The Cauchy-Kowalevski Theorem and Its Consequences

Candidate: Alessandro Pedone,
Advisor: Prof. Maurizio Grasselli

Politecnico di Milano

24 september 2024

Indice

Sofya Vasilyevna Kovalevskaya (1850-1891)

We assume the historical figure of Augustin-Louis Cauchy is well-known.
Kowalevski was:

- a Russian mathematician and student of Weierstrass
- the **first woman** to earn a doctorate (3 theses dating back to 1874) and to obtain a chair in Europe (in mathematics)

There are several **artistic representations** of her in both literature and cinema. The most notable are:

- An accurate biography: *Little Sparrow: A Portrait of Sophia Kovalevsky* (1983), Don H. Kennedy
- A short story: *Too Much Happiness* (2009), Alice Munro

Guiding Questions

Is it possible for an analytical solution
of a PDE system
with Cauchy conditions to exist?

The answer is affirmative, so we already ask:

- under what assumptions?
- is the solution unique?
- is the problem well-posed?
- what are the consequences of the obtained results?

Types of Equations (and Operators)

Equations of order k:

Linear	$\sum_{\lvert\alpha\rvert\leq k} a_\alpha\, D^\alpha u = f$
Quasi-linear	$\sum_{\lvert\alpha\rvert=k} a_\alpha(x, D^\beta u)\, D^\alpha u + a_0(x, D^\beta u) = f,$ $\lvert\beta\rvert < k$
Non-linear	$F(x, D^\alpha u) = 0, \quad \lvert\alpha\rvert \leq k$
In normal form	$D_t^k u = G(x, t, D_x^\alpha D_t^j u), \quad \lvert\alpha\rvert + j \leq k,\ j < k$

Tools

- Characteristic surfaces
- Method of characteristics
- Cauchy problems
- Power series

Characteristic Surfaces for Linear Operators

L linear differential operator.

Definizione 2.1

Characteristic form of L:
$\chi_L(x,\xi) = \sum_{|\alpha|=k} a_\alpha(x)\,\xi^\alpha \quad \text{with} \quad x,\xi \in \mathbb{R}^n$

Definizione 2.2

Characteristic variety of L at x:
$\mathrm{char}_x(L) = \{\xi \neq 0 : \chi_L(x,\xi) = 0\}$

Definizione 2.3

Γ characteristic surface for L at $x \iff \nu(x) \in \mathrm{char}_x(L)$

Osservazione

Case of 1st order operator: $A = (a_1, \ldots, a_n)$ tangent to Γ.
Useful for further generalizations.

Meaning

$$\xi \in \mathrm{char}_x(L)$$

at x L is not "properly" of order k in the direction ξ.

$$\Gamma \text{ non-characteristic}$$

given $D^i_\nu u\,(i < k)$ of a solution u on Γ it is possible to calculate all its partial derivatives on Γ.

1st Order Quasi-linear Operators

- $\gamma(s) : \mathbb{R}^{n-1} \to \mathbb{R}^n$ local parametrization of Γ
- $u = \phi$ on Γ Cauchy data

Definizione 2.4

Γ non-characteristic at $x_0 = \gamma(s_0)$

$$\iff \det \left[\underbrace{\begin{matrix} D_{s_1}\gamma_1 & \cdots & D_{s_{n-1}}\gamma_1 \\ \vdots & & \vdots \\ D_{s_1}\gamma_n & \cdots & D_{s_{n-1}}\gamma_n \end{matrix}}_{\text{span of the tangent plane}} \left| \begin{matrix} a_1(\gamma, \phi(\gamma)) \\ \vdots \\ a_n(\gamma, \phi(\gamma)) \end{matrix} \right. \right] (s_0) \neq 0$$

Method of Characteristics

The following problems[1] are **equivalent**.

$$PDE : \begin{cases} \sum a_j(x,u) D_{x_j} u = b(x,u) \\ u = \phi \text{ on } \Gamma \end{cases} \tag{1}$$

$$ODE : \begin{cases} D_t\, x = A(x,y)\ ^2 \\ D_t\, y = b(x,y) \\ x(0) = x_0 \\ y(0) = \phi(x_0) \quad \forall x_0 \in \Gamma \end{cases} \tag{2}$$

Where $y = u(x)$ and $A(x,y) = [a_1(x,y), \dots, a_n(x,y)]$.

