CHAPTER 2

The Cauchy-Kovalevskaya Theorem

This chapter deals with the only "general theorem" which can be extended from the theory of ODEs, the Cauchy-Kovalevskaya Theorem. This also will allow us to introduce the notion of *non-characteristic data*, *principal symbol* and the basic classification of PDEs. It will also allow us to explore why analyticity is not the proper regularity for studying PDEs most of the time.

Acknowledgements. Some parts of this chapter – in particular the four proofs of the Cauchy-Kovalevskaya theorem for ODEs – are strongly inspired from the nice lecture notes of Bruce Diver available on his webpage, and some other parts – in particular the organisation of the proof in the PDE case and the discussion of the counter-examples – are strongly inspired from the nice lectures notes of Tsogtgerel Gantumur available on his webpage.

1. Recalls on analyticity

Let us first do some recalls on the notion of *real analyticity*.

Definition 1.1. *A function f defined on some open subset U of the real line is said to be real analytic at a point $x_0 \in \mathcal{U}$ if it is an infinitely differentiable function such that the Taylor series at the point x_0*

$$T(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n \tag{1.1}$$

converges to $f(x)$ for x in a neighborhood of x_0.

A function f is real analytic on an open set $\mathcal{U}$ of the real line if it is real analytic at any point $x_0 \in \mathcal{U}$. The set of all real analytic functions on a given open set $\mathcal{U}$ is often denoted by $C^\omega(\mathcal{U})$.

Alternatively the notion of real analyticity can be understood as a growth control on the derivatives: f is real analytic on an open set $\mathcal{U}$ of the real line iff for any compact set $K \subset \mathcal{U}$ there are constants $C(K), r > 0$ so that

$$\forall\, x \in K, \quad |f^{(n)}(x)| \le C(K) \frac{n!}{r^n}.$$

Remark 1.2. *Recall that if a Taylor series simply converges for $|x - x_0| = r$, then the general term is bounded which implies by comparison that the series absolutely converges for all x' so that $|x' - x_0| < r$, and converges uniformly on any $\bar{B}(x_0, r')$, $r' < r$. This justifies defining the* radius of convergence $r_0 \in [0, +\infty]$ *of this particular entire series as the supremum of the radiuses where the series is absolutely converging.*

REMARK 1.3. *There is a possible third definition: a function f is real analytic on a open set $\mathcal{U}$ of $\mathbb{R}$ if it can be extended to a complex analytic function on open set of $\mathbb{C}$ around any point of $\mathcal{U}$. Prove this (easier) equivalence with the previous definitions as an exercise.*

REMARK 1.4. *There is a possible fourth definition based on Fourier calculus, saying essentially that analyticity is equivalent to the* exponential decay of the Fourier tranform at infinity*. We will not use this definition in this course.*

PROOF OF EQUIVALENCE OF THE DEFINITIONS. As it will important in this chapter, let us outline the proof of the equivalence of the two definitions.

(i) Assume first that f is real analytic in $\mathcal{U}$ in the sense of the first definition: $f(x) = T(x)$ and $T(x)$ absolutely-uniformly convergent in $\bar{B}(x_0, r) \subset \mathcal{U}$ (closed ball). We then define $f(z) = T(z)$ for $z \in \mathbb{C}$ on the ball $\bar{B}(x_0, r)$ in $\mathbb{C}$, which is well-defined thanks to the convergence of the series, complex analytic, and which extends f. Then for any $x \in \bar{B}(x_0, r/4) \cap \mathbb{R}$, we have by Cauchy's integral formula

$$f^{(n)}(x) = \frac{n!}{2\pi i} \int_{|z-x_0|=r/2} \frac{f(z)}{(z-x)^{n+1}} \, \mathrm{d}z$$

from which we deduce

$$\max_{x \in \bar{B}(x_0, r/4)} |f^{(n)}(x)| \le C \frac{n!}{r^n} \|f\|_{L^\infty(S(x_0, r/2))}.$$

This is the growth control in the second definition on $K = \bar{B}(x_0, r/4)$. The case of a general compact K is obtained by a finite covering by such closed balls.

(ii) Conversely, let us assume the growth control on $K = \bar{B}(x_0, r)$. Then for $x \in \bar{B}(x_0, r/2)$, we expand f into a Taylor series around x_0 at order n

$$f(x) = \sum_{k=0}^{n} f^{(k)}(x_0) \frac{(x-x_0)^k}{k!} + f^{(n+1)}(y_n(x)) \frac{(x-x_0)^{n+1}}{(n+1)!}$$

for some $y_n(x)$ in $\bar{B}(x_0, r/2)$. From the growth control the remainder satisfies

$$\left| f^{(n+1)}(y_n(x)) \frac{(x-x_0)^{n+1}}{(n+1)!} \right| \le \frac{C(K)}{2^{n+1}} \xrightarrow{n \to +\infty} 0$$

and the Taylor series is convergent. We deduce that on $B(x_0, r/2)$ the Taylor series is convergent and converges to f. □

REMARK 1.5. *The definition of a complex analytic function is obtained by replacing, in the definitions above, "real" with "complex" and "real line" with "complex plane".*

EXAMPLE 1.6. *Simple examples of analytic functions (on the real line or the complex plane) are polynomials and the exponential and trigonometric functions. The complex conjugate $z \mapsto \bar{z}$ is not complex analytic, although the restriction to the real line is the identity which is real analytic.*

EXERCISE 10. *Construct an example of a smooth non-analytic function on the real line.* Hint: Build a non-zero smooth function with compact support.

Exercise 11. *Recall the* Liouville theorem *for analytic function of the whole complex plane. Is it true on the real line?* Hint: Consider the function $f(x) = 1/(1+x^2)$.

Finally one can extend the notion of real analyticity to the case of several real variables. It requires the notion of *power series* of several variables: a power series is here defined to be an infinite series of the form

$$T(x_1,\dots,x_\ell) = \sum_{j_1,\dots,j_\ell=0}^{\infty} a_{j_1,\dots,j_\ell} \prod_{k=1}^{\ell} (x_k - x_{0,k})^{j_k}$$

at the base point $x_0 = (x_{0,1},\dots,x_{0,\ell})$.

Definition 1.7. *A function $f = f(x_1,\dots,x_\ell)$ of ℓ real variables is real analytic on an open set $\mathcal{U} \subset \mathbb{R}^\ell$ if it is an infinitely differentiable function such that at any point $x_0 \in \mathcal{U}$ there is a power series T in the form above, which is absolutely convergent in a neighbourhood of x_0 in $\mathcal{U}$, and which agrees with f in this neighbourhood.*

A vectorial function $\mathbf{f} = \mathbf{f}(x_1,\dots,x_\ell)$ of ℓ real variables and valued in $\mathbb{R}^m$ is real analytic if each component is real analytic as defined before.

Example 1.8. *The conjugate function $z \mapsto \bar{z}$ is real analytic from $\mathbb{R}^2$ to $\mathbb{R}^2$.*

2. The Cauchy-Kovalevskaya theorem for ODEs

2.1. Scalar ODEs. As a warm up we will start with the corresponding result for ordinary differential equations.

Theorem 2.1 (ODE Version of Cauchy–Kovalevskaya, I). *Suppose $b > 0$ and $F : (u_0 - b, u_0 + b) \to \mathbb{R}$ is real analytic, and $u(t)$ is the unique solution to the ODE*

$$u'(t) = \frac{\mathrm{d}}{\mathrm{d}t}u(t) = F(u(t)) \quad \textit{with} \quad u(0) = u_0 \in \mathbb{R} \tag{2.1}$$

on some neighborhood $(-a,a)$ of zero, with $u((-a,a)) \subset (u_0 - b, u_0 + b)$. Then u is also real analytic on $(-a,a)$.

Remark 2.2. *Since the construction and uniqueness of solutions, and their interval of existence, is already settled by the Picard-Lindelöf theorem, this is purely a* regularity *theorem saying that the solution is real analytic in the region where the force field is. This will not be the case in the case of PDE, where the construction of solutions will be part of the Cauchy-Kovalevskaya theorem.*

We will give four proofs. However it is the last proof that the reader should focus on for understanding the PDE version in Theorem 4.1.

Observe that

- the existence and uniqueness of solutions is granted by Picard-Lindelöf theorem;
- (however existence and uniqueness arguments could be devised using the arguments below as well;)
- moreover by induction the solution constructed by Picard-Lindelöf can be shown to be smooth, using that if u is C^k in some open set, then $F(u(t))$ is C^k as well by composition, and finally u' is C^k from the differential equation;

- it is enough to show the regularity in a smaller neighbourhood $(-a', a')$ of 0 for u, as then the argument can be performed again around any point x_0 where u is defined and F is analytic around $u(x_0$;
- finally without loss of generality we can restrict to $u_0 = 0$.

Proof 1 of Theorem 2.1. This details of the proof are left as an exercise. We go back to the proof of Picard-Lindelöf theorem by fixed-point argument, and replace the real variable by a complex variable and the real integral by a complex path integral:

$$u^{n+1}(z) = u_0 + \int_0^z F(u_n(z'))\,\mathrm{d}z' = u_0 + \int_0^1 F(u_n(zt))z\,\mathrm{d}t$$

with the initialisation $u^0 = u_0$ (constant function), and where F is a local complex analytic extension of F around u_0. Then we show that the sequence u^n is Cauchy for the supremum norm on the some neighborhood of zero, using the control on the complex derivative $F'(z)$ of F. Finally observe that u^n is complex analytic for any n by induction, then use Morera's theorem to deduce that the uniform limit u is complex analytic as well. Finally check that for z real, the solution u constructed is real (it coincides with the Picard-Lindelöf solution). □

Proof 2 of Theorem 2.1. We follow here the same strategy as for solving an ODE by "separation of variables". If $F(0) = 0$ then the solution is $u = 0$ which is clearly analytic and we are done. Assume $F(0) \neq 0$, then let us define the new function

$$G(y) = \int_0^y \frac{1}{F(x)}\,\mathrm{d}x, \quad y \in (-b', b') \subset (-b, b)$$

which is again real analytic for b' small enough (to make sure that $F(x)$ does not cancel in $(-b', b')$). Then we have by the chain rule in the (possibly smaller) neighbourhood $(-a', a') \subset (-a, a)$ where u is defined and which is mapped into the region where G is analytic $u((-a', a')) \subset (-b', b')$:

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[G(u(t))\right] = \frac{\dot{u}(t)}{F(u(t))} = 1$$

which implies, together with $G(u(0)) = G(0) = 0$, that $G(u(t)) = t$. Finally observe that since $G'(0) = 1/F(0) \neq 0$, there is $(-a'', a'') \subset (-a', a')$ where G^{-1} is defined and analytic. Then $u(t) = G^{-1}(t)$ is analytic on $(-a'', a'')$ which concludes the proof. □

Proof 3 of Theorem 2.1. Let us consider, for $z \in \mathbb{C}$, the solution $u_z(t)$ to

$$u_z'(t) = zF(u_z(t)), \quad u_z(0) = 0. \tag{2.2}$$

Then for any $|z| \leq 2$, one can construct by Picard-Lindelöf a solution to (2.2) on a neighbourhood $|t| \leq \varepsilon$: indeed the Lipschitz constant influences the size of the neighbourhood but can be uniformly bounded for $|z| \leq 2$, which yields a uniform neightborhood $|t| \leq \varepsilon$ for all $|z| \leq 2$. Assume that ε is also small enough so that one can construct a solution $u(t)$ to the original equation on $|t| < 2\varepsilon$. Observe then that by uniqueness in the Picard-Lindelöf theorem the solution of (2.2) is given by $u(zt)$ where u is the solution to the original problem, for $t \in (-\varepsilon, \varepsilon)$ and $|z| < 2$. This shows that $u_z(t)$ is smooth as a function of t and z.

