3D: Dirac, Delta & Distributionen

Talk at the UNO Mathematics Department
New Orleans, LA, Thursday, 3rd October 1991

Wiederholung im "Kleinen Kolloquium"
Innsbruck, Dienstag, 10. Oktober 1995

1. D'Alembert, Euler, and the Vibrating String

Small deflections $z(t,x)$ of a vibrating string are governed by the wave equation

$$\frac{\partial^2 z}{\partial t^2} - c^2 \frac{\partial^2 z}{\partial x^2} = \frac{F}{\rho}, \tag{1}$$

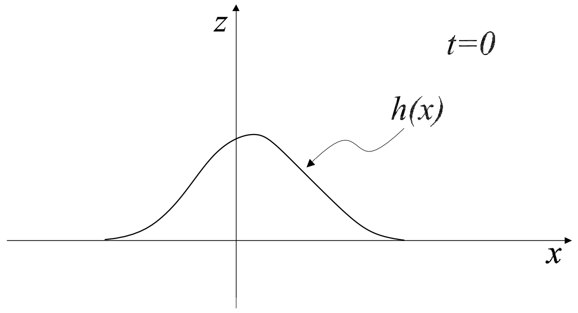
where $c^2 := \frac{T}{\rho}$, $T =$ tension in the string [kg m sec^{-2}], $\rho =$ density/length [kg m^{-1}], $F =$ force/length [kg sec^{-2}].

J. d'Alembert was the first to solve this equation for $F = 0$ and $c =$ constant in his paper "Recherches sur la courbe que forme une corde tendue mise en vibration", Mém. Acad. Sci. Berlin **3**(1747), 214–249. The trick is simple and by now well known: Introduce new co-ordinates $u = x - ct$, $v = x + ct \Longrightarrow \frac{\partial^2 z}{\partial u \partial v} = 0 \Longrightarrow$

$$z = f(u) + g(v) = f(x - ct) + g(x + ct) \tag{2}$$

where f and g are "arbitrary" functions. In order to give a meaning to the differential equation (1), d'Alembert required that f and g be "continuous" functions, which, in the language of that time, would mean that f, g be analytic. (We would be satisfied, of course, if f, g were twice differentiable.)

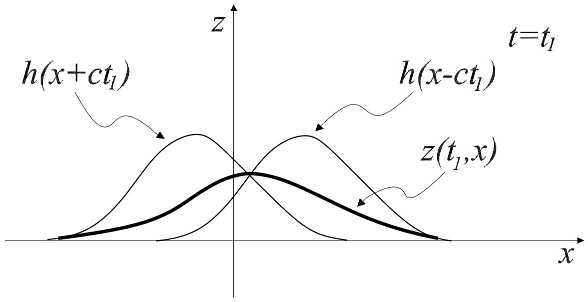
As an example, let us consider an infinite string, which initially has a given shape $h(x)$



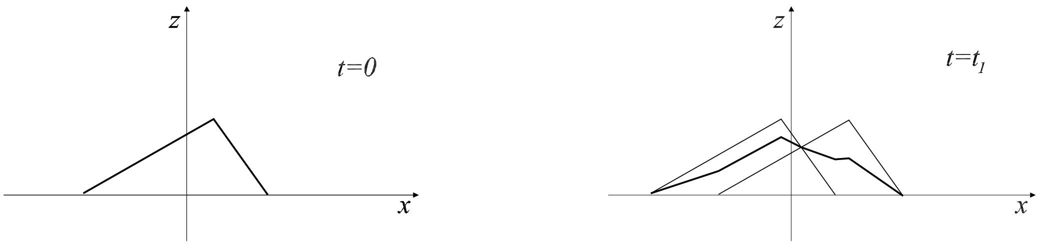
and velocity 0. Starting with initial deflection $z(0,x) = h(x)$ and initial velocity $\partial z/\partial t(0,x) = 0$ results in

$$f = \frac{1}{2}h = g, \text{ i.e., } z(t,x) = \frac{1}{2}[h(x - ct) + h(x + ct)]$$

This yields, at a time $t_1 > 0$, something like this:



Immediately after d'Alembert, Euler wrote a paper on the same subject: "Sur la vibration des cordes", Mém. Acad. Sci. Berlin **4**(1748), 63–72, where he arrived at the same formula (2), but allowed f, g to be "des fonctions discontinues", which in nowadays language would mean that f, g were continuous, but not necessarily differentiable functions. Let us see how this works with a "tugged" string:



Though it is, at first glance, a bit surprising that the edges are not smoothed out (but rather doubled), it looks reasonable. So the solution (2) still makes sense, though it cannot fulfil the differential equation (1), since it is not differentiable. This was the starting point for a long-lasting dispute between Euler and d'Alembert. Euler himself was not quite satisfied, and in 1756 he wrote:

> ...il faudra avouer que cette recherche nous ouvre une nouvelle carrière dans l'analyse, en nous mettant en état d'appliquer le calcul à des courbes, qui ne sont assujettis à aucune loi de continuité, et si cela a paru impossible jusqu'ici, la découverte sera d'autant plus importante.
>
> ...Cette partie dont nous ne connaissons presque encore que les premiers éléments mérite sans doute que tous les Géometres réunissent leurs forces pour la cultiver.

