THE CAUCHY PROBLEM VIA THE METHOD OF CHARACTERISTICS

ARICK SHAO

In this short note, we solve the Cauchy, or initial value, problem for general fully nonlinear first-order PDE. Throughout, our PDE will be defined by the function

$$F : \mathbb{R}^2_{x,y} \times \mathbb{R}_z \times \mathbb{R}^2_{p,q} \to \mathbb{R}.$$

We also fix an open interval $I \subseteq \mathbb{R}$, as well as functions $f, g, h : I \to \mathbb{R}$. In particular,

$$\Gamma := \{(f(r), g(r)) \mid r \in I\}$$

is the curve on which we impose the initial data, while h represents the initial data itself.

More specifically, our goal is to solve the following *Cauchy problem*:

$$F(x, y, u, \partial_x u, \partial_y u) = 0, \qquad u(f(r), g(r)) = h(r). \tag{1}$$

Our main result is as follows:

Theorem 1. *Let $F \in C^2(\mathbb{R}^2 \times \mathbb{R} \times \mathbb{R}^2)$, and let $f, g, h \in C^2(I)$. Moreover, let $r_0 \in I$ and $p_0, q_0 \in \mathbb{R}$ such that the following* admissibility conditions *are satisfied:*

$$f'(r_0) \cdot p_0 + g'(r_0) \cdot q_0 = h'(r_0), \tag{2}$$

$$F(f(r_0), g(r_0), h(r_0), p_0, q_0) = 0, \tag{3}$$

$$\det \begin{bmatrix} f'(r_0) & g'(r_0) \\ \partial_p F(f(r_0), g(r_0), h(r_0), p_0, q_0) & \partial_q F(f(r_0), g(r_0), h(r_0), p_0, q_0) \end{bmatrix} \neq 0. \tag{4}$$

Then, there exists a neighbourhood $\mathcal{U} \subseteq \mathbb{R}^2$ of $(f(r_0), g(r_0))$ such that the Cauchy problem (1) has a unique solution $u \in C^2(\mathcal{U})$ that satisfies the additional conditions

$$\partial_x u(f(r_0), g(r_0)) = p_0, \qquad \partial_y u(f(r_0), g(r_0)) = q_0. \tag{5}$$

Remark. *In the quasilinear case,*

$$F(x, y, z, p, q) = a(x, y, z)p + b(x, y, z)q - c(x, y, z),$$

one no longer needs to choose p_0, q_0 beforehand. To see this, note that:

- *The noncharacteristic condition (4) is independent of p_0 and q_0, as it reduces to*

$$\det \begin{bmatrix} f'(r_0) & g'(r_0) \\ a(f(r_0), g(r_0), h(r_0)) & b(f(r_0), g(r_0), h(r_0)) \end{bmatrix} \neq 0. \tag{6}$$

- *The remaining admissibility conditions (2), (3) are also unnecessary, as these now become*

$$f'(r_0) \cdot p_0 + g'(r_0) \cdot q_0 = h'(r_0),$$
$$a(f(r_0), g(r_0), h(r_0))p_0 + b(f(r_0), g(r_0), h(r_0))q_0 = c(x(r_0), y(r_0), z(r_0)),$$

  and (6) guarantees that the above yields exactly one possible pair (p_0, q_0).

Remark. *Furthermore, for the quasilinear case, we need only assume $a, b, c \in C^1$ and $f, g, h \in C^1$. This yields a unique solution $u \in C^1$. Less regularity is required here than in the fully nonlinear case, since if one proves the quasilinear analogue of Theorem 1 directly (without referring to the fully nonlinear case), then one can avoid altogether p, q, and second derivatives of u.*

Proof of Theorem 1: Existence. The first step for proving existence is to construct a set of similarly admissible data on Γ near $(f(r_0), g(r_0))$:

Lemma 1. *There exists an open interval $J \subseteq I$ containing r_0, and there exist unique $w, v \in C^1(J)$ such that $(w(r_0), v(r_0)) = (p_0, q_0)$, and such that the following hold for every $r \in J$:*

$$f'(r) \cdot w(r) + g'(r) \cdot v(r) = h'(r), \tag{7}$$

$$F(f(r), g(r), h(r), w(r), v(r)) = 0, \tag{8}$$

Furthermore, J, w, and v can be chosen such that for each $r \in J$,

$$\det \begin{bmatrix} f'(r) & g'(r) \\ \partial_p F(f(r), g(r), h(r), w(r), v(r)) & \partial_q F(f(r), g(r), h(r), w(r), v(r)) \end{bmatrix} \neq 0. \tag{9}$$