Let us recall the notation

$$\frac{\partial}{\partial z} = \frac{1}{2}\left(\frac{\partial}{\partial x} - i\frac{\partial}{\partial y}\right), \quad \frac{\partial}{\partial \bar z} = \frac{1}{2}\left(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y}\right).$$

Then the solution to (2.2) satisfies

$$\frac{\partial}{\partial t}\frac{\partial u_z(t)}{\partial \bar z} = zF'(u_z(t))\frac{\partial u_z(t)}{\partial \bar z}$$

(here we have used the chain rule, the fact that $\partial(z)/\partial\bar z = 0$ and the real differentiability of F). This implies that

$$\frac{\partial u_z(t)}{\partial \bar z} = \exp\left(\int_0^t zF'(u_z(s))\,\mathrm{d}s\right)\frac{\partial u_z(0)}{\partial \bar z},$$

which, together with the initial condition $u_z(0) = 0$, proves that $\partial u_z(t)/\partial\bar z = 0$ for $t \in (-\varepsilon, \varepsilon)$. These are the Cauchy-Riemann equations, and thus $z \mapsto u_z(t)$ is complex analytic with respect to $z \in B(0,2)$, for any t in $(-\varepsilon,\varepsilon)$.

We deduce the following absolutely converging Taylor series expansion

$$u_1(t) = \sum_{n=0}^{\infty}\frac{1^n}{n!}\left(\frac{\partial^n u_z(t)}{\partial z^n}\right)\Big|_{z=0}.$$

Now that $u_z(t) = u(zt)$ for $t \in (-\varepsilon,\varepsilon)$ and $|z| < 2$ we have

$$\left(\frac{\partial^n u_z(t)}{\partial z^n}\right)\Big|_{z=0} = \left(\frac{\partial^n u(zt)}{\partial z^n}\right)\Big|_{z=0} = t^n u^{(n)}(0)$$

which yields the convergence of the Taylor series of u at 0 on $|t| < \varepsilon$, together with the equality

$$u(t) = u_1(t) = \sum_{n=0}^{\infty}\frac{t^n}{n!}u^{(n)}(0),$$

which proves the real analyticity. □

Remark 2.3. *This argument relies on extending the original equation to a continuous family of equation by adding an additional parameter, this is a useful idea in the more general context of PDEs.*

Proof 4 of Theorem 2.1. This is the most important proof, as it is the historic proof of A. Cauchy (improved by S. Kovalevskaya) but also because it is beautiful and this is the proof we shall use in a PDE context. This is called the "method of majorants". Let us do first an *a priori examination of the problem*, assuming the analyticity (actually here it could be justified by using Picard-Lindelöf to contruct solutions, and then check by bootstrap that this solution is smooth). Then we compute the derivatives

$$\begin{cases} u^{(1)}(t) = F^{(0)}(u(t)), \\ u^{(2)}(t) = F^{(1)}(u(t))u^{(1)}(t) = F^{(1)}(u(t))F^{(0)}(u(t)), \\ u^{(3)}(t) = F^{(2)}(u(t))F^{(0)}(u(t))^2 + F^{(1)}(u(t))^2F^{(0)}(u(t)), \\ \dots \end{cases}$$

Remark 2.4. *The calculation of these polynomials is connected to a formula devised in the 19th century by Arbogast in France and Faà di Bruno in Italy. It is now known as Faà di Bruno's formula and it is good to keep it in one's analytic toolbox:*

$$\frac{\mathrm{d}^n}{\mathrm{d}t^n} F(u(t)) =$$
$$\sum_{m_1+2m_2+\cdots+nm_n=n} \frac{n!}{m_1!1!^{m_1} m_2!2!^{m_2} \dots m_n!n!^{m_n}} F^{(m_1+\cdots+m_n)}(u(t)) \prod_{j=1}^{n} \left(u^{(j)}(t)\right)^{m_j}.$$

Now the key observation is that there are universal (in the sense of being independent of the function F) polynomials p_n with **non-negative** integer coefficients, so that

$$u^{(n)}(t) = p_n\left(F^{(0)}(u(t)), \dots, F^{(n-1)}(u(t))\right).$$

This is proved by induction, or simply by direct use of Faà di Bruno's formula.

We deduce by monotonicity

$$|u^{(n)}(0)| \le p_n\left(|F^{(0)}(0)|, \dots, |F^{(n-1)}(0)|\right) \le p_n\left(G^{(0)}(0), \dots, G^{(n-1)}(0)\right)$$

for any function G with *non-negative* derivatives at zero and such that $G^{(n)}(0) \ge |F^{(n)}(0)|$ for all $n \ge 0$. Such a function is called a *majorant function* of F.

But for such a function G the RHS in the previous equation is exactly

$$p_n\left(G^{(0)}(0), \dots, G^{(n-1)}(0)\right) = v^{(n)}(0) = |v^{(n)}(0)|$$

where v solves the *auxiliary equation*

$$v'(t) = \frac{\mathrm{d}}{\mathrm{d}t} v(t) = G(v(t)), \quad v(0) = 0.$$

Observe that as a consequence of the inequalities between G and F we have $v^{(n)}(0) \ge |u^{(n)}(0)|$ for any $n \ge 0$. Hence if v is analytic near zero, the series

$$\mathcal{S}_v(t) := \sum_{n\ge 0} v^{(n)}(0) \frac{t^n}{n!}$$

has a positive radius of convergence and by comparison so does the series

$$\mathcal{S}_u(t) := \sum_{n\ge 0} |u^{(n)}(0)| \frac{t^n}{n!},$$

which shows the growth condition on the derivatives at zero.

We now construct the majorant function G. From the analyticity of F we have

$$\forall\, n \ge 0, \quad |F^{(n)}(0)| \le C \frac{n!}{r^n}$$

uniformly in $n \ge 0$, for some some constant $C > 0$ and some $r > 0$, and we then consider

$$G(z) := C \sum_{n=0}^{\infty} \left(\frac{z}{r}\right)^n = C \frac{1}{1 - z/r} = \frac{Cr}{r - z}$$

which is analytic on the ball centred at zero with radius $r > 0$. Since $G^{(n)}(0) = Cn!/r^n$ we have clearly the majoration $G^{(n)}(0) \geq |F^{(n)}(0)|$ for all $n \geq 0$.

To conclude the proof we need finally to compute the solution v to the auxiliary equation

$$v'(t) = G(v(t)) = \frac{Cr}{r - v(t)}, \quad v(0) = 0,$$

which can be solved by usual real differential calculus, using separation of variables:

$$(r - v)\,\mathrm{d}v = Cr\,\mathrm{d}t \quad \Longrightarrow \quad -\,\mathrm{d}(r - v)^2 = 2Cr\,\mathrm{d}t \quad \Longrightarrow \quad v(t) = r \pm r\sqrt{1 - \frac{2Ct}{r}}$$

and using the initial condition $v(0) = 0$ one finally finds

$$v(t) = r - r\sqrt{1 - \frac{2Ct}{r}}$$

which is analytic for $|t| < r/(2C)$.

This hence shows that the radius of convergence of $\mathcal{S}_v(t)$ is positive, which implies the growth control

$$\forall\, n \geq 1, \quad 0 \leq v^{(n)}(0) \leq C\frac{n!}{\varepsilon^n}$$

and in turn implies the growth control

$$\forall\, n \geq 1, \quad 0 \leq |u^{(n)}(0)| \leq C\frac{n!}{\varepsilon^n}.$$

Since this argument can be performed uniformly for any $t \in [-a', a'] \subset (-a, a)$, using the uniform growth control on the derivatives of F on the region $u([-a', a'])$, we deduced

$$\forall\, t \in [-a', a'], \quad \forall\, n \geq 1, \quad 0 \leq |u^{(n)}(t)| \leq C\frac{n!}{\varepsilon^n}.$$

This shows the real analyticity (second definition). □

Observe the important idea in this last proof:

- first one uncovers a general combinatorial structure at a universal level (not depending on F defining the ODE),
- then instead of coping with the combinatorial explosion in the calculation of the derivatives (due to the nonlinearity), one uses a monotonicity property encoded in the abstract structure to reduce the control to be established to a comparison with a simpler function,
- then one goes back to the equation to deduce the analytic control, without ever computing the derivatives.

2.2. Systems of ODEs. We now consider the extension of this theorem to *systems* of differential equations.

THEOREM 2.5 (ODE Version of Cauchy–Kovalevskaya, II). *Suppose $b > 0$ and $\mathbf{F} : \mathbf{u}_0 + (b, b)^m \to \mathbb{R}^m$, $m \in \mathbb{N}$ is real analytic and $\mathbf{u}(t)$ is the unique C^1 solution to the system of ODEs*

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}(t) = \mathbf{F}(\mathbf{u}(t)) \quad \textit{with} \quad \mathbf{u}(0) = \mathbf{u}_0 \in \mathbb{R}^m \tag{2.3}$$

on $(-a,a)$ *with* $\mathbf{u}((-a,a)) \subset \mathbf{u}_0 + (-b,b)^m$.

