Introduction to PDEs

YRD 2025, Domaine de la Tour

Lucas Davron

These notes are the transcription of a presentation given on the whiteboard at the occasion of the 2025's YRD. It is meant to be a pedestrian and biased introduction to the field of Partial Differential Equations (PDEs). Relevent references are: the survey article of Brezis and Browder [1], which gives a thorough and accessible account for the knowledge on PDEs by the 20th century; the book "Partial differential equations" by Evans [2], which is commonly accepted as the bible of PDEs for beginners or even working mathematicians.

1 Historical perspective

The study of PDEs began in the 18th century, motivated by the study of continuous media mechanics (see [1, §2]). These equations were derived by non rigorous heuristics, and then used to perform other non rigorous heuristics. The 3 paradigmatic PDEs are given in the following tabular.

Name	Equation	Associated person	Date
Wave equation	$\frac{\partial^2 w}{\partial t^2} = \frac{\partial^2 w}{\partial x^2}$	d'Alembert	~ 1752
Laplace equation	$\Delta u = 0$	Laplace	~ 1780
Heat equation	$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}$	Fourier	~ 1810

Most of the PDEs arising from physics can be obtained from the least action principle. We give a heuristic to do so, in the particular case the system does not depend on time (this can be thought of as an equilibrium configuration). For suppose we are given a Lagrangian

$$\mathcal{L} = \mathcal{L}(x, u, \xi) : \Omega \times \mathbb{R} \times \mathbb{R}^d \to \mathbb{R},$$

where $\Omega \subset \mathbb{R}^d$ is an open bounded subset of $\mathbb{R}^d$. We define the action functional

$$\mathcal{A}(u) := \int_\Omega \mathcal{L}(x, u(x), \nabla u(x)) dx,$$

for a function $u : \Omega \to \mathbb{R}$ (this is only formal!). The function u is seen as a configuration of a physical system, the laws of the physics are encoded in $\mathcal{L}$, and the quantity $\mathcal{A}(u)$ represents the energy associated to the state u. The least action principle then says: in reality we should observe a configuration u which minimizes the action $\mathcal{A}(u)$. Let us assume that u minimizes the action $\mathcal{A}$, pick $\varphi \in C_c^\infty(\Omega)$ and observe that the function

$$t \mapsto \mathcal{A}(u + t\varphi),$$

is well defined for $t \in \mathbb{R}$ and has a minimum at $t = 0$. From the first order rules for optimality, we therefore have

$$\begin{aligned}
0 &= \frac{d}{dt}\bigg|_{t=0} \mathcal{A}(u + t\varphi) \\
&= \frac{d}{dt}\bigg|_{t=0} \int_\Omega \mathcal{L}(x, u(x) + t\varphi(x), \nabla u(x) + t\nabla\varphi(x))dx \\
&= \int_\Omega \frac{d}{dt}\bigg|_{t=0} \mathcal{L}(x, u(x) + t\varphi(x), \nabla u(x) + t\nabla\varphi(x))dx \\
&= \int_\Omega \left\{ \frac{\partial \mathcal{L}}{\partial u}(x, u(x), \nabla u(x))\varphi(x) + \frac{\partial \mathcal{L}}{\partial \xi}(x, u(x), \nabla u(x)) \cdot \nabla\varphi(x) \right\} dx.
\end{aligned}$$

To proceed with the computations, we will use integration by parts to put the ∇ in the term $\nabla\varphi$, to the term $\partial_\xi \mathcal{L}$, leaving φ as a factor. This is done as follows: we first recall the divergence formula (called Stoke's formula in differential geometry).

