Partial Differential Equations – A. Visintin – 2018

PROGRAM OF THE COURSE

Outline: The modern theory of PDEs is based upon several analytical tools. In particular it has been and is still developed in function spaces, and thus rests upon functional analysis, besides some knowledge of ordinary differential equations.

This course will illustrate the main properties of linear PDEs of second order, keeping an eye on mathematical-physical applications. It will also introduce the Sobolev spaces, and formulate boundary- and initial-value problems for second-order PDEs.

Prerequisites (in parentheses the corresponding courses in Trento):

Differential and Integral Calculus, with Fourier series and ODEs *(Analisi I, II e III).*

Measure theory and Lebesgue integration *(Analisi III).*

Linear algebra *(Geometria I).*

General topology *(Geometria II).*

Elementary notions on ODEs and PDEs *(Fisica-Matematica).*

Banach and Hilbert spaces, functional spaces, linear and continuous operators *(Analisi Funzionale).*

Fourier Analysis at the first semester and the parallel course *Advanced Analysis* are recommended.

Main Issues:

1. Basic Linear Second-Order PDEs [Renardy-Rogers; chap. 1]

Review of ordinary differential equations (ODEs). Existence and uniqueness for initial-value problems. Eigenvalue problem for the homogeneous boundary-value problem for equations like $y'' + \lambda y = 0$.

Linear systems of ODEs. Fundamental solution and matrix function. Formula of variation of parameters. Gronwall's lemma.

Laplace/Poisson equation: boundary conditions. Solution by separation of variables.

Green identities. Variational formulation of the Laplace/Poisson equation. Fundamental lemma of the calculus of variations. Maximum principle.

Derivation of the diffusion and wave equations from balance laws and constitutive relations.

Heat equation: boundary and initial conditions. Solution by separation of variables. Backward heat equation. Energy inequality.

Duhamel principle for evolutionary PDEs. Variational formulation. Energy inequality. Maximum principle.

Wave equation: boundary and initial conditions. Solution by separation of variables. D'Alembert solution of the wave equation. Domain of dependence and domain of influence. Variational formulation. Energy conservation. Schrödinger equation.

Comparison between the qualitative properties of the heat, wave and Schrödinger equations.

Weak and strong maximum form of the maximum principle for linear elliptic and parabolic equations in nondivergence form in C^0. Monotone dependence of the solution on the data. Uniqueness of the solution. [Renardy-Rogers; chap. 4]

2. Characteristics [Renardy-Rogers; chap. 2], [Lecture Notes]

Multi-indices. Principal part and symbol of a linear differential operator. Classification of nonlinear PDEs: quasilinear, semilinear, fully-nonlinear equations.

Classification of linear second-order PDEs: elliptic, parabolic, hyperbolic.

Characteristics of linear and quasilinear equations.

Statement of the Cauchy-Kovalevskaya and Holmgren theorems (without proofs).

3. Sobolev Spaces [Lecture Notes]

Spaces of Hölder class.

Euclidean domains of Hölder class. Cone property.

Sobolev spaces of positive order. Characterization of the dual.

Extension operators. Theorem of Caldéron-Stein. The method of extension by reflection.

Density results: interior approximation (Mayer-Serrin theorem) and exterior approximation.

Sobolev inequality and imbedding between Sobolev spaces. Morrey theorem. Sobolev and Morrey indices. Rellich-Kondrachov compactness theorem.

L^p- and Sobolev spaces on manifolds. Traces.

4. Weak Formulation of Second-Order PDEs

Elliptic operators. Classical, strong and weak formulation.

Operators in divergence or nondivergence form. Conormal derivative.

Strong and weak formulation of second-order elliptic equations in divergence form, coupled with boundary-conditions (i.e., either Dirichlet or Neumann or mixed or periodic conditions).

Strong and weak formulation of the Cauchy problem for second-order parabolic/hyperbolic equations in divergence form, coupled with initial- and boundary-conditions.

Elliptic problems governed by a symmetric operator are equivalent to minimization problems.

Theorem of Lax-Milgram. Application to the analysis of PDEs in divergence form.

Reference textbook of the course:

M. Renardy, R. Rogers: An introduction to partial differential equations. Springer-Verlag, New York, 2004

Teacher's lecture notes (partly already available on the web).

Complementary textbooks:

Yu.V. Egorov, M.A. Shubin: Foundations of the classical theory of partial differential equations. Springer, Berlin 1992

L.C. Evans: Partial differential equations. American Mathematical Society, Providence, RI, 2010

D. Griffiths, J. Dold, D. Silvester: Essential partial differential equations. Springer, Cham, 2015

S. Salsa: Equazioni a derivate parziali: metodi, modelli e applicazioni. Springer Italia, Milano 2003 (English edition also available)

S. Salsa, G. Verzini: Equazioni a derivate parziali: complementi ed esercizi. Springer Italia, Milano 2005

S. Salsa et al.: A primer on PDEs. Springer Italia, Milano 2013. (This is a modified English translation of the above text of Salsa; it also includes the correction of several exercises.)

Modality of exam:

Written and oral examinations.

Contents of this part of the Lecture Notes:

1. Characteristics 2. Maximum Principle

By [Ex] we mean that the justification of a statement is left as an exercise.
By [] we mean that the justification is omitted, and is more than just an exercise.
By [ReRo] we indicate the book of Renardy and Rogers.
The asterisk is used to label complements in the text (in particular this applies to some slightly technical arguments), as well as more demanding exercises.

1 Characteristics

This is our plan in this section:
Symbol and classification of PDEs. The Cauchy-Kovalevskaya theorem, for equations in either normal or nonnormal form. Characteristics of PDEs of any order in $\mathbb{R}^N$ (in particular for second-order equations in $\mathbb{R}^2$). Characteristic hypersurfaces and characteristic curves. Integration of first-order PDEs via the method of characteristics.

1.1 PDEs and Symbols

Multi-indices. Let us set $D_j := \partial/\partial x_j$ for any i and $D = (D_1, ..., D_N)$ $(= \nabla)$. For any **multi-index** $\alpha := (\alpha_1, \dots, \alpha_N) \in \mathbb{N}^N$, let us set

$$\boxed{\begin{array}{l} D^\alpha f := D_1^{\alpha_1} \cdots D_N^{\alpha_N} f \qquad \forall f \in C^{|\alpha|}(\mathbb{R}^N), \\ h^\alpha := h_1^{\alpha_1} \cdots h_N^{\alpha_N} \qquad \forall h = (h_1, ..., \cdots h_N) \in \mathbb{R}^N, \\ |\alpha| := \sum_{j=1}^N \alpha_j, \qquad \alpha! := \alpha_1! \cdots \alpha_N! \qquad \forall \alpha \in \mathbb{N}^N, \\ \binom{\alpha}{\beta} := \dfrac{\alpha!}{\beta!(\alpha-\beta)!} \qquad \forall \alpha, \beta \in \mathbb{N}^N, \ \beta \leq \alpha \ \text{(i.e., } \beta_j \leq \alpha_j \ \forall j). \end{array}} \tag{1.1}$$

Linear operators. For any integers $M, N, m \geq 1$, the most general system of M PDEs of order m is of the form

$$F(x, \{D^\alpha u(x)\}_{|\alpha| \leq m}) = 0 \qquad \forall x \in \Omega \text{ (open subset of } \mathbb{R}^N), \tag{1.2}$$

for a given function F and an unknown function u, with F, u taking values in $\mathbb{R}^M$. Of course if $M = 1$ this is reduced to a single equation for a scalar unknown.

The solution can be represented either explicitly as $u = u(x)$, or locally and implicitly as $\phi(x, u) = 0$, for some function $\phi : \mathbb{R}^N \times \mathbb{R}^M \to \mathbb{R}^M$ such that $D_u \phi(x, u) \neq 0$ everywhere.

A linear PDE has the form

$$L(x, D)u := \sum_{|\alpha| \leq m} c_\alpha(x) D^\alpha u(x) = f(x) \qquad \forall x \in \Omega, \tag{1.3}$$

for prescribed functions $\{c_\alpha\}_{|\alpha| \leq m}$ and f. The operator

$$L_p(x, D) := \sum_{|\alpha| = m} c_\alpha(x) D^\alpha \qquad \forall x \in \Omega \tag{1.4}$$

is called the **principal part** of the linear operator L.

The order of the PDE may depend on the point. For instance, a linear PDE of at most second order in $\mathbb{R}^N$ is of the form

$$L(x,D)u := \sum_{i,j=1}^{N} c_{ij}(x) D_{x_i} D_{x_j} u(x) + \sum_{i=1}^{N} c_i(x) D_{x_i} u(x) + c_0(x) u = f(x) \qquad \forall x \in \Omega, \tag{1.5}$$

for prescribed functions c_{ij}, c_i, c_0 and f. The equation is of second order where $c_{ij} \neq 0$ for at least one pair (i,j); it is of first order where $c_{ij} = 0$ for any i,j and $c_i \neq 0$ for at least an index i.