More on this can be found in J. Lützen's book "The Prehistory of the Theory of Distributions", Springer 1981.

The problem raised is thus: How can we differentiate discontinuous functions?

2. Dirac, Delta, and Distributions

The 19th century was mainly concerned with giving rigour to analysis and the question of differentiating a non-differentiable function was therefore considered as meaningless.

Another strange object appearing already very early in mathematical papers was the delta function. Fourier used the formulae

$$\delta(x) = \frac{1}{\pi}\Big(\frac{1}{2} + \sum_{n=1}^{\infty} \cos nx\Big), \qquad -\pi < x < \pi, \tag{3}$$

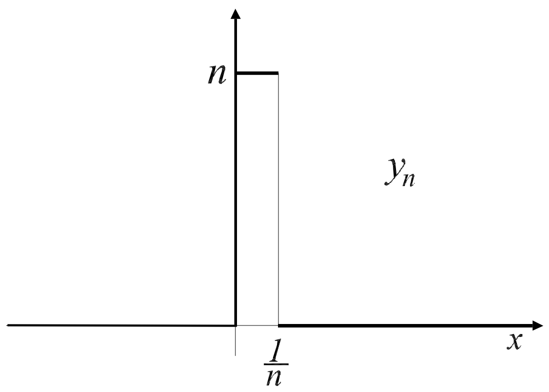
$$\delta(x) = \frac{1}{\pi} \lim_{n\to\infty} \frac{\sin nx}{x}, \qquad x \in \mathbf{R}. \tag{4}$$

Heaviside found

$$\delta(x) = \frac{1}{\pi}\int_0^{\infty} \cos sx \,\mathrm{d}s. \tag{4'}$$

In these formulae, which are non-sense from the viewpoint of classical analysis, δ denotes an idealization of the following function for $n = \infty$:

(5)



$$y_n(x) := \begin{cases} n : 0 < x < \dfrac{1}{n} \\ 0 : \text{else} \end{cases}$$

It was P. A. M. Dirac who gave it the name δ (for analogy with Kronecker's δ) and who gave a "warrant of apprehension" of this strange object (see "The Principles of Quantum Mechanics", Oxford 1930):

$$\int_{-\infty}^{\infty} \delta(x)\,\mathrm{d}x = 1, \qquad \delta(x) = 0 \text{ for } x \neq 0, \qquad \int_{-\infty}^{\infty} \phi(x)\delta(x)\,\mathrm{d}x = \phi(0) \tag{6}$$

$$\delta(-x) = \delta(x), \qquad x\delta(x) = 0, \qquad \delta(ax) = a^{-1}\delta(x) \quad (a > 0) \qquad \text{etc.} \tag{7}$$

$$\delta = \frac{\mathrm{d}H}{\mathrm{d}x} \qquad \text{where } H(x) := \begin{cases} 0 : x < 0 \\ 1 : x \geq 0 \end{cases} \tag{8}$$

is Heaviside's function.

After a first attempt of S. L. Sobolev in 1936, L. Schwartz finally showed the way how to accomplish both: to define differentiation for all functions and to define δ so as to give a meaning to (3)–(8) and similar identities (in his paper "Généralisation de la notion de fonction, de dérivation, de transformation de Fourier, et applications mathématiques et physiques", Ann. Univ. Grenoble **21** (1945), 57–74). His idea, though simple, is revolutionary: Define δ not as a function of x (which is impossible), but as a functional which assigns to every "reasonable" function ϕ the value $\phi(0)$, i.e.,

$$\delta : \{\text{reasonable functions } \phi(x) \text{ on } \mathbf{R}^1\} \longrightarrow \mathbf{C} : \phi \longmapsto \phi(0),$$

Note that

a) as "reasonable" functions one could take the functions which are finite at 0,

b) the functional δ is linear,

c) this definition is in accordance with the last equation in (6).