[1] it can be generalized to the nonlinear case (1st order!)

[2] the solutions x are called *characteristic curves*

Teorema 2.1

Hp	Problem (1) a_j, b, ϕ, $\Gamma \in C^1$ Γ non-characteristic
Ts	$\exists!$ unique C^1 solution in a neighborhood of Γ
Proof	using the local existence and uniqueness theorem for ODEs

Cauchy Problem

- Often used when the data surface is **not** a boundary.
- It also requires the **normal derivatives** $(D^j_\nu u)$ of the solution on the surface to uniquely determine it.
- It carries the risk of being **overdetermined** (good for uniqueness but less for the existence of the solution).

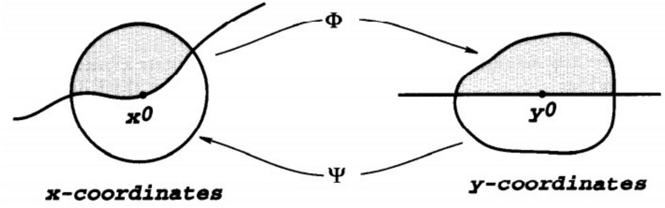
General Problem

$$
\begin{cases} F^*(x, D^\alpha u^*) = 0 & |\alpha| \leq k,\ F^* \text{ at least } C^1 \\ D_\nu^j u^* = \phi_j^* & \text{on } \Gamma^* \text{ for } j < k \end{cases}
$$

Mapping at $t = 0$

Let γ^* be the local parametrization of Γ^*, we apply the map:

$$\Phi(x) = \left[x_1 \quad \cdots \quad x_{n-1} \;\middle|\; x_n - \gamma^*(x_1, \ldots, x_{n-1})\right]$$



L. C. Evans, *Partial Differential Equations*

1. Select a privileged variable and call it "time":

$$t \leftarrow x_n$$
$$x \leftarrow (x_1, \dots, x_{n-1})$$

2. Call $\Gamma_0 = \{t = 0\}$.
3. Indicate the derivatives as follows: $D_x^\alpha D_t^j u$.
4. Obtain the problem ($u^* = u(\Phi)$):

$$\begin{cases} F(x, t, D_x^\alpha D_t^j u) = 0 & |\alpha| + j \leq k \\ D_t^j u(x, 0) = \phi_j(x) & \text{for } j < k \end{cases}$$

Non-characteristic Surfaces in General

Definizione 2.5

Γ^* (or Γ_0) is non-characteristic $\iff$ the equation on Γ_0 can be rewritten in **normal form** with respect to t.

Osservazione

It is shown to be consistent with previous definitions.

Osservazione

- Linear case $\rightarrow$ condition on the coefficients.
- Non-linear case $\rightarrow$ validity of implicit function theorem hypotheses on F.

Remarkable Power Series

Definizione 2.6

Majorizing function:

$$\mathcal{M}_{Cr}(x) = \frac{Cr}{r - (x_1 + \ldots + x_n)}$$

Osservazione

By the multinomial theorem, if $|x| < r/n$ we have that

$$\frac{Cr}{r - (x_1 + \ldots + x_n)} = C \sum_{\alpha} \frac{|\alpha|!}{\alpha! \, r^{|\alpha|}} x^{\alpha}.$$

Method of Majorants

Teorema 2.2 (utility of the majorant)

$$\begin{cases} g_\alpha \geq |f_\alpha| \\ \sum g_\alpha x^\alpha \text{ has conv. radius } R \end{cases} \implies \begin{array}{c} \sum f_\alpha x^\alpha \text{ has a radius} \\ \text{of at least } R \end{array}$$

In this case, we write: $\sum g_\alpha x^\alpha \gg \sum f_\alpha x^\alpha$.

