0 PRELUDE

0.1. General Intro

In many mathematical applications partial differential equations (PDE) play a main role. The fundamental laws of physics are formulated as PDE and frequently modelling in all of science results in formulating the dependencies found as a PDE. However, in many of these cases it is not really mandatory to write these dependencies in the form of a PDE; it is just a mathematically very elegant and convenient way to do so.
Continuous functions in many cases provide a sensible model of physical quantities, but the do not possess a derivative (at least within classical calculus). Also physical laws are often more general than their formulation as PDE. To cure these obvious drawback there are in general two possible options:

(i) Reformulate the respective physical law using integral equations, which also make perfect mathematical sense for continuous (or more general classes of) functions.

(ii) Keep the formulation as PDE but generalise the notion of derivative in such a way that also continuous (or more general classes of) functions possess derivatives.

The second possibility leads directly to the *theory of distributions* also called *generalised functions*.

However, in addition to the motivation from applications just given, there is also a clear interest from pure mathematics to define a notion of differentiation of classically non-differentiable functions. As a clear historic evidence for this 'wish' to differentiate more general classes of functions we quote Charles Hermite (1822-1901). In a letter to Thomas Stieltjes (1856-94) he wrote:

I recoil with dismay and horror at this lamentable plague of functions which do not have derivatives.

0.2. Some historical remarks[1]

We begin with a few historical remarks on the "prehistory" of distribution theory, the details of which may be found in the book of Lützen [L][2].

Some physicists and also mathematicians were using "generalized functions ideas" thereby for a long time anticipating the later rigorous theory. The names to mention here are most of all J.B. Fourier, G. Kirchhoff and O. Heaviside. P. Dirac in 1926 [D1] and later in his famous book [D2] introduced the concept of the δ-function which allowed him to draw an analogy between "discrete" and "continuous variables," thereby reaching a unified theory of quantum mechanics combining matrix mechanics and wave mechanics. Since these ideas were so beautiful and convincing the "Dirac-δ" very soon became a widespread tool for physicists. However, the notation δ does not stand for "Dirac" but was originally chosen by Dirac to put emphasis on the analogy between $\delta(p-q)dp$ and the unity matrix which, in physics, commonly is written as the "Kronecker-δ" δ_{pq} (cf. [L], p. 124). On the other hand, descriptions of the δ-distribution as a limit of a series of (smooth) functions go back as far as 1822 and Fourier [F]. Also Kirchhoff [K] already in 1882 fully captured the concept of δ calculating the fundamental solution of the wave operator in $\mathbb{R}^{1+3}$.

A rigorous theory which first (implicitly) used distributions was given in 1932 by S. Bochner [S] while the first definition of distributions in the modern sense (as functionals) appeared in S. Sobolev's 1936 paper [S]. Hence (according to Lützen [L], p. 159ff.) he may be be called the *inventor* of distributions, while finally L. Schwartz *created the theory* of distributions in his classical monograph [S], first published in 1950.

Schwartz' theory rapidly was well received both by mathematicians and physicists who now could use "improper functions" in a well-defined sense. In mathematics, distribution theory was the essential tool to build an elaborate solution concept for linear PDE. Most notable among the numerous contributions in that field is the famous theorem proved individually by L. Ehrenpreis [E] and B. Malgrange [M] guaranteeing the existence of a fundamental solution (in $\mathcal{D}'$) for every non zero constant coefficient linear partial differential operator. [...]

[B] Bochner, S. *Vorlesungen über Fouriersche Integrale.* Akad. Verlagsges., Leipzig, 1932.

[D1] Dirac, P. A. M. The physical interpretation of the quantum dynamics. *Proc. Roy. Soc. A*, 113:612–641, 1926.

[D2] Dirac, P. A. M. *The Principles of Quantum Mechanics.* Oxford University Press, 3 edition, 1947.

[E] Ehrenpreis, L. Solutions of some problems of division i. *Amer. J. Math.*, 76:883–903, 1954.

[1] taken from R.S., *Distributional Methods in General Relativity*, Sec. 2.1., Ph.D. Thesis, University of Vienna, (2000).

[2] The references cited in these remarks are mainly of historical interest. Therefore they will be listed below and *not* in the bibliography at the end of the notes.

[F] J. B. J. Fourier. *Théorie Analytique de la Chaleur.* (Oeuvers I.) ed. Darboux, Paris, 1882.

[K] Kirchhoff, G. Zur Theorie der Lichstrahlen. *Sitz. K. Preuss. Akad. Wiss.*, pages 641–669, 1882.

