THE METHOD OF CHARACTERISTICS FOR SYSTEMS AND APPLICATION TO FULLY NONLINEAR EQUATIONS

ARICK SHAO

While the method of characteristics is used for solving general first-order partial differential equations (with a single real-valued unknown), it fails to generalise to systems of first-order PDEs or to higher-order PDEs. However, here we show that the method of characteristics does extend directly to first-order systems for which all components propagate in the same direction. We then apply this extension to develop the method of characteristics for fully nonlinear first-order PDEs.

Co-Propagating Quasilinear Systems. Let Γ be a curve in $\mathbb{R}^2$,

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

where $I \subseteq \mathbb{R}$ is an open interval, and where the functions $f, g : I \to \mathbb{R}$ parametrise Γ. Consider the following Cauchy problem for a quasilinear *co-propagating* system of PDEs,

$$a(x, y, \vec{u}) \cdot \partial_x \vec{u} + b(x, y, \vec{u}) \cdot \partial_y \vec{u} = \vec{c}(x, y, \vec{u}), \qquad \vec{u}(f(r), g(r)) = \vec{h}(r). \tag{1}$$

where:

- The unknown is a function $\vec{u} : \Omega \to \mathbb{R}^d$, with Ω being some neighbourhood of Γ in $\mathbb{R}^2$.
- The quasilinear system is defined by the coefficients

$$a, b : \mathbb{R}^2_{x,y} \times \mathbb{R}^d_z \to \mathbb{R}, \qquad \vec{c} : \mathbb{R}^2_{x,y} \times \mathbb{R}^d_z \to \mathbb{R}^d.$$

- The initial value for $\vec{u}$ is given by the function $\vec{h} : I \to \mathbb{R}^d$.

The following theorem represents the direct analogue for the system (1) of the standard method of characteristics for a single quasilinear PDE:

Theorem 1. *Suppose $f, g, \vec{h}, a, b, \vec{c} \in C^1$, and suppose the noncharacteristic condition*

$$\det \begin{bmatrix} f'(r_0) & g'(r_0) \\ a(f(r_0), g(r_0), \vec{h}(r_0)) & b(f(r_0), g(r_0), \vec{h}(r_0)) \end{bmatrix} \neq 0 \tag{2}$$

holds for some $r_0 \in I_0$. Then, there is a neighbourhood $\mathcal{U} \subseteq \mathbb{R}^2$ of $(f(r_0), g(r_0)) \in \Gamma$ such that the Cauchy problem for the system (1) has a unique C^1-solution $\vec{u} : \mathcal{U} \to \mathbb{R}^d$.

Remark. *While the existence in Theorem 1 is over a neighbourhood of a point on Γ, one can apply Theorem 1 to every $r \in I$ and then patch the resulting solutions together via uniqueness to obtain a larger solution on a neighbourhood Ω of all of Γ.*

Proof of Theorem 1. Consider the following *characteristic equations*:

$$\begin{aligned} \partial_s x(r, s) &= a(x(r, s), y(r, s), \vec{z}(r, s)), & x(r, 0) &= f(r), \\ \partial_s y(r, s) &= b(x(r, s), y(r, s), \vec{z}(r, s)), & y(r, 0) &= g(r), \\ \partial_s \vec{z}(r, s) &= \vec{c}(x(r, s), y(r, s), \vec{z}(r, s)), & \vec{z}(r, 0) &= \vec{h}(r). \end{aligned} \tag{3}$$

Here, s represents the parameter along the characteristic curves, while r parametrises the characteristic curves corresponding to its starting point on Γ. For convenience, we also define

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

Since the right-hand sides of (3) are C^1-functions of $\tilde{\gamma}(r,s)$, and since $f, g, \vec{h}$ are at least C^1, then the standard ODE theory implies that there exists a neighbourhood $\mathcal{D} \subseteq \mathbb{R}^2$ of $(r_0, 0)$ on which one has a unique C^1-solution $\tilde{\gamma} : \mathcal{D} \to \mathbb{R}^2 \times \mathbb{R}^d$ of (3). (Here, $\tilde{\gamma}$ is C^1 in both r and s.)