Lemma 1. *Assume that Ω has Lipschitz boundary, let $G \in C^1(\overline{\Omega}; \mathbb{R}^d)$ and denote by ν the unit, outward pointing, normal vector field of $\partial\Omega$. Then*

$$\int_\Omega \operatorname{div} G(x)dx = \int_{\partial\Omega} G(x) \cdot \nu(x)d\sigma(x),$$

where σ stands for the surface measure on $\partial\Omega$ and

$$\operatorname{div} G(x) = \sum_{k=1}^d \frac{\partial G_k}{\partial x_k}(x).$$

In particular, if $G(x) = 0$ for x on the boundary $\partial\Omega$, then the above integrals vanish. Therefore, if $G \in C^1(\overline{\Omega}; \mathbb{R}^d)$, $g \in C^1(\overline{\Omega})$ vanishes on the boundary, then

$$0 = \int_\Omega \operatorname{div}(gG)(x)dx = \int_\Omega \nabla g(x) \cdot G(x)dx + \int_\Omega g(x) \operatorname{div} G(x)dx.$$

Coming back to our original computation, we find

$$0 = \int_\Omega \left\{ \frac{\partial \mathcal{L}}{\partial u}(x, u(x), \nabla u(x)) - \operatorname{div}\left[\frac{\partial \mathcal{L}}{\partial \xi}(\cdot, u(\cdot), \nabla u(\cdot))\right](x) \right\} \varphi(x)dx,$$

and this for all $\varphi \in C_c^\infty(\Omega)$. This should yield that

$$\forall x \in \Omega, \quad \frac{\partial \mathcal{L}}{\partial u}(x, u(x), \nabla u(x)) = \operatorname{div}\left[\frac{\partial \mathcal{L}}{\partial \xi}(\cdot, u(\cdot), \nabla u(\cdot))\right](x).$$

These are the so-called Euler-Lagrange equations, and the above computations give a heuristic on how to go from a Lagrangian $\mathcal{L}$ to a PDE.

Example 2. *If*

$$\mathcal{L} = \mathcal{L}(\xi) = \frac{1}{2}|\xi|^2,$$

we find Laplace equation

$$\Delta u = 0.$$

Example 3. *A slightly more sofisticated example in when*

$$\mathcal{L} = \mathcal{L}(\xi) = \sqrt{1 + |\xi|^2},$$

in which cas we find

$$\operatorname{div}\left(\frac{\nabla u}{\sqrt{1 + |\nabla u|^2}}\right) = 0.$$

This Lagrangian is of particular interest because the action functional

$$\mathcal{A}(u) = \int_\Omega \sqrt{1 + |\nabla u|^2} dx = \mathcal{H}^d(\operatorname{graph} u) = \textit{area of } \operatorname{graph} u.$$

Minimizing the action functional then amounts to find minimal surfaces.

The above computations are powerful heuristics, which allow us to make connection and find meaningful mathematical problems. A releveant question is thus to make such computations rigorous. In particular one should

- Specify what is the regularity of $\mathcal{L}$ and u.
- Prescribe boundary data for u.
- Define what it mean for u to solve the Euler-Lagrange equations.

2 Well-posedness

2.1 Hadamard well-posedness

An important topic in the study of PDEs are their well-posedness. We cannot give a general definition of what it mean for a PDE to be well-posed, as the definition should be tailored to each class of problems. For instance the definition should be different if the PDE has the time as a variable (as for the wave or heat equation) or not (as in the Laplace equation). To circumvent this difficulty we will borrow the framework of inverse problems to only give an analogy. Suppose we are interested in solving the equation

$$F(u) = y, \quad y \in Y \text{ given}, \quad u \in X \text{ to find}. \tag{1}$$

Definition 4 (Hadamard well-posedness)**.** *The problem* (1) *is well-posed if*

- *For all $y \in Y$, there is some $u \in X$ solving the equation.*
- *For fixed $y \in Y$, the solution $u \in X$ is unique, call it $u[y]$.*
- *The "inverse" mapping $y \mapsto u[y]$ is continuous $Y \to X$.*

For PDEs the main points to have in mind are

- Existence or uniqueness are never trivial, one should establish them.
- The equation $F(u) = y$ can be understood in a weak sense, which allows one to take X and Y larger, so that existence of a solution is easier (but the uniqueness is harder...)

- One should specify the topologies of X and Y.
- For time dependent problems of the form

$$\frac{\partial u}{\partial t} = Lu, \quad u(0) = u^0,$$

it is customary to work with $X = C([0,T];\mathcal{X})$ so that the third point of the above definition becomes

$$\exists C = C(T) > 0, \quad \forall u^0 \in \mathcal{X}, \quad \forall t \in [0,T], \quad \|u(t)\|_{\mathcal{X}} \leq C\|u^0\|_{\mathcal{X}}.$$

This is often referred to as "continuous dependency of the solution with respect to the initial data".

