Politecnico di Milano

Corso di Laurea Triennale in Ingegneria Matematica
Scuola di Ingegneria Industriale e dell'Informazione

The Cauchy-Kowalevski theorem and some of its consequences

Thesis by

Alessandro Pedone
Student ID 981105

Advisor:

Prof. Maurizio Grasselli

Graduation Session September 2024
Academic Year 2023/2024

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*

Abstract

In 1874, Sofya Kowalevski, the first woman to obtain a doctorate in mathematics in Europe, brought to light the proof of the Cauchy-Kowalevski theorem (CKT), the first general result for the existence of local analytic solutions to partial differential equations (PDEs) with Cauchy data.
The thesis aims to present this milestone of mathematics, highlighting the depth of detail, consequences, and the simplicity of the ideas it brought to light. To this end, recurring references to fundamental notions and results are made to address the topic, and all the main forms in which the CKT can be stated are also discussed.
Additionally, there is a section dedicated to three historically crucial examples for understanding PDEs and another dedicated to two fundamental applications of the CKT: the Holmgren theorem and the Cartan-Kähler theorem.

Keywords: PDE, characteristics, analyticity/holomorphy, power series, majorant method, Cauchy-Kowalevski, Holmgren, and Cartan-Kähler theorems

Contents

Chapter 1
Introduction

1.1 Who was Kowalevski?

Sofya Vasilyevna Kovalevskaya (1850-1891) was a Russian mathematician. For various reasons, including the theorem central to this discussion, she remains one of the most prominent female figures in the history of this discipline.
First of all, it is important to note that, from this point forward, we will often refer to her by the name she used to sign her publications, i.e., Kowalevski.
In order to leave Russia, she entered into a marriage of convenience, marrying a man with whom she had no real emotional relationship for several years and from whom she was often geographically distant. This allowed her to continue her studies in Germany, where she met Karl Weierstrass, one of the most influential mathematicians of his time. After their initial meeting in the professor's study, their relationship continued to develop due to Kowalevski's evident mathematical talents, which Weierstrass could not help but nurture. He continued to give her private lessons and eventually supervised her research work.
Regarding Kowalevski's political views, we can say with historical certainty that she was close to feminist movements and socialist and radical ideas, which can be traced back to her family background and the influences she encountered during her life in modern-day European states. It is certainly noteworthy that she received several copies of radical magazines from her sister Anna, which discussed the so-called Russian nihilism[1].
However, we want to focus on her contribution to mathematics rather than her political, social, and philosophical ideas. With the help of what we might call her mentor, Kowalevski made several important discoveries. After years of collaboration, she published three doctoral theses in a single year: 1874. This is remarkable not only for the sheer volume but also because she became the first woman to earn a PhD, thanks in part to the support of Weierstrass, as revealed in a letter he wrote to his colleague Fuchs at the University of Berlin regarding the approval of Kowalevski's theses.

[1]science, not religion or superstition, was considered by Russian nihilists to be the most effective means of helping the population lead better lives and thus represented truth and progress.

Moreover, her publications turned out to be milestones in mathematics. In particular, the topics covered are:

- Partial differential equations (PDE), Cauchy-Kowalevski theorem
- Mechanics, Kowalevski top
- Elliptic integrals

After the success, which was naturally followed by some awards, she returned to Russia for a period; however, this proved to be futile for the continuation of her academic career. Later, when the husband to whom she owed the opportunity to study in Germany passed away, she moved to Sweden, where she achieved another first: she became the world's first female professor of mathematics, obtaining a position at Stockholm University. Unfortunately, her life was cut short at the age of 41 by pneumonia, which, according to sources, prevented her from pursuing her great passion for literary production. Despite not being able to express herself in this field as she would have liked, there are numerous artistic representations of her, both in literature and cinema.
Here are some of the major cinematic works:

- *Sofya Kowalevski*(1985, Lenfilm, 3 episodes, 218 minutes), Ayan Gasanovna Shakhmaliyeva (1932-1999, originally from Azerbaijan).
- *A Hill on the Dark Side of the Moon* (Swedish: *Berget på månens baksida*)(1983), Lennart Hjulström (1938-2022)

Here are some of the major literary works:

- a biography: *Sonja Kovalevsky. What I Lived With Her and What She Told Me About Herself* (1892, Ed. Albert Bonniers, Stockholm), Anne Charlotte Leffler (a close friend of Kowalevski, sister of mathematician Gösta Mittag-Leffler and wife of the Italian algebraist Pasquale del Pezzo)
- an autobiography: *A Russian Childhood* (1978, Springer New York, NY), Sofya Kovalevskaya, translated and edited by Beatrice Stillman
- a biography: *Little Sparrow: A Portrait of Sophia Kovalevsky* (1983, Ohio University Press, Athens, Ohio), Don H. Kennedy
- a biographical novel: *Beyond the Limit: The Dream of Sofya Kovalevskaya* (2002, Tom Doherty Associates, LLC), Joan Spicci (mathematician and educator)
- a biographical short story: *Too Much Happiness*[2] (2009, Harper's Magazine), Alice Munro (1931-2024, Nobel Prize in Literature)

[2]this story recounts the last days of Kowalevski's life enriched with reminiscences of the past that Munro acquired from letters, diaries, and writings (documents she accessed through Don H. Kennedy's wife, a distant relative of Kowalevski)

1.2 The Cauchy-Kowalevski Theorem

Having introduced the historical figure, we can now take the first step toward exploring one of Kowalevski's major works: the Cauchy-Kowalevski theorem, which we will abbreviate as CKT from now on.
First, let us quickly describe the scientific context of the time concerning PDEs.
The father of the research carried out in the 19th century was Augustin-Louis Cauchy, a mathematician likely well-known to the reader. During those years, particularly between 1835 and 1842, Cauchy was developing the theory of holomorphic functions, already initiated by other major figures like Euler, Laplace, and Fourier.
Cauchy had the intuition to apply these results to differential equations.
It is important to grasp, with the mindset of that era, that classical theory and power series were promising tools, primarily for their simplicity and elegance, but also for the approximation potential of a simple truncation of a series.
Cauchy's attempt to apply the tools from his research to differential equations was successful, but only partially, for one simple reason: he was unable to go beyond the study of ordinary differential equations (ODEs) and linear PDEs.
The breakthrough came thanks to Kowalevski and Weierstrass. The latter was very optimistic about the results he thought could be achieved, perhaps even more so than Cauchy: he conjectured that it would be possible to define analytic functions via differential equations, using formal power series derived from the equations' expressions. For this reason, he encouraged Kowalevski, with her talent, to delve deeper into this subject.
However, it would be wrong to think that Kowalevski's guides were only Cauchy and Weierstrass: other mathematicians also worked on these topics, including Briot, Bouquet, and Fuchs, who further developed the concepts of singularities, and Jacobi, who was the first to define the normal form of an equation[3].
Building on these foundations, Kowalevski's key idea can be summarized as follows:

1. perform a variable change that allows a nonlinear equation to be written in normal form (see chapters 2 and 3 for the meaning of this term), maintaining regularity assumptions on the data, and address the existence of a solution for this system;

2. transform any equation in normal form into a particular quasi-linear system;

3. apply to this system the majorant method already used by Cauchy for his discoveries on ODEs and linear PDEs.

As often happens in mathematics, the proof was later simplified by E. Goursat in his mathematics textbook from around 1900. Moreover, over time, more abstract and general statements and proofs were proposed, thanks to the work of Ovsyannikov, Treves, and Nirenberg.
It should be noted that, during the same period, Darboux also achieved results very similar to Kowalevski, but with less generality.

[3]this concept, in particular, would prove crucial in Kowalevski's research

In light of what has been said so far, we pose some crucial questions, to which we seek the most comprehensive answers possible, and which will guide our discussion:

- is it possible for a system of PDEs with Cauchy data to have an analytic solution?

If so,

- under what assumptions?
- is the solution unique?
- does the solution depend continuously on the data?
- what are the consequences and applications of the results obtained?

Chapter 2
Essential Concepts and Tools

Before delving into the discussion of the theorem, let us recall some basic notions that will be essential for what we will discuss later. In particular, having a clear understanding of this information will be crucial to ensuring a thorough comprehension of the assumptions considered and the demonstration techniques employed.
First of all, to begin familiarizing ourselves with the notation, let us review the classification of partial differential equations of order k, and consequently the associated operators, with a summary table.

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

Remark. From here on, we will not always pay particular attention to the regularity assumptions on the data of the equations (f, a_α, F, G, and others), since for our purposes it is enough that the statements hold in the case where everything is assumed to be analytic (with a certain radius of convergence). The same applies to the data and the surfaces of the associated Cauchy problems. In any case, when not specified, the regularity can be considered as at least C^1.

Remark. In the case of an equation in normal form, the variables are split between space $x \in \mathbb{R}^{n-1}$ and time t, for a reason that will become clear by the end of this chapter.

We already anticipate that, later on, we will assume the coefficients and functions defining the equations to be very regular, specifically analytic (i.e., locally expandable in power series).

In light of what has been said so far, we realize that there are already some aspects that would be important to focus on. But to be more organized, let us summarize our topics of interest in four points, which reflect the structure of this chapter:

1. **Characteristic surfaces**: that is, those surfaces in $\mathbb{R}^n$ that are closely related to the form of the equation in observation and can pose problems when deciding to assign Cauchy data on them;

2. **Method of characteristics**: in the case of first-order, possibly nonlinear, equations, a PDE can be viewed as a system of ODEs dependent on a parameter;

3. **Cauchy problems**: the only type of problem we will deal with;

4. **Power series**: these form the foundation of the concept of an analytic function (and holomorphic in the case of complex numbers), which is the only type of function we will seek as a solution.

2.1 Characteristic Surfaces

In this first section, we introduce the concept of a characteristic surface in the simplest cases, to fully understand its meaning. Let us begin by considering the simplest situation of all, namely that of a **linear** equation. Such an equation is uniquely determined by the forcing term we called f and by a linear differential operator $L = \sum_{|\alpha|\leq k} a_\alpha D^\alpha$. Let us focus on the latter and provide three definitions.

Definition 2.1.1. The characteristic form of L is

$$\chi_L(x,\xi) = \sum_{|\alpha|=k} a_\alpha(x)\,\xi^\alpha \text{ where } x,\xi \in \mathbb{R}^n.$$

Definition 2.1.2. The characteristic variety of L at x is the set

$$\mathrm{char}_x(L) = \{\xi \neq 0 : \chi_L(x,\xi) = 0\}.$$

Definition 2.1.3. Γ is a characteristic surface for L at x if $\boldsymbol{\nu}(x) \in \mathrm{char}_x(L)$, where $\boldsymbol{\nu}(x)$ denotes the unit normal to Γ at x.

Let us now investigate the meaning of these definitions:

- First of all, notice that when $\xi \in \mathrm{char}_x(L)$, it is as if the operator were not "properly" of order k in the direction ξ.

