On the Cauchy-Kowalevski theorem for analytic nonlinear partial differential equations

Logan Meredith

April 12, 2018

1 Introduction

The name "Cauchy problem" is usually attributed to a class of boundary value problems associated to partial differential equations (PDE). The study of such problems began in earnest with Cauchy himself, who investigated the existence of solutions to analytic nonlinear PDE of the second order [1, 2, 3, 4, 5, 6]. This work was extended to general analytic nonlinear systems of PDE by Kowalevski in 1875 [11]. Both results are collectively known as the Cauchy-Kowalevski theorem, which is the primary focus of this paper. It is worth noting that in the same year as Kowalevski, Darboux published a similar result, which applied to less general problem [7]. In 1898, Goursat simplified Kowalevski's argument [10], and it is Goursat's proof that we present here.

The fact that properly defined Cauchy problems have unique analytic solutions is incredibly powerful. Equations such as the wave equation, Maxwell's equations, and the heat equation constitute Cauchy problems when paired with appropriate boundary conditions.

The applicability of the Cauchy-Kowalevski theorem is, however, limited. One major assumption for the theorem is that the functions describing the boundary data and the partial differential equation are all analytic (this term will be defined later). This is an unfortunately stringent requirement, and as the Lewy example in Section 4 shows, there is no way around it. In addition, the theorem only posits the existence and uniqueness of analytic solutions, thus not precluding the existence of nonanalytic ones.

The organization of this paper is as follows. In Section 2, we outline a number of definitions, theorems, and notational conventions, which will allow us to streamline the rather lengthy proof of the Cauchy-Kowalevski theorem. In Section 3, we provide the full proof of the theorem, which includes arguments for both the existence and uniqueness of solutions to the Cauchy problem. In Section 4, the Lewy counterexample is presented. This counterexample illustrates the necessity of requiring analyticity in our formulation of the Cauchy-Kowalevski theorem, as it is an example of a nonanalytic partial differential equation that admits no solutions.

2 Tools

In this section we present the notation used throughout the argument for the Cauchy-Kowalevski theorem. In addition, we provide a number of definitions and theorems which are also used in the proof. However, we do not provide proofs of the theorems described herein, as these are standard results in undergraduate mathematics courses.

2.1 Notation

We adopt the use of boldface letters to denote vectors, in order to differentiate them from scalars. To reduce size, $\partial_{x_i}^k$ indicates taking the k^{th} partial derivative with respect to x_i. The set of all partial derivatives of a function u of order up to and including k is denoted $D^k u = \{u, \partial_{x_1} u, \ldots, \partial_{x_n} u, \ldots, \partial_{x_n}^k u\}$.

The most important piece of notation used in the proof of the Cauchy-Kowalevski theorem is the multiindex, which will be denoted using the Greek letters α and β. A multiindex $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n$ is a vector whose entries are nonnegative integers. We must also define

$$\begin{aligned} \alpha! &\equiv \alpha_1! \alpha_2! \cdots \alpha_n!, \\ |\alpha| &\equiv \alpha_1 + \cdots + \alpha_n. \end{aligned}$$

Multiindices allow us to shorten expressions for partial derivatives by writing

$$\begin{aligned} \partial_{\mathbf{x}}^\alpha &= \partial_{x_1}^{\alpha_1} \partial_{x_2}^{\alpha_2} \cdots \partial_{x_n}^{\alpha_n}, \\ \mathbf{x}^\alpha &= x_1^{\alpha_1} \cdots x_n^{\alpha_n}. \end{aligned}$$

We also define σ^i to be the unique multiindex with $\sigma_i^i = 1$ and $\sigma_j^i = 0$ for all $j \neq i$. In other words, σ^i indicates the multiindex with a 1 as its i^{th} entry and 0 for all other entries.

Given a vector $\mathbf{w} \in \mathbb{R}^n$, we extend the notion of a derivative with respect to a vector to higher order derivatives by defining

$$\frac{\partial^k u}{\partial \mathbf{w}^k} = \sum_{|\alpha| = k} \partial_{\mathbf{x}}^\alpha u \cdot w_1^{\alpha_1} w_2^{\alpha_2} \cdot \cdots \cdot w_n^{\alpha_n}.$$

2.2 Definitions

A small number of definitions are required in order to precisely state and prove the Cauchy-Kowalevski theorem. These are presented here.