Then $\mathbf{u}$ *is also real analytic in* $(-a,a)$.

PROOFS OF THEOREM 2.5. All but the second proof of Theorem 2.1 can be adapted to cover this case of systems.

EXERCISE 12. *Extend the proofs 1 and 3 to this case.*

Let us give some more comments-exercises on the extension of the fourth proof, the method of majorants.

EXERCISE 13. *Suppose* $\mathbf{F} : (-a,a)^m \to \mathbb{R}^m$ *is real analytic near* $0 \in (-a,a)^m$, *prove that a majorant function is provided by*

$$\mathbf{G}(z_1,\ldots,z_m) := (G_1,\ldots,G_m), \quad G_1 = \cdots = G_m = \frac{Cr}{r - z_1 - \cdots - z_m}$$

for well-chosen values of the constants $r, C > 0$.

With this auxiliary result at hand, check that one can reduce the proof to proving the local analyticity of the solution to the system of ODE:

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{v}(t) = \mathbf{G}(\mathbf{v}(t)), \quad \mathbf{v}(t) = (v_1(t),\ldots,v_m(t)), \quad \mathbf{v}(0) = 0.$$

EXERCISE 14. *Prove that by symmetry one has* $v_j(t) = v_1(t) =: w(t)$ *for all* $1 \le j \le m$, *and that* $w(t)$ *solves the* scalar *ODE*

$$\frac{\mathrm{d}}{\mathrm{d}t}w(t) = \frac{Cr}{r - mw(t)}, \quad w(0) = 0$$

so that $w(t) = (r/m)(1 - \sqrt{1 - 2Cdt/r})$.

With the two last results it is easy to conclude the proof. $\square$

3. The analytic Cauchy problem in PDEs

3.1. The setting. We consider a k-th order scalar *quasilinear* PDE[1]

$$\sum_{|\alpha|=k} a_\alpha(\nabla^{k-1}u,\ldots,u,x)\partial_x^\alpha u + a_0(\nabla^{k-1}u,\ldots,u,x) = 0, \quad x \in \mathcal{U} \subset \mathbb{R}^\ell \tag{3.1}$$

where

$$\nabla^j u := \left(\partial_{x_{i_1}} \ldots \partial_{x_{i_j}} u\right)_{1 \le i_1,\ldots,i_j \le \ell}, \quad j \in \mathbb{N},$$

is the j-th iterated gradient, and

$$\partial_x^\alpha := \partial_{x_1}^{\alpha_1} \ldots \partial_{x_\ell}^{\alpha_\ell}$$

for a multi-index $\alpha = (\alpha_1,\ldots,\alpha_\ell) \in \mathbb{N}^\ell$, and $\mathcal{U}$ is some open set in $\mathbb{R}^\ell$ ($\ell \ge 2$ is the number of variables), and $u : \mathcal{U} \to \mathbb{R}$.

[1]Actually when $k = 1$ (first order) another simpler proof than the one we shall do here can be performed by the so-called *characteristics method* (see chapter 5 where we will come back to this). This method can be understood as the natural extension of the ODE arguments in proofs 1-2-3 above using trajectories for the PDE. However this method fails for systems, and therefore is unable to treat k-th order PDEs, which justifies the need for a more general proof, actually an extension of the fourth proof above.

TERMINOLOGY. The word “quasilinear” relates to the fact that the coefficient of the highest-order derivatives only depend on derivatives with strictly lower order. The equation would be *semilinear* if $a_\alpha = a_\alpha(x)$ does not depend on u and the nonlinearity is only in a_0. The equation is *linear* when of course both a_α and a_0 do not depend on u, and it is a *constant coefficient linear equation* when a_α and a_0 do not depend on x either.

DEFINITION OF THE HYPERSURFACE. We then consider a smooth $(\ell-1)$-dimensional hypersurface Γ in $\mathcal{U}$, equiped with a smooth unit normal vector $\mathbf{n}(x) = (n_1(x), \dots, n_\ell(x))$ for $x \in \Gamma$. More precisely:

(1) Γ is an *embedded smooth submanifold of* $\mathbb{R}^\ell$ *with dimension* $(\ell-1)$. This means that around any point $x \in \Gamma$ there a smooth non-singular bijective map $\Psi_x : B_\ell(0,\varepsilon) \subset \mathbb{R}^d \to \mathcal{U}_x \subset \mathcal{U}$ with $\Psi_x(B_\ell(0,\varepsilon) \cap \{y_\ell = 0\}) = \Gamma \cap \mathcal{U}_x$, and $\mathcal{U}_x$ some open neighbourhood of x in $\mathcal{U}$. The collection of all such pairs $(\mathcal{U}_x, \Psi_x)$ of neighbourhood and applications (called *charts*) forms an *atlas*.

(2) The smooth unit normal vector map is a smooth map when composed locally with the charts, and which is orthogonal at each point to the tangent hyperplane obtained by the chart. This tangent plan at a point $x = \Psi_x(0)$ is the affine vector hyperplane

$$\mathcal{T}_x = x + \operatorname{Span}\left(\frac{\partial \Psi_x}{\partial y_1}(0), \dots, \frac{\partial \Psi_x}{\partial y_{\ell-1}}(0) \right).$$

This is part of the differential geometry course, but we will not need more than what is introduced in this chapter.

CONSTRUCTION OF A PARTICULAR CHART. Then it shall prove useful to use the following construction, taking advantage of the unit vector:

(1) Locally in a neighbourhood of any point $x \in \Gamma$, by restriction of Ψ_x, we have an application $\bar\Psi_x : B_{\ell-1}(0,\varepsilon) \to (\mathcal{U}_x \cap \Gamma) \subset \mathcal{U} \subset \mathbb{R}^\ell$, that is again smooth from $\mathbb{R}^{\ell-1}$ to $\mathbb{R}^\ell$.

(2) By constructing then

$$\tilde\Psi_x(y) = \bar\Psi_x(\bar y) + y_\ell \mathbf{m}(\bar y), \quad \mathbf{m}(\bar y) := \mathbf{n}(\bar\Psi_x(\bar y)), \quad \bar y = (y_1, \dots, y_{\ell-1}),$$

we obtain a smooth bijective chart on some open set $B_\ell(0,\varepsilon')$ (possibly reducing $\varepsilon' < \varepsilon$ to avoid loosing injectivity) onto an open set $x \in \mathcal{V}_x \subset \mathcal{U}_x$ with

$$\frac{\partial \tilde\Psi_x}{\partial y_\ell}(y) = \mathbf{m}(\bar y) = \mathbf{n}(\bar\Psi_x(\bar y))$$

and

$$\mathcal{T}_{x'} = x' + \operatorname{Span}\left(\frac{\partial \tilde\Psi_x}{\partial y_1}(\bar y', 0), \dots, \frac{\partial \tilde\Psi_x}{\partial y_{\ell-1}}(\bar y', 0) \right), \quad x' \in \mathcal{V}_x \cap \Gamma.$$

(3) We then define $\tilde\Phi_x := (\tilde\Psi_x)^{-1}$ and the scalar smooth function $\varphi := (\Phi_x)_\ell : \mathcal{V}_x \to \mathbb{R}$ which satisfies $\Gamma \cap \mathcal{V}_x = \{\varphi_x = 0\}$ and $\nabla_x \varphi(x') = \mathbf{n}(x')$, $x' \in \mathcal{V}_x \cap \Gamma$.

Indeed starting from the equation $\varphi(\tilde\Psi_x(y)) = y_\ell$ and differentiating according to $y_1, y_2, \dots, y_{\ell-1}$ we get $\nabla_x \varphi(x') \perp \mathcal{T}_{x'}$ on $x' \in \mathcal{V}_x \cap \Gamma$ and therefore $\nabla_x \varphi(x')$ colinear to $\mathbf{n}(x')$. Finally differentiating according to y_ℓ we get $\nabla_x \varphi(x') \cdot \mathbf{n}(x') = 1$ which proves $\nabla_x \varphi(x') = \mathbf{n}(x')$.

Normal derivatives to the Cauchy hypersurface. We then define the *j-th normal derivative* (for $j \in \mathbb{N}$) of u at $x \in \Gamma$ as

$$\frac{\partial^j u}{\partial \mathbf{n}^j} := \sum_{|\alpha|=j} \nabla_x^\alpha u : \mathbf{n}^\alpha = \sum_{|\alpha|=j} \frac{\partial^\alpha u}{\partial^\alpha x} n^\alpha = \sum_{\alpha_1+\cdots+\alpha_\ell=j} \frac{\partial^j u}{\partial x_1^{\alpha_1} \dots \partial x_\ell^{\alpha_\ell}} n_1^{\alpha_1} \dots n_\ell^{\alpha_\ell}.$$

Now let $g_0, \dots, g_{k-1} : \Gamma \to \mathbb{R}$ be k given functions on Γ, and $x_0 \in \Gamma$. The **Cauchy problem** is then to find a function u solving (3.1) in some open set including x_0, subject to the boundary conditions

$$u = g_0, \quad \frac{\partial u}{\partial \mathbf{n}} = g_1, \quad \dots \quad , \quad \frac{\partial^{k-1} u}{\partial \mathbf{n}^{k-1}} = g_{k-1}, \quad x \in \Gamma. \tag{3.2}$$

We say that the equation (3.2) prescribes the *Cauchy data* g_0, ..., g_{k-1} on Γ.

If one wants to compute an entire series for the solution, certainly all the derivatives have to be determined from equations (3.1)-(3.2). In particular *all* partial derivatives of u on Γ should be computed from the boundary data (3.2).

The basic question is now: assuming first that we have a smooth solution and leaving aside the question of the convergence of the Taylor series, what kind of conditions do we need on Γ in order to so?