- Additionally, in the case of a first-order operator ($k = 1$), a surface Γ is characteristic when $A = (a_1, \ldots, a_n)$ is tangent to Γ point by point (i.e., for every $x \in \Gamma$).

- It is possible to show that a characteristic surface "carries more information" when Cauchy conditions are assigned on it. In fact, given the normal derivatives $D^j_\nu u$ ($j < k$) of a function u that we want to satisfy the equation, if Γ is non-characteristic at every point, it is possible to calculate all the partial derivatives of u on Γ.

Especially the last consideration, due to the lack of rigor, may be confusing at first reading. However, there exists a theorem that explicitly shows this result in the case of quasi-linear equations and can be found together with the proof in [Eva10, cap.4.6].

Given that we aim to prove a theorem that will turn out to be very general, we note that, unfortunately, linear equations will not be sufficient to solve all our problems. For this reason, we want to immediately generalize the concept of a non-characteristic surface to the **quasi-linear** case, even though we still remain in the first-order equation scenario. Now, suppose we have the Cauchy problem

$$\begin{cases} \sum a_j(x,u) D_{x_j} u = b(x,u) \\ u = \phi \text{ on } \Gamma \end{cases} \tag{2.1}$$

and that Γ has a local parametrization near $x_0 \in \Gamma$ given by the function $\gamma(s) : \mathbb{R}^{n-1} \to \mathbb{R}^n$, we provide the following generalization, clearly inspired by the case of first-order linear operators.

Definition 2.1.4. Γ is non-characteristic at $x_0 = \gamma(s_0)$ if

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

Now it is time to use these definitions to draw some useful conclusions.

2.2 Method of Characteristics

Let us consider an application of the notion of non-characteristic surface: the method of characteristics for first-order PDEs. This is a method for finding solutions to equations, possibly even fully nonlinear, which is based on the idea of transforming the problem into a system of ODEs that is equivalent.
We start directly from the case of a quasi-linear equation and consider again the corresponding Cauchy problem with data assigned on some surface Γ. We want to show that this problem is **equivalent** to another problem for a system of ODEs.

$$\text{PDE} : \begin{cases} \sum a_j(x,u) D_{x_j} u = b(x,u) \\ u = \phi \text{ on } \Gamma \end{cases} \tag{2.2}$$

$$\text{ODE} : \begin{cases} D_t\, x = A(x,y) \\ D_t\, y = b(x,y) \\ x(0) = x_0,\ y(0) = \phi(x_0) \quad \forall x_0 \in \Gamma \end{cases} \tag{2.3}$$

where $y = u(x)$ and $A(x,y) = (a_1(x,y), \ldots, a_n(x,y))$.

Remark. It is important to highlight three aspects:

- the solutions x are called **characteristic curves**;
- the second problem is parametric with respect to x_0, so the entire solution for u will be given by the union over all $x_0 \in \Gamma$ of all the y along the curves x;
- the case of linear equations is immediate to derive from what has been stated above, simply by assuming that the coefficients a_j depend only on x and that b is of the form $b(x,u) = f(x) - c(x)u$.

Without providing a precise statement, let us proceed with a reasoning that is still rigorous, which can be considered a proof of the equivalence.

Proof: In both directions, a simple derivation of a composed function:

1. Suppose we know, for every x_0, $y(t)$ and $x(t)$ that solve the problem (2.3). Then for every x_0 it holds that:

   $$b(x,y) = D_t y = \sum D_{x_j} y\; D_t x_j = \sum a_j(x,y) D_{x_j} y.$$

   From which it follows that the function $u(x)$ that has a graph given by the union of all the curves $(x(t), y(t))$ solves the problem (2.2).

2. Now assume instead that we know u, a solution of (2.2). We find x by solving $\forall j$:

   $$D_t\, x_j = a_j(x,y), \quad x_j(0) = (x_0)_j$$

   We define $y(t) = u(x(t))$ and finally, we use the same reasoning as before to conclude that y satisfies the equation of the ODE system:

   $$D_t y = \sum D_{x_j} u\; D_t x_j = \sum a_j(x,y) D_{x_j} u = b(x,y).$$

QED

At this point, one might wonder where the idea of verifying the equivalence with that specific system of ODEs arises. The answer to this question is interesting because it encompasses the geometric meaning of this method. In fact, recalling that the normal vector to the graph of a function u is proportional to the vector $(\nabla u, -1)$, we can state that the equation (2.2) tells us that the following vector field must be **tangent** to the graph of u.

$$(a_1(x, u(x)), \ldots, a_n(x, u(x)), b(x, u(x))) = (A(x, u(x)), b(x, u(x)))$$

Understanding this last aspect begins to outline the role of the characteristic property of a surface.
Now let us see a theorem to concretize this intuition.

Theorem 2.2.1.

Hp	*Problem* (2.2) a_j, b, ϕ, $\Gamma \in C^1$ Γ *is non-characteristic*
Ts	$\exists!$ *solution* C^1 *in a neighborhood of* Γ

The complete and detailed proof can be found in [Fol95, cap.1]; here we only mention the fundamental ideas. The uniqueness follows simply from the fact that the graph of the solution u can be seen as the union of the curves $(x(t), y(t))$, which do not intersect if one takes a sufficiently small neighborhood of Γ. To prove existence, the representation of the equation as a parametric set of systems of ODEs is used to carry out the following steps:

1. apply the local existence and uniqueness theorem for ODEs;
2. prove the invertibility of $x(s, t)$, where s is an auxiliary variable related to the local parameterization of Γ, thanks to the fact that it is non-characteristic;
3. thus define the solution $u(x)$ easily following the same idea from point 1 of the last proof made;
4. verify with the derivative of a composed function that u is a solution of the equation.

Both the definition of characteristic surface and the method of characteristics can be generalized to the case of a generic first-order equation. Moreover, there is also a generalization of theorem 2.2.1 for the fully nonlinear case, identical in both spirit and substance to the quasi-linear case. We will not address this topic in detail, as it does not add anything at a qualitative level of understanding of the subject and will not be useful in the subsequent discussion. For further reading, one can refer to [Fol95, cap.1] and [Eva10, cap.3].
However, the notion of non-characteristic surface is not sufficient for our purposes, and in the next paragraph, we want to extend it to the most general possible case: fully nonlinear equations of any order.

2.3 Cauchy Problems

So far, we have only seen the simplest case of a Cauchy problem, namely that for a first-order equation, where it is necessary to assign only the value of the function on a surface. For an equation of any order, this information is not sufficient to uniquely determine the solution; typically, what is done is to also assign the **normal derivatives** of the solution $D^j_\nu u$ with $j < k =$ order of the equation.

In light of what has been stated in the previous two paragraphs, we have already inferred that the notion of non-characteristic surface is useful for identifying those surfaces on which we want to assign Cauchy conditions in such a way as to have some guarantee of the existence of the solution in the neighborhood of the surface. We will now focus on understanding what is meant by a characteristic surface in the most general case we can imagine, following the simplest and most direct approach possible, as it does not require particular proofs. Let us consider the Cauchy problem:

$$\begin{cases} F^*(x, D^\alpha u^*) = 0 & |\alpha| \leq k,\ F^* \\ D^j_\nu u^* = \phi^*_j & \text{on } \Gamma^* \text{ for } j < k \end{cases} \tag{2.4}$$

Regardless of the form of the equation, it is possible to modify this problem in such a way as to locally flatten the boundary of the surface with respect to a variable. To achieve this result, a simple change of coordinates Φ, defined via γ^* (local parameterization of Γ^*), is sufficient:

$$\Phi(x) = \begin{pmatrix} x_1 & \cdots & x_{n-1} \mid x_n - \gamma^*(x_1, \ldots, x_{n-1}) \end{pmatrix}.$$

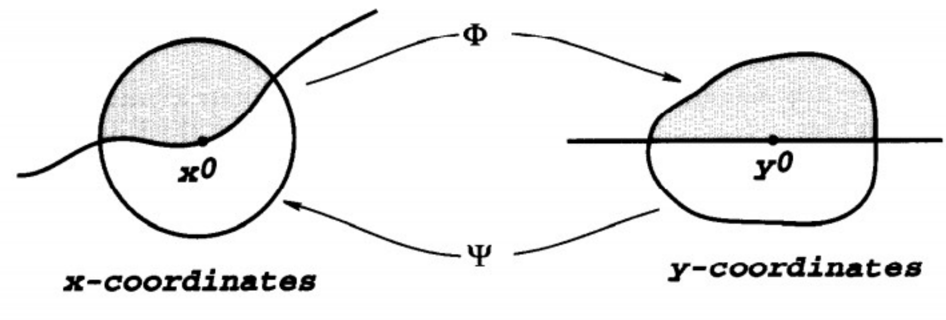


Image from [Eva10, cap.8]

Remark. Note that Φ preserves any analyticity of the surface Γ^*.

This transformation helps us understand how it is possible to choose to consider one variable as "privileged." From now on, this variable will be referred to as "time," and we will denote it with the letter t. To be more precise, we rename the variables as follows:

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

Furthermore, we introduce some notation that will be useful later:

- we denote $\Gamma_0 = \{t = 0\}$;
- we indicate the derivatives as follows: $D_x^\alpha D_t^j u$.

We conclude that thanks to the transformation Φ, we obtain the new problem:

$$\begin{cases} F(x, t, D_x^\alpha D_t^j u) = 0 & |\alpha| + j \leq k \\ D_t^j u(x, 0) = \phi_j(x) & \text{for } j < k \end{cases} \tag{2.5}$$

where $u^* = u(\Phi)$.

Definition 2.3.1. Γ^* (or Γ_0) is non-characteristic if the equation on Γ_0 can be rewritten in **normal form** with respect to t, that is, if the problem (2.5) can be rewritten as follows:

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

To make this definition more concrete, sufficient conditions, and possibly also necessary ones, are often sought so that the equation can be rewritten in normal form, as we did in the paragraph 2.1 for the simplest cases and as was done in [Eva10] and [Fol95]. Let us therefore examine what this involves, distinguishing the various cases and assuming we have already placed ourselves in the situation (2.5):

- linear and quasi-linear: it is required that $a_{(0,\dots,0,k)} \neq 0$ on Γ_0;
- fully nonlinear: it is required that the hypotheses of the implicit function theorem (also known as Dini's theorem) hold for F, that is, $D_{(D_t^k u)} F \neq 0$ on Γ_0.

Remark. Staying within the assumption that the surface is Γ_0, starting from these considerations, it is easy to see how the new definition of non-characteristic surface is consistent with the definitions in paragraph 2.1.

Finally, we recall that the notion of characteristic surface must guarantee us the ability to calculate all the partial derivatives of the solution on the surface. For this reason, the setup of this construction is partially inspired by [Eva10, cap.3], where the proof of this property is presented in two steps:

1. first, we reason assuming to be on Γ_0;
2. thanks to the transformation Φ, we obtain the property for a generic Γ^*.

2.4 Power Series

Assuming the theory of holomorphic functions is known, and consequently also the basic theory of analytic (real) functions, in this paragraph we want to discover, or better understand, only a few very specific tools that will allow us to prove the CKT.
Let's begin by studying a power series expansion of a function that we must not forget.