Definition 1. A function f is *real analytic* on an open set $U \subset \mathbb{R}^n$ if, for any $\mathbf{x}_0 \in U$,

$$f(\mathbf{x}) = \sum_{|\alpha| \geq 0} a_\alpha (\mathbf{x} - \mathbf{x}_0)^\alpha,$$

for $a_\alpha \in \mathbb{R}$ and the sum converges in a neighborhood of $\mathbf{x}_0$. Clearly, vector- and matrix-valued functions are analytic if each of their entries is analytic.

Care should be taken to avoid confusing *real analytic* functions and *complex analytic* functions, which are defined identically to real analytic functions with the word "real" replaced by "complex." This is because complex analytic functions enjoy a number of properties over real analytic functions by virtue of being defined on the complex numbers. Since there is likely no source of confusion in this paper, we shall refer to real analytic functions simply as *analytic*.

Definition 2. A hypersurface $S \subset \mathbb{R}^n$ is *noncharacteristic* for the equation $F(\mathbf{x}, D^k u) = 0$ if the values of $u, \partial u/\partial \mathbf{n}, \ldots, \partial^{k-1} u/\partial \mathbf{n}^{k-1}$ evaluated on S uniquely determine all derivatives of u on S, where $\mathbf{n}$ denotes the unit normal to S.

Intuitively, noncharacteristic hypersurfaces yield the useful property that equations for which they are noncharacteristic can be solved for $\partial^k u/\partial \mathbf{n}^k$. This fact is essential to reducing the Cauchy problem to a form more amenable to solution.

Definition 3. Given power series

$$\sum_{|\alpha| \geq 0} a_\alpha (\mathbf{x} - \mathbf{x}_0)^\alpha, \quad \sum_{|\alpha| \geq 0} b_\alpha (\mathbf{x} - \mathbf{x}_0)^\alpha,$$

with $a_\alpha \geq 0$, if $a_\alpha \geq |b_\alpha|$ for all α, the first power series is said to *majorize* the second.

The notion of majorization is crucial to proving that a given power series converges, since the direct comparison test can be immediately applied to the majorized series.

2.3 Theorems

In this section we enumerate several classical results used in the main argument.

Theorem 1. *Given a power series that converges absolutely, if it majorizes another power series, then that power series also converges absolutely.*

Theorem 2. *Let $f(\mathbf{x}) = \sum_{|\alpha| \geq 0} a_\alpha (\mathbf{x} - \mathbf{x}_0)^\alpha$ be convergent in a neighborhood of $\mathbf{x}_0 \in \mathbb{R}^n$, and let $\mathbf{x}(\mathbf{y}) = \sum_{|\alpha| \geq 0} \mathbf{b}_\alpha (\mathbf{y} - \mathbf{y}_0)^\alpha$ be convergent in a neighborhood of $\mathbf{y}_0 \in \mathbb{R}^m$ with $\mathbf{x}(\mathbf{y}_0) = \mathbf{b_0} = \mathbf{x}_0$. Then the composite function $g(\mathbf{y}) \equiv f(\mathbf{x}(\mathbf{y}))$ is analytic at $\mathbf{y}_0$. The power series expansions about $\mathbf{y}_0$ can be obtained by substituting the power series $\sum_{\alpha \neq 0} \mathbf{b}_\alpha (\mathbf{y} - \mathbf{y}_0)^\alpha$ for $\mathbf{x} - \mathbf{x}_0$ in the series for $f(\mathbf{x})$. The result is that $g(\mathbf{y}) = \sum_{|\alpha| \geq 0} c_\alpha (\mathbf{y} - \mathbf{y}_0)^\alpha$, where the c_α are polynomials of the $a_\alpha, \mathbf{b}_\alpha$ with nonnegative coefficients, since only addition and multiplication are used to compute each c_α.*

Theorem 3. *An analytic function is uniquely determined in the neighborhood of a point by its derivatives evaluated at that point.*

3 Cauchy-Kowalevski Theorem

The main objective of this section is the resolution of the Cauchy problem,

$$F(\mathbf{x}, D^k u) = 0, \tag{1a}$$

$$\left.\frac{\partial^i u}{\partial \mathbf{n}^i}\right|_S = \phi_i, \quad 0 \le i \le k-1, \tag{1b}$$

for which a slightly imprecise version of the Cauchy-Kowalevski theorem states

Theorem 4. *If $F, \phi_0, \phi_1, \ldots, \phi_{k-1}$ are all analytic and $S \in \mathbb{R}^n$ is non-characteristic with respect to* (1a)*, then the Cauchy problem admits a unique local analytic solution.*