Teorema 2.3 (construction of the majorant)

$\sum f_\alpha x^\alpha$ has radius $R \implies \exists\, r < R$, $C > 0$ such that

$$|f_\alpha| \leq C \frac{1}{r^{|\alpha|}} \leq C \frac{|\alpha|!}{\alpha!\, r^{|\alpha|}}$$

Outline of the Approach

Following the chronological order of discovery, we proceed by **progressive generalizations**:

1. ODEs
2. Quasi-linear PDEs
3. PDEs in normal form

ODE

Teorema 3.1

Hp

$A \subseteq \mathbb{C}$, $B \subseteq \mathbb{C}^n$ open

$\Omega \subseteq A$ open, connected

$f : A \times B \to \mathbb{C}^n$ holomorphic

Pb: $\begin{cases} y' = f(x, y) & \forall x \in \Omega \\ y(x_0) = y_0 \end{cases}$

Ts

locally there exists a unique holomorphic solution

Radius Estimate

Teorema 3.2

Hp

Assumptions of the previous theorem

$\exists\, \overline{B_a(x_0)} \subseteq A,\ \overline{B_b(y_0)} \subseteq B$

Ts

The solution converges with at least radius[3]

$\widetilde{r} = a\left[1 - \exp\left(-\frac{b}{aM(n+1)}\right)\right]$

[3] $M = \max_{B_a(x_0),\, B_b(y_0)} |f|$

Quasi-linear PDEs

Teorema 3.3

Hp

A_i, B analytic

Pb: $\begin{cases} D_t\, y = \sum_{i=1}^{n-1} A_i(x,y) D_{x_i} y + B(x,y) \\ y = 0 \quad \text{on } \Gamma_0 \end{cases}$

Ts

$\exists!\ y(x,t) : \mathbb{R}^n \to \mathbb{R}^m$ analytic solution in a neighborhood of the origin

Proof

1. Assume $y_h = \sum c^h_{\alpha j} x^\alpha t^j$
2. Inserting the series of y, A_j, B we get:

$$c^h_{\alpha j} = Q^h_{\alpha j}(\text{coeff. of series of } A_i, B)$$

   Q polynomial with non-negative coefficients
3. $\widetilde{A}_i \gg A_i$, $\widetilde{B} \gg B \implies \widetilde{y} \gg y$ thanks to Q
4. Choose $\widetilde{A}_i$, $\widetilde{B}$ so that $\widetilde{y}$ can be explicitly calculated as analytic with the method of characteristics

Majorizing System

As we already know, we majorize the series with

$$\mathcal{M}_{Cr}(x,y) \gg A_i(x,y),\ B(x,y)$$

and solve the problem[4]:

$$\begin{cases} D_t\,\widetilde{y}_h = \mathcal{M}_{Cr}(x,\widetilde{y}) \left[\sum_{i,j} D_{x_j}\widetilde{y}_i + 1 \right] \\ \widetilde{y}_h = 0 \quad \text{ on } \Gamma_0 \end{cases}$$

[4] with $h = 1, \ldots, m$

Majorant Solution

The previous system has the solution:

$$\widetilde{y}_h(x,t) = u(x_1 + \cdots + x_n,\, t) \quad \forall h$$

with

$$u(s,t) = \frac{r - s - \sqrt{(r-s)^2 - 2tCrmn}}{mn},$$

whose radius of convergence we can study.

Radius of Convergence Estimate

Teorema 3.4

The solution of theorem 3.3 converges with radius at least

$$\widetilde{r} = \frac{1}{n-1}\,\frac{r}{8Cmn} \text{ with } C \geq \frac{1}{2}$$

Let's observe its behavior[5] with respect to r, knowing that:

$$r < \min\{\textit{radii of conv. of the coefficients } a^i_{ml},\, b_m\}$$

$$C \geq \max\left\{\begin{array}{l} \max\limits_{i,m,l,\alpha} \left|(a^i_{ml})_\alpha\, r^{|\alpha|}\right| \\ \max\limits_{m,\alpha} \left|(b_m)_\alpha\, r^{|\alpha|}\right| \end{array}\right\}$$

[5] *trade-off* Cr

PDE in Normal Form

Teorema 3.5

The following two problems are equivalent

$$\text{nonlinear}: \begin{cases} D_t^k u = G(x,t,D_x^\alpha D_t^j u) & |\alpha| + j \leq k,\ j < k \\ D_t^j u = \phi_j & \text{on } \Gamma_0,\ j < k \end{cases}$$

$$\text{quasi-linear}: \begin{cases} D_t\, y = \sum_{i=1}^{n-1} A_i(x,y) D_{x_i} y + B(x,y) \\ y = 0 \quad \text{on } \Gamma_0 \end{cases}$$