Definition 2.4.1. The majorant function is given by

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

Using the multinomial theorem, we demonstrate that this function can be developed into a power series for $|x| < r/n$, deriving the expression for the coefficients c_α:

$$\begin{aligned}\mathcal{M}_{Cr}(x) &= \frac{Cr}{r-(x_1+\ldots+x_n)} = C\sum_{j=0}^{\infty}\left(\frac{x_1+\ldots+x_n}{r}\right)^j \\ &= C\sum_{j=0}^{\infty}\frac{1}{r^j}\sum_{\alpha}\binom{|\alpha|}{\alpha}x^\alpha = \sum_{\alpha}\underbrace{C\frac{|\alpha|!}{\alpha!\,r^{|\alpha|}}}_{c_\alpha}\,x^\alpha.\end{aligned}$$

From this result, we want to state two theorems, which constitute the backbone of the so-called majorant method, first devised by Cauchy, and which justify the terminology introduced earlier.

Theorem 2.4.1 (Utility of the Majorant)**.**

Hp	$g_\alpha \geq \lvert f_\alpha\rvert$ $\sum g_\alpha x^\alpha$ *has radius of convergence* R
Ts	$\sum f_\alpha x^\alpha$ *has radius at least* R

Theorem 2.4.2 (Construction of the Majorant)**.**

Hp	$\sum f_\alpha x^\alpha$ *has radius* R
Ts	$\exists\, r < R,\, C > 0 : \lvert f_\alpha\rvert \leq C\frac{\lvert\alpha\rvert!}{\alpha!\,r^{\lvert\alpha\rvert}}$

Proof: It is sufficient to note that taking $C \geq |f_\alpha r^{|\alpha|}|$ immediately implies that

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

QED

In the case where the hypotheses of theorem 2.4.1 hold, we will write: $\sum g_\alpha x^\alpha \gg \sum f_\alpha x^\alpha$.

Remark. The same theorems continue to hold in the case of complex numbers.

We conclude the paragraph and the chapter with some properties for manipulating power series. First of all, let's deal with the operation of composition.

Theorem 2.4.3.

$$\begin{array}{c|l} & y:\mathbb{R}^n\to\mathbb{R}^m \textit{ such that } y(x)=\sum y_\alpha(x-x_0)^\alpha \textit{ in a neighborhood of } x_0\\ Hp & g:\mathbb{R}^m\to\mathbb{R}^d \textit{ such that } g(y)=\sum g_\beta(y-y_0)^\beta \textit{ in a neighborhood of } y_0=y(x_0)\\ & f=g\circ y \end{array}$$

$$\begin{array}{c|l} & \exists\, f_\gamma=P_\gamma(g_\beta,\, y_\alpha \textit{ with } \alpha_i\le\gamma_i) \textit{ set of coefficients such that}\\ Ts & -\ P_\gamma \textit{ are polynomials with non-negative coefficients}\\ & -\ f(x)=\sum f_\gamma(x-x_0)^\gamma \end{array}$$

Remark. The form of the polynomials P_γ does not depend on g and y.

Proof: It is easy to convince oneself of this by explicitly writing the composition of the two series, especially regarding the fact that the coefficient f_γ depends only on the y_α such that $\alpha_i \le \gamma_i$. QED

Assuming, for simplicity, that we are at the origin and recovering the notation (x,t) from the previous paragraph, we state a simple rewriting of this theorem.

Theorem 2.4.4 (Composition)**.**

$$\begin{array}{c|l} & y:\mathbb{R}^n\to\mathbb{R}^m \textit{ such that } y(x,t)=\sum y_{\alpha j}\, x^\alpha t^j \textit{ in a neighborhood of the origin}\\ Hp & g:\mathbb{R}^m\to\mathbb{R}^d \textit{ such that } g(y)=\sum g_\beta\, y^\beta \textit{ in a neighborhood of the origin}\\ & f=g\circ y \end{array}$$

$$\begin{array}{c|l} & \exists\, f_{\gamma k}=P_{\gamma k}(g_\beta,\, y_{\alpha j} \textit{ with } j\le i) \textit{ set of coefficients such that}\\ Ts & -\ P_{\gamma k} \textit{ are polynomials with non-negative coefficients}\\ & -\ f(x,t)=\sum f_{\gamma k}\, x^\gamma t^k \end{array} \tag{2.6}$$

Another way to obtain a series of the type in (2.6) is to use differentiation with respect to any variable x_i. Let's see it with the following theorem.

Theorem 2.4.5 (Differentiation)**.**

$$\begin{array}{c|l} Hp & y:\mathbb{R}^d\to\mathbb{R}^m \textit{ such that } y(x,t)=\sum y_{\alpha j}\, x^\alpha t^j \textit{ in a neighborhood of the origin}\\ Ts & f=D_{x_i}y \textit{ is a power series as in } (2.6) \textit{ in a neighborhood of the origin} \end{array}$$

Proof: By differentiating term by term, we obtain

$$D_{x_i}\sum y_{\alpha j}\, x^\alpha t^j = \sum \underbrace{(\alpha_j+1)\, y_{(\alpha+e_i)j}}_{f_{\alpha j}}\, x^\alpha t^j.\ ^1$$

It is immediate to verify the other properties of $f_{\alpha j}$. QED

Finally, we are interested in seeing what happens when we multiply two series like in (2.6), obtained using one of the methods (theorems 2.4.4 and 2.4.5), starting from the same series y.

[1] e_i is the multi-index that equals 1 at its i-th component and 0 otherwise

Theorem 2.4.6.

Hp	f^1, f^2 *are series constructed with one of the two methods using the same* y
Ts	$f = f^1 f^2$ *is a power series as in* (2.6) *in a neighborhood of the origin*

Remark. The mixed case is also admissible, where f^1 is obtained with a composition and f^2 with a differentiation, and that is exactly what we will use.

Proof: By deriving the expression for the coefficients of f, we obtain that

$$f_{\gamma k} = \sum_{\substack{\omega+\theta=\gamma \\ l+h=k}} f^1_{\omega l}\, f^2_{\theta h}\,.$$

Consequently, $f_{\gamma k}$ will certainly be a polynomial with non-negative coefficients, since this property is preserved under sums and products. Moreover, noting that l, $h \leq k$, it can be shown that each individual polynomial $f_{\gamma k}$ inherits the property from $f^1_{\omega l}$ and $f^2_{\theta h}$, meaning that it depends only on the $y_{\alpha j}$ where $j \leq k$. QED

Chapter 3

The Cauchy-Kowalevski Theorem

Now that we have developed all the necessary tools, let us assume we have any Cauchy problem. As we showed in section 2.3, it can be rewritten in the form:

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

Consequently, we will only deal with this last case, where the conditions are assigned on $\Gamma_0 = \{t = 0\}$.
The fundamental assumption of this chapter is that the data (F and ϕ_j) are analytic in a neighborhood of the origin, a property that we will use to show the existence of a unique **analytic solution**, still in a neighborhood of the origin.
However, to guarantee existence, we are forced to make some assumptions about the structure of the equation. Considering what has been said in the previous chapter, especially regarding theorem 2.2.1, intuition suggests that it might be a good idea to consider the surface Γ_0 to be **non-characteristic**. This property allows us to rewrite the equation again in an even simpler form, namely:

$$\begin{cases} D_t^k u = G(x, t, D_x^\alpha D_t^j u) & |\alpha| + j \leq k,\ j < k \\ D_t^j u = \phi_j & \text{on } \Gamma_0,\ j < k \end{cases} \tag{3.1}$$

This idea will allow us to prove the CKT.

Cauchy-Kowalevski Theorem 3.0.1.

Hp	*Problem* (3.1) G, ϕ_j *analytic in a neighborhood of the origin*
Ts	$\exists!$ u *analytic solution in a neighborhood of the origin*

After keeping as general a view as possible, we want to understand how to prove the result we have in mind. The approach we will follow will be "in reverse," progressively generalizing the results. Indeed, we will start from the least general case until we reach that of an equation in normal form, effectively following the chronological order of the results.

3.1 ODEs

First, let us tackle a theorem very similar to the CKT, which deals with the case of a system of ODEs in normal form. We will start by stating the theorem.

Theorem 3.1.1.

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*

Remark. This does not exclude the possibility of finding other non-analytic solutions.

This result was the first application of the theory of holomorphic functions in combination with the method of majorants, which, as we already know, was proposed by Cauchy in the first half of the nineteenth century. We do not provide the full proof because it uses a different majorant from the one we introduced in section 2.4 (i.e., the one we will use to prove the CKT). In any case, it can be found in [Rou80], and the structure of the reasoning is the same as that of theorem 3.2.1.
Although we do not address the issue of existence in detail, it is worthwhile to discuss exhaustively the problem of **uniqueness** of analytic (or holomorphic) solutions. An analytic function is uniquely determined by all its derivatives at a point, which, in this case, are known due to the analyticity of the function f. We completely conclude the discussion by also addressing the situation of a PDE: here too, assuming the data are analytic, it is possible to know all the partial derivatives of the function, thanks to the fact that the surface on which the conditions are assigned is assumed to be non-characteristic. Since this result has been demonstrated by constructing a majorant for the solution y, it is possible to obtain an estimate of its radius of convergence by using theorem 2.4.1.

Theorem 3.1.2.

Hp	*Assumptions of theorem 3.1.1* $\exists\, \overline{B_a(x_0)} \subseteq A$, $\overline{B_b(y_0)} \subseteq B$ $M = \max_{B_a(x_0),\, B_b(y_0)} \lvert f \rvert$
Ts	*The solution has radius at least* $\widetilde{r} = a \left[1 - \exp\left(-\frac{b}{aM(n+1)} \right) \right]$

Remark. It is interesting to note what happens when $B = \mathbb{C}^n$.

3.2 Quasi-linear PDEs

Now it is time to address the cornerstone of the entire reasoning about PDEs, namely the theorem that shows the existence, and thus also the uniqueness, of an analytic solution to a quasi-linear system of PDEs in normal form.

Theorem 3.2.1.

Hp	A_i, B *analytic in a neighborhood of the origin* *Pb:* $\begin{cases} D_t\, y = \sum\limits_{i=1}^{n-1} A_i(x,y) D_{x_i} y + B(x,y) \\ y = 0 \quad \text{on } \Gamma_0 \end{cases}$
Ts	$\exists!\ y(x,t) : \mathbb{R}^n \to \mathbb{R}^m$ *analytic solution in a neighborhood of the origin*

Remark. This theorem can easily be modified by replacing analyticity with **holomorphy**, in order to obtain a statement similar to the case of ODEs, as the extension is immediate since no particular assumption distinguishes the real case from the complex one in the proof.

Proof: First of all, let us denote by a^i_{ml} the components of A_i and b_m those of B, while the coefficients of the series are respectively $(a^i_{ml})_\gamma$ and $(b_m)_\gamma$. Now let us proceed step by step.