In order to formulate a precise version of this result, we flatten the hypersurface S near a fixed point $\mathbf{x}_0 \in S$ by mapping it to a portion of the hyperplane $\{t = 0\} \subset \mathbb{R}^{n-1} \times \mathbb{R}$ near the origin $(\mathbf{x}, t) = (\mathbf{0}, 0)$. Moreover, the non-characteristic assumption on S is equivalent to the ability to solve (1a) for $\partial_t^k u$ in terms of $\mathbf{x}$, t, and the other derivatives of u. Thus, in the new setting, the Cauchy problem reads as

$$\partial_t^k u = G(\mathbf{x}, t, D^k u), \tag{2a}$$

$$\partial_t^i u(\mathbf{x}, 0) = \psi_i(\mathbf{x}), \quad 0 \le i \le k-1, \tag{2b}$$

where G does not depend on $\partial_t^k u$. We can now more precisely state the Cauchy-Kowalevski theorem applied to the reformulated Cauchy problem.

Theorem 5 (Cauchy-Kowalevski)**.** *If G, ψ_0, $\psi_1, \ldots, \psi_{k-1}$ are all analytic in a neighborhood of the origin, then there exists a unique analytic solution defined in a neighborhood of the origin.*

The strategy for the proof for the existence of solutions consists first of a series of reductions of the Cauchy problem to a quasilinear first order system with zero boundary data and whose coefficients do not depend on t. The final form of the system is treated through the method of formal power series jointly with sufficient conditions for the convergence of such series.

3.1 Reduction to first order system

In this section we adopt the classical method of transforming a high order partial differential equation to a first order system. Intuitively, we want to write a system for the vector-valued field $\mathbf{y}$ that satisfies $y_{\alpha i} = \partial_{\mathbf{x}}^\alpha \partial_t^i u$.

Consider the first order problem

$$\partial_t y_{\alpha i} = y_{\alpha(i+1)} \quad \text{for } 0 \le |\alpha| + i < k, \tag{3a}$$

$$\partial_t y_{\alpha i} = \partial_{x_j} y_{(\alpha-\sigma^j)(i+1)} \quad \text{for } |\alpha| + i = k, \quad i < k, \tag{3b}$$

$$\partial_t y_{\mathbf{0}k} = \frac{\partial G}{\partial t} + \sum_{|\alpha|+i<k} \frac{\partial G}{\partial y_{\alpha i}} y_{\alpha(i+1)} + \sum_{|\alpha|+i=k, i<k} \frac{\partial G}{\partial y_{\alpha i}} \partial_{x_j} y_{(\alpha-\sigma^j)(i+1)}, \tag{3c}$$

$$y_{\alpha i}(\mathbf{x}, 0) = \partial_{\mathbf{x}}^{\alpha} \psi_i(\mathbf{x}) \quad \text{for } i < k, \tag{3d}$$

$$y_{\mathbf{0}k}(\mathbf{x}, 0) = G(\mathbf{x}, 0, \psi), \tag{3e}$$

where, for each α, j is the smallest index with $\alpha_j \neq 0$, and G does not depend on $y_{\mathbf{0}k}$. We posit that if u is a solution of (2), then we can construct a solution to (3) by setting $y_{\alpha i} = \partial_{\mathbf{x}}^{\alpha} \partial_t^i u$. In fact, this can be readily seen by the construction of (3). A less obvious fact is that if $\mathbf{y}$ is a solution to (3), then $y_{\mathbf{0}0}$ is a solution to (2). We shall prove this now.

It can be clearly seen from (3a) that

$$y_{\alpha(i+l)} = \partial_t^l y_{\alpha i} \tag{4}$$

for $i + l \le k$. This fact and (3b) together imply that

$$\partial_t y_{\alpha i} = \partial_t \partial_{x_j} y_{(\alpha-\sigma^j)i} \tag{5}$$

for $|\alpha| + i = k$ and $i < k$. By integrating both sides of this equation with respect to t, we find that

$$y_{\alpha i}(\mathbf{x}, t) = \partial_{x_j} y_{(\alpha-\sigma^j)i}(\mathbf{x}, t) + f_{\alpha i}(\mathbf{x}) \tag{6}$$

for some function $f_{\alpha i} : \mathbb{R}^{n-1} \mapsto \mathbb{R}$. But we can use the initial data to determine the $f_{\alpha i}$. From (3d), we find