Classification of nonlinear PDEs. A PDE is called **semilinear** iff it can be written in the form

$$\sum_{|\alpha|=m} c_\alpha(x) D^\alpha u(x) = f(x, \{D^\beta u(x)\}_{|\beta|<m}) \qquad \forall x \in \Omega, \tag{1.6}$$

for prescribed functions $\{c_\alpha\}_{|\alpha|\leq m}$ and f. This means that the highest-order terms constitute a linear operator. A PDE is called **quasilinear** iff it can be written in the form

$$\sum_{|\alpha|=m} c_\alpha(x, \{D^\beta u(x)\}_{|\beta|<m}) D^\alpha u(x) = f(x, \{D^\beta u(x)\}_{|\beta|<m}) \qquad \forall x \in \Omega, \tag{1.7}$$

for prescribed functions $\{c_\alpha\}_{|\alpha|\leq m}$ and f. (In this case the highest-order terms thus do not constitute a linear operator.) The definition of principal part is extended to this quasilinear operator, by allowing the coefficients to depend upon derivatives of lower order than m:

$$L_p(x, \{D^\beta u(x)\}_{|\beta|<m}, D) := \sum_{|\alpha|=m} c_\alpha(x, \{D^\beta u(x)\}_{|\beta|<m}) D^\alpha \qquad \forall x \in \Omega. \tag{1.8}$$

A PDE is called **fully nonlinear** iff it is not quasilinear.

Remarks. (i) This classification also applies to equations in which the derivatives w.r.t. to some coordinates are missing in the principal part; this is the case e.g. for the diffusion equation (in which D_t^2 is missing). In this case one cannot apply the customary point of view, by which the selection of the appropriate boundary conditions and the qualitative behaviour of the solution only depends on the principal part of the PDE. For instance, for the diffusion equation either the initial- or the final-value problem is well-posed, depending on the occurrence of either $D_t u$ or $-D_t u$ (and this term is not comprised in the principal part).

(ii) This classification can be extended to systems; in this case $u(x) \in \mathbb{R}^M$ $(M > 1)$ for any x, and $c_\alpha \in \mathbb{R}^{M\times M}$ for any α. However for systems some care is needed in the classification, as it is discussed e.g. in [ReRo; Sect. 2.1.3].

Symbol of a linear operator. For a linear operator $L(x,D)$ of the form (1.3), we define the **symbol**

$$L(x, i\xi) = \sum_{|\alpha|\leq m} c_\alpha(x)(i\xi)^\alpha \qquad \forall x \in \Omega, \forall \xi \in \mathbb{R}^N \tag{1.9}$$

(notice that this is a function of $i\xi$), and its principal part (also named **principal symbol)**

$$L_p(x, i\xi) = \sum_{|\alpha|=m} c_\alpha(x)(i\xi)^\alpha \qquad \forall x \in \Omega, \forall \xi \in \mathbb{R}^N. \tag{1.10}$$

Notice that

$$L(x,D)e^{i\xi\cdot x} = L(x,i\xi)e^{i\xi\cdot x} \qquad \forall x \in \Omega, \forall \xi \in \mathbb{R}^N. \tag{1.11}$$

If $L(D) := \sum_{|\alpha|\le m} c_\alpha D^\alpha$ (with constant coefficients), then the notion of symbol has an obvious relation with the Fourier transform $\mathcal{F}$, since

$$\mathcal{F}(L(D)u) = \sum_{|\alpha|=m} c_\alpha \mathcal{F}\big(D^\alpha u\big) = \sum_{|\alpha|=m} c_\alpha (i\xi)^\alpha \mathcal{F}(u) = L(i\xi)\mathcal{F}(u).$$

For any $\xi \in \mathbb{R}^N$, the exponential function $\theta_\xi : x \mapsto e^{i\xi\cdot x}$ and the symbol $L(x, i\xi)$ are thus an eigenfunction and the corresponding eigenvalue of the operator $L(x, D)$.

The definition of principal symbol can be extended to quasilinear PDEs, by replacing $D^\alpha u$ by $(i\xi)^\alpha$ only in the derivatives of highest order, i.e., for $|\alpha| = m$. In this case the symbol also depends on $x, u(x)$ and on lower-order derivatives. For instance, the principal symbol of (1.7) reads

$$L_p(x, \{D^\alpha u(x)\}_{|\alpha|<m}, i\xi) = \sum_{|\alpha|=m} c_\alpha(x, \{D^\alpha u(x)\}_{|\alpha|<m})(i\xi)^\alpha \qquad \forall x \in \Omega, \forall \xi \in \mathbb{R}^N. \tag{1.12}$$

This can be justified via Fourier transform only if the operator does not explicitly depend on x (i.e., coefficients are constant).

1.2 The Cauchy-Kovalevskaya theorem

In this section we deal with scalar equations, i.e., single equations. The extension to vector equations, i.e., systems of equations, is however also feasible.

ODEs in normal form. Cauchy first proved existence and uniqueness of the solution for the initial-value problem

$$\begin{cases} y'(t) = f(y(t), t) & (t \in\,]-T, T[) \\ y(0) = y^0. \end{cases} \tag{1.13}$$

He assumed the function f to be **real analytic** (i.e., a real-valued function which is representable as a series of powers in a neighbourhood of each point of its domain), and searched for a real analytic solution of the form $y(t) = \sum_{n\ge 0} a_n t^n$. By calculating successive derivatives of $y(t)$ by means of $(1.13)_1$, restricting the ODE (= Ordinary Differential Equation) to $t = 0$, and using the initial condition, Cauchy identified the coefficients $\{a_n\}$ in cascade in terms of the data f, y^0. As by Taylor expansion $a_n = y^{(n)}(0)/n!$ for any $n \ge 0$, the first coefficients read as follows:

$$\begin{cases} a_0 = y(0) = y^0, \\ a_1 = y'(0) = f(y(t), t)\big|_{t=0} = f(y^0, 0), \\ 2a_2 = y''(0) = \{D_t f(y(t), t)\}\big|_{t=0} = D_1 f(y(t), t)\big|_{t=0} y'(0) + D_2 f(y(t), t)\big|_{t=0} \\ \qquad = D_1 f(y^0, 0) f(y^0, 0) + D_2 f(y^0, 0), \\ 6a_3 = y'''(0) = ... \\ ... \end{cases} \tag{1.14}$$

The last displayed equation is obtained by differentiating the ODE $(1.13)_1$ twice in time, and so on.

Cauchy then proved the convergence of the series $\sum_{n\ge 0} a_n t^n$ via the so-called method of *majorants;* this consists in exhibiting a convergent series $\sum_{n\ge 0} b_n t^n$ that *dominates* $\sum_{n\ge 0} a_n t^n$, that is, such that $|a_n| \le b_n$ for any $n \in \mathbb{N}$. []

The uniqueness of the solution within the class of analytic functions is straightforward: the above procedure determines the coefficients of the Taylor expansion, and these determine the analytic function.

PDEs in normal form. Throughout the present discussion, the independent variables (denoted either by x or by (x,t)), are assumed to be real; but the dependent variable (denoted by u) may be either real or complex.

Let us consider a Cauchy problem for a (possibly quasilinear) PDE of any order m of the form

$$\begin{cases} D_t^m u(x,t) = f(x,t,\{D^\alpha u(x,t)\}_{|\alpha|\le m}) & (x \in \mathbb{R}^N, t \in \mathbb{R}) \\ D_t^j u(x,0) = \varphi_j(x) \qquad j = 0,...,m-1, \end{cases} \tag{1.15}$$

with f that does not explicitly depend on $D_t^m u(x,t)$. In this case the equation $(1.15)_1$ is called a **PDE in normal form** (w.r.t. the variable t), [1] and might be fully nonlinear. E.g., the wave equation is in normal form, at variance with the heat equation.

Theorem 1.1 (Cauchy-Kovalevskaya) *In the equation* $(1.15)_1$ *let the functions* f *and* $\{\varphi_j\}$ *be (real) analytic,. Then the Cauchy problem* (1.15) *has one and only one (real) analytic solution in a suitable neighbourhood of* $(x_0,0)$ *for any* $x_0 \in \mathbb{R}^N$.

Nothing is stated concerning uniqueness in a larger class of functions, and about continuous dependence on the data. This is one of the drawbacks of this classical result.

The argument mimics that of (1.13): the initial data $(1.15)_2$ determine the derivatives w.r.t. x of any order of $D_t^j u(x,0)$ (for $j = 0,...,m-1$) in a neighbourhood U of x_0; that is, they determine $\{D^\alpha u(x,0)\}_{|\alpha|\le m}$ with the exception of $D_t^m u(x,0)$. Once these derivatives have been determined, the missing derivative $D_t^m u(x,0)$ is evaluated for $x \in U$ by restricting the PDE $(1.15)_1$ to $t=0$.