Proof. Consider the map $\Phi : \mathbb{R}_r \times \mathbb{R}^2_{P,Q} \to \mathbb{R}^2$ given by

$$\Phi(r, P, Q) := (f'(r) \cdot P + g'(r) \cdot Q - h'(r), F(f(r), g(r), h(r), P, Q)).$$

Note that (2) and (3) imply that $\Phi(r_0, p_0, q_0) = 0$. Moreover, since F, f, g, h are C^2, then Φ is C^1. We now compute the derivative of Φ with respect to P and Q at (r_0, p_0, q_0):

$$D_{P,Q}\Phi|_{(r_0,p_0,q_0)} = \begin{bmatrix} f'(r_0) & g'(r_0) \\ \partial_p F(f(r_0), g(r_0), h(r_0), p_0, q_0) & \partial_q F(f(r_0), g(r_0), h(r_0), p_0, q_0) \end{bmatrix}.$$

In particular, (4) implies $D_{P,Q}\Phi|_{(r_0,p_0,q_0)}$ is nonsingular. By the implicit function theorem, there exists an open interval $J \subseteq I$ and a unique C^1-function $\Psi = (w, v) : J \to \mathbb{R}^2$ such that

$$\Phi(r, \Psi(r)) = \Phi(r, w(r), v(r)) = 0, \qquad r \in J.$$

The definition of Φ now implies that (7) and (8) hold for $r \in J$. Finally, by reducing J if necessary, then continuity implies that (9) also holds for $r \in J$. ☐

Now that we have w and v set, we can set up the characteristic equations. Set

$$\tilde{\gamma}(r, s) := (x(r, s), y(r, s), z(r, s), p(r, s), q(r, s)), \qquad \gamma(r, s) := (x(r, s), y(r, s)),$$

and recall that the characteristic equations are precisely the following initial value problem:

$$\begin{aligned} \partial_s x(r, s) &= \partial_p F(\tilde{\gamma}(r, s)), & x(r, 0) &= f(r), \\ \partial_s y(r, s) &= \partial_q F(\tilde{\gamma}(r, s)), & y(r, 0) &= g(r), \\ \partial_s z(r, s) &= \partial_p F(\tilde{\gamma}(r, s)) \cdot p(r, s) + \partial_q F(\tilde{\gamma}(r, s)) \cdot q(r, s), & z(r, 0) &= h(r), \\ \partial_s p(r, s) &= -\partial_x F(\tilde{\gamma}(r, s)) - \partial_z F(\tilde{\gamma}(r, s)) \cdot p(r, s), & p(r, 0) &= w(r), \\ \partial_s q(r, s) &= -\partial_y F(\tilde{\gamma}(r, s)) - \partial_z F(\tilde{\gamma}(r, s)) \cdot q(r, s), & q(r, 0) &= v(r), \end{aligned} \tag{10}$$

Here, s represents the parameter along the characteristic curves, while r parametrises the characteristic curves corresponding to the point on Γ that they intersect.

Since the right-hand sides of the equations in (10) are C^1-functions of $\tilde{\gamma}(r, s)$, and since the initial data f, g, h, w, r are at least C^1, then standard ODE theory yields:

Lemma 2. *There exists an open rectangle $\mathcal{R} \subseteq \mathbb{R}^2$ about $(r_0, 0)$, on which there exists a unique C^1-function $\tilde{\gamma} : \mathcal{R} \to \mathbb{R}^2 \times \mathbb{R} \times \mathbb{R}^2$ which solves the system* (10)*.*

Since (7) and (8) represent consistency conditions for solutions, it is important to show that (7) and (8) are propagated along the characteristic curves:

Lemma 3. *The following conditions hold for each $(r, s) \in \mathcal{R}$:*

$$\partial_r x(r, s) \cdot p(r, s) + \partial_r y(r, s) \cdot q(r, s) = \partial_r z(r, s), \tag{11}$$

$$F(x(r, s), y(r, s), z(r, s), p(r, s), q(r, s)) = 0. \tag{12}$$

Proof. First, by the chain rule and then by (10),

$$\begin{aligned}\partial_s[F(\tilde\gamma(r,s))] &= \partial_x F(\tilde\gamma(r,s))\cdot\partial_s x(r,s) + \partial_y F(\tilde\gamma(r,s))\cdot\partial_s y(r,s) + \partial_z F(\tilde\gamma(r,s))\cdot\partial_s z(r,s)\\ &\qquad + \partial_p F(\tilde\gamma(r,s))\cdot\partial_s p(r,s) + \partial_q F(\tilde\gamma(r,s))\cdot\partial_s q(r,s)\\ &= 0.\end{aligned}$$

Since (8) is simply that $F(\tilde\gamma(r,0)) = 0$, we can conclude (12).