3.2. The case of a flat boundary. In order to gain intuition into the problem, we first examine the case where $\mathcal{U} = \mathbb{R}^\ell$ and $\Gamma = \{x_\ell = 0\}$ is a vector hyperplan. We hence have $\mathbf{n} = \mathbf{e}_\ell$ (the ℓ-th unit vector of the canonical basis) and the boundary prescriptions (3.2) read

$$u = g_0, \quad \frac{\partial u}{\partial x_\ell} = g_1, \quad \dots \quad , \quad \frac{\partial^{k-1} u}{\partial x_\ell^{k-1}} = g_{k-1}, \quad x \in \Gamma.$$

Which further partial derivatives can we compute on the hyperplan Γ? First since $u = g_0$ on Γ by differentiating tangentially we get that

$$\frac{\partial u}{\partial x_i} = \frac{\partial g_0}{\partial x_i}, \quad 1 \le i \le \ell - 1$$

is prescribed by the boundary data. Since we also know from (3.2) that

$$\frac{\partial u}{\partial x_\ell} = g_1$$

we can determine the full gradient on Γ. Similarly we can calculate inductively

$$\frac{\partial^\alpha u}{\partial \mathbf{x}^\alpha} = \frac{\partial^{\bar{\alpha}}}{\partial \mathbf{x}^{\bar{\alpha}}} \frac{\partial^j u}{\partial x_\ell^j} = \frac{\partial^{\bar{\alpha}}}{\partial \mathbf{x}^{\bar{\alpha}}} g_j, \quad \alpha_i \in \mathbb{N} \text{ for } 1 \le i \le \ell - 1, \quad |\alpha_\ell| \le k - 1$$

with $\bar{\alpha} := (\alpha_1, \dots, \alpha_{\ell-1}, 0)$. The difficulty then, in order to compute the full k-th derivative, is to compute the k-th order normal derivative $\partial^k u / \partial x_\ell^k$. This is not prescribed by the boundary conditions, but we shall recover this derivative from the PDE, under some non-degeneracy condition. Observe that if the coefficient a_α with $\alpha = (0, \dots, 0, k)$ is non-zero on Γ:

$$\begin{aligned} A(x) &:= a_{(0,\dots,0,k)}(\nabla^{k-1} u(x), \dots, u(x), x) \\ &= \text{Function}(g_{k-1}(x), g_{k-2}(x), \dots, g_0(x), x) \neq 0, \quad x \in \Gamma, \end{aligned}$$

(observe that it only depends on the boundary data) then we can compute for $x \in \Gamma$

$$\frac{\partial^k u}{\partial x_\ell^k} = -\frac{1}{A(x)} \left[\sum_{|\alpha|=k,\ \alpha_\ell \le k-1} a_\alpha(\nabla^{k-1} u(x), \dots, u(x), x) \partial_x^\alpha u(x) + a_0(\nabla^{k-1} u(x), \dots, u(x), x) \right]$$

where the coefficients in the RHS again only depend on the boundary data by the previous calculations, and consequently we can therefore compute $\nabla^k u$ on Γ.

Definition 3.1. *We say that the hypersurface $\Gamma = \{x_\ell = 0\}$ and boundary conditions $g_0, \dots, g_{k-1}$ are* non-characteristics *for the PDE* (3.1) *at $x \in \Gamma \cap \mathcal{U}$ if the function $A(x) = a_{(0,\dots,0,k)}(\nabla^{k-1} u(x), \dots, u(x), x)$ does not cancel on $\Gamma \cap \mathcal{U}_x$ for some neighborhood $\mathcal{U}_x$ of x. It is non-characteristic on $\Gamma \cap \mathcal{U}$ if the previous applies at any point $x \in \Gamma \cap \mathcal{U}$.*

Then can we calculate still higher derivatives on Γ, assuming this non-degeneracy condition? The answer is yes, here is a concise inductive way of iterating the argument:

Let us denote

$$g_k(x) := \frac{\partial^k u}{\partial x_\ell^k}(x) = -\frac{1}{A(x)} \left[\sum_{|\alpha|=k,\ \alpha_\ell \le k-1} a_\alpha(\dots) \partial_x^\alpha u(x) + a_0(\dots) \right], \quad x \in \Gamma,$$

as computed before. We now differentiate the equation along x_ℓ (we already know how to compute all the derivatives along the other coordinates, provided we have less than k derivatives along x_ℓ), which results into a new equation

$$\sum_{|\alpha|=k} a_\alpha(\nabla^{k-1} u(x), \dots, u(x), x) \partial_x^\alpha \partial_{x_\ell} u(x) + \tilde{a}_0(\nabla^k u(x), \dots, u(x), x) = 0, \quad x \in \mathcal{U} \subset \mathbb{R}^\ell,$$

with

$$\begin{aligned} \tilde{a}_0(\nabla^k u(x), \dots, u(x), x) := \partial_{x_\ell} &\left[\sum_{|\alpha|=k} a_\alpha(\nabla^{k-1} u(x), \dots, u(x), x) \right] \partial_x^\alpha u(x) \\ &+ \partial_{x_\ell} \left[a_0(\nabla^{k-1} u(x), \dots, u(x), x) \right], \end{aligned}$$

which results following the same argument into

$$\begin{aligned} g_{k+1}(x) &:= \frac{\partial^{k+1} u}{\partial x_\ell^{k+1}}(x) \\ &= -\frac{1}{A(x)} \left[\sum_{|\alpha|=k,\ \alpha_\ell \le k-1} a_\alpha(\nabla^{k-1} u(x), \dots, u(x), x) \partial_x^\alpha \partial_{x_\ell} u(x) + \tilde{a}_0(\nabla^k u(x), \dots, u(x), x) \right] \end{aligned}$$

(observe that the RHS only involves derivatives in x_ℓ of order less than k), which allows to calculate the $k+1$-derivative in x_ℓ from the boundary data. One can then continue inductively and calculate all derivatives.

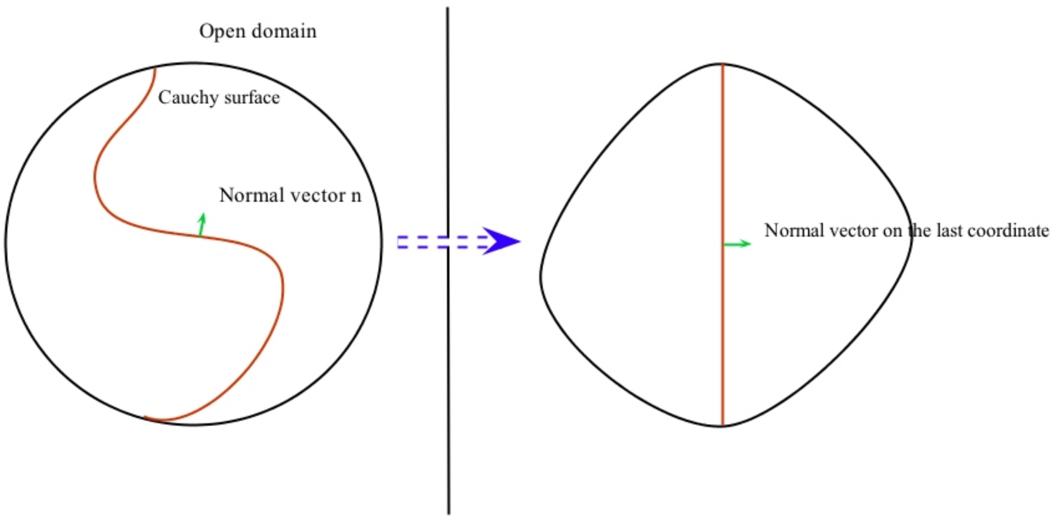

FIGURE 3.1. Local rectification of the flow.

3.3. General hypersurfaces. We shall now generalize the results and definitions above to the general case, when Γ is a smooth hypersurface with normal vector field $\mathbf{n}$.

DEFINITION 3.2. *We say that the hypersurface* Γ *and boundary conditions* $g_0, \ldots, g_{k-1}$ *are* non-characteristic *for the PDE* (3.1) *if*

$$A(x) := \sum_{|\alpha|=k} a_\alpha(\nabla^{k-1}u(x), \ldots, u(x), x)\mathbf{n}^\alpha(x) \neq 0, \quad x \in \Gamma \cap \mathcal{U}$$

(where the RHS only depends on the boundary data).

Let us prove the theorem corresponding the calculation of the partial derivatives

THEOREM 3.3 (Cauchy data and non-characteristic surfaces). *Assume that* Γ *is smooth and non-characteristic for the PDE* (3.1) *and boundary data* (3.2)*. Then if* u *is a* C^∞ *solution to* (3.1) *in* $\mathcal{U}$ *with the boundary data* (3.2)*, we can uniquely compute all the partial derivatives of* u *on* Γ *in terms of the charts of* Γ*, the functions* g_0*, ...,* g_{k-1}*, and the coefficients* a_α*,* a_0*.*

PROOF OF THEOREM 3.3. We consider a base point $x \in \Gamma$ and using the previous discussion we find C^∞ maps Φ, Ψ defined on open sets $\mathcal{U}_x$ and $\mathcal{V}_0$ of $\mathbb{R}^\ell$ so that

$$\Phi(\Gamma \cap B(x,r)) = \Theta \subset \{y_\ell = 0\}, \quad \Phi(x) = y, \quad \Psi = \Phi^{-1}$$

where Θ is the new Cauchy surface in the new coordinates, and with the two properties

$$\frac{\partial \Psi}{\partial y_\ell}(y) = \mathbf{m}(\bar{y}) \quad \text{and} \quad \mathbf{m}(\bar{y}) = \mathbf{n}(\Psi(\bar{y}, 0)) = \mathbf{m}(y_1, \ldots, y_{\ell-1}), \quad y \in \mathcal{V}_0 \tag{3.3}$$

$$\nabla_x \Phi_\ell(x') = \mathbf{n}(x'), \quad x' \in \mathcal{U}_x \cap \Gamma. \tag{3.4}$$

Then we define $v(y) := u(\Psi(y))$ and calculations show that v satisfies a new equation of the form

(3.5)
$$\sum_{|\alpha|=k} b_\alpha(\nabla_y^{k-1} v(y), \dots, v(y), y)\partial_y^\alpha v(y) + b_0(\nabla_y^{k-1} v(y), \dots, v(y), y) = 0, \quad y \in \mathcal{V}_0 \subset \mathbb{R}^\ell.$$

We drop the indeces denoting points on symbols of open sets and maps from now on.