Generally, a distribution T is a linear functional (with some continuity property)

$$T : \mathcal{D} \longrightarrow \mathbf{C} : \phi \longmapsto T(\phi)$$

on the space of "test functions" $\mathcal{D}$, which consists of all infinitely differentiable functions with compact support. In choosing as space of "reasonable functions" the rather small space $\mathcal{D}$, we have more possibilities to define distributions $T \in \mathcal{D}'$. For $T(\phi)$ one usually writes $\langle \phi, T \rangle$.

Measures μ (to be precise, Borel measures, which are finite on every compact set) can immediately be interpreted as distributions:

$$\mu : \mathcal{D} \longrightarrow \mathbf{C} : \phi \longmapsto \int \phi(x)\, \mathrm{d}\mu(x).$$

Especially, a function (to be precise, a locally integrable one) is identified with the measure $f(x)\,\mathrm{d}x$, $\mathrm{d}x =$ Lebesgue's measure, such that

$$f : \mathcal{D} \longrightarrow \mathbf{C} : \phi \longmapsto \int \phi(x) f(x)\, \mathrm{d}x.$$

(In analogy with this, $\langle \phi, T \rangle$ is sometimes written as $\int \phi(x)T(x)\,\mathrm{d}x$, which must not be interpreted as an integral (since $T(x)$ does not make sense), but just as a symbolic expression for $\langle \phi, T \rangle$.) In this way, functions and measures are identified with special distributions. This approach was well known at the time to measure theorists (Stieltjes, F. Riesz, Lebesgue). However, they did not go further and define differentiation so as to obtain a connexion to classical analysis.

In order that differentiation coincides with the usual notion for differentiable functions, one has to define

$$\frac{\partial T}{\partial x_j} : \mathcal{D} \longrightarrow \mathbf{C} : \phi \longmapsto -\langle \frac{\partial \phi}{\partial x_j}, T \rangle.$$

(Indeed, if $\left(\frac{\partial}{\partial x_j}\right)_{\text{new}}$ and $\left(\frac{\partial}{\partial x_j}\right)_{\text{old}}$ denote the new and old definitions, respectively, then, for a differentiable function f, we have:

$$\begin{aligned}\langle \phi, \left(\frac{\partial}{\partial x_j}\right)_{\text{new}} f \rangle &= -\langle \frac{\partial \phi}{\partial x_j}, f \rangle = -\int \frac{\partial \phi}{\partial x_j}(x) f(x)\, \mathrm{d}x \\ &= \text{(partial integration)} = \int \phi(x) \frac{\partial f}{\partial x_j}(x)\, \mathrm{d}x = \langle \phi, \left(\frac{\partial}{\partial x_j}\right)_{\text{old}} f \rangle.)\end{aligned}$$

Now it can easily be verified that (1) is satisfied in the distributional sense for any z given by (2) with f, g being "arbitrary" functions, and that (8) holds true.
(In the co-ordinates u, v, (1) is equivalent with $\frac{\partial}{\partial v} f(u) = 0$, which follows from

$$\langle \phi(u,v), \frac{\partial}{\partial v} f(u)\rangle = -\langle \frac{\partial \phi}{\partial v}, f(u)\rangle = -\iint \frac{\partial \phi}{\partial v}(u,v) \mathrm{d}v f(u)\, \mathrm{d}u = -\int 0 \cdot f(u)\, \mathrm{d}u = 0,$$

whereas (8) is proven by

$$\langle \phi, \frac{\mathrm{d}H}{\mathrm{d}x}\rangle = -\langle \phi', H\rangle = -\int_0^\infty \phi'(x)\, \mathrm{d}x = -\ \phi(x)|_{x=0}^{\infty} = \phi(0) = \langle \phi, \delta\rangle.)$$

Note that $\mathcal{D}'$ includes distributions not expressible as measures, such as δ'. Therefore, the invention of $\mathcal{D}'$ is in some sense comparable to the transition from the set of natural numbers to that of all integers. In this analogy, 1 would correspond to continuous functions which are not continuously differentiable, 0 to derivatives of functions in class 1 (e.g., H), -1 to derivatives of functions resp. measures in class 0 (e.g., δ), and so on.

Finally, to verify identities such as (3), (4), or $\delta = \lim_{n\to\infty} y_n$ as in (5), the concept of pointwise (so-called "weak") convergence is introduced in $\mathcal{D}'$:

$$T_n \to T \text{ in } \mathcal{D}' :\Longleftrightarrow \forall \phi \in \mathcal{D} : \langle \phi, T_n\rangle \to \langle \phi, T\rangle \text{ in } \mathbf{C}.$$

Hence (4) would mean

$$\lim_{n\to\infty} \int_{-\infty}^{\infty} \phi(x) \frac{\sin nx}{x}\, \mathrm{d}x = \pi\phi(0), \qquad \phi \in \mathcal{D},$$

and is, in this form, known to Fourier analysts as Riemann's localization principle. (5) is easy to prove directly:

$$\lim_{n\to\infty} \langle \phi, y_n\rangle = \lim_{n\to\infty} \int_0^{1/n} \phi(x)\, \mathrm{d}x \Big/ \frac{1}{n} = \phi(0) = \langle \phi, \delta\rangle,$$

by the fundamental theorem of calculus.