Next, for (11), we define

$$A(r,s) := \partial_r x(r,s)\cdot p(r,s) + \partial_r y(r,s)\cdot q(r,s) - \partial_r z(r,s).$$

A direct computation using (10) and (12) yields

$$\begin{aligned}\partial_s A(r,s) &= \partial_r\partial_p F(\tilde\gamma(r,s))\cdot p(r,s) + \partial_r x(r,s)\cdot[-\partial_x F(\tilde\gamma(r,s)) - \partial_z F(\tilde\gamma(r,s))\cdot p(r,s)]\\ &\qquad + \partial_r\partial_q F(\tilde\gamma(r,s))\cdot q(r,s) + \partial_r y(r,s)\cdot[-\partial_y F(\tilde\gamma(r,s)) - \partial_z F(\tilde\gamma(r,s))\cdot q(r,s)]\\ &\qquad - \partial_r[\partial_p F(\tilde\gamma(r,s))\cdot p(r,s) + \partial_q F(\tilde\gamma(r,s))\cdot q(r,s)]\\ &= -\partial_r x(r,s)\cdot\partial_x F(\tilde\gamma(r,s)) + \partial_z F(\tilde\gamma(r,s))\cdot\partial_r x(r,s)\cdot p(r,s)\\ &\qquad - \partial_r y(r,s)\cdot\partial_y F(\tilde\gamma(r,s)) + \partial_z F(\tilde\gamma(r,s))\cdot\partial_r y(r,s)\cdot q(r,s)\\ &\qquad - \partial_p F(\tilde\gamma(r,s))\cdot\partial_r p(r,s) - \partial_q F(\tilde\gamma(r,s))\cdot\partial_r q(r,s)\\ &= -\partial_r[F(\tilde\gamma(r,s))] + \partial_z F(\tilde\gamma(r,s))\cdot A(r,s)\\ &= \partial_z F(\tilde\gamma(r,s))\cdot A(r,s).\end{aligned}$$

Since $A(r,0) = 0$ by (7), then either Gronwall's inequality or solving the above directly yields $A(r,s) = 0$ for all $(r,s)\in\mathcal{R}$, which is the remaining relation (11). □

In particular, Lemma 2 implies that the projected characteristic map γ defines a C^1 change of variables from (r,s) to $(x,y) = (x(r,s), y(r,s))$. Next, we show that γ can be locally inverted:

Lemma 4. *There exists an open rectangle $\mathcal{R}'\subseteq\mathcal{R}$ about $(r_0,0)$ on which $\gamma|_{\mathcal{R}'}$ is one-to-one and $D\gamma|_{\mathcal{R}'}$ is everywhere nonsingular. Furthermore, this local inverse ϕ of $\gamma|_{\mathcal{R}'}$ is C^1.*

Proof. A direct computation using (10) shows that at $(r_0,0)$, we have

$$D\gamma|_{(r_0,0)} = \begin{bmatrix} f'(r_0) & g'(r_0)\\ \partial_p F(f(r_0),g(r_0),h(r_0),p_0,q_0) & \partial_q F(f(r_0),g(r_0),h(r_0),p_0,q_0)\end{bmatrix}$$

Since (4) implies the above is nonsingular, the inverse function theorem yields the desired ϕ. □

We can now construct our solution u by

$$u = z\circ\phi\in C^1(\mathcal{U}),\qquad \mathcal{U} = \gamma(\mathcal{R}'),$$

which by the invertibility in Lemma 4 is equivalent to

$$u(x(r,s),y(r,s)) = z(r,s),\qquad (r,s)\in\mathcal{R}'. \tag{13}$$

By the chain rule, we compute from (13) that

$$\begin{aligned}\partial_r z(r,s) &= \partial_r x(r,s)\cdot\partial_x u(x(r,s),y(r,s)) + \partial_r y(r,s)\cdot\partial_y u(x(r,s),y(r,s)),\\ \partial_s z(r,s) &= \partial_s x(r,s)\cdot\partial_x u(x(r,s),y(r,s)) + \partial_s y(r,s)\cdot\partial_y u(x(r,s),y(r,s)).\end{aligned} \tag{14}$$