Proof

1 The system is constructed so that $y_{\alpha j} = D_x^{\alpha} D_t^{j} u$

The matrices A_i and B will then be derived from the expressions[6]:

$$
\begin{aligned}
D_t y_{\alpha j} =& y_{\alpha(j+1)} && |\alpha| + j < k \\
D_t y_{\alpha j} =& D_{x_l} y_{(\alpha - e_l)(j+1)} && |\alpha| + j = k,\ j < k \\
D_t y_{0k} =& D_t G + \sum_{|\alpha|+j<k} D_{y_{\alpha j}} G y_{\alpha(j+1)} \\
& + \sum_{|\alpha|+j=k,\ j<k} D_{y_{\alpha j}} G \, D_{x_l} y_{(\alpha - e_l)(j+1)}
\end{aligned}
$$

The Cauchy data will be:

$$
\begin{aligned}
y_{\alpha j}(x,0) =& D_x^\alpha \phi_j(x) && j < k \\
y_{0k}(x,0) =& G\left(x, 0, D_x^\alpha \phi_j(x)\right) && |\alpha| + j \le k,\ j < k
\end{aligned}
$$

[6] $l(\alpha) = \min\{l : \alpha_l \neq 0\}$

2. removing ϕ : $y(x,t) \leftarrow y(x,t) - \phi(x)$
3. removing t : the variable $y^0 = t$ is added (with its corresponding equation)

Holomorphic Version

As in the case of ODEs, everything extends in an **immediate** way to the complex case by assuming holomorphic data.

Examples

We now answer the questions with three examples:

- Lewy's example: importance of analyticity
- Kowalevski's example: importance of non-characteristicity
- Hadamard's example: the problem might not be well-posed

Lewy's Example

Definizione 4.1

$$\mathcal{L} = D_x + iD_y - 2i(x + iy)D_t$$

is called Lewy's operator.

Teorema 4.1

Hp	f continuous real-valued function depending only on t $u \in C^1 \ : \ \mathcal{L}u = f$ in a neighborhood of the origin
Ts	f analytic in a neighborhood of $t = 0$
Proof	Schwarz reflection principle

The previous statement can be generalized as follows:

Teorema 4.2

Hp	$A \subseteq \mathbb{R}^3$ open
Ts	$\exists\, F \in C^\infty(\mathbb{R}^3, \mathbb{R}) \;:\; \nexists\, u \in C^1(A, \mathbb{R})$ such that $\begin{cases} \mathcal{L}u = F \text{ in } A \\ u_x,\, u_y,\, u_t \text{ satisfy the Hölder condition} \end{cases}$

Proof

1. Translate the problem of the previous theorem so as to reduce it to the case of a generic point (x_0, y_0, t_0), using the function $g(x, y, t) = f(t - 2xy_0 + 2x_0y)$ as the forcing function.
2. Construct a function $S_a \in C^\infty$ for each $a \in l^\infty$ using a series.
3. Construct closed sets $E_{j,n} \subseteq l^\infty$ with no interior using S_a and the Ascoli-Arzelà theorem.
4. Conclude the proof of the new theorem using the aforementioned lemmas to derive, by a contradiction argument, the equality $l^\infty = \bigcup E_{j,n}$, allowing the application of Baire's argument.

Kowalevski's Example

This problem admits no analytic solutions[7] in a neighborhood of the origin:

$$\begin{cases} u_t - u_{xx} = 0 \\ u(x,0) = \frac{1}{1+x^2} \quad \forall\, x \in \mathbb{R} \end{cases}$$

Osservazione

The surface is characteristic!

[7] proof by contradiction

Hadamard's Example

$$\begin{cases} u_{xx} + u_{yy} = 0 \\ u(x,0) = 0 \\ u_y(x,0) = n\sin(nx)e^{-\sqrt{n}} \text{ with } n \in \mathbb{N} \end{cases}$$

The solution to this problem is:

$$u_n(x,y) = \sin(nx)\underbrace{\sinh(ny)e^{-\sqrt{n}}}_{\xrightarrow{n\to\infty}\infty}$$

Alternative Versions

Abstract Version
(Ovsyannikov classes)

$$\Downarrow$$

Classical Version
(similar to local existence and uniqueness for ODEs)

$$\Downarrow$$

Invariant Version
(non-characteristic surfaces)