2.2 A counter-example by Tychonoff

To exemplify the subtelty of the above definition, we will work out a counter-example due to Tychonoff [3]. We introduce the heat equation on the whole real line

$$\left\{ \begin{array}{rcl} \partial_t u &=& \partial_{xx} u, \quad t > 0, \quad x \in \mathbb{R}, \\ u(0,x) &=& u^0(x). \end{array} \right. \tag{2}$$

Let us agree on the following concept of "clasical" solutions.

Definition 5. *Given $u^0 \in C^0(\mathbb{R})$, a solution of* (2) *is $u \in C^2((0,\infty)\times\mathbb{R}) \cap C^0([0,\infty)\times\mathbb{R})$ which satisfies* (2) *pointwise.*

Let us stress that this concept of solution is natural as it seems to be the minimal one so that (2) makes sense. We claim that with this definition, the problem (2) is ill-posed because its solutions are not unique. We first observe that taking advantage of the linearity of the equation, and the so-called superposition principle, uniqueness of the solutions of (2) simply means that: when $u^0 = 0$ then any solution is identically 0. We will construct a family of non-zero solutions of (2) when $u^0 = 0$. To do so the idea is actually very simple: we search a solution which is a power series of x. Let us perform *formal* computations: we *assume* that u is a solution of the form

$$u(t,x) = \sum_{k=0}^{\infty} a_k(t) \frac{x^k}{k!},$$

and we find relations beween the $a_k's$. At least formally we have

$$\begin{aligned} 0 &= (\partial_t - \partial_{xx})\, u \\ &= (\partial_t - \partial_{xx}) \sum_{k=0}^{\infty} a_k(t) \frac{x^k}{k!} \\ &= \sum_{k=0}^{\infty} (\partial_t - \partial_{xx})\, a_k(t) \frac{x^k}{k!} \\ &= \sum_{k=0}^{\infty} \left(a_k'(t) \frac{x^k}{k!} - 1_{k\geq 2} a_k(t) \frac{x^{k-2}}{(k-2)!} \right) \\ &= \sum_{k=0}^{\infty} (a_k'(t) - a_{k+2}(t)) \frac{x^k}{k!}, \end{aligned}$$

hence

$$\forall t > 0, \quad \forall k \in \mathbb{N}, \quad a_k'(t) = a_{k+2}(t).$$

We therefore find, still formally,

$$u(t,x) = \sum_{k=0}^{\infty} a_0^{(k)}(t)\frac{x^{2k}}{(2k)!} + \sum_{k=0}^{\infty} a_1^{(k)}(t)\frac{x^{2k+1}}{(2k+1)!}. \tag{3}$$

Now we can reverse the computations, and start from the coefficients a_0 and a_1 to build a candidate solution. To do so we first understand what are appropriate conditions for the above series to converge pointwisely. Fix $a_0 \in C^\infty[0,\infty)$ and $t \geq 0$, consider the series

$$\sum_{k=0}^{\infty} a_0^{(k)}(t)\frac{x^{2k}}{(2k)!}.$$

It is a power series of x, so that the Cauchy-Hadamard rule brings that its radius of convergence $R(t)$ satisfies

$$\frac{1}{R(t)} = \limsup_{k\infty} \left|\frac{a_0^{(k)}(t)|}{(2k)!}\right|^{\frac{1}{2k}}.$$

We wish that the solution u is defined for all real x, hence we wish for $R(t) = \infty$, hence

$$\limsup_{k\infty} \left|\frac{a_0^{(k)}(t)|}{(2k)!}\right|^{\frac{1}{2k}} = 0.$$

This is equivalent to

$$\forall \epsilon > 0, \quad \exists C > 0, \quad k \in \mathbb{N}, \quad |a_0^{(k)}(t)| \leq C(2k)!\epsilon^{2k},$$

which motivates to impose the somewhat stronger condition

$$\forall T < \infty, \quad \forall R > 1, \quad \exists C > 0, \quad \forall k \in \mathbb{N}, \quad \forall t \in [0,T], \qquad |a_0^{(k)}(t)| \leq C\frac{(2k)!}{R^{2k}} \tag{4}$$

It is elementary but lenghty to show the following.