Moreover, from (11) and (10), we have

$$\begin{aligned}\partial_r z(r,s) &= \partial_r x(r,s)\cdot p(r,s) + \partial_r y(r,s)\cdot q(r,s),\\ \partial_s z(r,s) &= \partial_s x(r,s)\cdot p(r,s) + \partial_s y(r,s)\cdot q(r,s).\end{aligned} \tag{15}$$

Since $D\gamma$ is invertible on $\mathcal{R}'$, one then concludes that

$$\partial_x u(x(r,s),y(r,s)) = p(r,s),\qquad \partial_y u(x(r,s),y(r,s)) = q(r,s). \tag{16}$$

Note that (16) can be restated as

$$\partial_x u = p \circ \phi, \qquad \partial_y u = q \circ \phi.$$

Since the right-hand sides above are C^1, it follows that $u \in \mathcal{C}^2(\mathcal{U})$.

Finally, we show that u indeed solves (1). The initial condition holds, since by (13),

$$u(f(r), g(r)) = u(\gamma(r, 0)) = z(r, 0) = h(r).$$

In addition, u solves the PDE, since by (12), (13), and (16),

$$F(\gamma(r, s), u(\gamma(r, s)), \nabla u(\gamma(r, s))) = F(\tilde{\gamma}(r, s)) = 0.$$

Since setting $(r, s) = (r_0, 0)$ in (16) yields

$$\partial_x u(f(r_0), g(r_0)) = p_0, \qquad \partial_y u(f(r_0), g(r_0)) = q_0,$$

this completes the proof of existence in Theorem 1.

Proof of Theorem 1: Uniqueness. Suppose $\bar{u}$ is another C^2-solution to (1) on $\mathcal{U}$. Since

$$F(x, y, \bar{u}(x, y), \nabla \bar{u}(x, y)) = 0, \qquad (x, y) \in \mathcal{U},$$

taking partial derivatives of the above yields the following relations:

$$\begin{aligned} 0 &= \partial_x F(x, y, \bar{u}(x, y), \nabla \bar{u}(x, y)) + \partial_z F(x, y, \bar{u}(x, y), \nabla \bar{u}(x, y)) \cdot \partial_x \bar{u}(x, y) \\ &\quad + \partial_p F(x, y, \bar{u}(x, y), \nabla \bar{u}(x, y)) \cdot \partial^2_{xx} \bar{u}(x, y) + \partial_q F(x, y, \bar{u}(x, y), \nabla \bar{u}(x, y)) \cdot \partial^2_{xy} \bar{u}(x, y), \\ 0 &= \partial_y F(x, y, \bar{u}(x, y), \nabla \bar{u}(x, y)) + \partial_z F(x, y, \bar{u}(x, y), \nabla \bar{u}(x, y)) \cdot \partial_y \bar{u}(x, y) \\ &\quad + \partial_p F(x, y, \bar{u}(x, y), \nabla \bar{u}(x, y)) \cdot \partial^2_{yx} \bar{u}(x, y) + \partial_q F(x, y, \bar{u}(x, y), \nabla \bar{u}(x, y)) \cdot \partial^2_{yy} \bar{u}(x, y). \end{aligned} \tag{17}$$

Next, we define the functions $\lambda = (\bar{x}, \bar{y}) : \mathcal{R}' \to \mathbb{R}^2$ via the initial value problem

$$\begin{aligned} \partial_s \bar{x}(r, s) &= \partial_p F(\lambda(r, s), \bar{u}(\lambda(r, s)), \nabla \bar{u}(\lambda(r, s))), \qquad \bar{x}(r, 0) = f(r) = x(r, 0), \\ \partial_s \bar{y}(r, s) &= \partial_q F(\lambda(r, s), \bar{u}(\lambda(r, s)), \nabla \bar{u}(\lambda(r, s))), \qquad \bar{y}(r, 0) = g(r) = y(r, 0). \end{aligned} \tag{18}$$