$$\begin{aligned} y_{\alpha i}(\mathbf{x}, 0) &= \partial_{\mathbf{x}}^{\alpha} \psi_i(\mathbf{x}) \\ &= \partial_{x_j} \partial_{\mathbf{x}}^{\alpha-\sigma^j} \psi_i(\mathbf{x}) \\ &= \partial_{x_j} y_{(\alpha-\sigma^j)i}(\mathbf{x}, 0), \end{aligned}$$

which implies that $f_{\alpha i}(\mathbf{x}) = 0$ everywhere. Hence

$$y_{\alpha i} = \partial_{x_j} y_{(\alpha-\sigma^j)i} \tag{7}$$

for $|\alpha| + i = k$ and $i < k$. By induction on $k - |\alpha| - i = l$, we can show that (7) holds for all $\alpha \neq \mathbf{0}$. Suppose that (7) holds for $l - 1$; then we have

$$\begin{aligned} \partial_t y_{\alpha i} &= y_{\alpha(i+1)} \\ &= \partial_{x_j} y_{(\alpha-\sigma^j)(i+1)} \\ &= \partial_t \partial_{x_j} y_{(\alpha-\sigma^j)i}. \end{aligned}$$

By integrating both sides of this equation with respect to t, we find that

$$y_{\alpha i}(\mathbf{x}, t) = \partial_{x_j} y_{(\alpha-\sigma^j)i}(\mathbf{x}, t) + f_{\alpha i}(\mathbf{x}) \tag{8}$$

for some function $f_{\alpha i} : \mathbb{R}^{n-1} \mapsto \mathbb{R}$. But we can use the initial data to determine $f_{\alpha i}$. From (3d), we find

$$\begin{aligned} y_{\alpha i}(\mathbf{x}, 0) &= \partial_{\mathbf{x}}^{\alpha} \psi_i(\mathbf{x}) \\ &= \partial_{x_j} \partial_{\mathbf{x}}^{\alpha - \sigma^j} \psi_i(\mathbf{x}) \\ &= \partial_{x_j} y_{(\alpha - \sigma^j) i}(\mathbf{x}, 0), \end{aligned}$$

which implies that $f_{\alpha i}(\mathbf{x}) = 0$ everywhere. Hence (7) holds for all $\alpha \neq \mathbf{0}$.

Now we solve for $y_{\mathbf{0}k}$. From (3c), (4), and (7), we have

$$\begin{aligned} \partial_t y_{\mathbf{0}k} &= \frac{\partial G}{\partial t} + \sum_{|\alpha|+i<k} \frac{\partial G}{\partial y_{\alpha i}} y_{\alpha(i+1)} + \sum_{|\alpha|+i=k, i<k} \frac{\partial G}{\partial y_{\alpha i}} \partial_{x_j} y_{(\alpha-\sigma^j)(i+1)} \\ &= \frac{\partial G}{\partial t} + \sum_{|\alpha|+i \leq k, i<k} \frac{\partial G}{\partial y_{\alpha i}} \partial_t y_{\alpha i} \\ &= \partial_t (G(\mathbf{x}, t, \mathbf{y})). \end{aligned}$$

By integrating both sides of this equation with respect to t, we find that

$$y_{\mathbf{0}k}(\mathbf{x}, t) = G(\mathbf{x}, t, \mathbf{y}(\mathbf{x}, t)) + f_{\mathbf{0}k}(\mathbf{x})$$

for some function $f_{\mathbf{0}k} : \mathbb{R}^{n-1} \mapsto \mathbb{R}$. But we can use the initial data to determine $f_{\mathbf{0}k}$. From (3e), we find

$$\begin{aligned} y_{\mathbf{0}k}(\mathbf{x}, 0) &= G(\mathbf{x}, 0, \psi(\mathbf{x})) \\ &= G(\mathbf{x}, 0, \mathbf{y}(\mathbf{x}, 0)), \end{aligned}$$

which implies that $f_{\mathbf{0}k} = 0$ everywhere. Hence

$$y_{\mathbf{0}k} = G(\mathbf{x}, t, \mathbf{y}). \tag{9}$$

From (4) and (7), it is easy to see that

$$y_{\alpha i} = \partial_t^i \partial_{\mathbf{x}}^{\alpha} y_{\mathbf{0}0}. \tag{10}$$

Thus, by (3d) and (9), $y_{\mathbf{0}0}$ is a solution to (2).

Finally, we make one further reduction in the sense that (3) can be shown to be equivalent to a similar system with zero boundary data and coefficients that do not depend on t.