3. Fundamental Solutions

In order to state at least one theorem, to give one non-trivial calculation, and to give a connexion with my own work, I shall speak about fundamental solutions. If $P(\partial)$ denotes a linear partial differential operator, then a distribution E is called a *fundamental solution* (FS) of $P(\partial)$ if $P(\partial)E = \delta$. If $P(\partial)$ models a physical system, then E can be interpreted as the answer of this system to a unit disturbance at the origin. The notion of FS is therefore of paramount interest in physics.

Theorem. *(Malgrange–Ehrenpreis, 1954) Every non-vanishing linear partial differential operator with constant coefficients has a fundamental solution in* $\mathcal{D}'$.

For a proof see [4], Theorem 7.3.10, p. 189 (see also the historical notes on p. 249). This theorem is one of the reasons that distribution theory is now world-wide accepted as the appropriate "language" for the theory of *linear* PDE-s. For non-linear PDE-s, however, the impossibility of generally defining a multiplication of distributions lead to different approaches (comp. M. Oberguggenberger's book "Multiplication of distributions and applications to PDE-s").

As an illustration of the notion of FS, let us consider the following table of FS-s ($r := \sqrt{x^2+y^2+z^2}$) :

$$\begin{array}{lccc} \text{a)} & \Delta_3 := \frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2} & : & \frac{-1}{4\pi r} \\ \text{b)} & \frac{\partial}{\partial t}-\Delta_3 & : & \frac{H(t)}{(4\pi t)^{3/2}}\,\mathrm{e}^{-r^2/4t} \\ \text{c)} & \frac{\partial^2}{\partial t^2}-\Delta_3 & : & \frac{1}{4\pi t}\,\delta(t-r). \end{array}$$

Note that a FS is not unique unless additional conditions are imposed (e.g., for a) $\lim_{|x|\to\infty} E(x) = 0$, for b) and c) $\mathrm{e}^{-t}E \in \mathcal{S}'$).

As an example, let us check a), which, in a physical context, expresses the fact that the electrostatic potential of a point charge in $\mathbf{R}^3$ is given by $-\frac{1}{4\pi r}$. Usually, this is derived with the help of Green's formula; in contrast, we shall regularize $\frac{1}{r}$ and use the fact that, in opposition to classical calculus, limits and differentiation can be interchanged in $\mathcal{D}'$.

$$\begin{aligned}\Delta_3\frac{1}{r} = \Delta_3 \lim_{n\to\infty}\frac{1-\mathrm{e}^{-nr}}{r} &= \lim_{n\to\infty}\Delta_3\frac{1-\mathrm{e}^{-nr}}{r} \\ &= \lim_{n\to\infty}\Big(\partial_r+\frac{2}{r}\Big)\partial_r\frac{1-\mathrm{e}^{-nr}}{r} = \lim_{n\to\infty}\frac{-n^2\,\mathrm{e}^{-nr}}{r};\end{aligned}$$

application to a test function ϕ yields:

$$\begin{aligned}&\langle\phi,\Delta_3\frac{1}{r}\rangle = -\lim_{n\to\infty} n^2\int\phi(x)\frac{\mathrm{e}^{-nr}}{r}\,\mathrm{d}x = -\lim_{n\to\infty}n^2\iint_{|\omega|=1}\mathrm{d}\omega\int_{r=0}^{\infty}\phi(r\omega)\,\mathrm{e}^{-nr}r\,\mathrm{d}r \\ &= (\text{setting } nr=u)\ -\lim_{n\to\infty}\iint_{|\omega|=1}\mathrm{d}\omega\int_{u=0}^{\infty}\phi(u\omega/n)\,\mathrm{e}^{-u}\,\mathrm{d}u \\ &= (\text{by Lebesgue's theorem})\ -\phi(0)\cdot 4\pi\cdot\int_0^\infty \mathrm{e}^{-u}\,\mathrm{d}u = \langle\phi,-4\pi\delta\rangle.\end{aligned}$$