Classical Version

Teorema 5.1

Hp

$\overline{\mathcal{O}}_0 \subseteq \mathcal{O}_1 \subseteq \mathbb{C}^n$ open connected bounded

A_j, f, y_0 holomorphic wrt z

A_j, f continuous wrt t

$$\text{Pb: } \begin{cases} D_t y = \sum A_j(z,t) D_{z_j} y + A_0(z,t) y + f(z,t) \\ y(z,0) = y_0(z) \end{cases}$$

Ts

$\exists\, \delta \in (0,T) : \exists!\, y$ solution when $|t| < T$

$-$holomorphic wrt z

$- C^1$ wrt t $\quad \rightarrow (\neq$ Holmgren$)$

Consequences

The consequences of this theorem can be observed in various fields, including the main ones:

- theory of differential equations
- mathematical physics: emergence of numerous questions (what happens in reality if a local analytic solution exists?)
- differential geometry
- economic theory

Impact on the theory of differential equations:

- refuting Weierstrass's conjecture
- Holmgren's theorem
- research on necessary and/or sufficient conditions for the existence of local solutions by Treves and Nirenberg
- Hörmander's theory of linear differential operators

Holmgren's Theorem

Result of **uniqueness** of solutions for linear PDEs.

Osservazione

Cauchy-Kowalevski theorem does not exclude the existence of other solutions that are not analytic!

$$
\begin{array}{r|lcccc}
\text{CK} & \text{abstract} & \Longrightarrow & \text{classical} & \Longrightarrow & \text{invariant} \\
 & \Downarrow & & & & \\
\text{H} & \text{abstract} & \Longrightarrow & \text{classical} & \Longrightarrow & \text{invariant}
\end{array}
$$

Abstract Version

Any linear equation can be reduced to a **first-order system**.
We focus on this case.

Teorema 6.1

Hp

- y distribution on $(\mathcal{O}_0 \cap \mathbb{R}^n) \times (-T, T)$:
- $K \subseteq \mathcal{O}_0 \cap \mathbb{R}^n$ compact: $y = 0$ in $\mathcal{O}_0 \cap \mathbb{R}^n \setminus K$
- $\begin{cases} D_t y = \sum A_j(z,t) D_{z_j} y + A_0(z,t) y \\ y = 0 \text{ for } t < 0 \end{cases}$

Ts

$y = 0$ in $(\mathcal{O}_0 \cap \mathbb{R}^n) \times (-T, T)$

Classical Version

Teorema 6.2

Hp	$\Omega \subseteq \mathbb{R}^n$ open A_j analytic $y \in C^1(\Omega \times (-T,T))$: $\begin{cases} D_t y = \sum A_j(x,t) D_{x_j} y + A_0(x,t) y \\ y = 0 \text{ for } t = 0 \end{cases}$
Ts	$y = 0$ in a neighborhood of $\Omega \times \{0\}$

Proof

It is an application of the abstract version to the function

$$\widetilde{y}(x,t) = H(t)\, y(x,t),$$

which always satisfies a system of the same type.

Cartan-Kähler Theorem

A very important theorem in differential geometry:

- on the integrability of **exterior differential systems**
- which is proved using the Cauchy-Kowalevski theorem
- which has an application in the economic field (I. Ekeland, P.A. Chiappori)

Quoting Ekeland regarding the paper written in 1999 with Chiappori:

This paper solves a basic problem in economic theory, which had remained open for ***thirty years****, namely the characterization of market demand functions. The method of proof consists of reducing the problem to a system of nonlinear PDEs, for which convex solutions are sought. This is rewritten as an exterior differential system, and is solved by the Cartan-Kähler theorem, together with some algebraic manipulations to achieve* ***convexity****.*

Despite the research conducted in those years

- wasn't guided by immediate applications
- led to **disappointing** results compared to the expectations of Cauchy and Weierstrass

it has had a gigantic impact thanks to the understanding of solutions of PDE systems it allowed us to achieve.

In conclusion, a quote about the relationship between Weierstrass and Kowalevski:

All his life – he had difficulty saying this, as he admitted, being always wary of too much enthusiasm – all his life he had been waiting for such a student to come into this room. A student who would challenge him completely, who was not only capable of following the strivings of his own mind but perhaps of flying beyond them.

— Alice Munro, *Too Much Happiness*