Let m be the length of $\mathbf{y}$ from the previous step; in other words, m is the number of derivatives that can be taken of a function of n variables of order up to and including k. By construction, our first order problem (3) can be rewritten as

$$\partial_t \mathbf{y} = \sum_{i=1}^{n-1} A_i(\mathbf{x}, t, \mathbf{y}) \partial_{x_i} \mathbf{y} + \mathbf{c}(\mathbf{x}, t, \mathbf{y}), \tag{11a}$$

$$\mathbf{y}(\mathbf{x}, 0) = \Psi(\mathbf{x}), \tag{11b}$$

where Ψ is a vector-valued function containing the analytic functions $\partial_{\mathbf{x}}^{\alpha}\psi_i$ and $G(\mathbf{x}, 0, \psi(\mathbf{x}))$, and the $A_i : \mathbb{R}^{n-1} \times \mathbb{R} \times \mathbb{R}^m \mapsto \mathbb{M}^{m\times m}$ and $\mathbf{c} : \mathbb{R}^{n-1} \times \mathbb{R} \times \mathbb{R}^m \mapsto \mathbb{R}^m$ are analytic. The equivalence of (11) to (3) can be seen by observing that the right sides of (3a), (3b), and (3c) are quasilinear combinations of derivatives of $\mathbf{y}$ with respect to $\mathbf{x}$ whose coefficients are analytic functions of $\mathbf{x}$, t, and $\mathbf{y}$. In addition, the right sides of (3d) and (3e) are analytic functions of $\mathbf{x}$ alone.

This first-order system can be further simplified. If the A_i or $\mathbf{c}$ depend on t, we can add a new component $y_{m+1} \equiv t$ to $\mathbf{y}$ to eliminate this dependence. Thus, by setting

$$\tilde{\mathbf{y}}(\mathbf{x}, t) = \mathbf{y}(\mathbf{x}, t) - \Psi(\mathbf{x}), \tag{12a}$$

$$\tilde{A}_i(\mathbf{x}, \tilde{\mathbf{y}}) = A_i(\mathbf{x}, \tilde{\mathbf{y}} + \Psi), \tag{12b}$$

$$\tilde{\mathbf{c}}(\mathbf{x}, \tilde{\mathbf{y}}) = \mathbf{c}(\mathbf{x}, \tilde{\mathbf{y}} + \Psi) + \sum_{i=1}^{n-1} A_i(\mathbf{x}, \tilde{\mathbf{y}} + \Psi)\partial_{x_i}\Psi(\mathbf{x}), \tag{12c}$$

we can rewrite (11) as

$$\partial_t\tilde{\mathbf{y}} = \sum_{i=1}^{n-1} \tilde{A}_i(\mathbf{x}, \tilde{\mathbf{y}})\partial_{x_i}\tilde{\mathbf{y}} + \tilde{\mathbf{c}}(\mathbf{x}, \tilde{\mathbf{y}}), \tag{13a}$$

$$\tilde{\mathbf{y}}(\mathbf{x}, 0) = \mathbf{0}, \tag{13b}$$

where $\mathbf{y}$ can be obtained from $\tilde{\mathbf{y}}$ with (12a). Importantly, since (2) is equivalent to (11), we know that $u = \tilde{y}_{\mathbf{0}0} + \psi_0$ is a solution to (2) if and only if $\tilde{\mathbf{y}}$ is a solution to (13).

3.2 Solution by formal power series

Following the work in the previous section, we proceed to solve (13) by formally computing a power series near the origin.

We have thus far been indexing the entries of $\tilde{\mathbf{y}}$ and related vectors using two indices. Now we rewrite $\tilde{\mathbf{y}} = (\tilde{y}_1, \ldots, \tilde{y}_m)$, $\tilde{\mathbf{c}} = (\tilde{c}_1, \ldots, \tilde{c}_m)$, and $\tilde{A}_i = ((\tilde{a}_1^{st}), \ldots, (\tilde{a}_m^{st}))_{s,t=1}^m$. Since $\tilde{\mathbf{y}}$ is analytic, we can write

$$\tilde{\mathbf{y}} = \sum_{|\alpha|, i\geq 0} \mathbf{v}_{\alpha,i}\mathbf{x}^{\alpha}t^i, \tag{14}$$

where

$$\mathbf{v}_{\alpha,i} = \frac{\partial_{\mathbf{x}}^{\alpha}\partial_t^i\tilde{\mathbf{y}}(\mathbf{0}, 0)}{\alpha! i!}. \tag{15}$$

From (13b), it can be seen that $\mathbf{v}_{\alpha,0} = \mathbf{0}$ for all $|\alpha| \geq 0$.