The Malgrange–Ehrenpreis theorem asserts that every linear partial differential operator with constant coefficients has a FS. Though there exist constructive proofs of this theorem—extending some work of H. König I recently found a new one: For $P(\partial)\in\mathbb{C}[\partial_1,\ldots,\partial_n]$,

$$E(x) := \frac{1}{\overline{P_m(\eta)}}\int_{\lambda\in\mathbb{C},\,|\lambda|=1}\lambda^m\mathrm{e}^{\lambda\eta x}\mathcal{F}^{-1}_{\xi\to x}\left(\frac{\overline{P(\mathrm{i}\xi+\lambda\eta)}}{P(\mathrm{i}\xi+\lambda\eta)}\right)\frac{\mathrm{d}\lambda}{2\pi\mathrm{i}\lambda},$$

is a FS of $P(\partial)$ if $\deg P = m$, $\deg(P - P_m) < m$, $\eta \in \mathbb{R}^n$, $P_m(\eta) \neq 0$, $\mathcal{F} =$ Fourier transform, (cf. "A short proof of the Malgrange–Ehrenpreis theorem", joint work with N. Ortner, Proceedings of a Workshop on Functional Analysis, Trier 1994, ed. by S. Dierolf, to appear)—they do not furnish representations that are suitable from the practical viewpoint. Whereas the FS of the classical operators of mathematical physics (as e.g. a), b), c)) are known long since, this is not the case for more complicated operators. This is the main field of my own (rather small) contribution to mathematics.

4. Guide to Textbooks on Distributions

The standard reference is of course Schwartz' book [7], which, in its first version, appeared already in 1950/51 and had an enormous impact on the development of the theory of linear partial differential equations. The relation with this subject is stressed in Hörmander's old book [3], which contains only a brief summary of distribution theory. A greatly expanded treatment of this is given in [4] (which is the first of four volumes on linear partial differential operators).

The beginner is probably better off if he starts with an introductory text such as [8], [9], [10], [11] or the first volume of Gel'fand's treatise [1]. The second volume of this series treats the topological foundations of distribution theory (namely the theory of locally convex spaces), which is also done, from a more abstract viewpoint, in Horváth's book [5]. The third one is devoted to applications in differential equations and is similar in its aim to the quite unknown, but excellent textbook [9] of Shilov. The fourth volume presents the theory of nuclear spaces and gives a good account of Schwartz' kernel theorem. Moreover it contains some applications to random processes (Minlos' theorem). The fifth volume treats applications in the representation theory of Lie groups (especially $\mathrm{Sl}_2(\mathbf{C})$).

Zuily's book [12], which contains 100 non-trivial problems (with solutions) in distribution theory is also highly recommended to the beginner. The emphasis of [2] is on Fourier analysis. I enjoyed reading it because of the author's elegant style and because of its broad perspective. Last not least, I mention de Rham's classical and fascinating book [6] which thoroughly treats distributions on differentiable manifolds ("currents"), its author having been very much involved in the creation of the theory of distributions.

References

1. Gel'fand, I. M., Shilov, G. E., Graev, M. I. and N. Ya. Vilenkin, *Generalized Functions. Vol. I–V*, Academic Press, New York 1964.
2. Donoghue, W. F., Jr., *Distributions and Fourier Transforms*, Academic Press, New York, 1969.
3. Hörmander, L., *Linear Partial Differential Operators*, Grundlehren, 116, 3rd ed., Springer, Berlin, 1969.
4. Hörmander, L., *The Analysis of Linear Partial Differential Operators. Vol. I. Distribution Theory and Fourier Analysis*, Grundlehren, 256, 2nd ed., Springer, Berlin, 1990.
5. Horváth, J., *Topological Vector Spaces and Distributions. Vol. I*, Addison-Wesley, Reading, Mass., 1966.
6. Rham (de), G., *Variétés différentiables. Formes, courants, formes harmoniques*, (1ère éd.: Act. sc. ind. 1222: 1955) 3ème éd., Hermann, Paris, 1973.

7. Schwartz, L., *Théorie des distributions*, (1ère èd.: Act. Sc. Ind. 1091, 1122: 1950/51) Nouv. éd., Hermann, Paris, 1966.
8. Schwartz, L., *Méthodes mathématiques pour les sciences physiques*, 2nde éd., Hermann, Paris, 1965.
9. Shilov, G. E., *Generalized Functions and Partial Differential Equations*, Gordon and Breach, New York, 1968.
10. Vladimirov, V. S., *Equations of Mathematical Physics*, Dekker, New York, 1971.
11. Vladimirov, V. S., *Generalized Functions in Mathematical Physics*, MIR, Moscow, 1979.
12. Zuily, C., *Problèmes de distributions*, Hermann, Paris, 1978.

Date: 6. 10. 1995