1. Considering component by component, we assume $y_h = \sum c^h_{\alpha j} x^\alpha t^j$ with $h = 1, \ldots, m$.

2. The Cauchy condition tells us that $c^h_{\alpha 0} = 0$.

3. By inserting the series for y, A_i, B into the equation and using theorems 2.4.4, 2.4.5 and 2.4.6, we obtain for each row h the equation:

$$\sum_{\alpha,\, j} (j+1) c^h_{\alpha(j+1)}\, x^\alpha t^j = \sum_{\alpha,\, j} P_{\alpha j} \left((c^h_{\alpha l})_{l \leq j}, (a^i_{ml})_\gamma, (b_m)_\gamma \right)\, x^\alpha t^j$$

   where the polynomials $P_{\alpha j}$ are naturally non-negative coefficients.

4. Thanks to this operation, we derive a recursive formula for the coefficients:

$$c^h_{\alpha(j+1)} = (j+1)^{-1} P_{\alpha j} \left((c^h_{\alpha l})_{l \leq j}, (a^i_{ml})_\gamma, (b_m)_\gamma \right),$$

   which allows us to conclude that $c^h_{\alpha j} = Q_{\alpha j}\left((a^i_{ml})_\gamma, (b_m)_\gamma\right)$ where $Q_{\alpha j}$ are always polynomials with non-negative coefficients, the form of which does not depend on A_i and B. Thus, we can also say that it is always possible to construct a power series that satisfies the equation. We are left to understand whether it converges with a positive radius.

5. Now suppose we have another problem with the same structure defined by the functions $\widetilde{A}_i$ and $\widetilde{B}$ and we know a local analytic solution $\widetilde{y}$. We want to show that

$$\widetilde{A}_i \gg A_i,\ \widetilde{B} \gg B \implies \widetilde{y} \gg y.$$

Considering that for both problems the same considerations hold up to point 4 (the polynomials $Q_{\alpha j}$ are the same!), we can write the following chain of inequalities:

$$\begin{aligned}
\left|c^h_{\alpha j}\right| &= \left|Q_{\alpha j}\left((a^i_{ml})_\gamma, (b_m)_\gamma\right)\right| \\
&\leq Q_{\alpha j}\left(\left|(a^i_{ml})_\gamma\right|, \left|(b_m)_\gamma\right|\right) && \text{non-negative coeff.} \\
&\leq Q_{\alpha j}\left((\widetilde{a}^i_{ml})_\gamma, (\widetilde{b}_m)_\gamma\right) = \widetilde{c}^h_{\alpha j} && \widetilde{A}_i \gg A_i,\ \widetilde{B} \gg B
\end{aligned}$$

6. The last step consists in choosing $\widetilde{A}_i$, $\widetilde{B}$, in such a way as to explicitly calculate $\widetilde{y}$ and show that it is analytic. Given our knowledge about the power series that come from theorem 2.4.2, we know how to construct an upper bound for A_i and B. Thus, we select two constants C and r such that

$$\frac{Cr}{r - (x_1 + \ldots + x_{n-1}) - (y_1 + \ldots + y_m)} = \mathcal{M}_{Cr}(x, y) \gg A_i(x, y),\ B(x, y)$$

and that satisfy the inequalities (3.4) and (3.5). We then define the problem

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

where $h = 1, \ldots, m$. At this point it is possible to show, using the method of characteristics for quasi-linear first-order equations (paragraph 2.2), that it has the solution

$$\widetilde{y}_h(x, t) = u(x_1 + \cdots + x_{n-1},\ t) \quad \forall h \tag{3.2}$$

where

$$u(s, t) = \frac{r - s - \sqrt{(r - s)^2 - 2tCrmn}}{mn}, \tag{3.3}$$

which is clearly analytic in a neighborhood of the origin. See [Fol95, cap.1] for the complete unfolding of the calculation.

7. We conclude by observing an interesting fact: there is nothing specific that guarantees that the solution found continues to ensure the upper bound. In fact, it is important to verify that $\widetilde{y}$ satisfies the inequality $|(x, \widetilde{y}(x, t))| < r$ (so that $\mathcal{M}_{Cr}$ is indeed a power series). For more details, see proposition 3.2.3.

QED

As in the case of the system of ODEs, if we utilize theorem 2.4.1, we can estimate the radius of convergence by studying the radius of the dominating solution in (3.2).

Theorem 3.2.2. *The solution of Theorem 3.2.1 converges with a radius of at least*

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

Remark. $n \geq 2$ for the system to be truly in partial derivatives.

Remark. It is interesting to focus on the behavior with respect to r, knowing that:

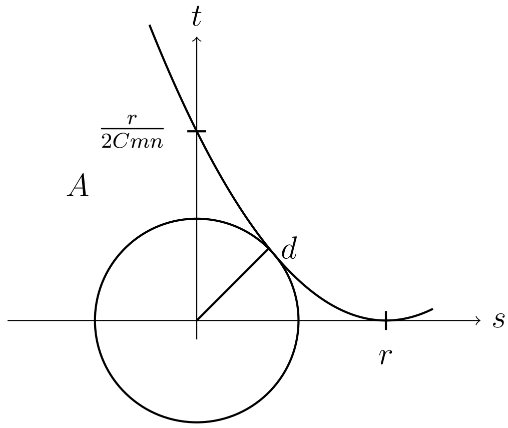
$$r < \min\{\textit{radii of convergence of the coefficients } a^i_{ml},\, b_m\} \tag{3.4}$$

$$C \geq \max \begin{Bmatrix} \max\limits_{i,m,l,\alpha} \left|(a^i_{ml})_\alpha\, r^{|\alpha|}\right| \\ \max\limits_{m,\alpha} \left|(b_m)_\alpha\, r^{|\alpha|}\right| \end{Bmatrix} \tag{3.5}$$

Proof: Let us fix r and C as above, and furthermore assume that $C \geq 1/2$ (we can do this without any problem). Initially, we focus on the function in (3.3), which is analytic in a neighborhood of the origin, particularly in the set

$$A = \left\{(s,t) \in \mathbb{R}^2 : t < \frac{(r-s)^2}{2Crmn}\right\}.$$

That is, it can be expanded in a power series in $B_l(0)$ with $0 < l < d = \operatorname{dist}(0, \partial A)$.



Choosing $l_1 = (n-1)\,\widetilde{r}$, it can be shown that, for $C > 1/8$, we indeed have that $l_1 < d$. The goal would therefore be to verify that

$$\sqrt{l_1^2 - s^2} < \frac{(r-s)^2}{2Crmn}$$

for every $|s| < l_1$, but this is implied by

$$l_1 < \frac{(r-l_1)^2}{2Crmn}\,, \tag{3.6}$$

an inequality that holds true if and only if $C > 1/(4mn) \leq 1/8$.

Now, we will generalize this for the function $\widetilde{y}_h$ in 3.2 with h fixed. It is analytic in the region

$$A = \left\{ (x,t) \in \mathbb{R}^n : t < \frac{(r - (x_1 + \ldots + x_{n-1}))^2}{2Crmn} \right\}.$$

The structure of the problem remains the same, so, naturally, the definition of d remains unchanged. The only aspect we need to take care of is that, in this situation, it will be necessary to choose $l_2 = \widetilde{r}$. Thus, we want to show that

$$\mathcal{L} = \sqrt{l_2^2 - (x_1^2 + \ldots + x_{n-1}^2)} < \frac{(r - (x_1 + \ldots + x_{n-1}))^2}{2Crmn} = \mathcal{R}$$

when $|x| = |(x_1, \ldots, x_{n-1})| < l_2$. But this is implied by the inequality 3.6, which we know to be true. We will prove it in two steps.

- It holds that $\mathcal{L}^2 < l_2^2 \le l_1^2$ for $|x| < l_2$.
- It holds that

  $$\left(\frac{(r - l_1)^2}{2Crmn} \right)^2 \le \min \left\{ \mathcal{R}^2 : |x| \le l_2 \right\} < \mathcal{R}^2 \text{ for } |x| < l_2.$$

  This is shown knowing that

  $$\max \{ x_1 + \ldots + x_{n-1} : |x| \le l_2 \} = (n-1) \frac{l_2}{\sqrt{n-1}} \le l_1$$

  and showing that $r - (x_1 + \ldots + x_{n-1}) > 0$ for $|x| \le l_2$ with the triangle inequality.

QED

A careful reader will surely wonder about the reason behind choosing a constant $C \ge 1/2$. Well, this issue arises from what was left open in the proof of Theorem 3.2.1, namely the fact that the solution $\widetilde{y}$ is indeed dominating only if $|(x, \widetilde{y}(x,t))| < r$. This property is precisely guaranteed in a ball of radius $\widetilde{r}$. Let us see this with a proposition that, in addition to clarifying, completes the logical framework of the proofs.

Proposition 3.2.3. *The domination of $\widetilde{y}$ holds in $B_{\widetilde{r}}(0)$, i.e.*

$$|(x,t)| < \widetilde{r} \implies |(x, \widetilde{y}(x,t))| = x_1^2 + \ldots + x_{n-1}^2 + m\, u^2(x_1 + \ldots + x_{n-1}, t) < r^2$$

Proof: For simplicity, we demonstrate that

$$|(s,t)| < l = (n-1)\, \widetilde{r} \implies s^2 + m\, u^2(s,t) < r^2.$$

The generalization is trivial if we take inspiration from the proof of Theorem 3.2.2.

Considering that $|t|, |s| < l < r$ and that $s^2 + t^2 < l^2 = r/(8Cmn)$, we write the following chain of inequalities.

$$
\begin{aligned}
& s^2 + m\left[\frac{r - s - \sqrt{(r-s)^2 - 2tCrmn}}{mn}\right]^2 \\
& \leq s^2 + \frac{1}{mn^2}\left[(r-s)^2 + |(r-s)^2 - 2tCrmn|\right] && \begin{cases} |s| < r \Rightarrow r - s > 0 \\ \sqrt{(r-s)^2 - 2tCrmn} > 0 \end{cases} \\
& \leq s^2 + \frac{2}{mn^2}(r-s)^2 + \frac{2|t|Crmn}{mn^2} \\
& \leq l^2 + \frac{2}{mn^2}\left(r^2 + l^2 + 2rl\right) + \frac{2lCrmn}{mn^2} && \begin{cases} |s| < l \Rightarrow (r-s)^2 < (r+l)^2 \\ |t| < l \end{cases} \\
& = \left(\frac{r}{8Cmn}\right)^2 + \frac{2r^2}{mn^2}\left[1 + \frac{1}{(8Cmn)^2} + \frac{1}{4Cmn} + \frac{1}{8}\right] \\
& < r^2 \underbrace{\frac{2}{mn^2}\left[\frac{r}{8} + \frac{1}{(8C)^2mn^2} + \frac{r}{(8Cmn)} + \frac{r}{4Cmn} + \frac{1}{8}\right]}_{<1} < r^2
\end{aligned}
$$

In particular, the last statement holds because

$$
\begin{aligned}
n \geq 2 \Rightarrow \frac{2}{mn^2}(\ldots) &\leq \frac{1}{2}\left(\frac{9}{8} + \frac{2}{(8C)^2} + \frac{1}{8C}\right) \\
&\leq \frac{3}{16}\left(3 + \frac{1}{C}\right) < 1 && \Leftarrow C \geq \frac{1}{2}
\end{aligned}
$$

QED

3.3 EDP in Normal Form

Now we will utilize the results from the previous section to generalize that result to the case of an equation in normal form. To do this, it is sufficient to state and prove the following theorem.