The $\mathbf{v}_{\alpha,i}$ for $i > 0$ can be computed by substituting (14) into (13). From this substitution, we get

$$\sum_{|\alpha|, i\geq 0} (i+1)\mathbf{v}_{\alpha,i+1}\mathbf{x}^{\alpha}t^i = \sum_{|\alpha|, i\geq 0} \mathbf{q}_{\alpha,i}(\tilde{A}_1, \ldots, \tilde{A}_{n-1}, \tilde{\mathbf{c}}, (\mathbf{v}_{\beta,j})_{j\leq i})\mathbf{x}^{\alpha}t^i,$$

where the $\mathbf{q}_{\alpha,i}$ are polynomials with nonnegative coefficients. This is guaranteed by Theorem 2. Since $\mathbf{q}_{\alpha,i}$ depends on $\mathbf{v}_{\beta,j}$ for $j \leq i$, and we know $\mathbf{v}_{\beta,0}$, we can inductively solve for the rest of them to find

$$\mathbf{v}_{\alpha,i} = \mathbf{p}_{\alpha,i}(\tilde{A}_1, \ldots, \tilde{A}_{n-1}, \tilde{\mathbf{c}}), \tag{16}$$

where the $\mathbf{p}_{\alpha,i}$ are polynomials with nonnegative coefficients.

All that is left is to argue for the convergence of (14). The idea is to build a system which is easily solved using the method of characteristics and whose solution written as a power series majorizes the one for the system of interest.

Suppose we have A_i^* for $1 \leq i \leq n-1$ and $\mathbf{c}^*$, both analytic, whose power series majorize those of $\tilde{A}_i$ and $\tilde{\mathbf{c}}$, respectively. Further suppose that the problem given by

$$\partial_t \mathbf{y}^* = \sum_{i=1}^{n-1} A_i^*(\mathbf{x}, \mathbf{y}^*) \partial_{x_i} \mathbf{y}^* + \mathbf{c}^*(\mathbf{x}, \mathbf{y}^*), \tag{17a}$$

$$\mathbf{y}^*(\mathbf{x}, 0) = \mathbf{0} \tag{17b}$$

yields an analytic solution in a neighborhood of the origin. Then

$$\mathbf{y}^* = \sum_{|\alpha|, i \geq 0} \mathbf{v}_{\alpha,i}^* \mathbf{x}^\alpha t^i \tag{18}$$

for

$$\mathbf{v}_{\alpha,i}^* = \mathbf{p}_{\alpha,i}(A_1^*, \ldots, A_{n-1}^*, \mathbf{c}^*), \tag{19}$$

where the $\mathbf{p}_{\alpha,i}$ are precisely the same polynomials with nonnegative coefficients as in the previous step. The fact that $\mathbf{p}_{\alpha,i}$ have nonnegative coefficients is essential to the proof, since it implies that

$$\begin{aligned} |\mathbf{v}_{\alpha,i}| &= |\mathbf{p}_{\alpha,i}(\tilde{A}_1, \ldots, \tilde{A}_{n-1}, \tilde{\mathbf{c}})| \\ &\leq \mathbf{p}_{\alpha,i}(|\tilde{A}_1|, \ldots, |\tilde{A}_{n-1}|, |\tilde{\mathbf{c}}|) \\ &\leq \mathbf{p}_{\alpha,i}(A_1^*, \ldots, A_{n-1}^*, \mathbf{c}^*) \\ &= \mathbf{v}_{\alpha,i}^*. \end{aligned}$$

By definition, the power series of $\mathbf{y}^*$ therefore majorizes the power series of $\tilde{\mathbf{y}}$, and so they both converge in a neighborhood of the origin.

The final step is to show that such a majorizing power series $\mathbf{y}^*$ exists. Let

$$\mathbf{c}^* = \frac{Mr}{r - \sum_{i=1}^{n-1} x_i - \sum_{i=1}^{m} y_i^*} \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix},$$

$$A_i^* = \frac{Mr}{r - \sum_{i=1}^{n-1} x_i - \sum_{i=1}^{m} y_i^*} \begin{pmatrix} 1 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \cdots & 1 \end{pmatrix}$$

for all i. Then (17) can be written as

$$\partial_t \mathbf{y}^* = \frac{Mr}{r - \sum_{i=1}^{n-1} x_i - \sum_{i=1}^{m} y_i^*} \left(\sum_{i=1}^{n-1} \sum_{j=1}^{m} \partial_{x_i} y_j^* + \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \right), \tag{20a}$$

$$\mathbf{y}^*(\mathbf{x}, 0) = \mathbf{0}. \tag{20b}$$

Given the solution u^* to the problem

$$\partial_t u^* = \frac{Mr}{r - s - Nu^*}(N(n-1)\partial_x u^* + 1), \tag{21a}$$

$$u^*(x, 0) = 0, \tag{21b}$$

a solution to (20) can be found as

$$\mathbf{y}^*(\mathbf{x}, t) = u^* \left(\sum_{i=1}^{n-1} x_i, t \right) \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}.$$