Indeed, standard ODE theory indicates that $\bar{x}$ and $\bar{y}$ exist, at least locally near $\mathcal{R}' \cap \{s = 0\}$, and are C^1 (since the right-hand sides of (18) are C^1-functions of λ). Given λ, we next define

$$\bar{z}(r, s) = \bar{u}(\lambda(r, s)), \qquad \bar{p}(r, s) = \partial_x \bar{u}(\lambda(r, s)), \qquad \bar{q}(r, s) = \partial_y \bar{u}(\lambda(r, s)). \tag{19}$$

and, for convenience, the shorthands

$$\tilde{\lambda}(r, s) := (\bar{x}(r, s), \bar{y}(r, s), \bar{z}(r, s), \bar{p}(r, s), \bar{q}(r, s)).$$

Lemma 5. *The following relations hold:*

$$\bar{z}(r, 0) = h(r) = z(r, 0), \qquad \bar{p}(r, 0) = w(r) = p(r, 0), \qquad \bar{q}(r, 0) = v(r) = q(r, 0). \tag{20}$$

Proof. The first relation in (20) is an immediate consequence of the assumption that $\bar{u}$ solves (1). For the remaining relations, we note from (1) that on the initial data curve Γ,

$$\begin{aligned} F(\tilde{\lambda}(r, 0)) &= F(f(r), g(r), \bar{u}(f(r), g(r)), \nabla \bar{u}(f(r), g(r))) = 0, \\ f'(r) \cdot \bar{p}(r, 0) + g'(r) \cdot \bar{q}(r, 0) &= d_{(f'(r), g'(r))} \bar{u}(f(r), g(r)) = h'(r). \end{aligned}$$

Since we have assumed

$$\bar{p}(r_0, 0) = \partial_x \bar{u}(f(r_0), g(r_0)) = p_0, \qquad \bar{q}(r_0, 0) = \partial_y \bar{u}(f(r_0), g(r_0)) = q_0,$$

the relations for $\bar{p}$ and $\bar{q}$ in (20) follow from the uniqueness of w, v in Lemma 1. □

Lemma 6. *The following identities hold for any* $(r, s) \in \mathcal{R}'$*:*

$$\begin{aligned} \partial_s \bar{z}(r,s) &= \partial_p F(\tilde{\lambda}(r,s)) \cdot \bar{p}(r,s) + \partial_q F(\tilde{\lambda}(r,s)) \cdot \bar{q}(r,s), \\ \partial_s \bar{p}(r,s) &= -\partial_x F(\tilde{\lambda}(r,s)) - \partial_z F(\tilde{\lambda}(r,s)) \cdot \bar{p}(r,s), \\ \partial_s \bar{q}(r,s) &= -\partial_y F(\tilde{\lambda}(r,s)) - \partial_z F(\tilde{\lambda}(r,s)) \cdot \bar{q}(r,s). \end{aligned} \tag{21}$$

Proof. The first relation in (21) follows immediately from the chain rule and (18):

$$\begin{aligned} \partial_s \bar{z}(r,s) &= \partial_s \bar{x}(r,s) \cdot \partial_x \bar{u}(\lambda(r,s)) + \partial_s \bar{y}(r,s) \cdot \partial_y \bar{u}(\lambda(r,s)) \\ &= \partial_p F(\tilde{\lambda}(r,s)) \cdot \bar{p}(r,s) + \partial_q F(\tilde{\lambda}(r,s)) \cdot \bar{q}(r,s). \end{aligned}$$

Applying similar computations to $\bar{p}$, we see that

$$\begin{aligned} \partial_s \bar{p}(r,s) &= \partial_s \bar{x}(r,s) \cdot \partial_{xx} \bar{u}(\lambda(r,s)) + \partial_s \bar{y}(r,s) \cdot \partial_{yx} \bar{u}(\lambda(r,s)) \\ &= \partial_p F(\tilde{\lambda}(r,s)) \cdot \partial_{xx} \bar{u}(\lambda(r,s)) + \partial_q F(\tilde{\lambda}(r,s)) \cdot \partial_{yx} \bar{u}(\lambda(r,s)). \end{aligned}$$

Recalling the first equation in (17) results in the first relation in (21). The remaining relation for $\bar{q}$ in (21) can be derived using analogous methods. □

Finally, combining (18), (20), and (21), we see that $\tilde{\lambda}$ solves the same initial value problem (10) as $\tilde{\gamma}$. Thus, by standard uniqueness results for ODEs, we see that $\tilde{\lambda} = \tilde{\gamma}$. It then follows that

$$\bar{u}(x,y) = \bar{z}(\phi(x,y)) = z(\phi(x,y)) = u(x,y), \qquad (x,y) \in \mathcal{U},$$

which completes the proof of uniqueness.