Theorem 6. *Assume that $a_0 \in C^\infty[0,\infty)$ satisfies the estimate (4), and that $a_1 \in C^\infty[0,\infty)$ satisfies the estimate (4) with $2k+1$ in place of $2k$. Then the candidate (3) is defined by an absolutely converging series, for all $(t,x) \in [0,\infty)\times\mathbb{R}$. It defines a function of class $C^\infty([0,\infty)\times\mathbb{R})$ which is a solution of (2) for*

$$u^0(x) := \sum_{k=0}^{\infty} a_0^{(k)}(0)\frac{x^{2k}}{(2k)!} + \sum_{k=0}^{\infty} a_1^{(k)}(0)\frac{x^{2k+1}}{(2k+1)!}$$

The estimate (4) is understood as follows: the successive derivatives of a_0 are allowed to grow to $+\infty$, but not too fast. The scale $(2k)!$ dominates and dictates the order of the growth, while the parameter R gives a refinment of the size of the growth. Such an estimate is commonly referred to as a Gevrey type estimate. Observe that if we replace the terms $2k$ by k we obtain an estimate of the form

$$|f^{(k)}(t)| \leq C \frac{k!}{R^k},$$

where the right-hand side grows more slowly, hence f is more regular. In fact such an f is analytic, so that Gevrey functions can be thought of as an intermediate between analytic and C^∞ functions. This is fortunate for us because in order to impose $u(0, \cdot) = 0$ we will ask that a_0 satisfies in addition

$$\forall k \in \mathbb{N}, \quad a_0^{(k)}(0) = 0,$$

and similarly for a_1. This is possible for an analytic function only if a_0 is constant to 0. For a Gevrey function this is possible even though a_0 is non constant. The proto-typical example is $f(t) = e^{-1/t}$ which satisfies

$$f^{(k)}(0) = 0, \quad |f^{(k)}(t)| \leq C(2k)!2^k.$$

This f do not satisfy (4) because the above only allow $R = 1/\sqrt{2}$ (one can show it is sharp). We modify this f by considering

$$f(t) = \exp\left(-\frac{1}{t^\gamma}\right), \quad \gamma > 0,$$

which satisfies

$$\exists R, C > 0, \quad \forall t \geq 0, \quad \forall k \in \mathbb{N}, \quad |f^{(k)}(t)| \leq C \frac{(k!)^s}{R^k},$$

where $s = 1 + 1/\gamma$. Taking $\gamma > 1$ we find $s < 2$ so that such an f satisfies the Gevrey estimate (4), and all its derivative at $t = 0$ vanish.

We then take $a_0 = f$ as above and $a_1 = 0$, we find that u defined by (3) solves (2) with $u^0 = 0$ but is not constant to 0. Hence (2) do not possess a unique solution.

2.3 Conclusive remarks

In order to make (2) well-posed we usually require the function u to be of moderate growth with respect to x. In [3] Tychonoff shows that

$$\exists C > 0, \quad \forall t \geq 0, \quad \forall x \in \mathbb{R}, \quad |u(t,x)| \leq Ce^{Cx^2},$$

is enough (and sharp in some sense) to obtain uniqueness. In practice it is more convenient to require that

$$u(t, \cdot) \in L^p(\mathbb{R}),$$

for some $1 \leq p < \infty$.

References

[1] Haïm Brezis and Felix Browder. "Partial differential equations in the 20th century". In: *Adv. Math.* 135.1 (1998), pp. 76–144. ISSN: 0001-8708,1090-2082. DOI: `10.1006/aima.1997.1713`. URL: `https://doi.org/10.1006/aima.1997.1713`.

[2] Lawrence C. Evans. *Partial differential equations*. Second. Vol. 19. Graduate Studies in Mathematics. American Mathematical Society, Providence, RI, 2010, pp. xxii+749. ISBN: 978-0-8218-4974-3. DOI: `10.1090/gsm/019`. URL: `https://doi.org/10.1090/gsm/019`.

[3] A. Tychonoff. "Théorèmes d'unicité pour l'équation de la chaleur". In: *Sb. Math.* (1935).