; the limit is

$$u(\phi) = - \int \log(x+i0) \phi'(x) dx,$$

where $\log(x+i0) = \log|x| + i\pi H(-x)$, with H the step function $H(x) = 1$ if $x > 0$, $H(x) = 0$, if $x < 0$.

If one wants to, one can integrate by parts once more to get

$$\begin{aligned} u(\phi) &= \lim_{\epsilon \rightarrow 0} \int u_\epsilon(x) \phi(x) dx \\ &= \int (x + i\epsilon)(\log(x + i\epsilon) - 1) \phi''(x) dx \\ &= \int x(\log(x + i0) - 1) \phi''(x) dx, \end{aligned}$$

with the integrand continuous now even at $\epsilon = 0$.

The distribution u is called $(x + i0)^{-1}$.

A simple and interesting calculation gives

$$(x + i0)^{-1} - (x - i0)^{-1} = -2\pi i \delta_0.$$

This is all well, but has the goal been achieved, namely can we differentiate any distribution? Yes! We could see this by approximating distributions by differentiable functions, whose derivative we thus already know, and show that the limit exists. But this requires first proving that every distribution can be approximated by such functions. So we proceed more directly.

If $u = \iota_f$, and f is C^1 , we want $u' = \iota_{f'}$. That is, we want

$$\begin{aligned} u'(\phi) &= \iota_{f'}(\phi) = \int f'(x)\phi(x) dx \\ &= - \int f(x)\phi'(x) dx = -\iota_f(\phi') = -u(\phi'). \end{aligned}$$

So for any $u \in \mathcal{D}'(\mathbb{R})$, we define $u' \in \mathcal{D}'(\mathbb{R})$ by $u'(\phi) = -u(\phi')$.

It is easy to see that u' is indeed a distribution. In particular, it can be differentiated again, etc. It is also easy to check that if $u_n \rightarrow u$ in $\mathcal{D}'(\mathbb{R})$ then $u'_n \rightarrow u'$ in $\mathcal{D}'(\mathbb{R})$.

Example: $u = \delta_a$. Then $u'(\phi) = -u(\phi') = -\phi'(a)$, i.e. δ'_a is the distribution $\phi \mapsto -\phi'(a)$.

Example: $u = \iota_H$, H the step function. Then

$$\begin{aligned} u'(\phi) &= -u(\phi') = -\int_{-\infty}^{\infty} H(x)\phi'(x) dx \\ &= -\int_0^{\infty} \phi'(x) dx = \phi(0) = \delta_0(\phi) \end{aligned}$$

by the fundamental theorem of calculus, so $H' = \delta_0$.

Now it is easy to check that $u(x, t) = H(x - ct)$ solves the wave equation!

Another good feature is that all standard identities hold for distributional derivatives, e.g. $\frac{\partial^2 u}{\partial x \partial y} = \frac{\partial^2 u}{\partial y \partial x}$, since they hold for test functions ϕ .

The downside: multiplication does not extend to $\mathcal{D}'(\mathbb{R})$, e.g. $\delta_0 \cdot \delta_0$ makes no sense. To see this, consider a sequence u_n of continuous functions converging to δ_0 , and check that u_n^2 does not converge to *any* distribution. Actually, there are algebraic problems as well: the product rule gives an incompatibility for differentiation and multiplication when applied to 'bad' functions.

This is why solving non-linear PDE's can be hard: differentiation and multiplication fight against each other: e.g. $u_{tt} = u_{xx}^2$.

However, one can still multiply distributions by C^∞ functions f : $(fu)(\phi) = u(f\phi)$, motivated as for differentiation. Thus, distribution theory is ideal for solving variable coefficient linear PDE's: e.g. $u_{tt} = c(x)^2 u_{xx}$.

Also note that

$$(x + i0)^{-1} \cdot (x + i0)^{-1} = (x + i0)^{-2}$$

makes perfectly good sense, as does $(x - i0)^{-2}$. The problem is with the product $(x + i0)^{-1} \cdot (x - i0)^{-1}$. A more general perspective that distinguishes $(x + i0)^{-1}$ and $(x - i0)^{-1}$, by saying that they are both singular at 0 but in different 'directions', is microlocal analysis.

As an application, consider the fundamental theorem of calculus.

Suppose that $u' = f$, and f is a given distribution. What is u ?

Since $f(\psi) = u'(\psi) = -u(\psi')$, we already know what u is applied to the derivative of a test function. But we need to know what $u(\phi)$ is for any test function ϕ .

So let ϕ_0 be a fixed test function with $\int_{\mathbb{R}} \phi_0(x) dx = 1$. If $\phi \in C_c^\infty(\mathbb{R})$, define $\tilde{\phi} \in C_c^\infty(\mathbb{R})$ by

$$\tilde{\phi}(x) = \phi(x) - \left(\int_{\mathbb{R}} \phi(x') dx' \right) \phi_0(x).$$

Then $\int_{\mathbb{R}} \tilde{\phi}(x) dx = 0$, hence $\tilde{\phi}$ is the derivative of a test function ψ , namely we can let

$$\psi(x) = \int_{-\infty}^x \tilde{\phi}(x') dx'.$$

Thus, $\phi(x) = \psi'(x) + \left(\int_{\mathbb{R}} \phi(x') dx' \right) \phi_0(x)$, so

$$\begin{aligned}
u(\phi) &= u(\psi') + \left(\int_{\mathbb{R}} \phi(x') dx' \right) u(\phi_0) \\
&= -f(\psi) + \int_{\mathbb{R}} c\phi(x') dx'
\end{aligned}$$

with $c = u(\phi_0)$ a constant independent of ϕ . Thus, u is determined by $u' = f$, plus the knowledge of $u(\phi_0)$.

In particular, if $f = 0$, we deduce that $u = \iota_c$, i.e. u is a constant function!

This is a form of the fundamental theorem of calculus: if u is C^1 , $a, b \in \mathbb{R}$, $a < b$, we can take ϕ_0 approach δ_a , ϕ approach δ_b , in which case ψ will approach a function that is -1 between a and b , 0 elsewhere, so we recover $u(b) = u(a) + \int_a^b f(x) dx$.

More examples: electrostatics. The electrostatic potential u generated by a charge density ρ satisfies

$$-\Delta u = \rho, \quad \Delta u = u_{xx} + u_{yy} + u_{zz}.$$

If $\rho = \delta_0$, i.e. we have a point charge, what is u ? We need conditions at infinity, such as $u \rightarrow 0$ at infinity, to find u . In fact, $u = \frac{1}{4\pi r}$, $r(X) = |X|$, $X = (x, y, z)$, as a direct calculation shows: to evaluate $-\Delta u$, consider

$$\begin{aligned} -\Delta u(\phi) &= u(-\Delta\phi) = - \int_{\mathbb{R}^3} \frac{1}{4\pi|X|} \Delta\phi(X) dX \\ &= \lim_{\epsilon \rightarrow 0} \int_{|X| > \epsilon} \frac{1}{4\pi|X|} \Delta\phi(X) dX, \end{aligned}$$

and use the divergence theorem to show that the right hand side converges to $\phi(0) = \delta_0(\phi)$!

This also solves the PDE $-\Delta u = f$ for any f (with some decay at infinity), by

$$u(x) = \int E(X - Y) f(Y) dY, \quad E(X) = \frac{1}{4\pi|X|};$$

this integral actually makes sense even if f is a distribution (with some decay at infinity).