Solution of (21) proceeds via the method of characteristics. We begin by solving the ordinary differential equations

$$\begin{aligned} \frac{dx}{d\mu} &= -Mrm(n-1), \\ \frac{dt}{d\mu} &= r - x - mu^*, \\ \frac{du^*}{d\mu} &= Mr, \end{aligned}$$

along with initial conditions given by

$$\begin{aligned} x(0) &= 0, \\ t(0) &= \eta, \\ u(0) &= 0. \end{aligned}$$

The solutions to this problem are

$$\begin{aligned} x &= -Mrm(n-1)\mu + \eta, \\ t &= \frac{1}{2}Mrm(n-2)\mu^2 + (r-\eta)\mu, \\ u^* &= Mr\mu. \end{aligned}$$

We can invert the first two equations to find μ and η as functions of x and t. Hence u^* can be found in terms of x and t to be

$$u^*(x, t) = \frac{r - x - \sqrt{(r-x)^2 - 2mnMrt}}{Mn}. \tag{22}$$

The solution to (20) is therefore given by

$$\mathbf{y}^*(\mathbf{x},t) = \frac{r - \sum_{i=1}^{n-1} x_i - \sqrt{\left(r - \sum_{i=1}^{n-1} x_i\right)^2 - 2mnMrt}}{Mn} \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}. \tag{23}$$

This is analytic for $|\mathbf{x}| < r$. Furthermore, the A_i^* and $\mathbf{c}^*$ are analytic for $|\mathbf{x}| < r$ and majorize $\tilde{A}_i$ and $\tilde{\mathbf{c}}$, respectively, for sufficiently large $M > 0$ and sufficiently small $r < 0$. Hence $\tilde{\mathbf{y}}$ is analytic in a neighborhood of the origin, and so the proof of existence of solutions is complete.

3.3 Uniqueness of solutions

In this section, we prove uniqueness of solutions to the Cauchy problem in order to complete the proof of the Cauchy-Kowalevski theorem. Uniqueness can be precisely stated with the following theorem.

Theorem 6. *If G, ψ_0, $\psi_1, \ldots, \psi_{k-1}$ are all analytic in a neighborhood of the origin, then there is at most one analytic solution to* (2) *defined in a neighborhood of the origin.*

Proof. According to Theorem 3, an analytic function is uniquely determined by its derivatives evaluated at a single point. For solutions to the Cauchy problem, the derivatives are uniquely determined by (2). $\square$

4 The Lewy Example

The Cauchy-Kowalevski theorem requires that every function in (2) be analytic. However, analyticity is a rather strong requirement. It is natural to ask whether it is necessary for the functions in (2) to be constrained to be analytic. It was not until 1957 that Hans Lewy [12] provided an example of a partial differential equation defined in terms of nonanalytic functions that yielded no solution:

$$\partial_x u + i\partial_y u - 2i(x+iy)\partial_t u = f(t), \tag{24}$$

where $f \in C^1(\mathbb{R})$. A precise statement of the theorem that no solution exists is given as follows.

Theorem 7. *If there exists a solution $u \in C^1(\mathbb{R}^3)$ to* (24) *in a neighborhood of the origin, then f is analytic at $t = 0$.*

The contrapositive in particular states that if f is not analytic, then there does not exist a solution to (24).