[L] Lützen, J. *The Prehistory of the Theory of Distributions.* Springer, New York, 1982.

[M] Malgrange, B. Existence et approximation des solutions des équations aux dérivées partielles et des équations de convolution. *Ann. Inst. Fourier*, 6:271–335, 1955-56.

[S] Schwartz, L. *Théorie des Distributions.* Hermann, Paris, 1966.

[So] Sobolev, S. L. Méthode nouvelle á résoudre le problème de cauchy pour les équations linéaires hyperboliques normale. *Mat. Sb.*, 1(43):39–71, 1936.

0.3. Motivating examples from PDE

(i) *A transport equation*

In this first example we deal with the equation of continuity in a 1-dimensional medium. Our main focus is to exemplify the discussion in 0.1.

Generally speaking a continuity equation in physics is an equation that describes the transport of some quantity. Here we will concentrate on the transport of mass in a 1-dimensional medium—you can think of a 1-dimensional river.

We denote the position (length) in the 1-dimensional medium by x and time by t, see Figure 0.1. Moreover will use the following quantities and notations:

- the mass density $\rho(t, x)$ (with physical units mass per length)
- the mass flow $q(t, x)$ (with physical units mass per time)
- the velocity (of a point particle) $\nu(t, x)$ (with physical units length per time)

We have the following fairly obvious correlation between the mass flow on the one hand and the mass density and velocity on the other hand

$$q = \rho \nu . \tag{0.1}$$

Figure 0.1: A 1-dimensional medium.

We will suppose that $\rho \geqslant 0$ which is physically reasonable and implies that q and ν share their orientation and we fix ν to be positive for growing x.

We next consider a space interval $[x, x_1]$ and a time interval $[t, t_1]$ and formulate the physical law of mass conservation, see Figure 0.2.

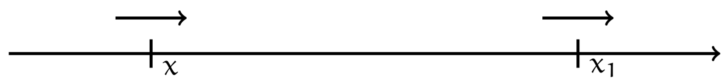

Figure 0.2: Mass conservation in a 1-dimensional medium.

$$
\underbrace{\int_{x}^{x_1} \rho(t_1, y)\, dy}_{\text{total mass in } [x, x_1] \text{ at time } t_1} - \underbrace{\int_{x}^{x_1} \rho(t, y)\, dy}_{\text{total mass in } [x, x_1] \text{ at time } t}
$$

$$
(0.2) \qquad = \underbrace{\int_{t}^{t_1} q(s, x)\, ds}_{\text{ingoing mass in } [t, t_1] \text{ at } x} - \underbrace{\int_{t}^{t_1} q(s, x_1)\, ds.}_{\text{outgoing mass in } [t, t_1] \text{ at } x_1}
$$

Dividing by $t_1 - t$ leads to

$$
(0.3) \qquad \int_{x}^{x_1} \frac{\rho(t_1, y) - \rho(t, y)}{t_1 - t}\, dy = \frac{1}{t - t_1} \int_{t}^{t_1} \big(q(s, x) - q(s, x_1)\big)\, ds.
$$

Note that up to now assuming ρ to be locally integrable was sufficient to justify all the calculations. But now we assume ρ to be a $\mathcal{C}^1$-function to be able to take the limit $t_1 \to t$. Applying the fundamental theorem of calculus gives

$$
(0.4) \qquad \int_{x}^{x_1} \partial_t \rho(t, y)\, dy = q(t, x) - q(t, x_1)\,.
$$

Analogously assuming that $q \in \mathcal{C}^1$ we may also take the limit $x_1 \to x$ after dividing by $x_1 - x$ to obtain

$$
(0.5) \qquad \partial_t \rho(t, x) = -\partial_x q(t, x),
$$

which we will aslo write in the form $q_t + q_x = 0$. If we now suppose that the velocity is a constant, $v = c$ and insert (0.1) we finally obtain the continuity equation in one dimension to be

$$
(0.6) \qquad \rho_t + c\rho_x = 0.
$$

We note that for $\rho \in \mathcal{C}^1$ the PDE (0.6) follows from the integral equation (0.2). However, the integral equation (0.2) also makes sense for more general ρ, e.g. ρ continuous or (locally) integrable. Hence in accordance with the spirit of 0.1 we find that *the physical law of mass-conservation* (0.2) *is strictly more general than the PDE* (0.6)*, which was derived under the additional regularity assumptions* $\rho, q \in \mathcal{C}^1$. Also note the general 'slogan' that PDE are local while integral equations take into account mean values.