Theorem 3.3.1. *The following two problems are equivalent*

$$\textit{non-linear:} \quad \begin{cases} D_t^k u = G(x,t,D_x^\alpha D_t^j u) & |\alpha|+j \le k,\ j<k \\ D_t^j u = \phi_j & \textit{on } \Gamma_0,\ j<k \end{cases}$$

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

Proof: The reasoning is divided into three steps:

1. We construct the system such that $y_{\alpha j} = D_x^\alpha D_t^j u$.
   Then, the matrices A_i and B can be obtained from the expressions

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

   where $l(\alpha) = \min\{l : \alpha_l \neq 0\}$, and the Cauchy data will be

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

2. We remove the conditions ϕ, redefining $y(x,t) \leftarrow y(x,t) - \phi(x)$.

3. We eliminate the dependence on t, adding the variable $y^0 = t$, together with the equation $D_t y^0 = 1$ and the data $y^0(x,0) = 0$.

We conclude by stating that, obviously, if u is a solution of the problem in normal form, the $y_{\alpha j}$ will be solutions of the newly constructed problem. However, to demonstrate that $y_{(0,\dots,0)}$ (solution of the latter) is also a solution of the problem in normal form, various calculations are necessary, which can be found in detail in [Fol95, cap.1]. QED

Remark. There are three aspects, which also emerge from the proof, that are worth briefly reflecting upon:

- Bringing together the considerations made at the beginning of the chapter and the theorems 3.2.1 and 3.3.1, the CKT follows immediately;
- The estimate of the radius of convergence continues to hold;
- This equivalence theorem can be readily generalized to the case of a system in normal form.

Chapter 4
Esempi

Dopo aver visto il CKT nella sua forma più nota, concentriamo ora lo sguardo su tre esempi importanti che aiutano a inquadrare meglio i limiti di questo teorema e il ruolo che giocano le ipotesi.
Tale discussione risulta particolarmente di rilievo, poiché per molto tempo si ritenne ragionevole pensare che un'equazione differenziale con coefficienti piuttosto regolari, come ad esempio C^∞, dovesse avere almeno una soluzione. Questo, però, oltre al caso di analiticità trattato dal CKT, in generale non accade.

4.1 Esempio di Lewy

Questo primo esempio è decisamente il più importante ed interessante tra quelli qui trattati, proprio perché permette di introdurre, in modo più rigoroso, il problema appena citato.
Nel 1957 Hans Lewy propose un semplice controesempio, volto a mostrare come l'ipotesi di **analiticità** nel teorema di Cauchy-Kowalevski fosse cruciale, portando un caso di un operatore differenziale lineare con coefficienti analitici che necessita della presenza di una forzante anch'essa analitica per possedere delle soluzioni almeno C^1.
Ciò mostra come sia cruciale, non solo una discussione sulle condizioni sufficienti per l'esistenza di soluzioni locali, ma anche una sulle condizioni necessarie. Infatti, Hörmander, matematico che contribuì ampiamente alla teoria delle equazioni lineari, rispose all'emersione di questo problema proprio con delle condizioni necessarie per l'esistenza di soluzioni locali (e quindi anche globali!) per equazioni lineari, le quali ispirarono poi, a loro volta, il lavoro di Treves e Nirenberg volto alla ricerca di condizioni necessarie e sufficienti.

Preliminarmente si riportano qui sotto gli enunciati di due teoremi che torneranno utili nella discussione:

Formula di Green in $\mathbb{C}$ 4.1.1.

$$\begin{array}{c|l} Hp & \begin{array}{l} D \subseteq \mathbb{C} \textit{ dominio regolare} \\ f : D \to \mathbb{C} \\ f \in H(\mathring{D}) \end{array} \\ & \\ Ts & \displaystyle\oint\limits_{\partial^+ D} f(z)\, dz = 2i \iint\limits_{D} \frac{\partial f}{\partial \overline{z}}(x+iy)\, dxdy \end{array}$$

Remark. La definizione di dominio regolare non ci tornerà particolarmente utile, infatti ai nostri scopi è sufficiente sapere che una qualsiasi palla chiusa è regolare (questo verrà utilizzato nella dimostrazione del teorema 4.1.3). Per una formalizzazione di questo concetto si veda [FMS20, cap.8], dove è presente una trattazione dell'analogo teorema in $\mathbb{R}^2$ che va sotto il nome di "Formule di Gauss-Green" e "Formula di Stokes", di quale la generalizzazione in $\mathbb{C}$ è immediata.

Principio di riflessione di Schwarz 4.1.2.

$$\begin{array}{c|l} Hp & \begin{array}{l} D \subseteq \mathbb{C} \textit{ dominio regolare e simmetrico rispetto a } \mathbb{R} \\ D \cap \mathbb{R} \textit{ è un intervallo} \\ f : D \to \mathbb{C} \\ f(\mathbb{R} \cap D) \subseteq \mathbb{R} \\ f \in H(\mathring{D}) \end{array} \\ Ts & f(\overline{z}) = \overline{f(z)} \quad \forall z \in \mathring{D} \end{array}$$

Remark. La definizione di insieme simmetrico rispetto a $\mathbb{R}$ è data in modo naturale: esso deve soddisfare la condizione $z \in D \implies \overline{z} \in D$.

Per entrare nel vivo dell'esempio, definiamo il seguente operatore:

$$L = D_x + iD_y - 2i(x+iy)D_t$$

che ha dei coefficienti C^∞ e il cui comportamento peculiare emerge dal teorema che enunciamo di seguito.

Theorem 4.1.3.

$$\begin{array}{c|l} Hp & \begin{array}{l} f \textit{ funzione continua a valori reali che dipende solo da } t \\ u \in C^1 \ : \ Lu = f \textit{ in un intorno dell'origine} \end{array} \\ Ts & f \textit{ analitica in un intorno di } t = 0 \end{array}$$

Proof: Innanzitutto fissiamo un $R > 0$ tale che $\{(x,y,t) : x^2 + y^2 < R^2, |t| < R\}$ sia contenuto nell'intorno dell'origine delle ipotesi (ovviamente questo R esiste sempre) e procediamo seguendo cinque passi.

1. Definiamo la funzione:
$$V(t,s) = \int_{\gamma_r} u(x,y,t)\,dz \quad \text{con} \quad \begin{cases} t \in (-R,R) \\ r^2 = s \in [0,R^2) \\ \gamma_r = \partial^+ B_r(0,0) \\ z = x+iy \end{cases}$$

2. Troviamo una relazione tra V_s e V_t:
$$\begin{aligned} V &= i \iint_{B_r(0,0)} (u_x + iu_y)(x,y,t)\,dx\,dy && \text{per formula di Green} \\ &= i \int_0^r \int_0^{2\pi} (u_x + iu_y)(\rho\cos\theta,\ \rho\sin\theta,\ t)\,\rho\,d\rho\,d\theta && \text{in coordinate polari} \\ V_r &= i \int_0^{2\pi} (u_x + iu_y)(\rho\cos\theta,\ \rho\sin\theta,\ t)\,r\,d\theta && \text{derivando} \\ &= \int_{\gamma_r} (u_x + iu_y)(x,y,t)\,r\,\frac{dz}{z} \\ V_s &= \frac{1}{2r} V_r = \int_{\gamma_r} (u_x + iu_y)(x,y,t)\,\frac{dz}{2z} \\ &= \int_{\gamma_r} u_t(x,y,t)\,dz + \int_{\gamma_r} f(t)\,\frac{dz}{2z} && \text{usando } Lu = f \\ &= iV_t + \pi i f(t) \end{aligned} \tag{4.1}$$

3. Definiamo le funzioni:
$$F(t) = \int_0^t f(\tau)\,d\tau$$
$$U(t,s) = V(t,s) + \pi F(t)\ .$$
e osserviamo le seguenti proprietà di U, vista come funzione di $w = t + is$:
   - si può verificare che soddisfa l'equazione di Cauchy-Riemann $U_t + iU_s = 2U_{\overline{z}} = 0$ utilizzando la relazione (4.1),
   - è olomorfa per $(s,t) \in (0,R^2) \times (-R,R)$ per la proprietà precedente,
   - è continua per $(s,t) \in [0,R^2) \times (-R,R)$ perché lo è V,
   - $U(0,t) = \pi F(t)$ per $t \in (-R,R)$, ovvero assume valori reali sull'asse reale.

4. Possiamo ora prolungare analiticamente U in un intorno dell'origine, infatti, date le proprietà appena osservate, valgono le ipotesi del principio di riflessione di Schwarz, che ci permette di definire U per $s \in (-R^2, 0)$ con la seguente formula:
$$U(t,s) = \overline{U(t,-s)}.$$

5. Concludiamo il ragionamento notando che, se il prolungamento di U è analitico in un intorno dell'origine, lo deve essere anche $U(t,0) = \pi F(t)$ e anche $f = F'$. QED

Generalization. The theorem we just discussed can actually be extended to an interesting generalization, and the idea is as follows: we aim to show that, despite the characteristic form of L having no singular points, it is possible to choose a forcing term $F \in C^\infty(\mathbb{R}^3, \mathbb{R})$ such that, **everywhere**, the differential equation $Lu = F$ admits no solutions.

Remark. Given two matrix spaces (X, d_X) and (Y, d_Y), the notation $C(X,Y)$ with $k \in \mathbb{N} \cup \{\infty\}$ denotes the set of continuous functions of the type $h : X \to Y$. In the case where $X = \mathbb{R}^n$ and $Y = \mathbb{R}^m$, we will naturally use the notation $C^k(\mathbb{R}^n, \mathbb{R}^m)$ for C^k functions.

Before delving into the specifics of this second part of the discussion on Lewy's example, it is useful to recall three definitions:

Definition 4.1.1. A subset D of a topological space X is dense if for every open set $A \in X$, $D \cap A \neq \emptyset$.

Definition 4.1.2. A subset E of a metric space has no interior if $\mathring{E} = \emptyset$.

Definition 4.1.3. A topological space is called a "Baire space" if the countable union of any family of closed sets with empty interior has empty interior.

The reason we have mentioned these concepts is that we are interested in a theorem, or rather a corollary, that allows us to develop an argument by contradiction when dealing with complete metric spaces. The following statements are provided.

Baire Category Theorem 4.1.4.