Afterwards, higher-order derivatives are determined by differentiating the PDE w.r.t. x,t, and the initial conditions w.r.t. x. For instance, let us differentiate the system (1.15) w.r.t. x_k, and set $v_k := D_{x_k} u$ for $k = 1,..,N$. Denoting by $D_t f$ the partial derivative of f w.r.t. t and by a_β (with $|\beta| \le m$) other appropriate first-order partial derivatives of f, one gets another Cauchy problem of the form

$$\begin{cases} D_t^m v_k(x,t) = \displaystyle\sum_{|\beta|\le m} a_\beta(x,t,\{D^\alpha u(x,t)\}_{|\alpha|\le m}) D^\beta v_k(x,t) \\ \qquad\qquad\qquad + D_{x_k} f(x,t,\{D^\alpha u(x,t)\}_{|\alpha|\le m}) \qquad (x \in \mathbb{R}^N, t \in \mathbb{R}) \\ D_t^j v_k(x,0) = D_{x_k}\varphi_j(x) \ (=: \psi_k(x)) \qquad j = 0,...,m-1, k = 1,...,N, \end{cases} \tag{1.16}$$

for known functions $a_\beta(...)$ and ψ_k. Notice that some of the terms $\{a_\beta\}_{|\beta|\le m}$ include the partial derivatives of f w.r.t. its last argument.

So $(1.15)_2$ determines the derivatives $\{D^\alpha v_k(x,0)\}_{|\alpha|\le m}$ for any x in a neighbourhood $U \subset \mathbb{R}^N$ of x_0, with the exception of $D_t^m v_k(x,0)$. The known derivatives determine the terms

$$\{a_\beta(x,0,\{D^\alpha u(x,t)\}_{|\alpha|\le m, t=0})\}_{|\beta|\le m}, \qquad D_{x_k} f(x,0,\{D^\alpha u(x,t)\}_{|\alpha|\le m, t=0}).$$

The missing derivative $D_t^m v_k(x,0)$ is then evaluated for $x \in U$ by restricting the PDE $(1.16)_1$ to $t=0$. Higher-order derivatives of u are then similarly determined near $x=0$, for $t=0$. This completes the formulation of a Cauchy problem for $v_k := D_{x_k} u$ for $k = 1,..,N$.

Knowing all partial derivatives of u at $(x_0,0)$, we can represent $u(x,t)$ as a power series in a neighbourhood of $(x_0,0)$. As for (1.13), the convergence of the series is then proved via the method of majorants, namely, by exhibiting a dominating convergent series. []

[1] w.r.t. = with respect to.

The uniqueness of the solution within the class of analytic functions is straightforward: the above procedure determines the coefficients of the Taylor expansion, and obviously these uniquely determine the analytic function — just as we saw for the Cauchy problem (1.13).

This analysis can be extended to systems of PDEs.

*** PDEs in nonnormal form.** Let us denote by Γ an analytic hypersurface of $\mathbb{R}^N$ (namely, an analytic manifold of codimension one), by ν a unit normal field on Γ, and consider a Cauchy problem for a fully-nonlinear PDE:

$$\begin{cases} F(x, \{D^\alpha u(x)\}_{|\alpha|\le m}) = 0 \quad (x \in \mathbb{R}^N) \\ D_\nu^j u(x) = \varphi_j(x) \qquad \text{on } \Gamma, j = 0, ..., m-1 \; (D_\nu := \partial/\partial\nu). \end{cases} \tag{1.17}$$

If not otherwise specified, all given functions will be assumed to be at least continuous.

Let us fix a point $\bar{x} \in \Gamma$. All derivatives of u at $\bar{x}$ in the tangential directions to Γ are determined by the Cauchy data on Γ. All derivatives of u up to order m are thus determined at $\bar{x}$, with the exception of the mth normal derivative $D_\nu^m u(\bar{x})$. In order to evaluate the latter derivative, let us (nonlinearly) transform the coordinates from x to $\widetilde{x}$ in a neighbourhood of $\bar{x}$, in such a way that the axis $\widetilde{x}_N$ is orthogonal to Γ. The functions u and F are accordingly transformed to $\widetilde{u}$ and $\widetilde{F}$, and the transformed PDE reads

$$\widetilde{F}(\widetilde{x}, \{D^\alpha \widetilde{u}(\widetilde{x})\}_{|\alpha|\le m}) = 0 \qquad \forall \widetilde{x} \in U. \tag{1.18}$$

The conditions $(1.17)_2$ allow one to evaluate all derivatives $\{D^\alpha u\}_{|\alpha|\le m}$ in a suitable neighbourhood $U \subset \Gamma$ of $\widetilde{x}$, with the exception of $D^m_{\widetilde{x}_N} u$. By the implicit function theorem, in U (or in a smaller neighbourhood of $\bar{x}$) $D^m_{\widetilde{x}_N} \widetilde{u}$ can then be expressed as a function of the data, provided that (displaying just the arguments of interest, denoting by ξ the final argument of $\widetilde{F}$, and dropping the tildas)

$$\frac{\partial}{\partial\xi} F(x, ..., D^m_{x_{N-1}} u(x), D^m_{x_N} u(x)) \neq 0 \qquad \forall x \in U. \tag{1.19}$$

The solution occurs in this formula, but $u(x), ..., D^m_{x_{N-1}} u(x)$ are provided by the Cauchy data $(1.17)_2$. Obviously, (1.19) holds whenever

$$\frac{\partial}{\partial\xi} F(x, ..., D^m_{x_{N-1}} u(x), \xi) \neq 0 \qquad \forall x \in U, \forall \xi \in \mathbb{R}. \tag{1.20}$$

If (1.19) is fulfilled, then one says that the hypersurface Γ is **noncharacteristic** for the differential operator F at the point $(x, u(x), ..., D^m_{x_N} u(x))$ (at the point x if the equation is linear). Knowing all the partial derivatives of u at x, we can represent $u(x)$ as a power series in a neighbourhood of x, as we did above for equations in normal form. [2] By the method of majorants, the convergence of the series is then proved by exhibiting a dominating convergent series. []

In conclusion,

$$\boxed{\begin{array}{l}\text{the Cauchy-Kovalevskaya theorem is extended to PDEs in nonnormal form,} \\ \text{in a neighbourhood of any noncharacteristic point of } \Gamma \text{ (assumed analytic).}\end{array}} \tag{1.21}$$

[2] Freely speaking, in a neighbourhood of x one can actually approximate the Cauchy problem (1.17) by a problem of the form (1.15).

(Of course here the functions f and $\{\varphi_j\}_{j=0,\dots,m-1}$ are still assumed to be analytic.) Ahead we shall illustrate an example for second-order equations. This analysis can be extended to systems of PDEs.

On the Holmgren theorem. The Cauchy-Kovalevskaya theorem provides uniqueness of the solution (of *nonlinear* equations) just in the restricted class of analytic functions. For *linear* equations with analytic coefficients, with second member and Cauchy data of class C^k for some $k \geq 0$ (not smaller than the order of the equation) on a noncharacteristic analytic hypersurface, Holmgren proved that the solution is unique among functions of class C^k, for the same k; see e.g. [ReRo; Sect. 2.3].

On the Lewy counterexample. Hans Lewy considered the following linear operator with constant (complex) coefficients:

$$L(x, y, z, D)u = -u_x - iu_y + 2i(x + iy)u_z \qquad \forall(x, y, z) \in \mathbb{R}^3 \ (i^2 = -1). \tag{1.22}$$

He was able to exhibit a nonanalytic function $f \in C^\infty(\mathbb{R}^3)$, such that the linear equation

$$L(x, y, z, D)u = f(x, y, z)$$

has no solution in any open subset of $\mathbb{R}^3$; see e.g. [ReRo; Sect. 2.2]. This settled in the negative a long-dating question concerning the solvability of *all* linear PDEs with constant coefficients.

So far we dealt with equations with real coefficients. However the equation (1.22) with complex coefficients can easily be reduced to a system with real coefficients.

A trivial example. If characteristic points form an open subset of the hypersurface on which the Cauchy datum is prescribed, then the thesis of the Cauchy-Kovalevskaya theorem fails: for certain (analytic) data the Cauchy problem has no (analytic) solution, whereas for other data it is uniqueness that fails. If characteristic points do not form an open subset of that hypersurface, in general one cannot establish a priori whether the solution exists and is unique, as the following simple example shows.

Let us consider the trivial PDE

$$u_x(x, y) = 1 \qquad \forall(x, y) \in \mathbb{R}^2, \tag{1.23}$$

and the analytic curves $\Gamma_1 = \{(x, x^3) : x \in \mathbb{R}\}$ and $\Gamma_2 = \{(x, x^2) : x \in \mathbb{R}\}$. (This is actually a degenerate PDE: it is rather on ODE parameterized by y.)