The new boundary data are (for $y \in \Theta = \{y_\ell = 0\} \cap \mathcal{V}_0$)

$$\begin{cases} v(y) = g_0(\Psi(y)), \\ \dfrac{\partial v}{\partial y_\ell}(y) = (\nabla_x u)(\Psi(y))\dfrac{\partial \Psi}{\partial y_\ell}(y) = g_1(\Psi(y)), \\ \dfrac{\partial^2 v}{\partial y_\ell^2}(y) = (\nabla_x^2 u)(\Psi(y)) : \left(\dfrac{\partial \Psi}{\partial y_\ell}(y)\right)^{\otimes 2} = g_2(\Psi(y)) \\ \dfrac{\partial^3 v}{\partial y_\ell^3}(y) = \cdots = g_3(\Psi(y)) \\ \cdot \\ \cdot \\ \cdot \\ \dfrac{\partial^{k-1} v}{\partial y_\ell^{k-1}}(y) = \cdots = g_{k-1}(\Psi(y)) \end{cases}$$

where we have used repeatedly the first condition (3.3) we have imposed on our rectification map, and one checks by induction that they can be computed only in terms of the boundary conditions.

Then we want to impose the non-characteristic property in the new coordinates on the new Cauchy surface that we have computed in the flat case in the previous section:

$$b_{(0,\dots,0,k)}(\nabla_y^{k-1} v(y), \dots, v(y), y) \neq 0, \quad y \in \Theta.$$

Let us compute it more precisely.

Indeed we calculate with the chain rule on $u = v \circ \Phi = v(\Phi_1(x), \dots, \Phi_\ell(x))$ that for $|\alpha| = k$ and $x \in \Gamma$, we have

$$\frac{\partial^\alpha u}{\partial x^\alpha}(x) = \frac{\partial^k v}{\partial y_\ell^k}(y)\,(\nabla_x \Phi_\ell(x))^\alpha + \text{ lower order terms} = \frac{\partial^k v}{\partial y_\ell^k}(y)\mathbf{n}^\alpha(x) + \text{ lower order terms}$$

where the lower-order terms only involve partial derivatives with order less than $k-1$ in y_ℓ, and where we have used the second condition (3.4) we have imposed on our rectification map. The other derivatives with $|\alpha| < k$ only create terms involving partial derivatives with order less than $k-1$ in y_ℓ. We hence deduce that

$$b_{(0,\dots,0,k)}(\nabla_y^{k-1} v, \dots, v, y) = \sum_{|\alpha|=k} a_\alpha(\nabla_x^{k-1} u \circ \Psi(y), \dots, u \circ \Psi(y), x)\mathbf{n}^\alpha(\Psi(y))$$

and the non-characteristic condition reads in the original parametrisation

$$A(x) := \sum_{|\alpha|=k} a_\alpha(\nabla_x^{k-1} u(x), \dots, u(x), x)\mathbf{n}^\alpha(x) \neq 0, \quad x \in \Gamma.$$

Exercise 15. *Redo carefully the previous calculation.*

From the assumption of the theorem we therefore satisfy the definition of a non-characteristic surface in the rectified parametrisation. Then using the previous case of a flat boundary we can compute all derivatives of v, and then all derivatives of u, which concludes the proof. □

Remark 3.4. *The choice of the conditions imposed on the rectification map is important in this proof. The condition* (3.3) *is not the simplest one in terms of simplicity of the new boundary data, but the second condition* (3.4) *simplifies the computation of the new non-characteristic condition. The final non-characteristic statement is of course intrinsic to* Γ *and* $g_0, \dots, g_{k-1}$ *and the original PDE and should not depend on the choice of the rectification.*

Exercise 16. *Check that all the results in this section are still valid for* systems *of PDEs, when imposing boundary data as above on each coordinate of the solution* $\mathbf{u} = (u_1, \dots, u_m)$.

4. The Cauchy-Kovalevskaya Theorem for PDEs

4.1. The statement. We shall now prove the following result:

Theorem 4.1 (Cauchy-Kovalevakaya Theorem for PDEs). *We now consider a* real analytic *Cauchy surface* Γ *in an open set* $\mathcal{U}$. *Under real analyticity assumptions on all coefficients on* $\mathcal{U}$ *and boundary data functions in* Γ, *and the non-characteristic condition, there is a unique local analytic solution* u *to the equations* (3.1)-(3.2). *This means that for any* $x \in \Gamma \subset \mathcal{U}$, *there is* $\mathcal{U}_x \subset \mathcal{U}$ *an open set including* x *so that there is a unique analytic solution* u *to* (3.1) *on* $\mathcal{U}_x$ *satisfying the boundary data* (3.2) *on* $\Gamma \cap \mathcal{U}_x$.

Remark 4.2. *The meaning of a* real analytic *hypersurface is: it is implicitly defined by an analytic function* φ: $\Gamma = \{\varphi = 0\}$, *or equivalently it is locally rectifiable with analytic maps* Φ, Ψ *as before. One can check that it is still possible to make the construction of a* real analytic *rectification with the two conditions imposed before by the same arguments.*

4.2. Reduction to a first-order system with flat boundary. We make the sucessive following reductions of the problem:

(1) First, upon flattening out the boundary by an analytic mapping, we can reduce to the situation $\Gamma = \{x_\ell = 0\} \cap \mathcal{U}$ and work around 0, by reasoning as before.

(2) Second, upon dividing by $a_{(0,\dots,0,k)}$ locally around Γ, we can assume that $a_{(0,\dots,0,k)} = 1$ by changing the coefficients to new (still analytic) coefficients.

(3) Third, by substracting off appropriate analytic functions, we may assume that the Cauchy data are identically zero.

(4) Fourth, we transform the equation to a first-order *system* by defining

$$\mathbf{u} := \left(u, \frac{\partial u}{\partial x_1}, \dots, \frac{\partial u}{\partial x_\ell}, \frac{\partial^2 u}{\partial x_i \partial x_j}, \dots \right)$$

where the vector includes all partial derivatives with total order less than $k-1$. Let m denotes the number of components of this vector. It results into the following system with boundary conditions

$$\begin{cases} \dfrac{\partial \mathbf{u}}{\partial x_\ell} = \displaystyle\sum_{j=1}^{\ell-1} \mathbf{b}_j(\mathbf{u}, \bar{x}) \frac{\partial \mathbf{u}}{\partial x_j} + \mathbf{b}_0(\mathbf{u}, \bar{x}), \quad x \in \mathcal{U} \\ \\ \mathbf{u} = 0 \quad \text{on} \quad \Gamma = \{x_\ell = 0\} \cap \mathcal{U}, \end{cases} \tag{4.1}$$

with matrix-valued functions $\mathbf{b}_j : \mathbb{R}^m \times \mathbb{R}^{\ell-1} \mapsto \mathcal{M}_{m\times m}$ and vector-valued function $\mathbf{b}_0 : \mathbb{R}^m \times \mathbb{R}^{\ell-1} \mapsto \mathbb{R}^m$ which are locally analytic around $(0,0)$, and where $\bar{x} = (x_1, \dots, x_{\ell-1})$. Observe that we have assume that the coefficients $\mathbf{b}_j$, $j = 0, \dots, \ell-1$, do not depend on x_ℓ. This can obtained by adding a further component $\mathbf{u}_{m+1} = x_\ell$ if necessary.

Remark 4.3. *Observe that the reduction in this subsection uses crucially the non-characteristic condition. As is made clear by the reduction of the problem, it means at a physical level that we have been able to use one of the variables as a time variable in order to reframe the problem as an evolution problem. However finding a non-characteristic Cauchy surface to start with can be difficult, this is for instance one of the issues in solving the Einstein equations in general relativity, as in the Choquet-Bruhat Theorem.*

4.3. The proof in the reduced case. We now consider a base point on Γ, say 0 w.l.o.g., and, as in the ODE case, we calculate the partial derivatives at this point by repeatedly differentiating the equation.

We have first obviously $\mathbf{u}(0) = 0$.

Second by differentiating the boundary data in $\bar{x}$ we get

$$\partial_x^\alpha \mathbf{u}(0) = 0, \qquad \text{for any} \quad \alpha \quad \text{with} \quad \alpha_\ell = 0.$$

Then for α with $\alpha_\ell = 1$ we calculate using the PDE (4.1) (denoting $\alpha' = (\alpha_1, \dots, \alpha_{\ell-1}, 0)$):

$$\partial_x^\alpha \mathbf{u} = \sum_{j=1}^{\ell-1} \partial_x^{\alpha'} \left(\mathbf{b}_j(\mathbf{u}, \bar{x}) \frac{\partial \mathbf{u}}{\partial x_j} \right) + \partial_x^{\alpha'} \mathbf{b}_0(\mathbf{u}, \bar{x})$$

which yields at $x = 0$ (using the previous step):

$$\partial_x^\alpha \mathbf{u}(0) = 0 + \left(\partial_x^{\alpha'} \mathbf{b}_0(\mathbf{u}, \bar{x}) \right)_{|x=0} = \left(D_2^{\alpha'} \mathbf{b}_0 \right)(0,0)$$

where D_2 means the partial derivatives according the second argument of $\mathbf{b}_0$.

Then for α with $\alpha_\ell = 2$ we calculate again (denoting $\alpha' = (\alpha_1, \dots, \alpha_{\ell-1}, 1)$):

$$\partial_x^\alpha \mathbf{u} = \sum_{j=1}^{\ell-1} \partial_x^{\alpha'} \left(\mathbf{b}_j(\mathbf{u}, x') \frac{\partial \mathbf{u}}{\partial x_j} \right) + \partial_x^{\alpha'} \mathbf{b}_0(\mathbf{u}, x')$$

which yields at $x = 0$:

$$\partial_x^\alpha \mathbf{u}(0) = \dots = \text{polynomial}(\partial \mathbf{b}_j(0,0), \partial \mathbf{u}(0))$$

where is in the RHS it only involves derivatives of $\mathbf{u}$ with $\alpha_\ell \leq 1$, which then can be expressed in terms of derivatives of $\mathbf{b}_j$ again.

We can continue the calculation inductively, and prove by induction that there are universal (independent of $\mathbf{u}$) polynomials with integer non-negative coefficients so that

$$\frac{\partial^\alpha \mathbf{u}_i}{\partial x^\alpha} = p_{\alpha,i}\left(D_2^\beta \mathbf{b}_j(0,0), |\beta| \leq |\alpha| - 1, 0 \leq j \leq \ell - 1\right).$$

This means: (1) all derivatives at the Cauchy hypersurface can be determined thanks to the coefficients and boundary data (here zeros), (2) we can amplify the idea of Cauchy to provide bounds on these derivatives.