Hp
- (X, d) *complete metric space*
- $\{A_n\}_{n\in\mathbb{N}} \subseteq 2^X$ *family of dense open sets in* X
- $\{E_n\}_{n\in\mathbb{N}} \subseteq 2^X$ *family of closed sets with no interior*

Th 1
- $\displaystyle\bigcap_{n\in\mathbb{N}} A_n$ *is dense in* X

Th 2
- $\displaystyle\bigcup_{n\in\mathbb{N}} E_n$ *has no interior*

Remark. This theorem shows that complete metric spaces are indeed Baire spaces under the topology induced by the metric. See [RF10, Ch.10] for the proof and more details.

Corollary (Baire's argument by contradiction) 4.1.5.

Hp
- (X, d) *complete metric space*
- $\{E_n\}_{n\in\mathbb{N}} \subseteq 2^X$ *family of closed sets*
- $X = \displaystyle\bigcup_{n\in\mathbb{N}} E_n$

Ts
- $\exists\, n \in N$ *such that* $\mathring{E}_n \neq \emptyset$

Remark. This statement is the contrapositive of the second claim of theorem 4.1.4, and, as we mentioned earlier, it can be used to derive a contradiction by exhibiting a complete metric space equal to the union of a family of closed sets with no interior.

The second important result from functional analysis, which will play a crucial role in achieving the stated goal, is the Ascoli-Arzelà theorem: a "compactness" theorem, which replaces the Heine-Borel theorem in the search for a convergent subsequence in cases where the compactness property of the metric spaces is not known. In particular, we will use it to show that a certain set (whose structure will be understood later) is closed, by exploiting the uniform convergence property guaranteed by the theorem.
To fully understand the statement of this theorem, we recall two definitions along with it.

Definition 4.1.4. A sequence of functions $\{f_n : X \to \mathbb{R}\}_{n\in\mathbb{N}_0}$ is said to be uniformly bounded in X if $\exists\, M \geq 0$ such that $|f_n| \leq M$ in X.

Definition 4.1.5. A sequence of functions $\{f_n : X \to \mathbb{R}\}_{n\in\mathbb{N}_0}$ is said to be equicontinuous in X if $\forall \varepsilon > 0\ \exists\, \delta > 0$ such that $d(x, y) < \delta \implies |f_n(x) - f_n(y)| < \varepsilon \quad \forall x, y \in X, \forall n \in \mathbb{N}_0$.

Ascoli-Arzelà Theorem 4.1.6.

$$
\begin{array}{c|l}
 & (X, d) \textit{ complete metric space} \\
\textit{Hp} & \{f_n : X \to \mathbb{R}\}_{n\in\mathbb{N}_0} \textit{ sequence of functions} \\
 & \quad - \quad \textit{uniformly continuous} \\
 & \quad - \quad \textit{uniformly bounded} \\
\textit{Ts} & \exists\, f \in C(X, \mathbb{R}), n_k \textit{ such that } f_{n_k} \to f \textit{ uniformly}
\end{array}
$$

After reviewing these tools, it is time to delve into the discussion, and we do so by outlining the reasoning to be followed step by step:

1. We will shift the problem of theorem 4.1.3 to refer to a generic point (x_0, y_0, t_0), using the function $g(x, y, t) = f(t - 2xy_0 + 2x_0y)$ as a forcing term (lemma 4.1.7);

2. We will construct a function $S_a \in C^\infty$ for each $a \in l^\infty$ (lemma 4.1.8);

3. We will build sets $E_{j,n} \subseteq l^\infty$ that are closed and have no interior using S_a and the Ascoli-Arzelà theorem (lemma 4.1.9);

4. We will conclude the proof of theorem 4.1.10 by using the aforementioned lemmas to derive, through a contradiction argument, the equality $l^\infty = \bigcup E_{j,n}$, which allows us to apply Baire's argument.

Now we will detail the steps just outlined with statements and proofs.

Lemma 4.1.7.

Hp	$F \in C^\infty(\mathbb{R}, \mathbb{R})$ $(x_0, y_0, t_0) \in \mathbb{R}^3$ $u \in C^1 \ : \ Lu(x, y, t) = F'(t - 2xy_0 + 2x_0y)$ *in a neighborhood of* (x_0, y_0, t_0)
Ts	F *and* F' *are analytic in a neighborhood of* $t = t_0$

Proof: By exploiting the invariance of the operator L with respect to

$$T(x, y, t) = (x + x_0, y + y_0, t + t_0 + 2xy_0 - 2x_0y),$$

i.e., the validity of the identity (easy to verify) $L(u \circ T) = (Lu) \circ T$, we deduce that, if u is a solution to the equation in the hypothesis, it also holds in a neighborhood of the origin:

$$L(u \circ T)(x, y, t) = f(t + t_0) \text{ with } f = F'. \tag{4.2}$$

Clearly, $u \circ T \in C^1$, and $g(t) = f(t + t_0)$ satisfies the conditions of theorem 4.1.3, so by applying it to the second equation, the thesis is proved. QED

Remark. The analyticity of F follows from the last step in the proof of theorem 4.1.3, considering that it takes the form $F(t) = \int_0^t f(\tau) + c$ with $c \in \mathbb{R}$.

Remark. Equation (4.2) holds in a neighborhood of the origin because the operator T makes $\mathbb{R}^3$ a group, generally known as the Heisenberg group, and in this context, it acts like a translation.

Lemma 4.1.8.

Hp	$\{(x_j, y_j, t_j)\}_{j=1}^\infty$ *dense in* $\mathbb{R}^3$ $c_j = 2^{-j} e^{-\rho_j}$ *with* $\rho_j = \lvert x_j\rvert + \lvert y_j\rvert \quad \forall j \in \mathbb{N}_0$ $a = \{a_n\}_{n=1}^\infty \in l^\infty$ $F \in C^\infty(\mathbb{R}, \mathbb{R})$ *periodic and non-analytic* $f_j(x, y, t) = F'(t + 2xy_j - 2x_jy)$ (4.3)
Th 1	$S_a = \sum_{j=1}^\infty a_j c_j f_j$ *converges uniformly in* $\mathbb{R}^3$
Th 2	*the same holds for the formal derivatives* $D^\alpha S_a = \sum_{i=1}^\infty a_j c_j D^\alpha f_j$

Remark. Naturally, S_a is a C^∞ function.

Proof: Since F is C^∞ and periodic, we define $M_k = \sup_t \left|F^{(k)}(t)\right| \in \mathbb{R}$ for every $k \in \mathbb{N}$. This allows us to write, for each multi-index α and $j \in \mathbb{N}_0$, the following inequalities:

$$\begin{aligned} |a_j c_j D^\alpha f_j| &\leq \|a\|_\infty \, 2^{-j} \, e^{-\rho_j} \, M_{|\alpha|+1} \, \rho_j^{|\alpha|} \\ &\leq \|a\|_\infty \, 2^{-j} \, M_{|\alpha|+1} \left(\frac{|\alpha|}{e}\right)^{|\alpha|} \qquad \text{because } \max_{x\geq 0} \frac{x^{|\alpha|}}{e^x} = \left(\frac{|\alpha|}{e}\right)^{|\alpha|} \end{aligned} \tag{4.4}$$

$D^\alpha S_a$ converges absolutely, and therefore also uniformly, as the series

$$\sum_{j=1}^{\infty} \sup_{\mathbb{R}^3} |a_j c_j D^\alpha f_j|$$

has a general term that is less than or equal to the right-hand side of inequality (4.4), whose corresponding numerical series is obviously convergent. QED

Remark. Before continuing, let's briefly pause on two noteworthy points:

- l^∞ is a Banach space when equipped with the norm: $\|b\|_\infty = \sup_n |b_n|$ for every $b \in l^\infty$;
- there exists a function f with the properties mentioned in the hypotheses: for instance, the function
$$F(x) = \sum_{n=1}^{\infty} \frac{\cos(n!\,x)}{(n!)^n}$$
is defined by a pointwise convergent series and is $C^\infty(\mathbb{R}, \mathbb{R})$. Additionally, it is periodic with period 2π and can be shown not to be analytic at any point $x \in \mathbb{R}$. For more on this, see problem 4 in [Joh82, cap.3].

Notation. $A_{j,n} = B_{n^{-1/2}}(x_i, y_i, t_i)$ where (x_i, y_i, t_i) are the points in the hypotheses of lemma 4.1.8.

Lemma 4.1.9.

$$Hp \left| \begin{array}{l} \textit{Same hypotheses as lemma 4.1.8} \\ \{E_{j,n}\}_{j,n\in\mathbb{N}_0} \subseteq l^\infty \textit{ such that} \\ a \in E_{j,n} \textit{ if and only if } \exists\, u \in C^1(A_{j,n}) \textit{ such that} \\ \quad - \quad Lu = S_a \textit{ in } A_{j,n} \\ \quad - \quad u(x_j, y_j, t_j) = 0 \\ \quad - \quad |D^\alpha u| \leq n \textit{ for } |\alpha| \leq 1 \textit{ in } A_{j,n} \\ \quad - \quad |D^\alpha u(v) - D^\alpha u(w)| \leq n|v-w|^{1/n} \textit{ for } \begin{cases} |\alpha| = 1 \\ v, w \in A_{j,n} \end{cases} \end{array} \right. \tag{4.5}$$

$$Ts \;\; \Big| \;\; \{E_{j,n}\} \textit{ are closed and nowhere dense sets}$$

Proof: We will prove the two properties separately.

1. Regarding the property of closure, we want to show that if $\{a^k\} \subseteq E_{j,n}$ is such that $a^k \xrightarrow{l^\infty} a$, then $a \in E_{j,n}$. This, in turn, reduces to showing the existence of a function u with the properties in (4.5).

   We immediately deduce that $S_{a^k} \to S_a$ uniformly, since $|S_a - S_{a^k}| \leq M_1 \left\| a - a^k \right\|$ from (4.4) with $\alpha = 0$. Additionally, due to the hypotheses on a^k, there exists a function u_k that solves the equation $Lu_k = S_{a^k}$ and satisfies the other properties in (4.5). Due to these latter properties, u_k satisfies the hypotheses of the Ascoli-Arzelà theorem with $X = A_{j,n}$, so for some u it holds that $u_{k_h} \to u$ uniformly .

   In particular, exploiting the fact that L is a first-order operator, it is easy to deduce that $Lu = S_a$ in $A_{j,n}$ since

   $$\begin{array}{lr} Lu_{k_h} \to Lu & \text{uniformly for the properties of } u_k \\ \| & \\ S_{a^{k_h}} \to S_a & \text{uniformly} \end{array}$$

   and that u inherits all other properties in (4.5) from u_k due to uniform convergence.

2. Finally, we will show that $\mathring{E}_{j,n} = \emptyset$ by reasoning by contradiction. Thus, suppose there exists a sequence a inside $\mathring{E}_{j,n}$. Then, by defining

   $$\delta_j = \frac{1}{c_j} \mathbb{1}_{\{j\}} \in l^\infty,$$

   we observe that there exists a $\theta \in \mathbb{R}$ small enough such that $a' = a + \theta\delta_j \in E_{j,n}$. Now, let u and u' be the solutions to $Lu = S_a$ and $Lu = S_{a'}$ respectively, satisfying the properties in (4.5), and let

   $$u'' = \frac{u' - u}{\theta}.$$

   Clearly, $u'' \in C^1$; moreover, using the linearity of L and the definition of the series S, it is immediate to see that the relation

   $$Lu'' = S_{\delta_j} = f_j$$

   holds. However, this contradicts lemma 4.1.7 (whose hypotheses are all satisfied), since F is not analytic.