In either case the point $(0, 0)$ is characteristic. Anyway the Cauchy problem with datum on Γ_1 has one and only one solution in a neighbourhood of $(0, 0)$, whereas in the case of Γ_2 either existence or uniqueness fails. This is easily checked, since the above PDE has the obvious general solution $u(x, y) = x + c(y)$, for any function $c : \mathbb{R} \to \mathbb{R}$. For instance, if we prescribe $u(x, y) = x$ on Γ_j ($j = 1$ or 2), then the problem has one and only one solution for $j = 1$, that is, $u(x, y) = x$ for all $(x, y) \in \mathbb{R}^2$. For $j = 2$ instead even a local solution misses in a neighbourhood of $(0, 0)$.

*** Two apparent failures of the Cauchy-Kovalevskaya and Holmgren theorems.**

Example (i). Let $u = u(x, t)$ be such that

$$\begin{cases} D_t u = D_x^2 u \qquad \forall x \in \,]-1, 1[, \forall t > 0, \\ u(x, 0) = (1 - x)^{-1} = \sum_{n \geq 0} x^n \qquad \forall x \in \,]-1, 1[. \end{cases} \tag{1.24}$$

If u were analytic in a neighbourhood U of $(0,0)$, then it would coincide with the *semigroup solution*

$$u(x,t) = \exp(tD_x^2)\, u(x,0) := \sum_{k\geq 0} \frac{(tD_x^2)^k}{k!}\, u(x,0) \qquad \forall (x,t) \in U. \tag{1.25}$$

As $\big(D_x^m \sum_{n\geq 0} x^n\big)|_{x=0} = m!$ for any $m \in \mathbb{N}$, we would then have

$$u(0,t) = \sum_{k\geq 0} \left(\frac{(tD_x^2)^k}{k!} u(x,0) \right)\bigg|_{x=0} = \sum_{k\geq 0} \left(\frac{t^k D_x^{2k}}{k!} \sum_{n\geq 0} x^n \right)\bigg|_{x=0} = \sum_{k\geq 0} \frac{(2k)!}{k!} t^k \qquad \forall t > 0.$$

But this series diverges for any $t \neq 0$.

This shows that the solution of the Cauchy problem for the heat equation with an analytic initial datum need not be analytic. Indeed here the Cauchy-Kovalevskaya theorem does not apply, since the curve of equation $t = 0$ is characteristic for the heat equation, which actually is not in normal form w.r.t. t.

Incidentally notice that this pathology would be removed, if the heat equation $D_t u = D_x^2 u$ were replaced by a first-order equation, e.g. by the transport equation $D_t u = D_x u$. [Ex]

Example (ii). For any $g \in C^\infty(\mathbb{R})$, it is easy to check that the Tychonov function

$$u(x,t) := \sum_{k=0}^{\infty} g^{(k)}(t) \frac{x^{2k}}{(2k)!} \qquad \forall x \in \mathbb{R}, \forall t \in \mathbb{R} \tag{1.26}$$

solves the one-dimensional heat equation $D_t u = D_x^2 u$.

Let us now select the Cauchy function $g(t) = \exp\{-1/t^2\}$ for any $t \in \mathbb{R} \setminus \{0\}$, $g(0) = 0$. It is well known that $g \in C^\infty(\mathbb{R})$ but is not analytic. Moreover, since $g^{(k)}(t) \to g^{(k)}(0) = 0$ as $t \to 0$ for any $k \in \mathbb{N}$,

$$u(x,t) \to 0 \qquad \text{as } t \to 0, \forall x \in \mathbb{R}. \tag{1.27}$$

The Cauchy problem for the one-dimensional heat equation with homogeneous (i.e., identically vanishing) source term and homogeneous initial datum thus has the above nontrivial solution, besides the identically vanishing solution. But here the Holmgren theorem of uniqueness does not apply, because the curve of equation $t = 0$ is characteristic for the heat equation.

For any t, the Tychonov function $u(x,t)$ has a high order of growth as $|x| \to \infty$, at variance with the solution that is constructed via the fundamental solution (that we introduce ahead). This suggests that the solution of the Cauchy problem for the heat equation may be unique only in a class of functions that do not grow too fast at infinity (this is the class of tempered distributions on $\mathbb{R}$, which indeed does not include the Tychonov function).

1.3 Characteristics hypersurface and classification of linear PDEs

*** About the terminology.** In the literature there is some ambiguity in the use of the term **characteristic** in connection with PDEs. In $\mathbb{R}^N$ one defines

characteristic surfaces, characteristic curves, and characteristic directions.

For $N = 2$ characteristic *surfaces* are reduced to curves (!), and coincide with characteristic curves; but for $N > 2$ this is not the case. The traditional term *surface* is questionable: it should be better replaced by *hypersurface,* or *manifold of codimension one* (i.e., manifold of dimension $N-1$), thus a curve for $N = 2$. In these notes we shall often refer to 2-dimensional characteristic hypersurfaces as *characteristic surfaces,* and to 1-dimensional characteristic hypersurfaces as *characteristic curves.*

Second order PDEs in $\mathbb{R}^2$. Let us consider the linear PDE

$$A(x,y)u_{xx} + B(x,y)u_{xy} + C(x,y)u_{yy} = G(x,y) \qquad (x,y) \in \mathbb{R}^2, \tag{1.28}$$

with A, B, C, G prescribed functions of (x,y), and at least one among A, B, C not identically vanishing, so that the equation has order two. Let us prescribe the Cauchy data on a regular curve $\Gamma \subset \mathbb{R}^2$ parameterized by $x = \xi(t), y = \eta(t)$ for $t \in I \subset \mathbb{R}$:

$$\begin{aligned} u(\xi(t), \eta(t)) &= v(t) \\ u_x(\xi(t), \eta(t)) &= \varphi(t) \qquad\qquad \forall t \in I. \\ u_y(\xi(t), \eta(t)) &= \psi(t) \end{aligned} \tag{1.29}$$

Note that just two conditions of this triplet are independent, since

$$v'(t) = \frac{d}{dt} u(\xi(t), \eta(t)) = u_x(...)\xi'(t) + u_y(...)\eta'(t) = \varphi(t)\xi'(t) + \psi(t)\eta'(t) \qquad \forall t.$$

In alternative, as a Cauchy datum one might also prescribe $(1.29)_1$ and the normal derivative to Γ, that is,

$$D_\nu u(\xi(t), \eta(t)) = -\eta'(t) u_x(\xi(t), \eta(t)) + u_y(\xi(t), \eta(t))\xi'(t) = \text{ a prescribed function of } t.$$

Let us see whether all second-order derivatives of u can be computed on Γ. The PDE and the Cauchy data yield [3]

$$\begin{cases} Au_{xx} + Bu_{xy} + Cu_{yy} = G \\ \xi' u_{xx} + \eta' u_{xy} = \varphi' \\ \xi' u_{xy} + \eta' u_{yy} = \psi' \end{cases} \qquad \text{on } \Gamma. \tag{1.30}$$

Setting

$$\mathcal{A} = \begin{pmatrix} A & B & C \\ \xi' & \eta' & 0 \\ 0 & \xi' & \eta' \end{pmatrix}, \quad U = \begin{pmatrix} u_{xx} \\ u_{xy} \\ u_{yy} \end{pmatrix}, \quad F = \begin{pmatrix} G \\ \varphi' \\ \psi' \end{pmatrix}, \tag{1.31}$$

the system (1.30) also reads

$$\mathcal{A}U = F \qquad \text{on } \Gamma. \tag{1.32}$$

In a neighbourhood of any point $(\xi(\bar{t}), \eta(\bar{t})) = (\bar{x}, \bar{y}) \in \Gamma$, the latter system can be (uniquely) solved w.r.t. U iff

$$\text{Det } \mathcal{A}(\bar{x}, \bar{y}) \neq 0 \quad \text{ i.e. } \quad A(\eta')^2 - B\xi'\eta' + C(\xi')^2 \neq 0 \qquad \text{at } t = \bar{t}. \tag{1.33}$$

Because of the Cauchy-Kovalevskaya and Holmgren theorems, in this case the Cauchy problem (1.28), (1.29) is then uniquely solvable in a neighbourhood of Γ (for analytic data).

Characteristic curves. A point $(\bar{x}, \bar{y}) = (\xi(\bar{t}), \eta(\bar{t})) \in \Gamma$ is called a characteristic point for the Cauchy problem above if the condition (1.33) is violated at $t = \bar{t}$. Γ is called a characteristic hypersurface (here a curve since u depends on tow variables) whenever (1.33) is violated everywhere on Γ, i.e.,

$$A(\eta')^2 - B\xi'\eta' + C(\xi')^2 = 0 \qquad \text{on } \Gamma. \tag{1.34}$$

[3] Whenever we shall differentiate a function, we shall implicitly assume that the function is differentiable and has continuous derivatives.