Unlike the case of ODEs we do not have a priori a solution to the PDE. We thus shall construct it by defining the Taylor series

$$u(x_1, \dots, x_\ell) = \sum_{\alpha \geq 0} \frac{\mathbf{x}^\alpha}{\alpha!} \partial_x^\alpha \mathbf{u}(0), \quad \mathbf{x}^\alpha := x_1^{\alpha_1} \dots x_\ell^{\alpha_\ell}, \ \alpha! := \alpha_1! \dots \alpha_\ell! \tag{4.2}$$

If the radius of convergence of the entire series is non-zero, then it defines a real analytic function in some open set $\mathcal{U}_0$ including zero, which satisfies all the previous equations on the derivatives at zero. Moreover the PDE can be rewritten by composition of real analytic functions as another entire series

$$\frac{\partial \mathbf{u}}{\partial x_\ell} - \sum_{j=1}^{\ell-1} \mathbf{b}_j(\mathbf{u}, \bar{x}) \frac{\partial \mathbf{u}}{\partial x_j} + \mathbf{b}_0(\mathbf{u}, \bar{x}) =: \mathbf{H}(x) = \sum_{\alpha \geq 0} \frac{\mathbf{x}^\alpha}{\alpha!} \mathbf{h}_\alpha$$

for some coeffecients $\mathbf{h}_\alpha \in \mathbb{R}$, with a non-zero radius of convergence. Hence by reducing if necessary the open set, we have in $\mathcal{U}_0$ that (1) u is defined by the previous Taylor series, and (2) the equation writes $\mathbf{H}(x) = 0$ with some well-defined real analytic function $\mathbf{H}(x)$. But from the previous equations for the derivatives at zero, we have $\mathbf{h}_\alpha = 0$ for all $\alpha \geq 0$, and therefore $\mathbf{H}(x) = 0$ in the open set. We hence deduce that u satisfies the PDE in $\mathcal{U}_0$.

It remains therefore to prove that the initial Taylor series for u has a non-zero radius of convergence. We perform then the same argument as for ODEs with the majorant function

$$\mathbf{b}_j^* = \frac{Cr}{r - (x_1 + \dots x_{\ell-1}) - (z_1 + \dots + z_m)} \mathbf{M}_1, \quad j = 1, \dots, \ell - 1,$$

where $\mathbf{M}_1$ is the $m \times m$-matrix with 1 in all entries, and

$$\mathbf{b}_0^* = \frac{Cr}{r - (x_1 + \dots x_{\ell-1}) - (z_1 + \dots + z_m)} \mathbf{U}_1$$

where $\mathbf{U}_1$ is the m-vector with 1 in all entries, resulting in the solution

$$\mathbf{v} = \frac{1}{m\ell} \left(r - (x_1 + \dots + x_{\ell-1})\right) - \left[\left(r - (x_1 + \dots + x_{\ell-1})\right)^2 - 2m\ell C r x_\ell\right]^{1/2} \mathbf{U}_1$$

which yields the analyticity in all variables by using the growth condition characterisation of analyticity, as in the fourth ODE proof.

Let us decompose this last calculations into several steps, given in exercises.

EXERCISE 17. *Using the exercise 13 on all entries of* $\mathbf{b}_j$, $j = 0, \dots, \ell - 1$ *(which depend on* $m + \ell - 1$ *variables), find* $C, r > 0$ *so that*

$$g(z_1, \dots, z_m, x_1, \dots, x_{\ell-1}) = \frac{Cr}{r - (x_1 + \dots x_{\ell-1}) - (z_1 + \dots + z_m)}$$

is a majorant of all these entries.

EXERCISE 18. *Defining* $\mathbf{b}_j^* = g\mathbf{M}_1$, $j = 1, \dots, \ell - 1$, *and* $\mathbf{b}_0^* = g\mathbf{U}_1$, *check that the solution* $\mathbf{v} = (v_1, \dots, v_m)$ *to*

$$\begin{cases} \dfrac{\partial \mathbf{v}}{\partial x_\ell} = \displaystyle\sum_{j=1}^{\ell-1} \mathbf{b}_j^*(\mathbf{v}, x') \frac{\partial \mathbf{u}}{\partial x_j} + \mathbf{b}_0^*(\mathbf{v}, x') \\ \mathbf{v} = 0 \quad \text{on} \quad \Gamma, \end{cases}$$

can be searched in the form $v_1 = \dots = v_m =: w$, *and*

$$w = w(x_1 + x_2 + \dots + x_{\ell-1}, x_\ell) = w(\xi, x_\ell), \quad \xi := x_1 + \dots + x_{\ell-1}.$$

Then it reduces the problem to solving the following scalar simple transport equation (relabeling $x_\ell = t$ for conveniency)

$$\partial_t w = \frac{Cr}{r - \xi - \gamma_1 w} \left(\gamma_2 \partial_\xi w + 1 \right), \quad w(\xi, 0) = 0, \quad t, \xi \in \mathbb{R}. \tag{4.3}$$

EXERCISE 19. *Show that the* w *defining the solution* $\mathbf{v}$ *to the majorant problem above satisfies the equation* (4.3) *with* $\gamma_2 = (\ell - 1)m$ *and* $\gamma_1 = m$.

Finally we can solve the equation (4.3) by the so-called *characteristic method* (which we shall study in much more details in the chapter on hyperbolic equations). Let us sketch the method in this case: if we can find $\xi(t)$ and $\eta(t)$ solving

$$\begin{cases} \xi'(t) = \dfrac{-Cr\gamma_2}{r - \xi(t) - \gamma_1 \eta(t)}, \quad \xi(0) = \xi_0, \\ \eta'(t) = \dfrac{Cr}{r - \xi(t) - \gamma_1 \eta(t)}, \quad \eta(0) = \eta_0, \end{cases}$$

then if we set $\eta_0 = 0$ and now define $w(\xi, t)$ by the implicit formula $w(\xi(t), t) = \eta(t)$, it solves by the chain-rule

$$\eta'(t) = (\partial_t w)(\xi(t), t) + \xi'(t)(\partial_\xi w)(\xi(t), t) = \frac{Cr}{r - \xi(t) - \gamma_1 \eta(t)}$$

which writes

$$(\partial_t w)(\xi(t), t) - \frac{Cr\gamma_2}{r - \xi - \gamma_1 w(\xi(t), t)} (\partial_\xi w)(\xi(t), t) = \frac{Cr}{r - \xi(t) - \gamma_1 w(\xi(t), t)}$$

with the initial data $w(\xi_0, 0) = \eta_0 = 0$. This is exactly the desired equation at the point $(t, y(t))$. Hence as long as the map $\xi_0 \mapsto \xi(t)$ is invertible and smooth, we have a solution to the original PDE problem.[2]

[2]The first time where this maps stops being invertible is called a *caustic* of *shock wave* depending on the context, and will be studied in details in the chapter on hyperbolic equations.

Therefore let us solve locally in time the ODEs for $\xi(t)$ and $\eta(t)$. Since obviously $\xi'(t) + \gamma_2 \eta'(t) = 0$, we deduce the key a priori relation

$$\forall\, t \geq 0, \quad \xi(t) + \gamma_2 \eta(t) = \xi_0.$$

Whe thus replace $\xi(t)$ in the ODE for $\eta(t)$:

$$\eta'(t) = \frac{Cr}{r - \xi_0 + (\gamma_2 - \gamma_1)\eta(t)}, \quad \eta(0) = 0.$$

Here observe that $\gamma_2 \geq \gamma_1$ (remember that $\ell \geq 2$). If $\gamma_1 = \gamma_2$ then

$$w(\xi(t), t) = \eta(t) = \frac{Crt}{r - \xi_0}, \quad \xi(t) = \xi_0 - \frac{C\gamma_2 rt}{r - \xi_0}$$

which provides analyticity of the solution and concludes the proof. If $(\gamma_2 - \gamma_1) = (\ell - 2)m > 0$, then we calculate (using $\eta(0) = 0$ to decide on the root)

$$\eta(t) = w(\xi(t), t) = \frac{1}{\gamma_1 + \gamma_2} \left((r - \xi(t)) - \sqrt{(r - \xi(t))^2 - 2(\gamma_1 + \gamma_2)Crt} \right)$$

which gives

$$w(\xi, t) = \frac{1}{\ell m} \left((r - \xi) - \sqrt{(r - \xi)^2 - 2\ell m Crt} \right)$$

and concludes the proof.

Exercise 20. *Check that the previous formula for w indeed provides a solution.*

5. Examples, counter-examples, and classification

5.1. Failure of the Cauchy-Kovalevaskaya Theorem and evolution problems. If we consider the heat equation with initial conditions (this counter-example is due to S. Kovalevskaya)

$$\partial_t u = \partial_x^2 u, \quad u = u(t, x), \; (t, x) \in \mathbb{R}^2$$

around the point $(0, 0)$, with the initial condition

$$u(0, x) = g(x)$$

we have, in the previous setting, $\Gamma = \{t = 0\}$, and the normal unit vector in $\mathbb{R}^2$ is simply $(1, 0) = \mathbf{e}_1$. The non-characteristic condition writes $a_{k,0} \neq 0$, where k is the order of the equation (here $k = 2$), which is not true here (independently of the choice of the boundary data g). Hence the initial value problem for the heat equation is always characteristic. This reflects the fact that the equation cannot be reversed in time, or in other words, the Cauchy problem is ill-posed for negative times. In particular, consider the following initial data (considered by S. Kovalevskaya)

$$g(x) = \frac{1}{1 + x^2}$$

which are clearly analytic. Then let us search for an analytic solution

$$u(t, x) = \sum_{m,n \geq 0} a_{m,n} \frac{t^m}{m!} \frac{x^n}{n!}.$$

Then the PDE imposes the following relation on the coefficients

$$\forall\, m, n \geq 0, \quad a_{m+1,n} = a_{m,n+2},$$

with the initialization

$$\forall\, n \geq 0, \quad a_{0,2n+1} = 0, \quad a_{0,2n} = (-1)^n (2n)!$$

We deduce that

$$\forall\, m, n \geq 0, \quad a_{m,2n+1} = 0$$

using the inductive relation, and then

$$\forall\, m, n \geq 0, \quad a_{m,2n} = (-1)^{m+n} (2(m+n))!$$

Now since

$$\frac{(2(m+n))!}{m!(2n)!} = \frac{(4n)!}{n!(2n)!} \sim \sqrt{\frac{1}{\pi n}} n^n 2^{2n} e^{-n} \longrightarrow +\infty$$

(using the Stirling formula) as $m = n \to \infty$, in a way which cannot be damped by geometric factors $t^n x^{2n}$, and we deduce that the entire series defining u has a radius of convergence equal to zero.