QED

Theorem 4.1.10.

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* $\begin{cases} Lu = F \text{ in } A \\ u_x,\, u_y,\, u_t \text{ satisfy} \\ \text{the Hölder condition} \end{cases}$

Remark. The conclusion naturally implies that there are no C^k solutions either, for any $k \geq 1$, since $C^k \subseteq C^1$.

Proof: By reasoning by contradiction, we conclude with the following three steps (of which the second deserves the most attention).

1. $E_{j,n} \subseteq l^\infty$ for every $j, n \in \mathbb{N}_0$, of course.

2. $a \in l^\infty \implies a \in E_{j,n}$ for some $j, n \in \mathbb{N}_0$ (which depend on a).

   Assuming the thesis is false, we can assert that $\forall a \in l^\infty\ \exists\, A \in \mathbb{R}^3$, $u^* \in C^1(A, \mathbb{R})$ such that $Lu^* = S_a$ and that u^* has first derivatives continuous according to Hölder in A.

   Moreover, due to the density of the set of points in (4.3), there exists a $(x_j, y_j, t_j) \in A$, and since A is open, there exists a k (chosen large enough) such that $A_{j,k} \subseteq A$.

   Now consider the function $u = u^* - u^*(x_j, y_j, t_j)$, so that u retains the properties of u^*, but at the same time satisfies the condition $u(x_j, y_j, t_j)$ as required in one of the properties in (4.5).

   Finally, it is clear that, since the first derivatives of u are continuous according to Hölder, there exists an m large enough such that the remaining conditions in (4.5) hold with m instead of the subscript n, and then taking $n = \max\{k, m\}$, the implication is proven.

3. From the first two steps, we conclude that

   $$l^\infty = \bigcup_{j,n \in \mathbb{N}_0} E_{j,n},$$

   but, therefore, since l^∞ is a Banach space and due to the properties of the sets $E_{j,n}$, both the hypotheses of Corollary 4.1.5 and the negation of the thesis hold. This is absurd.

QED

4.2 Kowalevski Example

The example we now focus on is due to Kowalevski herself, and it was useful at the time to gain a deeper, more essential understanding of the importance, or rather the necessity, of assuming that the surface chosen to assign the Cauchy data is **non-characteristic** for the differential equation under observation. Moreover, it constitutes a counterexample to the conjecture proposed by Weierstrass, which suggested the possibility of defining analytic functions through differential equations.
All of this is mentioned in a letter addressed to Fuchs (a German mathematician from the University of Berlin) written by Weierstrass (who supervised Kowalevski's research), in which he requested the acceptance of Kowalevski's doctoral thesis. The letter is fully reported in [Ken83, app.C].
Following in Kowalevski's footsteps, we consider the following Cauchy problem for the heat equation in one dimension:

$$u_t - u_{xx} = 0 \tag{4.6}$$

$$u(x,0) = \frac{1}{1+x^2} \quad \forall\, x \in \mathbb{R} \tag{4.7}$$

Remark. In fact, the initial data actually chosen by Kowalevski during her research is $\frac{1}{1-x}$, but we have decided not to use it here for simplicity, avoiding some issues related to the singularity of the function while keeping the reasoning unchanged.

Our goal is to prove that it admits no analytic solutions in a neighborhood of the origin.

1. To begin, note that in this case, the surface on which the Cauchy data (1.2) is assigned is $\Gamma = \{(x,t) \in \mathbb{R}^2 : t = 0\}$. At every point, its normal vector is $(0,1)$ and, therefore, it is characteristic for the equation (4.6), since[1]

$$\sum_{|\alpha|=2} a_\alpha \boldsymbol{\nu}^\alpha = a_{(2,0)} \boldsymbol{\nu}^{(2,0)} = 0.$$

2. By contradiction, suppose we have a solution of the problem u analytic in a neighborhood of the origin, that is:

$$u(x,t) = \sum_{\alpha=(\alpha_1,\alpha_2)} c(\alpha)\, x^{\alpha_1}\, t^{\alpha_2}, \quad c(\alpha) = \frac{D^\alpha u(0,0)}{\alpha!}$$

   where $|(x,t)| < r$ for some $r > 0$.

3. We calculate the values of the coefficients $c(2n,0)$ $\forall n \in \mathbb{N}$.
   To do this, we need to expand in power series, centered at the origin, the function of the Cauchy problem:

$$\frac{1}{1+x^2} = \frac{d}{dx}\arctan(x) = \frac{d}{dx}\sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1}\, x^{2n+1} = \sum_{n=0}^{\infty} (-1)^n\, x^{2n} \quad \forall x \in \mathbb{R}.$$

[1] see definition 2.1.3

From this series, we obtain the relations:

$$D_x^{2n}u(0,0) = \frac{d^{2n}}{dx^{2n}}\left.\frac{1}{1+x^2}\right|_{x=0} = (-1)^n(2n)!$$
$$D_x^{2n+1}u(0,0) = \frac{d^{2n+1}}{dx^{2n+1}}\left.\frac{1}{1+x^2}\right|_{x=0} = 0$$

from which we derive: $c(2n,0) = (-1)^n$ and $c(2n+1,0) = 0$.

4. We calculate the values of the coefficients $c(2n,n)$ and show that $c(2n,n) \xrightarrow{n\to\infty} +\infty$. For this purpose, we use the equation (4.6) to obtain the following relation between the coefficients:

$$c(\alpha_1, \alpha_2+1) = \frac{(\alpha_1+2)(\alpha_1+1)}{(\alpha_2+1)}\, c(\alpha_1+2, \alpha_2). \tag{4.8}$$

And we use this as follows:

$$\begin{aligned}
c(2n,n) &= \frac{(2n+2)(2n+1)}{n}\, c(2n+2, n-1) && (4.8) \quad \text{with} \quad \begin{cases}\alpha_1 = 2n\\ \alpha_2+1 = n\end{cases}\\
&= \ldots = \frac{(2n+2n)\cdots(2n+1)}{n!}\, c(2n+2n, 0) && \text{iterating over } n\\
&= \frac{(4n)!}{(2n)!\,n!}(-1)^{2n}\\
&\sim \frac{1}{\sqrt{\pi n}}\left(\frac{64n}{e}\right)^n \xrightarrow{n\to\infty} +\infty && \text{using Stirling's formula}
\end{aligned}$$

5. We complete the reasoning by immediately observing that

$$c(2n,n)\,x^{2n}\,t^n \xrightarrow{n\to\infty} +\infty \quad \forall (x,t) \neq (0,0),$$

since this directly implies that the power series does not converge at any point other than the origin.

4.3 Hadamard's Example

The final example we address, due to Hadamard (1932), helps to understand an important limitation of the Cauchy-Kowalevski Theorem (CKT), namely the fact that it provides no control over the **relationship** between the Cauchy data and the form of the analytic solution: the problem may become unstable, meaning that small variations in the data may not correspond to small variations in the solution.
To observe this behavior, let us consider the following Cauchy problem for the two-dimensional Laplace equation as n varies:

$$\begin{aligned} u_{xx} + u_{yy} &= 0 \\ u(x,0) &= 0 \\ u_y(x,0) &= n \sin(nx) e^{-\sqrt{n}} \quad \text{with} \quad n \in \mathbb{N} \end{aligned} \tag{4.9}$$

What we want to show is how, as n increases, a blow-up of the solution u_n of the problem (4.9) occurs.

1. The problem, as in the previous example, is assigned on $\Gamma = \{(x,y) \in \mathbb{R}^2 : y = 0\}$, which is naturally a non-characteristic surface for the Laplace equation (in fact, note that it is elliptic and thus possesses no characteristic surfaces).

2. It is easy to verify that the function $u_n(x,y) = \sin(nx)\sinh(ny)e^{-\sqrt{n}}$ satisfies (4.9) and that it is analytic, hence it is also the only possible solution with this property.

3. Finally, we observe that $\sinh(ny)e^{-\sqrt{n}} \xrightarrow{n\to\infty} \infty$.

As a conclusion to this discussion, let us consider the problem for $n = \infty$, that is, with data $u(x,0) = u_y(x,0) = 0$, and we immediately notice that the only analytic solution is $u \equiv 0$, which is naturally different from the asymptotic behavior of u_n. Therefore, we have just discovered that the solution does not continuously depend on the data.
Building on these considerations, Hadamard continued his studies, first defining the concept of well-posedness of a Cauchy problem[2], and then discovering that problems constructed with hyperbolic equations with constant coefficients always satisfy this condition in the class of C^∞ functions.

[2] a Cauchy problem is well-posed if there exists a unique solution and it depends continuously on the initial data

Chapter 5
Subsequent Developments

This result has led to several subsequent developments in various fields throughout the 20th century. First of all, it is worth noting that, despite Cauchy's expectations, and later those of Weierstrass, Kowalevski's result revealed the inherently more complicated nature of PDEs. In particular, it contributed to definitively disproving the conjecture formulated by Weierstrass, which we have already mentioned in paragraph 1.2.
The most immediate consequences, on which we will focus in the next two paragraphs, correspond to two aspects related to the theory of differential equations:

- the alternative and more abstract versions of the theorem;
- Holmgren's theorem, which is a result of existence and uniqueness (of a solution) for a system of linear PDEs in the class of C^1 functions.

Furthermore, this knowledge stimulated and inspired further research. In particular, the works worth mentioning are those of

- François Trèves and Louis Nirenberg, on the search for necessary and/or sufficient conditions for the existence of local solutions in broader classes of functions;
- Lars Hörmander, on a specific theory for linear differential operators, thanks to which necessary conditions for the existence and uniqueness of solutions were found (for further details see [Hö63]).

5.1 Alternative Versions

Although the underlying meaning remains virtually unchanged, if we restrict ourselves to the case of a linear system of PDEs, there are three main possible statements for the CKT, and to distinguish which version we are referring to, we use three different adjectives: abstract, classical, and invariant. In particular, with the last term, we refer precisely to the version we discussed in chapter 3.
The order in which the three names have been listed is not accidental; there is indeed a logical dependence among these statements, which can be represented with the following scheme.

$$\text{Abstract version} \implies \text{Classical version} \implies \text{Invariant version}$$

Thus, assuming these relationships to be true, we can say that there exists a different way to prove the theorem we have already extensively discussed. As the names suggest, the fundamental idea of the approach is to analyze the problem by placing it within a more general and abstract theoretical framework, so that we can deduce the theorem in its most common (invariant) version as a corollary. Following this path entails, first of all, a substantial increase in difficulty, and then also a loss of the direct and immediate link with the idea of characteristic surface.
The fundamental notion from which this alternative path originates is that of the Ovsyannikov classes (that is, sets of Banach spaces composed of holomorphic functions), which were introduced for the first time by the Russian mathematician L. V. Ovsyannikov between 1960 and 1970 (see [Ovs65]).
In this thesis, we refer to the treatment present in [Tre75, cap.17-19], reporting only the salient points. Therefore, we do not focus on the construction of the classes just mentioned and do not provide the statement of the theorem in its most abstract version, but we dwell only on the statement of the classical version, which already allows us to grasp and appreciate all the observations made so far.