In a neighbourhood V of any point of Γ in which the curve can be represented in the form $y = y(x)$, the differential $(d\xi, d\eta)$ is replaced by $(1, y'(x))dx$, i.e. $\eta'/\xi' = y'$, so that (1.34) is tantamount to $Ay'(t)^2 - By'(t) + C = 0$, that is,

$$y' = \frac{B \pm \sqrt{B^2 - 4AC}}{2A} \qquad \text{on } \Gamma. \tag{1.35}$$

At $(\bar{x}, \bar{y})$ thus:
(i) if $B^2 - 4AC > 0$, then the PDE (1.28) has two characteristic curves,
(ii) if $B^2 - 4AC = 0$, then the PDE (1.28) has one characteristic curve,
(iii) if $B^2 - 4AC < 0$, then the PDE (1.28) has no characteristic curve.

The equation is accordingly called hyperbolic, parabolic or elliptic (at $(\bar{x}, \bar{y})$), respectively.

Equivalently, as the quadratic form $t \mapsto At^2 - Bt + C$ is associated to the matrix

$$\mathcal{A}(\bar{x}, \bar{y}) = \begin{pmatrix} A & B/2 \\ B/2 & C \end{pmatrix}, \tag{1.36}$$

(i) if $\mathcal{A}$ indefinite (i.e., it has two eigenvalues of opposite sign), then the PDE (1.28) has two characteristic curves at $(\bar{x}, \bar{y})$,

(ii) if $\mathcal{A}$ is semidefinite (i.e., it has a vanishing eigenvalue), then the PDE (1.28) has just one characteristic curve at $(\bar{x}, \bar{y})$,

(iii) if $\mathcal{A}$ is definite (i.e., it has two eigenvalues of the same sign), then the PDE (1.28) has no characteristic curve at $(\bar{x}, \bar{y})$.

As the coefficients A, B, C depend on (x, y), then the property of being characteristic, and accordingly the type of the equation, may depend on the point. If A, B, C also depend on (u, u_x, u_y) (i.e., the equation is quasilinear), then this property will depend on the solution, too.

As we already saw, characteristics are of interest because, if Cauchy data are prescribed on a characteristic curve Γ, then not all second derivatives of u on Γ are determined by the equation and by the data. The Cauchy problem need not be locally solvable in a neighbourhood of a characteristic point $(x, y) \in \Gamma$; moreover, if it is locally solvable, then the solution need not be unique.

*** Characteristic hypersurfaces of linear second-order equations in $\mathbb{R}^N$.** Let us consider a quasilinear second-order equation in N variables:

$$A_{ij} u_{x_i x_j} + \{\text{lower-order terms}\} = G \qquad x \in \mathbb{R}^N \tag{1.37}$$

(we imply the sum over repeated indices), with $\mathcal{A} = \{A_{ij}\}$ and G prescribed functions of $(x, u, \nabla u)$. As the Hessian matrix $\{u_{x_i x_j}\}$ is symmetric, we can assume that $\mathcal{A}$ is symmetric without loss of generality.

Let $\Gamma \subset \mathbb{R}^N$ be a regular hypersurface: $\Gamma = \{x \in \mathbb{R}^N : \varphi(x) = 0\}$, with $\varphi \in C^1$ and $\nabla\varphi(x) \neq 0$ for any $x \in \Gamma$. We claim that this hypersurface is characteristic at some $x \in \Gamma$ iff

$$\begin{aligned} &[\nabla\varphi(x)]^\tau \cdot A(x, u(x), \nabla u(x)) \cdot \nabla\varphi(x) = 0 \qquad \text{i.e.,} \\ &\varphi_{x_i}(x)\, A_{ij}(x, u(x), \nabla u(x))\, \varphi_{x_j}(x) = 0. \end{aligned} \tag{1.38}$$

Indeed, as $\nabla\varphi(x)/|\nabla\varphi(x)|$ coincides with the unit normal ν_x to Γ at x,

$$D^2_\nu u(x) = |\nabla\varphi(x)|^{-2}\, \varphi_{x_i}(x)\, u_{x_i x_j}(\varphi(x))\, \varphi_{x_j}(x). \tag{1.39}$$

Therefore $D^2_\nu u(x)$ is determined by the PDE (1.37) iff the equality (1.38) fails (see e.g. [ReRo] p. 42).

Hyperbolic, parabolic, elliptic equations. For the sake of simplicity, let us next assume that the matrix $\{A_{ij}\}$ does not depend on u and ∇u, so that the equation (1.37) is semilinear. Otherwise the property of Γ being characteristic would depend on the solution, too.

There are three cases of interest:

(i) if $N-1$ eigenvalues of the matrix $\mathcal{A}$ have the same sign and the remaining one has the opposite sign (thus $\mathcal{A}$ is indefinite), then the equation is called **hyperbolic;**

(ii) if an eigenvalue of the matrix $\{A_{ij}\}$ vanishes and all the others have the same sign (thus $\mathcal{A}$ is semi-definite), then the equation is called **parabolic;**

(iii) if all the eigenvalues of the matrix $\{A_{ij}\}$ have the same sign (i.e., $\mathcal{A}$ is definite), then the equation is called **elliptic.**

There are also other cases; but they are of little interest, since they occur in applications only exceptionally.

Let us now retrieve (1.34) for $N=2$. In this case we can represent $\Gamma = \{(x,y) \in \mathbb{R}^2 : \varphi(x,y) = 0\}$ parametrically by $x = \xi(t), y = \eta(t)$ for $t \in I \subset \mathbb{R}$. Thus $\varphi(\xi(t),\eta(t)) = 0$ for any $t \in I$, whence $\nabla\varphi(\xi(t),\eta(t))\cdot(\xi'(t),\eta'(t)) = 0$, that is,

$$\xi'(t)\,\varphi_x(\xi(t),\eta(t)) + \eta'(t)\,\varphi_y(\xi(t),\eta(t)) = 0 \qquad \forall t \in I. \tag{1.40}$$

For $N=2$ (1.38) also reads

$$A_{11}(\varphi_x)^2 + 2A_{12}\varphi_x\,\varphi_y + A_{22}(\varphi_y)^2 = 0 \qquad \text{on } \Gamma, \tag{1.41}$$

that is, by (1.40),

$$A_{11}(\xi(t),\eta(t))\,\eta'(t)^2 - 2A_{12}(\xi(t),\eta(t))\,\eta'(t)\,\xi'(t) + A_{22}(\xi(t),\eta(t))\,\xi'(t)^2 = 0 \qquad \forall t \in I. \tag{1.42}$$

Selecting $\mathcal{A}$ as in (1.36), the latter equation is reduced to (1.34).

Characteristic hypersurfaces of linear PDEs of any order in $\mathbb{R}^N$. The boundary conditions that are appropriate for a specific PDE depend on the order and on the type of the PDE. The type of the equation and the characteristic hypersurfaces just depend on the principal part of the PDE.

For any linear differential operator

$$L(x,D) = \sum_{|\alpha|\le m} a_\alpha(x) D^\alpha \qquad (x \in \mathbb{R}^N), \tag{1.43}$$

$\Gamma = \{x \in \mathbb{R}^N : \varphi(x) = 0\}$ is a **characteristic hypersurface** relative to the operator $L(x,D)$ iff

(i) $\varphi \in C^1(\mathbb{R}^N)$,

(ii) $\nabla\varphi \neq 0$ everywhere,

(iii) φ solves the nonlinear first-order PDE

$$L_p(x, D\varphi(x)) = \sum_{|\alpha|=m} a_\alpha(x)[D\varphi(x)]^\alpha = 0 \qquad \forall x \in \Gamma. \tag{1.44}$$

Indeed, denoting by ν_x a unit normal field to Γ, and recalling that $\nabla\varphi(x)$ is paraller to ν_x, the equation $\sum_{|\alpha|=m} a_\alpha(x) D^\alpha u = f$ determines $D^m_{\nu_x} u$ in terms of the other derivatives iff this equality fails.

The equation (1.44) is linear if $m=1$; otherwise it is fully nonlinear. We shall call **characteristic direction** any direction that is orthogonal to a characteristic hypersurface at a point of that hypersurface: [4]

$$\boxed{\text{characteristic directions are orthogonal to characteristic hypersurfaces.}} \tag{1.45}$$

[4] Note that characteristic directions are not the directions of characteristic curves: they are orthogonal to them (!)

More generally, the hypersurface Γ is characteristic at a point $\bar{x} \in \Gamma$ iff $\Gamma = \{x \in \mathbb{R}^N : \varphi(x) = 0\}$ with $\varphi \in C^1(V)$ for some neighbourhood V of $\bar{x}$, $\nabla\varphi(\bar{x}) \neq 0$, and φ solves (1.44) at $\bar{x}$.