In words, what we have exploited in this proof is that the equation implies $\partial_t^k u = \partial_x^{2k} u$ for all $k \in \mathbb{N}$, and the strongest bound on the x-derivatives for general analytic initial data $u(0,x)$ are of the form $\mathrm{cst}(2k)!/r^k$, whereas on the LHS the t-derivatives should grow at most, in order to recover analyticity, as $\mathrm{cst} k!/\rho^k$, and these two things are contradictory. Hence by equating more spatial derivatives on the right hand side with less derivatives on the left hand side, one generates faster growth in the right hand side than is allowed for the left hand side to be analytic.

This example shows how the notion of characteristic boundary condition highlights some key physical and mathematical aspects of the equation at hand. It can easily be seen that a necessary conditions for an *evolution problem*

$$\partial_t^k u = \sum_{|\alpha| = l} a_\alpha \partial_x^\alpha u$$

to be solvable by Cauchy-Kovalevskaya Theorem for the Cauchy surface $\Gamma = \{t = 0\}$, is that $l \leq k$.

EXERCISE 21. *Check the last point, and formulate a similar conditions for systems.* Hint*: For k_i time derivatives on the i-th component, no spatial derivatives on this component should be of order higher than k_i.*

5.2. Principal symbol and characteristic form. Let P be a scalar linear differential operator of order k:

$$Pu := \sum_{|\alpha| \leq k} a_\alpha(x) \partial_x^\alpha u, \quad u = u(x),\ x \in \mathbb{R}^\ell.$$

For convenience let us assume here that the $a_\alpha(x)$ are smooth functions.

Then the *total symbol* of the operator is defined as

$$\sigma(x,\xi) := \sum_{|\alpha| \leq k} a_\alpha(x) \xi^\alpha, \quad \xi^\alpha := \xi_1^{\alpha_1} \dots \xi_\ell^{\alpha_\ell}$$

and the *principal symbol* of the operator is defined as

$$\sigma_p(x,\xi) := \sum_{|\alpha|=k} a_\alpha(x)\xi^\alpha,$$

where $\xi \in \mathbb{C}$. This principal symbol is also called the *characteristic form* of the equation.

REMARK 5.1. *For geometric PDEs in an open set $\mathcal{U}$ of some manifold $\mathcal{M}$, the principal symbol is better thought of as a function on the cotangent bundle: $\sigma_p : T^*\mathcal{U} \to \mathbb{R}$.*

EXERCISE 22. *Show that the principal symbol is an* homogeneous function of degree *k in ξ, i.e.*

$$\sigma_p(x,\lambda\xi) = \lambda^k\sigma_p(x,\xi), \quad \lambda \in \mathbb{R}.$$

Then the non-characteristic condition at $x \in \Gamma$ becomes in this context

$$\sigma_p(x,\mathbf{n}(x)) \neq 0.$$

REMARK 5.2. *With a more geometric intrinsic formulation, we could say that for any $\xi \in T^*_x\mathcal{U} \setminus \{0\}$, $x \in \Gamma \subset \mathcal{U}$, with $\langle \xi, w\rangle = 0$ for all $w \in T^*_x\mathcal{U}$ tangent to Γ at $x \in \mathcal{U}$, then $\sigma_p(x,\xi) \neq 0$.*

We also introduce the *characteristic cone*[3] of the PDE at $x \in \mathbb{R}^\ell$:

$$\mathcal{C}_x := \left\{\xi \in \mathbb{R}^\ell \ : \ \sigma_p(x,\xi) = 0\right\}.$$

Then a surface is characteristic at a point if the normal to the surface at that point belongs to the characteristic cone at the same point.

5.3. The main linear PDEs and their characteristic surfaces. Let us go through the main linear PDEs and study their characteristic surfaces. We shall study in the next chapters some paradigmatic examples:

- Laplace's equation and Poisson's equation

$$\Delta u = 0 \quad \text{or} \quad \Delta u = f, \quad u = u(x_1,\dots,x_\ell), \qquad \left(\Delta = \sum_{j=0}^{\ell} \partial^2_{x_j}\right).$$

  In this case as we discussed the characteristic form is $\sigma_p(x,\xi) = |\xi|^2$ and the characteristic cone is $\mathcal{C}_x = \{0\}$ for any $x \in \mathbb{R}^\ell$, and any real surface cannot be characteristic to the Laplace equation. Equation without real characteristic surfaces are called *elliptic equations.*
- The wave equation

$$\Box u = 0, \quad u, \quad u = u(x_1,\dots,x_\ell), \qquad \left(\Box := -\partial^2_{x_\ell} + \sum_{j=1}^{\ell-1} \partial^2_{x_j} = -\partial^2_{x_\ell} + \Delta_{x_1,\dots,x_{\ell-1}}\right).$$

  The wave equation is obtained from the Laplace equation by the so-called *Wick rotation* $x_\ell \mapsto ix_\ell$. Its characteristic form is $\sigma_p(x,\xi) = \xi_1^2+\dots+\xi_{\ell-1}^2-\xi_\ell^2$

[3]The name "cone" is related to the homogeneity property described above: the characteristic cone is hence invariant by multiplication by a real number.

and its characteristic cone is the so-called *light cone* $\mathcal{C}_x = \{\xi_\ell^2 = \xi_1^2 + \cdots + \xi_{\ell-1}^2\}$ for any $x \in \mathbb{R}^\ell$. Any surface whose normal makes an angle $\pi/4$ with the direction $\mathbf{e}_\ell$ is a characteristic surface. The variable $x_\ell = t$ represents the time.

- The heat equation

$$\partial_{x_\ell} u - \Delta_{x_1,\ldots,x_{\ell-1}} u = 0, \quad u = u(x_1, \ldots, x_\ell), \quad \left(\Delta_{x_1,\ldots,x_{\ell-1}} := \sum_{j=1}^{\ell-1} \partial_{x_j}^2 \right).$$

  Its characteristic form is $\sigma_p(x,\xi) = \xi_1^2 + \cdots + \xi_{\ell-1}^2$ and its characteristic cone is $\mathcal{C}_x = \{\xi_1 = \cdots = \xi_{\ell-1} = 0\}$ for any $x \in \mathbb{R}^\ell$, and so the characteristic surfaces are the horizontal planes $\{x_\ell = \mathrm{cst}\}$ (hence corresponding to an initial condition). The variable $x_\ell = t$ represents the time.
- The Schrödinger equation is a system of PDEs (the two components of the solution are written as a complex number)

$$i\partial_t u + \Delta u = 0, \quad u = u(t, x_1, \ldots, x_\ell) \in \mathbb{C}, \quad \left(\Delta := \sum_{j=1}^{\ell-1} \partial_{x_j}^2 \right).$$

  The Schrödinger equation is obtained from the heat equation by the *Wick rotation* $x_\ell \mapsto ix_\ell$. Its characteristic form is again $\sigma_p(x,\xi) = \xi_1^2 + \cdots + \xi_{\ell-1}^2$ and its characteristic cone is again $\mathcal{C}_x = \{\xi_1 = \cdots = \xi_{\ell-1} = 0\}$ for any $x \in \mathbb{R}^\ell$, and so the characteristic surfaces are again the horizontal planes $\{x_\ell = \mathrm{cst}\}$ (hence corresponding to an initial condition).
- The transport (including Liouville) equation

$$\sum_{j=1}^{\ell} c_j(x) \partial_{x_j} u = 0, \quad u = u(x_1, \ldots, x_\ell).$$

Then its characteristic form is

$$\sigma_p(x,\xi) = \left(\sum_{j=1}^{\ell} c_j(x) \xi_j \right)$$

and its characteristic cone is

$$\mathcal{C}_x = \mathbf{c}(x)^\perp, \quad \mathbf{c} = (c_1, \ldots, c_\ell)$$

for any $x \in \mathbb{R}^\ell$. This means that a characteristic surface is everywhere tangent to $\mathbf{c}(x)$. Then all our transport equation tells us is the behaviour of u along the characteristic surface, and what u does in the transversal direction is completely "free". This means that the existence is lost unless the initial condition on the surface satisfies certain constraints, and if a solution exists, it will not be unique. The situation is reminiscent to solving the linear system $Ax = b$ with a non-invertible square matrix A.

REMARK 5.3. *The equations all share the property that they are* linear, *and they often occur when linearising more complicated equations which play a role in Mathematical Physics, or by other types of limits.*

5.4. What is wrong with analyticity? First remark that we could have exposed the theory of analytic solutions to PDEs in $\mathbb{C}^\ell$ with the real analytic theory as a special case. We rather emphasized the real analytic case which is more common in practice.

Let us explain why if we allowed only analytic solutions, we would be missing out on most of the interesting properties of partial differential equations. For instance, since analytic functions are completely determined by its values on any open set however small, it would be extremely cumbersome, if not impossible, to describe phenomena like wave propagation, in which initial data on a region of the initial surface are supposed to in influence only a specific part of spacetime. A much more natural setting for a differential equation would be to require its solutions to have just enough regularity for the equation to make sense. For example, the Laplace equation $\Delta u = 0$ already makes sense for twice differentiable functions. Actually, the solutions to the Laplace equation, i.e. harmonic functions, are automatically analytic, which has a deep mathematical reason that could not be revealed if we restricted ourselves to analytic solutions from the beginning. In fact, the solutions to the Cauchy–Riemann equations, i.e. holomorphic functions, are analytic by the same underlying reason, and complex analytic functions are nothing but functions satisfying the Cauchy–Riemann equations. From this point of view, looking for analytic solutions to a PDE in $\mathbb{R}^\ell$ would mean coupling the PDE with the Cauchy–Riemann equations and solving them simultaneously in $\mathbb{R}^{2\ell}$. In other words, if we are not assuming analyticity, $\mathbb{C}^\ell$ is better thought of as $\mathbb{R}^{2\ell}$ with an additional algebraic structure. Hence the real case is more general than the complex one, and from now on, we will be working explicitly in real spaces such as $\mathbb{R}^\ell$, unless indicated otherwise.