The general solution of the PDE (0.6) is given by

$$\rho(t,x) \;=\; a(x-ct) \quad \text{with } a \in \mathcal{C}^1(\mathbb{R}) \tag{0.7}$$

which intuitively amounts to a peak travelling to the right, see Figure 0.3. (To derive (0.7) one can either use the method of characteristics from PDE-theory or transform variables to $\xi = x + ct$, $\eta = x - ct$ and solve the resulting ordinary differential equation (ODE).)

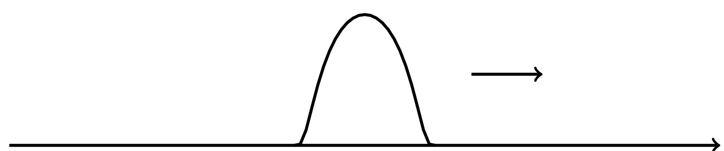

Figure 0.3: A (smooth) peak travelling to the right.

Let now $\varphi \in \mathcal{C}^\infty(\mathbb{R}^2)$ (i.e., φ has continuous partial derivatives of all orders, hence is a so-called *smooth* function) be vanishing outside a bounded set. We denote the vector space of all such functions[3] by $\mathcal{D}(\mathbb{R}^2)$. Then we have for all $\varphi \in \mathcal{D}(\mathbb{R}^2)$

$$\int_{\mathbb{R}^2} \underbrace{\big(\rho_t(t,x) + c\rho_x(t,x)\big)}_{=0 \text{ by } (0.6)} \,\varphi(t,x)\,dt\,dx \;=\; 0. \tag{0.8}$$

Conversely one may show that for any $\rho \in \mathcal{C}^1(\mathbb{R}^2)$ satisfying equation (0.8) for all $\varphi \in \mathcal{D}(\mathbb{R}^2)$, also the PDE (0.6) holds. (Suppose not then there exists (t_0, x_0) such that (0.6) is violated there. By continuity (0.6) is violated in a neighbourhood of (t_0, x_0). Now choose a smooth function $\rho \geqslant 0$ and $\rho \neq 0$ only in this neighbourhood[4]. Then (0.8) does not hold for this particular φ.)
On the other hand we find by integration by parts that (0.8) holds iff (i.e., if and only if)

$$\int_{\mathbb{R}^2} \rho\,(\varphi_t + c\varphi_x)\,dtdx \;=\; 0 \quad \text{for all } \varphi \in \mathcal{D}(\mathbb{R}^2). \tag{0.9}$$

[3]The study of this space will be one of the main objectives of Chapter 1.
[4]The existence of such 'bump functions' will be established in Chapter 1.

Observe that—similar to the case of (0.2)—the latter equation makes sense for more general ρ than the PDE (0.6). We will call any solution of (0.9) a *weak solution* of (0.6). Keep in mind that this means the integral equation holds for all functions $\varphi \in \mathcal{D}(\mathbb{R}^2)$. With this terminology we have for all $\rho \in \mathcal{C}^1(\mathbb{R}^2)$ that

$$\rho \text{ solves (0.6)} \iff \rho \text{ is a weak solution of (0.6).} \tag{0.10}$$

However, there are more weak solutions to (0.6) than there are 'real' solutions to (0.6), called *calssical* solutions in this context. E.g. any weak solution ρ that is merely continuous or (locally) integrable is *not* a classical solution: One cannot even calculate the derivatives ρ_x and ρ_t which enter the PDE (0.6)!
A weak solution graphically amounts to having e.g. a box shaped bump travelling to the right, see Figure 0.4. One may even show that (0.7) with the function a being continuous or (locally) integrable, gives rise to a weak solution of (0.6).

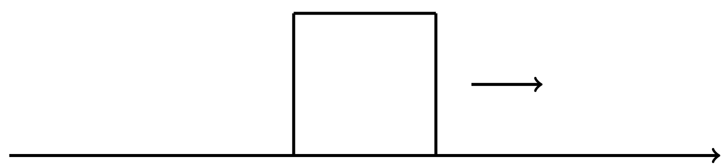

Figure 0.4: A continuous, non-smooth peak travelling to the right.