We stress that

$$\boxed{\begin{array}{l}\text{a characteristic hypersurface is } \textit{characteristic} \text{ at all of its points;}\\ \text{however a hypersurface is called } \textit{noncharacteristic} \text{ iff}\\ \text{all of its points fail to be characteristic.}\end{array}} \tag{1.46}$$

Thus, for a hypersurface, being noncharacteristic is a much stronger property than just not being characteristic. (Otherwise stated, a hypersurface that is not characteristic need not be noncharacteristic.) For instance, the curve C_a of equation $t = ax^2$ ($a \in \mathbb{R}$) is characteristic for the equation $D_t u = D_{xx} u$ iff $a = 0$. For any $a \neq 0$, $(0,0)$ is the only characteristic point for that equation; thus C_a is not characteristic, although it is *not noncharacteristic.*

These definitions are extended to quasilinear equations, just by allowing the coefficients to depend on lower-order derivatives.

Simple waves and characteristic hypersurfaces of first-order PDEs. Let us now consider a linear first-order operator $L(x, D) = \sum_{j=1}^N a_j(x)\, D_j + a_0(x) I$; its principal part reads

$$L_p(x, D) = \sum_{j=1}^N a_j(x)\, D_j \qquad (x \in \mathbb{R}^N). \tag{1.47}$$

Let Ω be a domain of $\mathbb{R}^N$, and $\varphi \in C^1(\mathbb{R}^N)$ be such that $\nabla\varphi \neq 0$ everywhere and

$$L_p(x, D)\varphi(x) = \sum_{j=1}^N a_j(x)\, D_j\varphi(x) = 0 \qquad \forall x \in \Omega. \tag{1.48}$$

This means that the hypersurface $\Gamma = \{x \in \mathbb{R}^N : \varphi(x) = 0\}$ is characteristic. (Obviously, the same holds for $\Gamma_C = \{x \in \mathbb{R}^N : \varphi(x) = C\}$ for any $C \in \mathbb{R}$.) For any function $w \in C^1(\mathbb{R})$, then

$$L_p(x, D)(w \circ \varphi) := \sum_{j=1}^N a_j(x)\, D_j[w(\varphi(x))] = w'(\varphi(x)) \sum_{j=1}^N a_j(x)\, D_j\varphi(x) = 0 \qquad \forall x \in \Omega. \tag{1.49}$$

Thus $u = w \circ \varphi$ is a solution of the principal-part PDE $L_p(x, D)u = 0$. This solution is called a **simple wave.** Thus [5]

$$\boxed{\text{a simple wave } \ u \ \text{ is a solution of the } \textit{principal-part PDE:} \ L_p(x, D)u = 0.} \tag{1.50}$$

Note that characteristic hypersurfaces are level hypersurfaces of simple waves. [6]

Characteristic hypersurfaces of quasilinear PDEs. The notion of characteristic hypersurface, that we defined for linear PDEs, can be extended to quasilinear PDEs. In this case the principal part of the PDE depends on the solution itself.

Let us represent the quasilinear operator canonically as a polynomial of D, with coefficients that depend nonlinearly on the unknown function. Once the solution is inserted into the coefficients, the operator becomes linear, so that one can define characteristic hypersurfaces also for quasilinear PDEs.

[5] As it is illustrated by [Renardy-Rogers p. 41], for higher order equations the same occurs only "at leading order".

[6] We can thus retrieve a property that is at the basis of the Cauchy-Kovalevskaya theorem: the Cauchy problem is well posed only if the Cauchy datum is prescribed on a noncharacteristic hypersurface... Why?

Although these hypersurfaces depend on the solution, in this way some properties can be derived without knowing the solution. On the other hand, for semilinear PDEs the principal part is linear, so that these characteristics do not depend on the solution.

From quasilinear to linear (for first-order PDEs). Let us consider a quasilinear first-order PDE in N variables:

$$\sum_{n=1}^{N} a_n(x,u) D_{x_n} u = f(x,u). \tag{1.51}$$

Let

$$\varphi(x,u) = 0 \tag{1.52}$$

be a solution in implicit form, with $D_u\varphi(x,u) \neq 0$ where $\varphi(x,u) = 0$. Hence, assuming that $\varphi(x,u(x)) = 0$ for any x,

$$D_{x_n}\varphi(x,u(x)) + D_u\varphi(x,u(x)) D_{x_n}u(x) = 0 \qquad \text{for } n = 1,..,N, \forall x.$$

Thus, as $D_u\varphi(x,u) \neq 0$,

$$D_{x_n}u(x) = -\frac{D_{x_n}\varphi(x,u(x))}{D_u\varphi(x,u(x))} \qquad \text{for } n = 1,..,N, \forall x,$$

so that, multiplying (1.51) by $D_u\varphi(x,u)$ ($\neq 0$), this equation also reads

$$\sum_{n=1}^{N} a_n(x,u) D_{x_n}\varphi(x,u(x)) + f(x,u) D_u\varphi(x,u(x)) = 0. \tag{1.53}$$

By setting

$$x_{N+1} = u, \quad a_{N+1} = f(x,u), \quad y = (x_1, ..., x_N, x_{N+1}), \tag{1.54}$$

the quasilinear PDE (1.51) is thus reduced to a (homogeneous) [7] *linear* equation (!):

$$\sum_{n=1}^{N+1} a_n(y) D_{y_n}\varphi(y) = 0. \tag{1.55}$$

As we pointed out, the solution of the original equation (1.51) is then represented in implicit form by $\varphi(x,u) = 0$ (or by "$\varphi(x,u) =$ constant").

Conversely, if φ solves (1.55), then $\varphi(x,u) = 0$ implicitly defines a solution $u = u(x)$ of (1.51) in the neighbourhood of any (x,u) such that $D_u\varphi(x,u) \neq 0$ [8]

We thus conclude with the following *reduction principle:*

$$\boxed{\begin{array}{l}\text{every quasilinear first-order PDE in } N \text{ variables is equivalent} \\ \text{to a homogeneous linear first-order PDE in } N+1 \text{ variables,} \\ \text{and conversely.}\end{array}} \tag{1.56}$$

Ahead we shall use this property to solve quasilinear first-order PDEs.

Here is a quicker derivation of (1.55). Let us set $\varphi(x,z) = u(x) - z$, whence $z = u(x)$ iff $\varphi(x,u(x)) = 0$, and

$$D_{x_i}\varphi(x,z) = D_{x_i}u(x), \qquad D_z\varphi(x,z) = -1.$$

[7] i.e., with identically vanishing second member.

[8] This condition is crucial: e.g. $\varphi \equiv 0$ solves any homogeneous linear PDE!

Therefore $\sum_{n=1}^N a_n(x,u) D_{x_n} u = f(x,u)$ is equivalent to

$$\sum_{n=1}^N a_n(x,z) D_{x_n}\varphi(x,z) + D_z\varphi(x,u(x)) f(x,u(x)) = 0 \qquad \forall x.$$

Setting (1.54), this coincides with (1.55).

1.4 Characteristic curves for first-order PDEs

Linear first-order operators have principal part of the form $L_p(x,D) = \sum_{j=1}^N a_j(x)\, D_j$ and are hyperbolic. Quasilinear first-order operators of the form

$$L(x,u,D)u = \sum_{j=1}^N a_j(x,u)\, D_j u + a_0(x,u)$$

are also hyperbolic.

The equation of characteristic hypersurfaces of linear first-order PDEs concides with the original PDE, since

$$(D\varphi)^\alpha = D^\alpha\varphi \qquad \forall\varphi \in C^\infty(\mathbb{R}^N), \forall\alpha \in \mathbb{R}^N \text{ such that } |\alpha| = 1. \tag{1.57}$$

In the literature one also uses the term *characteristic* with a different meaning from that of characteristic hypersurface. One says that a curve of equation $x = \widehat{x}(s)$ with $\widehat{x} \in C^1([a,b])$ is a (projected) **characteristic curve** [9] for the linear first-order operator (1.47) iff it is a line of the field $(a_1, ..., a_N)$, that is,

$$\widehat{x}'(s) \quad \text{is parallel to} \quad (a_1(\widehat{x}(s)), ..., a_N(\widehat{x}(s))) \qquad \forall s \in\,]a,b[. \tag{1.58}$$

By possibly rescaling the parameter s, we can assume that these two vector functions coincide, so that along characteristic curves the operator $L_p(x,D) = \sum_{j=1}^N a_j(x)\, D_j$ reads

$$[L_p(x,D)u]_{x=\widehat{x}(s)} = \sum_{j=1}^N \widehat{x}_j'(s)\, D_j u(\widehat{x}(s)) = \frac{d}{ds} u(\widehat{x}(s)) \qquad \forall s \in\,]a,b[. \tag{1.59}$$

Along characteristic curves the principal part of the PDE (also called the "principal part PDE") is thus reduced to the so-called **characteristic ODE:**

$$\frac{d}{ds} u(\widehat{x}(s)) = F(\widehat{x}(s)) \qquad \forall s \in\,]a,b[. \tag{1.60}$$

In particular,

for semilinear first-order PDEs, any solution of the *homogeneous* principal-part PDE is constant along any projected characteristic curve (in $\mathbb{R}^N$).	(1.61)

If we prescribe Cauchy data along a characteristic curve, then these data are constrained to fulfill the characteristic ODE. Because of these constraints on the data, these curves lie on characteristic hypersurfaces; more specifically, in $\mathbb{R}^N$ with $N > 2$,

characteristic hypersurfaces are unions of (projected) characteristic curves.	(1.62)

[9] Ahead we shall also meet *unprojected* characteristic curves, which are contained in the ambient space $\mathbb{R}^{N+1}$. In the literature projected and unprojected characteristic curves are sometimes both named characteristic curves tout court (of course this does not help to understand!).