As soon as we allow non-analytic data and/or solutions, many interesting questions arise surrounding the Cauchy-Kovalevskaya theorem. First, assuming a setting to which the Cauchy-Kovalevskaya theorem can be applied, we can ask if there exists any (necessarily non-analytic) solution other than the solution given by the Cauchy-Kovalevskaya theorem. In other words, is the uniqueness part of the Cauchy-Kovalevskaya theorem still valid if we now allow non-analytic solutions?

For linear equations an affirmative answer is given by *Holmgren's uniqueness theorem*. For ODEs the Cauchy-Kovalevskaya theorem can be interpreted merely as an a posteriori regularity theorem: all solutions constructed by Picard-Lindelöf are indeed analytic as soon as the coefficients are. A natural question is whether the same occurs for PDEs. The answer is yes for linear systems, and this is the content of Holmgren's uniqueness Theorem (which can be again interpreted as an a posteriori regularity result).

The second question is whether existence holds for smooth but non-analytic data, and the answer is negative in general. A large class of counter-examples can be constructed, by using the fact that some equations, such as the Laplace and the Cauchy-Riemann equations, have only analytic solutions, therefore their boundary data, as restrictions of the solutions to an analytic hypersurface, cannot be non-analytic. Hence such equations with non-analytic initial data do not have solutions. Remark that in some cases, this can be interpreted as one having "too many" initial conditions that make the problem overdetermined, since in those cases the situation can be remedied by removing some of the initial conditions. For example, with sufficiently regular closed surfaces as initial surfaces, one can remove either one of the two Cauchy data in the Laplace equation, arriving at the Dirichlet or Neumann problem.

The third question is whether we still have existence of some solutions when we relax the analyticity assumption on the coefficients (think to the passage from Picard-Lindelöf to Cauchy-Peano in the ODE theory). Hans Lewy's celebrated counter-example of 1957 (and subsequent examples in the same line) exhibited examples of PDEs where the inhomogeneous part of a linear equation is smooth but not analytic and which have no solutions, regardless of boundary data. The lesson to be learned from these examples is that the existence theory in a non-analytic setting is much more complicated than the corresponding analytic theory, and in particular one has to carefully decide on what would constitute the boundary data for the particular equation.

Indeed, there is an illuminating way to detect the poor behaviour of some equations discussed in the previous paragraph with regard to the Cauchy problem, entirely from within the analytic setting, that runs as follows. Suppose that in the analytic setting, for a generic initial datum ψ it is associated the solution $u = S(\psi)$ of the equation under consideration, where $S : \psi \mapsto u$ is the solution map. Now suppose that the datum is non-analytic, say, only continuous. Then by the Weierstrass approximation theorem, for any $\varepsilon > 0$ there is a polynomial ψ_ε that is within an ε distance from ψ. Taking some sequence $\varepsilon \to 0$, if the solutions $u_\varepsilon = S(\psi_\varepsilon)$ converge locally uniformly to a function u, we could reasonably argue that u is a solution (in a generalized sense) of our equation with the (non-analytic) datum ψ. The counter-examples from the preceding paragraph suggest that in those cases the sequence u_ε cannot converge. Actually, the situation is much worse, as the following example due to J. Hadamard shows.

Consider the Cauchy problem for the Laplace equation

$$\partial_{tt}^2 u + \partial_{xx}^2 u = 0, \quad u(x,0) = 0, \quad \partial_t u(x,0) = a_\omega \cos(\omega x)$$

for some parameter $\omega > 0$, whose solution is explicitely given by

$$u(x,t) = a_\omega \sinh(\omega t) \cos(\omega x).$$

Then if we choose $a_\omega = 1/\omega$ and $\omega >> 1$, we see that the initial data is small: $u(x,0) = \mathcal{O}(1/\omega)$, $\partial_t u(x,0) = 0$, whereas the solution grows arbitrarily fast as ω tends to infinity: $u(x,1) = O(e^w/w)$. Hence the relation between the solution and the Cauchy data becomes more and more difficult to invert as we go to higher and higher frequencies $\omega \to \infty$. For instance if the initial data is the error of an approximation of non-analytical data in the uniform norm as $\omega \to \infty$, then the solutions with initial data given by the approximations diverge unless a_ω and b_ω decay faster than exponential. But functions that can be approximated by analytic functions with such small error form a severely restricted class, being between the smooth functions C^∞ and the analytic functions.

Exercise 23. *In this exercise we give a slightly amplified version of the example of J. Hadamard: consider the problem*

$$\partial_{tt}^2 u + \partial_{xx}^2 u = 0, \quad u(x,0) = \phi(x), \quad \partial_t u(x,0) = \psi(x).$$

For a given $\varepsilon > 0$ and an integer $k > 0$, construct initial data ϕ and ψ so that

$$\|\phi\|_\infty + \left\|\phi^{(1)}\right\|_\infty + \cdots + \left\|\phi^{(k)}\right\|_\infty + \|\psi\|_\infty + \left\|\psi^{(1)}\right\|_\infty + \cdots + \left\|\psi^{(k)}\right\|_\infty < \varepsilon$$

and

$$\|u(\cdot,\varepsilon)\| \geq \frac{1}{\varepsilon}.$$

Repeat the exercise with the condition on the initial data replaced by

$$\forall\, k \geq 0, \quad \left\|\phi^{(k)}\right\|_\infty + \left\|\psi^{(k)}\right\|_\infty < \varepsilon.$$

Let us contrast the previous (elliptic) example with the following (hyperbolic) one: consider the Cauchy problem for the wave equation

$$\partial^2_{tt} u - \partial^2_{xx} u = 0, \quad u(x,0) = \phi(x), \quad \partial_t u(x,0) = \psi(x),$$

whose solution is given by d'Alembert's formula

$$u(t,x) = \frac{\phi(x-t)+\phi(x+t)}{2} + \frac{1}{2}\int_{x-t}^{x+t} \psi(y)\,\mathrm{d}y.$$

We deduce that

$$|u(x,t)| \leq \sup_{[x-t,x+t]} |\phi| + |t| \sup_{[x-t,x+t]} |\psi|$$

showing that small initial data lead to small solutions (and also showing domain of dependency). The explicit solution constructed can also be shown to be the unique solution by energy methods (see later in the next chapters).

This is in response to these considerations that Hadamard introduced the concept of *well-posedness* of a problem that we have introduced in the first chapter.

5.5. Basic classification. Roughly speaking, the (1) *hyperbolic*, (2) *elliptic*, (3) *parabolic*, and (4) *dispersive* classes arise as one tries to identify the equations that are similar to, and therefore can be treated by extensions of techniques developed for, the (1) wave (and transport), the (2) Laplace (and the Cauchy-Riemann), the (3) heat, and the (4) Schrödinger equations, respectively.

Indeed, the idea of *hyperbolicity* is an attempt to identify the class of PDEs for which the Cauchy-Kovalevskaya theorem can be rescued in some sense when we relax the analyticity assumption. The simplest examples of hyperbolic equations are the wave and transport equations. In contrast, trying to capture the essence of the poor behaviour of the Laplace and Cauchy-Riemann equations in relation to their Cauchy initial-time problems leads to the concept of ellipticity. Features of elliptic equations are: no real characteristic surfaces, smooth solutions for smooth data, overdeterminacy of the Cauchy data hence boundary value problems, and being associated to stationary phenomena.

The class of *parabolic equations* is a class for which the evolution problem is well-posed for positive times, but is ill-posed for negative times. The initial condition is characteristic and the Cauchy-Kovalevskaya Theorem fails. The informations is transmitted at infinite speed, and there is instantaneous regularisation: the solution becomes analytic for positive times. The latter phenomenon is extremely important and obviously cannot be captured in analytic setting.

The class of *dispersive equations* is a class which is close to transport-wave equations in the sense that their "extension" in the space-frequency phase space has the structure of a transport equation. Moreover the evolution is reversible and well-posed at the level of the linear equation, but the Cauchy-Kovalevskaya is not adapted again. And they transport information at finite speed. However let us discuss the crucial difference between these two classes which justifies the name “dispersive”.

Consider a *plane wave* solution $u(t,x) = \cos(k(t+x))$ to the wave equation

$$\partial_{tt}^2 u = \partial_{xx}^2 u, \quad t, x \in R,$$

with the initial data $u(0,x) = \cos(kx)$, $\partial_t u(0,x) = 0$. Then the information travels at speed 1, whatever the frequency $k \in \mathbb{R}$ of the wave. Next consider again a plane wave solution $u(t,x) = e^{i(kx-|k|^2 t)}$ to the Schrödinger equation

$$i\partial_t u + \partial_{xx}^2 u = 0, \quad t, x \in R,$$

with the initial data $u(0,x) = e^{ikx}$. The information then travels at speed $|k|$ which now *depends* on the frequency! In physics words, the *dispersion relation* is $\omega(k) = k$ for the wave equation, and $\omega(k) = k^2/2$ for the Schrödinger equation. In the first case, the dispersion relation is linear and there is no wave packet dispersion, while there is in the second case. This dispersive feature results in numerous mathematical consequences (like so-called Strichartz estimates) which are key to many Cauchy theorems...

Let us remark to finish with that they are of course many equations that are of mixed type, e.g. the Euler-Tricomi equation $\partial_{xx}^2 u = x\partial_{yy}^2 u$ for $u = u(x,y)$ which is hyperbolic in the region $\{x > 0\}$ and elliptic in the region $\{x < 0\}$. There are variants of these classes where some properties are weakened, e.g. the hypoelliptic equations as pionnered by Kolmogorov and Hörmander.

6. Historical notes

This Cauchy-Kovalevskaya theorem was first proved by A. Cauchy in 1842 on first order quasilinear evolution equations, and formulated in its most general form by S. Kovalevskaya in 1874. At about the same time, G. Darboux also reached similar results, although with less generality than Kovalevskaya's work. Both Kovalevskaya's and Darboux's papers were published in 1875, and the proof was later simplified by E. Goursat in his influential calculus textbook around 1900. Nowadays these results are collectively known as the *Cauchy-Kovalevskaya Theorem*.