Theorem 5.1.1.

$$\mathit{Hp}\ \left|\ \begin{array}{l} \overline{\mathcal{O}}_0 \subseteq \mathcal{O}_1 \subseteq \mathbb{C}^n \ \textit{open connected bounded} \\ A_j, f, y_0 \ \textit{holomorphic in } z \\ A_j, f \ \textit{continuous in } t \\ \textit{Pb:} \begin{cases} D_t y = \sum A_j(z,t) D_{z_j} y + A_0(z,t) y + f(z,t) \\ y(z,0) = y_0(z) \end{cases} \end{array}\right.$$

$$\mathit{Ts}\ \left|\ \begin{array}{l} \exists\, \delta \in (0,T) : \exists!\, y \ \textit{solution for } |t| < \delta \\ - \textit{holomorphic in } z \\ - C^1 \ \textit{in } t \end{array}\right.$$

Remark. Any equation or linear system can be reduced to a first-order system; for this reason, we focus only on the latter case.

We do not provide the proof, as it is a simple application of the more abstract version, but we want to understand how such abstraction proves to be a useful tool for proving Holmgren's theorem.

5.2 Holmgren's Theorem

Let us begin by recalling that the result obtained by Kowalevski does not provide any information about the existence of non-analytic solutions, which may therefore either exist or not. Consequently, in this paragraph, we want to investigate what happens, under the hypotheses of the CKT, when we expand the class of functions in which we seek the solutions to the problem. In particular, we ask: for which equations and under what additional conditions is the **uniqueness** of the solution guaranteed in a class of functions larger than the analytic ones?

An interesting and general answer to this question is provided by Holmgren's theorem, which, as mentioned in the previous paragraph, can be seen as a consequence of the CKT. However, the logical relationship between this theorem and the invariant version of the CKT is not direct; indeed, for the proof, it is necessary to rely on the more abstract framework that we introduced earlier.

To be more precise, as with the CKT, there are three versions of Holmgren's theorem, which we will similarly call: abstract, classical, and invariant. Thus, we will integrate the previously proposed scheme, adding Holmgren's theorem.

$$
\begin{array}{r||ccccc}
\text{Cauchy-Kowalevski} & \text{abstract} & \Longrightarrow & \text{classical} & \Longrightarrow & \text{invariant} \\
 & \Downarrow & & & & \\
\text{Holmgren} & \text{abstract} & \Longrightarrow & \text{classical} & \Longrightarrow & \text{invariant}
\end{array}
$$

This result guarantees the uniqueness of the solution in the class C^1, in the case of linear equations. To better understand its significance, we will first look at the statement in its most abstract version, then we will state the classical version and prove the latter using the abstract version. We will omit the statement and proof of the invariant version, which can be found in full in [Tre75, cap.21]; however, it is worth noting that in the latter case, the notion of characteristic surface will explicitly appear in the hypotheses.

Theorem 5.2.1.

$$
\begin{array}{c|l}
 & \mathcal{O}_0 = \{z \in \mathbb{C}^n : |z| < r_0\} \textit{ with } r_0 > 0 \\
 & A_j \textit{ analytic in } x \textit{ and continuous in } t \\
 & y \textit{ distribution on } (\mathcal{O}_0 \cap \mathbb{R}^n) \times (-T, T) \textit{ such that} \\
Hp & -K \subseteq \mathcal{O}_0 \cap \mathbb{R}^n \textit{ compact: } y = 0 \textit{ in } \mathcal{O}_0 \cap \mathbb{R}^n \setminus K \\
 & - \begin{cases} D_t y = \sum A_j(x,t) D_{x_j} y + A_0(x,t) y \\ y = 0 \textit{ for } t < 0 \end{cases} \\
Ts & y = 0 \textit{ in } (\mathcal{O}_0 \cap \mathbb{R}^n) \times (-T, T)
\end{array}
$$

Theorem 5.2.2.

$$
\begin{array}{c|l}
Hp & \Omega \subseteq \mathbb{R}^n \textit{ open} \\
 & A_j \textit{ analytic in } x \textit{ and continuous in } t \\
 & y \in C^1(\Omega \times (-T,T)) \textit{ such that} \\
 & \begin{cases} D_t y = \sum A_j(x,t) D_{x_j} y + A_0(x,t) y \\ y = 0 \textit{ for } t = 0 \end{cases} \\
Ts & y = 0 \textit{ in a neighborhood of } \Omega \times \{0\}
\end{array}
$$

Remark. The crucial difference observed between this statement and that of the classical version of the CKT lies in the fact that this solution can be C^1 with respect to all variables and not just with respect to time.

Proof: Let us start by considering the function

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

and observe that

- $D_t \widetilde{y} = H(t) D_t y$ since $\widetilde{y}(x,0) = 0$, hence $\widetilde{y}$ satisfies the same equation as y;
- it obviously vanishes when $t < 0$.

To conclude the proof, we want to show that by applying the change of variables

$$x' = x, \quad t' = t + |x - x_0|^2 \text{ with } x_0 \in \Omega,$$

we obtain a function $\widetilde{y}(x', t')$ that satisfies the hypotheses of the abstract version of Holmgren's theorem. Let us verify them one by one, assuming for simplicity and without loss of generality that $x_0 = 0 \in \Omega$:

1. considering that x and x' vary over the same set, we can choose r_0 such that $\Omega \subseteq \mathcal{O}_0$;
2. obviously, $\widetilde{y}(x', t')$ can be seen as a distribution on $(\mathcal{O}_0 \cap \mathbb{R}^n) \times (-T, T)$;
3. $\widetilde{y}$ vanishes when $t' < |x'|^2$ and this condition can be rewritten in the form $\mathcal{O}_0 \cap \mathbb{R}^n \setminus K$, if we define K as the projection onto the x space of the set
   $$\{(x', t') : x' \in \Omega,\ |t'| < T,\ |x'|^2 \le t'\},$$
   assuming that T is sufficiently small such that $K \Subset \Omega$;
4. Explicitly calculating the derivatives with respect to x' and t' (see [Tre75, cap.21]) shows that $\widetilde{y}$ satisfies an equation of the type
   $$D_{t'} \widetilde{y} = \sum C_j(x', t') D_{x'_j} y + C_0(x', t') \widetilde{y},$$
   where the coefficients C_j are analytic and therefore have a unique holomorphic extension in a neighborhood of $\mathcal{O}_0$;
5. $\widetilde{y}$ obviously vanishes when $t' < 0$.

QED

5.3 Other Applications

The consequences of Kowalevski's work do not stop at the field of differential equations, but can also be found in the following areas.

- Mathematical physics: for example, in cases where there is a model represented by a system of PDEs that satisfies the hypotheses of the CKT, one may ask whether having a local analytic solution has any physical significance.

- Differential geometry: thanks to the CKT, it has been possible to prove the Cartan-Kähler theorem on the integrability of exterior differential systems, which is very similar in spirit to the CKT.

- Economic theory: utilizing the Cartan-Kähler theorem, I. Ekeland and P.A. Chiappori resolved a problem that had been open for several decades between 1999 and 2009; Ekeland summarizes the work (found in [CE99], [CE06], [CE09a], and [CE09b]) done together with Chiappori with the words quoted below.

*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**. The introduction of exterior differential calculus proved to be a breakthrough, and was the starting point of a long collaboration with P.A. Chiappori. We realized that the mathematical structure we had discovered in this problem was to be found also in one of the major problems of econometrics: given a group (a household, for instance), can one characterize and identify the preferences of each member if one observes only the collective demand? I am happy to say that this research program is now concluded [...].*

Bibliography

[CE99] Pierre A. Chiappori and Ivar Ekeland. Aggregation and market demand: an exterior differential calculus viewpoint. *Econometrica*, 67:1435–1458, 1999.

[CE06] Pierre A. Chiappori and Ivar Ekeland. The micro economics of group behavior: General characterization. *Journal of Economic Theory*, 130:1–26, 2006.

[CE09a] Pierre A. Chiappori and Ivar Ekeland. The economics and mathematics of aggregation: Formal models of efficient group behavior. *Foundations and Trends® in Microeconomics*, 5:1–2, 2009.

[CE09b] Pierre A. Chiappori and Ivar Ekeland. The Microeconomics of Efficient Group Behavior: Identification. *Econometrica*, 77:763 – 799, 2009.

[Eke] Ivar Ekeland. Some applications of the Cartan-Kähler theorem to economic theory.

[Eva10] Lawrence C. Evans. *Partial Differential Equations*. American Mathematical Society, 2010.

[FMS20] Nicola Fusco, Paolo Marcellini, and Carlo Sbordone. *Lectures on Mathematical Analysis 2* (in Italian). Zanichelli, 2020.

[Fol95] Gerald B. Folland. *Introdution to Partial Differential Equations*. Princeton University Press, 1995.

[Hö63] Lars Hörmander. *Linear Partial Differential Operators*. Spinger-Verlag, 1963.

[Joh82] Fritz John. *Partial Differential Equations*. Springer-Verlag, 1982.

[Ken83] Don H. Kennedy. *Little Sparrow: A Portrait of Sophia Kovalevsky*. Ohio University Press, 1983.

[Kov78] Sofya Kovalevskaya. *A Russian Childhood*. Springer New York, 1978. Translated and edited by Beatrice Stillman.

[Lef92] Anne Charlotte Leffler. *Sonja Kovalevsky. What I Experienced with Her and What She Told Me About Herself* (in Russian). Ed. Albert Bonniers, 1892.

[Luc24] Sandra Lucente. Cauchy-Kovalevskaja Theorem for Differential Equations (in Italian). *Le Scienze*, 2024. Collona *Rivoluzioni matematiche*.

[Ovs65] L.V. Ovsyannikov. *Singular operators in Banach spaces scales* (in Russian). Doklady Acad. Nauk., 1965. p. 819–822.

[RF10] Halsey L. Royden and Patrick M. Fitzpatrick. *Real Analysis*. Pearson, 2010.

[Rou80] Delfina Roux. *Foundations of Advanced Analysis - PART I* (in Italian). Edizioni la Viscontea, 1980.

[Spi02] Joan Spicci. *Beyond the Limit: The Dream of Sofya Kovalevskaya*. Tom Doherty Associates, 2002. Translated and edited by Beatrice Stillman.

[Tre70] François Treves. On local solvability of partial differential equations. 1970.

[Tre75] François Treves. *Basic Linear Partial Differntial Equations*. Amacademic Press, 1975.

[Tre22] François Treves. *Analytic Partial Differential Equations*. Springer Nature, 2022.