(In $\mathbb{R}^2$ hypersurfaces are curves, so that what here we referred to as characteristic hypersurfaces are actually projected characteristic curves.) Notice that characteristic directions are orthogonal to characteristic hypersurfaces, namely to projected characteristic curves for $N = 2$.

Solutions of a semilinear first-order equation

$$L(x, D)u = \sum_{j=1}^{N} a_j(x)\, D_j u = f(x, t, u)$$

can thus be calculated along projected characteristic curves, by integrating the characteristic ODE (1.60). One may express this in a pictorial way by saying that

$$\boxed{\text{the Cauchy datum } \textit{propagates} \text{ along the projected characteristic curves.}}$$

This allows one to construct the solution of first-order PDEs by solving suitable ODEs; this procedure is known as the **method of characteristics.** We have shown this for semilinear first-order equations; anyway because of (1.56) this can be extended to quasilinear first-order equations.

A simple example. Let us consider the advection [10] problem

$$D_t u + a D_x u = f(x,t) \quad (\text{with } a \in \mathbb{R}), \qquad u(x,0) = u^0(x). \tag{1.63}$$

The integral curves of the field $(a, 1)$ are the graphs of the parametric curves

$$\widehat{x}(t) = x_0 + at \qquad \forall t \in \mathbb{R}, \forall x_0 \in \mathbb{R}.$$

(This includes the trivial case of $a = 0$.) Along each of these curves the PDE is reduced to the characteristic ODE

$$\frac{d}{dt} u(x_0 + at, t) = [D_t u + a D_x u]\,(x_0 + at, t) = f(x_0 + at, t) \qquad \forall x_0 \in \mathbb{R}, \forall t \in \mathbb{R}, \tag{1.64}$$

whence

$$u(x_0 + at, t) = u(x_0, 0) + \int_0^t f(x_0 + as, s)\, ds \qquad \forall x_0 \in \mathbb{R}, \forall t \in \mathbb{R}. \tag{1.65}$$

By replacing x_0 by $x - at$, we get

$$u(x,t) = u^0(x - at) + \int_0^t f(x + a(s-t), s)\, ds \qquad \forall x \in \mathbb{R}, \forall t \in \mathbb{R}. \tag{1.66}$$

These two equivalent formulae respectively represent the solution along the characteristics and as a field. [11]

More generally, if the PDE includes lower-order terms (in particular in the case of first-order semilinear PDEs), its solution can be reduced to that of a family of ODEs, as we shall illustrate ahead.

This is trivially extended if a depends on t.

[10] Advection = transport. If $u(x,t)$ represents the concentration of a substance which is translated along a curve $x = x(t)$, then $u(x(t), t)$ is constant in time; hence $\frac{d}{dt} u(x(t), t) = 0$. Setting $a(t) = x'(t)$, this yields $D_t u + a(t) D_x u = 0$.

[11]* The reader may notice that the function f need not be differentiable, although the separate evaluation of the derivatives of u involves the derivatives of f. This paradox finds an explaination in the framework of the theory of distributions.

About terminology. Let us still deal with first-order PDEs. We already saw that

$$\boxed{\begin{array}{l}\text{for } N=2 \text{ characteristic hypersurfaces and projected characteristic curves coincide;}\\ \text{for } N>2 \text{ characteristic hypersurface are unions of projected characteristic curves.}\end{array}} \tag{1.67}$$

For any linear first-order operator

$$L(x,D) = \sum_{|\alpha|\le 1} a_\alpha(x) D^\alpha + a_0(x) I \qquad (x \in \mathbb{R}^N), \tag{1.68}$$

summing up we have defined:

$$\boxed{\begin{array}{l}\text{the principal symbol } L_p(x,i\xi) = \displaystyle\sum_{|\alpha|=1} a_\alpha(x)(i\xi)^\alpha,\ \forall\xi\in\mathbb{R}^N,\\ \text{the characteristic hypersurface } \{x\in\mathbb{R}^N : \varphi(x)=0\}, \text{ with } \displaystyle\sum_{|\alpha|=1} a_\alpha(x)(\nabla\varphi(x))^\alpha = 0,\\ \text{the characteristic curve } x=\widehat{x}(t), \text{ with } \widehat{x}'(t) \text{ parallel to } \big(a_{(1,\dots,0)},\dots,a_{(0,\dots,1)}\big)(\widehat{x}(t)),\\ \text{the characteristic direction } \nabla\varphi(x) \text{ at any point } x\in\varphi^{-1}(0).\end{array}} \tag{1.69}$$

(Characteristic directions are orthogonal to characteristic curves.)

Finally, notice that all first order PDEs are hyperbolic and thus have characteristic curves and characteristic hypersurfaces. On the other hand elliptic PDEs have neither characteristic curves nor characteristic hypersurfaces. Parabolic equations typically miss a time derivative in the principal part; their characteristic hypersurfaces are that of the form $\{(x_1,\dots,n_N,t) : t = C\}$, and are of no use for constructing the solution.

2 Maximum Principle

What follows is essentially a synthesis of the results of Sects. 4.1 and 4.4 of [ReRo].

Weak maximum principle for elliptic operators. [12] Let Ω be a bounded Euclidean domain of Lipschitz class, $a_{ij}, b_i \in C^0(\bar\Omega)$ for any $i,j\in\{1,\dots,N\}$, with $\{a_{ij}(x)\}$ positive definite for any $x\in\Omega$. Let us define two linear elliptic operators of second order (in *non-divergence form):*

$$\begin{aligned}&L_0 := a_{ij}(x)D_iD_j + b_i(x)D_i, \quad L := L_0 + c(x)I,\\ &\text{with } a_{ij}, b_i, c \in C^0(\bar\Omega)\ \forall i,j, \text{ and } A=\{a_{ij}\} \text{ everywhere positive definite}\end{aligned} \tag{2.1}$$

(we still imply the sum over repeated indices). Without loss of generality, we can assume that the matrix $\{a_{ij}(x)\}$ is symmetric for a.e. x. For instance, $L_0 = \Delta$ (the Lapace operator) and $L := \Delta + cI$ for some real c. Throughout we shall assume that $u \in C^2(\Omega)\cap C^0(\bar\Omega)$.

The weak maximum principle states that

$$(i) \qquad -L_0 u \le 0 \quad \text{in } \Omega \quad \Rightarrow \quad \max_{\bar\Omega} u \le \max_{\partial\Omega} u, \tag{2.2}$$

that is, the maximum of u is attained on the boundary (possibly in the interior, too); and

$$(ii) \qquad -Lu \le 0, \quad c \le 0 \quad \text{in } \Omega \quad \Rightarrow \quad \max_{\bar\Omega} u \le \max_{\partial\Omega} u^+, \tag{2.3}$$

[12] This does not refer to the strong/weak formulation of the problem! (Here we just deal with the strong formulation.)

that is, a positive maximum of u is attained on the boundary (possibly in the interior, too).

Let $f \in C^0(\bar{\Omega})$ and $g \in C^0(\partial\Omega)$ be prescribed, and consider the nonhomogeneous Dirichlet problem

$$-Lu = f \quad \text{in } \Omega, \qquad u = g \quad \text{on } \partial\Omega. \tag{2.4}$$

It is promptly checked that (2.3) entails the *monotone dependence on the data:* if $c \le 0$ in Ω then

$$\begin{cases} -Lu_1 \le -Lu_2 & \text{in } \Omega \\ u_1 \le u_2 & \text{on } \partial\Omega \end{cases} \quad \Rightarrow \quad u_1 \le u_2 \quad \text{in } \Omega. \tag{2.5}$$

Obviously, this yields the uniqueness of the solution. [13]

On the strong maximum principle. Let $x_0 \in \partial\Omega$ be an element of a (closed) ball $B \subset \bar{\Omega}$ (this means that there exists a ball contained in $\bar{\Omega}$ that is tangent to $\partial\Omega$ at x_0). We shall denote by ν the outward-oriented unit normal vector field on $\partial\Omega$, and by $\partial/\partial\nu$ the corresponding normal derivative. If $u \le u(x_0)$ in a neighborhood U of x_0, then it is clear that $\partial u(x_0)/\partial\nu \ge 0$. Following Hopf, it can be proved that

$$\boxed{u(x) < u(x_0), \quad -L_0 u(x) \le 0 \qquad \forall x \in U \cap \Omega \quad \Rightarrow \quad \partial u(x_0)/\partial\nu > 0.} \tag{2.6}$$

For $N = 1$ this implication is easily checked. For $N > 1$ the argument is not trivial; see see e.g. [EgSh] p. 86.