(ii) *The wave equation*
We next consider the wave equation in $1+1$ dimensions,

$$u_{tt} = u_{xx} \qquad \text{for } u : \mathbb{R}^2 \to \mathbb{R}, \tag{0.11}$$

which models e.g. the vibration of a string and belongs to the trinity[5] of basic linear PDE. It is well-known and shown in every lecture course on PDE that every $\mathcal{C}^2$-solution of (0.11) is of the form

$$u(t,x) = f(x+t) + g(x-t), \tag{0.12}$$

with $f, g \in \mathcal{C}^2(\mathbb{R})$. (Either use the method of characteristics or transform variables as above.)
Weak solutions of (0.11) are defined analogous to the above as solutions of

$$\int_{\mathbb{R}^2} u\,(\varphi_{tt} - \varphi_{xx})\,dt\,dx = 0 \quad \text{for all } \varphi \in \mathcal{D}(\mathbb{R}^2). \tag{0.13}$$

Now in complete analogy to (i) it holds that

[5]The other two members are the Laplace and the heat equations.

(a) Every $\mathcal{C}^2$- (i.e., classical) solution of (0.11) is also a weak solution but there are more weak solutions than classical ones.

(b) Every function of the form (0.12) with f and g continuous or (locally) integrable is a weak solution.

0.4. Point charges in electrostatics—the delta-function

Next we consider a motivating example from physics. In electrostatics the potential V of a point charge Q at a point $x_0 \in \mathbb{R}^3$ is given by

$$V(x) = \frac{Q}{|x - x_0|}, \tag{0.14}$$

see Figure 0.5. (The electric force is then given by $F = -\nabla V$.)

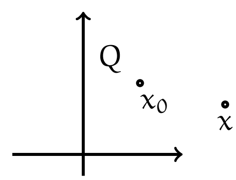

Figure 0.5: A point charge Q at x_0.

Here a point charge is an idealisation where some finite charge Q sits at a single point x_0 is space. On the other hand the potential of a charge density ρ is given by the Newton potential

$$V(x) = \int_{\mathbb{R}^3} \frac{\rho(y)}{|x - y|}\, dy\,, \tag{0.15}$$

where the charge density heuristically is given by

$$\rho(y) = \lim_{r \to 0} \frac{\text{charge contained in } B_r(y)}{\text{volume of } B_r(y)}\,,$$

where $B_r(x)$ denotes the (open) ball of radius r centered at x, c.f. Figure 0.6.

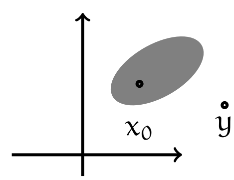

Figure 0.6: A smeared out charge density around a point x_0.

An interesting question is now: What is the charge density of a point charge in x_0. Most intuitively we would like to write

$$\rho(y) = \begin{cases} 0 & y \neq x_0 \\ \lim_{r \to 0} \dfrac{Q}{4/3\,\pi r^3} = \infty & y = x_0\,. \end{cases} \tag{0.16}$$

But this implies that $\rho = 0$ Lebesgue almost everywhere and so $\int \rho(y)dy = 0$. However to obtain the 'correct' potential (0.15) we would need

$$V(x) \;=\; \int_{\mathbb{R}^3} \frac{\rho(y)}{|x-y|}\,dy \;=\; \frac{Q}{|x-x_0|}\,. \tag{0.17}$$

That is we would need a 'function' δ_{x_0} (with $\rho(y) = Q\delta_{x_0}(y)$) satisfying (cf. Figure 0.7)

(i) $\delta_{x_0}(y) = 0$ for all $y \neq x_0$, and

(ii) $\int \delta_{x_0}(y)f(y)\,dy = f(x_0)$, in particular $\int \delta_{x_0}(y)dy = 1$.

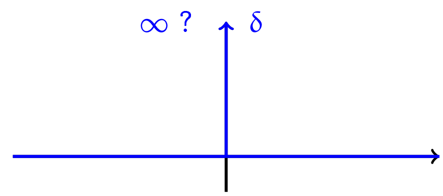

Figure 0.7: The delta-function.

These properties are often also informally formulated as follows: δ_{x_0} vanishes everywhere but at the single point x_0 where it has a value so large that the integral over δ_{x_0} is unity. Clearly such an object cannot be a 'classical function'.
Moreover in electrostatics the Newtonian potential (0.15) is obtained as a solution of the Poisson equation

$$\triangle V(x) \;=\; -4\pi\rho(x)\,. \tag{0.18}$$

Assuming now for simplicity that $x_0 = 0$ and $Q = 1$ and hence $\rho(y) = \delta_0(y) =: \delta(y)$, equation (0.18) takes the form

$$\triangle \underbrace{\left(\int_{\mathbb{R}^3} \frac{\rho(y)}{|x-y|}\,dy\right)}_{\frac{1}{|x|}} \;=\; -4\pi\rho(x)\,, \tag{0.19}$$

and hence

$$\triangle\Big(\frac{1}{|x|}\Big) \;=\; -4\pi\rho(x)\,. \tag{0.20}$$

Summing up we are looking for a theory where the delta-function is a sensible object and moreover equation (0.19) is a valid equation. Indeed distribution theory provides all those features as we will preview next.