Proof. Let $u = u(x,y,t) \in C^1(\mathbb{R}^3)$ solve (24) in $B_R(\mathbf{0})$ for some $R > 0$. Let

$$V(r,t) = \int_0^{2\pi} u(r\cos\theta, r\sin\theta, t) ire^{i\theta}\, d\theta.$$

By taking a derivative with respect to $s = r^2$ and using Green's theorem, this reduces to

$$\begin{aligned}
\frac{\partial V}{\partial s} &= \frac{1}{2r}\frac{\partial V}{\partial r} \\
&= \frac{1}{2r}\frac{\partial}{\partial r}\left[\int_0^r \int_0^{2\pi} \left(\frac{\partial u}{\partial x}(\xi\cos\theta, \xi\sin\theta, t) + i\frac{\partial u}{\partial y}(\xi\cos\theta, \xi\sin\theta, t)\right) i\xi \, d\theta \, d\xi\right] \\
&= \frac{1}{2r}\int_0^{2\pi} \left(\frac{\partial u}{\partial x}(r\cos\theta, r\sin\theta, t) + i\frac{\partial u}{\partial y}(r\cos\theta, r\sin\theta, t)\right) ir \, d\theta \\
&= \frac{1}{2}\int_{|z|=r} \left(\frac{\partial u}{\partial x}(x, y, t) + i\frac{\partial u}{\partial y}(x, y, t)\right) \frac{1}{z} \, dz \\
&= \int_{|z|=r} \frac{\partial u}{\partial t} i \, dz + \frac{1}{2}\int_{|z|=r} f(t)\frac{1}{z} \, dz \\
&= i\frac{\partial}{\partial t}\int_0^{2\pi} u(r\cos\theta, r\sin\theta, t) \, ire^{i\theta} \, d\theta + \frac{1}{2}f(t)2\pi i \\
&= i\frac{\partial V}{\partial t} + \pi i f.
\end{aligned}$$

If we introduce $F(t) = \int_0^t f(\xi)\, d\xi$, then the equation

$$\frac{\partial U}{\partial t} + i\frac{\partial U}{\partial s} = 0$$

is solved by $U(s, t) = V(s, t) + \pi F(t)$. However, this equation is the classical Cauchy-Riemann equation, which describes holomorphic functions. Since $V(0, t) = 0$, $U(0, t) = \pi F(t)$ is real. By the Schwarz reflection principle, we can analytically continue U to a neighborhood of the origin, since it is only defined for $0 < s < R^2$. The analytic continuation of U satisfies $U(0, t) = \pi F(t)$, and so F is analytic, and therefore so is f. $\square$

References

[1] Augustin Louis Cauchy. Mémoire sur l'application du calcul des limites à l'intégration d'un système d'équations aux dérivées partielles. *Comptes Rendus Acad. Sci*, 40:85–101, 1842.

[2] Augustin Louis Cauchy. Mémoire sur l'emploi du calcul des limites dans l'intégration des équations aux dérivées partielles. *Comptes Rendus Acad. Sci*, 40:44–59, 1842.

[3] Augustin Louis Cauchy. Mémoire sur les intégrales des systèmes d'équations différentielles ou aux dérivées partielles, et sur les développements de ces intégrales en séries ordonnées suivant les puissances ascendantes d'un paramètre que renferment les équations proposées. *Comptes Rendus Acad. Sci*, 40:141–146, 1842.

[4] Augustin Louis Cauchy. Mémoire sur les systèmes d'équations aux dérivées partielles d'ordres quelconques, et sur leur réduction à des systèmes d'équations linéaires du premier ordre. *Comptes Rendus Acad. Sci*, 40:131–138, 1842.

[5] Augustin-Louis Cauchy. Mémoire sur un théorème fondamental, dans le calcul intégral. *Comptes Rendus Acad. Sci*, 14:1020–1026, 1842.

[6] Augustin Louis Cauchy. Note sur divers théorèmes relatifs aux calcul des limites. *Comptes Rendus Acad. Sci*, 40:138–139, 1842.

[7] G Darboux. Sur l'existence de l'intégrale dans les équations aux derivées partielles d'ordre quelconque. *CR Acad. Sci. Paris*, 80:317–318, 1875.

[8] L.C. Evans and American Mathematical Society. *Partial Differential Equations*. Graduate studies in mathematics. American Mathematical Society, 2010.

[9] G.B. Folland. *Introduction to Partial Differential Equations*. Princeton University Press, 1995.

[10] E Goursat. Sur l'existence des fonctionne inégrales d'un systeme d'équations aux dérivées partielles. *Bulletin de le Société Mathématique de France*, 26:129–134, 1898.

[11] Sophie Kowalevski. Zur Theorie der partiellen Differentialgleichungen. *J. Reine Angew. Math.*, 80:1–32, 1875.

[12] Hans Lewy. An example of a smooth linear partial differential equation without solution. *Annals of Mathematics*, pages 155–158, 1957.

[13] R.C. McOwen. *Partial Differential Equations: Methods and Applications*. Prentice Hall, 1996.