Notice that if $c \le 0$ in U and $u(x_0) > 0$, then, possibly restricting the neighborhood U of x_0, $u > 0$ in U; hence $cu \le 0$ in U. Therefore

$$-L_0 u := -Lu + cu \le -Lu \qquad \text{in } U \cap \Omega. \tag{2.7}$$

By (2.6) we then extend (2.6) as follows:

$$\begin{gathered} u(x_0) > 0, \quad u(x) < u(x_0), \quad c(x) \le 0, \quad -Lu(x) \le 0 \qquad \forall x \in U \cap \Omega, \\ \Rightarrow \quad \partial u(x_0)/\partial\nu > 0. \end{gathered} \tag{2.8}$$

The implication (2.6) (and then also (2.8)) can be refined. For instance, the hypothesis $u(x) < u(x_0)$ can be replaced by the weaker condition

$$\begin{gathered} \text{there exists a neighborhood } U \text{ of } x_0 \text{ such that} \\ u \text{ is not constant in } U \text{ and } u(x) \le u(x_0) \text{ in } U. \end{gathered} \tag{2.9}$$

This refinement entails the following strong maximum principle (for the argument, see e.g. [ReRo] p. 109):

Theorem 2.1 (Hopf) *Let Ω be a (possibly unbounded) connected Euclidean domain of Lipschitz class, and $u \in C^2(\Omega) \cap C^0(\bar{\Omega})$. Then:*

(i) if $-L_0 u \le 0$ in Ω, then u can attain its maximum in Ω only if it is constant;

(ii) if $-Lu \le 0$ and $c \le 0$ in Ω, then u can attain a positive maximum in Ω only if it is constant.

[13] Here we explain why one speaks of *monotone dependence on the data.* Let us consider the operator $A : C^0(\bar{\Omega}) \cap C^2(\Omega) \to C^0(\bar{\Omega}) \times C^0(\partial\Omega) : u \mapsto (-Lu, u|_{\partial\Omega})$. The above assertion states the monotonicity of the inverse operator A^{-1}, if it exists. (In passing notice that the operator A fails to be monotone, for any choice of the coefficients).

By applying the weak/strong maximum principle to $-u$, analogous weak/strong minimum principles are easily derived.

Maximum principle for parabolic operators. Let L_0 and L be as above, fix any $T > 0$, and set

$$\mathcal{L}_0 := L_0 - D_t, \quad \mathcal{L} := L - D_t \qquad \text{in } Q := \Omega \times]0,T[. \tag{2.10}$$

Let us also set $\Sigma := (\partial\Omega)\times]0,T]$, and define the *parabolic boundary* $\partial_p Q := (\Omega\times\{0\}) \cup \Sigma$.

The weak maximum principle states that

$$(i) \qquad -\mathcal{L}_0 u \leq 0 \quad \text{in } Q \quad \Rightarrow \quad \max_{\bar{Q}} u \leq \max_{\partial_p Q} u, \tag{2.11}$$

that is, the maximum of u is attained on the parabolic boundary (possibly in the interior, too); and

$$(ii) \qquad -\mathcal{L}u \leq 0, \quad c \leq 0 \quad \text{in } Q \quad \Rightarrow \quad \max_{\bar{Q}} u \leq \max_{\partial_p Q} u^+, \tag{2.12}$$

that is, a positive maximum of u is attained on the parabolic boundary (possibly in the interior, too).

Let $f \in C^0(\bar{Q})$, $u^0 \in C^0(\bar{\Omega})$ and $g \in C^0(\partial_p Q)$ be prescribed, and consider the *initial- and boundary-value problem*

$$-\mathcal{L}u = f \quad \text{in } Q, \qquad u = u^0 \quad \text{on } \Omega\times\{0\}, \qquad u = g \quad \text{on } \Sigma. \tag{2.13}$$

This may be labelled as a nonhomogeneous Cauchy-Dirichlet problem. Still assuming that $c \leq 0$ in Ω, (2.12) entails the *monotone dependence on the data:*

$$\begin{cases} -\mathcal{L}u_1 \leq -\mathcal{L}u_2 & \text{in } Q \\ u_1^0 \leq u_2^0 & \text{in } \Omega \\ u_1 \leq u_2 & \text{on } \Sigma \end{cases} \quad \Rightarrow \quad u_1 \leq u_2 \quad \text{in } Q. \tag{2.14}$$

In turn this yields the uniqueness of the solution.

A strong maximum principle can also be derived for parabolic equations.

Theorem 2.2 *Let Ω be a (possibly unbounded) connected Euclidean domain of Lipschitz class, and $u \in C^2(Q) \cap C^0(\bar{Q})$. Then:*

(i) if $-L_0 u \leq 0$ in Q, and u attains its maximum at some $(x_0,t_0) \in Q$, then $u(x,t) = u(x_0,t_0)$ for any $(x,t) \in \Omega\times[0,t_0]$;

(ii) if $-Lu \leq 0$ and $c \leq 0$ in Q, and u attains a positive maximum at some $(x_0,t_0) \in Q$, then $u(x,t) = u(x_0,t_0)$ for any $(x,t) \in \Omega\times[0,t_0]$.

By applying the weak/strong maximum principle to $-u$, analogous weak/strong minimum principles are easily derived also in the parabolic case.

A priori estimates via maximum principle. Let Ω be a bounded domain and consider the Dirichlet problem

$$\begin{cases} -Lu := -\Delta u + \tilde{c}(x)u = f(x) & \text{in } \Omega \\ u = g \quad \text{on } \partial\Omega, \end{cases} \tag{2.15}$$

with $\tilde{c} \in C^0(\bar{Q})$, $\tilde{c} \geq 0$, $f \in C^0(\bar{Q})$ and $g \in C^0(\partial Q)$. Let us set

$$\begin{aligned} &\delta := \max\nolimits_{\bar{\Omega}} x_1, \quad M := \max\nolimits_{\bar{\Omega}} f, \quad G := \max\nolimits_{\partial\Omega} g, \\ &w := Me^{\delta - x_1} + G \qquad \forall x \in \Omega. \end{aligned} \tag{2.16}$$

It is promptly checked that then

$$\begin{cases} -Lw = Me^{\delta - x_1} + \tilde{c}(x)Me^{\delta - x_1} + \tilde{c}(x)G \geq M \geq f(x) = -Lu & \text{in } \Omega \\ u \leq w \quad \text{on } \partial\Omega. \end{cases} \tag{2.17}$$

By the weak maximum principle, we conclude that $u \leq w$ in Ω.

If we redefine M and G as $M := \max_{\bar{\Omega}} |f|$ and $G := \max_{\partial\Omega} |g|$, then we also have $-u \leq w$ in Ω; therefore $|u| \leq w$ in Ω.

A priori of knowledge of the solution, we thus conclude with the following estimate.

Proposition 2.3 *Under the previous assumptions, there exists a constant $C > 0$ such that*

$$\max_{\bar{\Omega}} |u| \leq C \max_{\bar{\Omega}} |Lu| + \max_{\partial\Omega} |u|. \tag{2.18}$$

It suffices to define δ as above and set

$$C = \max_{x \in \bar{\Omega}} \; e^{\delta - x_1} \; (= e^{\delta - \inf_{x \in \bar{\Omega}} x_1}).$$

This result can be extended to more general second-order operators, see e.g. [ReRo] p. 109.

It is promptly checked that the latter inequality is tantamount to

$$\frac{1}{C} u(y) \leq u(x) \leq Cu(y) \qquad \forall x, y \in V. \tag{2.19}$$

2.1 Some References

(The following books are available at the Povo library)

[EgSh] Yu.V. Egorov, M.A. Shubin: Foundations of the classical theory of partial differential equations. Springer , Berlin 1992

[Ev] L.C. Evans: Partial differential equations. American Mathematical Society, Providence, RI, 1998

[ReRo] M. Renardy, R. Rogers: An introduction to partial differential equations. Springer-Verlag, New York, 2004

[Sa] S. Salsa: Equazioni a derivate parziali: metodi, modelli e applicazioni. Springer Italia, Milano 2003 (English edition also available)

[SaVe] S. Salsa, G. Verzini: Equazioni a derivate parziali: complementi ed esercizi. Springer Italia, Milano 2005.