0.5. Preview

It will turn out that a further development of the notion of weak solution (see 0.3) also makes the issues of 0.4 tractable. The basic features of such a theory might already be guessed from 0.3. Instead of considering a function f as classical mappings that take points to points, i.e.,

$$f:\ \mathbb{R}^n \to \mathbb{C}, \quad x \mapsto f(x) \tag{0.21}$$

we consider the mapping

$$f:\ \mathcal{D}(\mathbb{R}^n) \to \mathbb{C}, \quad \varphi \mapsto \int\limits_{\mathbb{R}^n} f(x)\varphi(x)dx\,. \tag{0.22}$$

Equation (0.22) defines by the properties of the integral a *linear map* from the space of so-called *test functions* into the complex numbers, hence a *linear form on* $\mathcal{D}(\mathbb{R}^n)$. Driven by 0.4 we also allow more general maps of that kind, e.g.

$$\delta:\ \mathcal{D}(\mathbb{R}^n) \to \mathbb{C}, \quad \varphi \mapsto \langle \delta, \varphi \rangle = \varphi(0)\,. \tag{0.23}$$

Clearly this assignment defines a linear form on $\mathcal{D}(\mathbb{R}^n)$.
Using the notation $\langle f, \varphi \rangle$ for the integral in (0.22) (one says "f acts distributionally on φ".) we have for all $f \in \mathcal{C}^1(\mathbb{R}^n)$ (by integration by parts and recalling that φ vanishes outside a bounded set)

$$\langle \partial_i f, \varphi \rangle \ = \int\limits_{\mathbb{R}^n} \partial_i f(x)\varphi(x)\,dx \ = \ -\int\limits_{\mathbb{R}^n} f(x)\partial_i\varphi(x)\,dx \ = \ -\langle u, \partial_i\varphi \rangle\,. \tag{0.24}$$

Modelled on that equation we define more generally for any linear form u on $\mathcal{D}(\mathbb{R}^n)$

$$\langle \partial_i u, \varphi \rangle \ := \ -\langle u, \partial_i \varphi \rangle\,. \tag{0.25}$$

Using this trick many classically non-differentiable functions become differentiable (in this sense). As an example we consider the *Heaviside function* also sometimes called *step function*

$$H(x) := \begin{cases} 0 & \text{if} \quad x < 0 \\ 1 & \text{if} \quad x > 0\,, \end{cases} \tag{0.26}$$

see Figure 0.8. (Note that H is only defined almost everywhere, i.e., outside $x = 0$.)

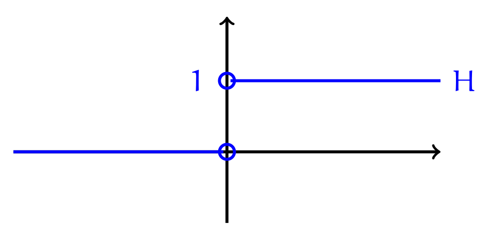

Figure 0.8: The Heaviside step function.

Its distributional action is given by

$$(0.27) \qquad \langle H, \varphi \rangle = \int_{-\infty}^{\infty} H(x)\varphi(x)\,dx = \int_{0}^{\infty} \varphi(x)\,dx\,.$$

And we have for its derivative

$$(0.28) \qquad \langle H', \varphi \rangle = -\langle H, \varphi' \rangle = -\int_{0}^{\infty} \varphi'(x)\,dx = \varphi(0) = \langle \delta, \varphi \rangle\,,$$

and hence $H' = \delta$, i.e., the derivative of the step function is the delta-function.

Finishing this preview we mention that distributions (generalised functions) will be defined in Chapter 1 as *continuous* linear functionals on the space of test functions $\mathcal{D}$, where a suitable notion of continuity will have to be supplied. We will then see throughout this course that persistently using the ideas outlined here, one can develop a theory which can answer satisfactorily all the problems and questions raised in 0.3 and 0.4 and a whole lot more issues in applications e.g. in physics and in PDE. In particular, it will provide a very powerful set of tools to study linear PDE.