The projected characteristics γ define a C^1-change of variables from (r, s) to $(x, y) = \gamma(r, s)$. Moreover, a direct computation using (3) shows that at $(r_0, 0)$,

$$D\gamma|_{(r_0,0)} = \begin{bmatrix} f'(r_0) & g'(r_0) \\ a(f(r_0), g(r_0), \vec{h}(r_0)) & b(f(r_0), g(r_0), \vec{h}(r_0)) \end{bmatrix},$$

which is nonsingular by (2). By the inverse function theorem, there is some open rectangle

$$\mathcal{R} := J \times (-s_0, s_0), \qquad r_0 \in J, \quad s_0 > 0,$$

for which $\gamma|\mathcal{R}$ has a C^1-inverse ϕ.

We can now construct our solution $\vec{u}$ by

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

which can be equivalently stated as

$$\vec{u}(x(r,s), y(r,s)) = \vec{z}(r,s), \qquad (r,s) \in \mathcal{R}. \tag{4}$$

The initial condition in (1) holds for $\vec{u}$, since by (4),

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

In addition, $\vec{u}$ satisfies the PDE in (1), since for $(r, s) \in \mathcal{R}$,

$$\begin{aligned} \vec{c}(\gamma(r,s), \vec{u}(\gamma(r,s))) &= \vec{c}(\gamma(r,s), \vec{z}(r,s)) \\ &= \partial_s \vec{z}(r,s) \\ &= \partial_s[\vec{u}(\gamma(r,s))] \\ &= \partial_x \vec{u}(\gamma(r,s)) \cdot \partial_s x(r,s) + \partial_y \vec{u}(\gamma(r,s)) \cdot \partial_s y(r,s) \\ &= a(\gamma(r,s), \vec{u}(r,s)) \cdot \partial_x \vec{u}(\gamma(r,s)) + b(\gamma(r,s), \vec{u}(r,s)) \cdot \partial_y \vec{u}(\gamma(r,s)), \end{aligned}$$

where we applied (3) and (4). This completes the proof of existence.

Suppose now that $\vec{v}$ is another C^1-solution to (1) on $\mathcal{U}$. Consider the initial value problem,

$$\begin{aligned} \partial_s \bar{x}(r,s) &= a(\bar{x}(r,s), \bar{y}(r,s), \vec{v}(x(r,s), y(r,s))), \qquad \bar{x}(r,0) = f(r) = x(r,0), \\ \partial_s \bar{y}(r,s) &= b(\bar{x}(r,s), \bar{y}(r,s), \vec{v}(x(r,s), y(r,s))), \qquad \bar{y}(r,0) = g(r) = y(r,0), \end{aligned} \tag{5}$$

which has a unique C^1-solution locally near $\mathcal{R} \cap \{s = 0\}$. Given $(\bar{x}, \bar{y})$, we next define

$$\bar{z}(r,s) = \vec{v}(\bar{x}(r,s), \bar{y}(r,s)). \tag{6}$$

Using (5) and that $\vec{v}$ satisfies (1), we see that

$$\begin{aligned} \bar{z}(r,0) &= \vec{v}(f(r), g(r)) = \vec{h}(r) = z(r,0), \\ \partial_s \bar{z}(r,s) &= \partial_x \vec{v}(\bar{x}(r,s), \bar{y}(r,s)) \cdot \partial_s \bar{x}(r,s) + \partial_y \vec{v}(\bar{x}(r,s), \bar{y}(r,s)) \cdot \partial_s \bar{y}(r,s) \\ &= \partial_x \vec{v}(\bar{x}(r,s), \bar{y}(r,s)) \cdot a(\bar{x}(r,s), \bar{y}(r,s), \vec{v}(x(r,s), y(r,s))) \\ &\quad + \partial_y \vec{v}(\bar{x}(r,s), \bar{y}(r,s)) \cdot b(\bar{x}(r,s), \bar{y}(r,s), \vec{v}(x(r,s), y(r,s))) \\ &= \vec{c}(\bar{x}(r,s), \bar{y}(r,s), \vec{v}(x(r,s), y(r,s))). \end{aligned} \tag{7}$$

Combining (5)-(7), we see that $(\bar{x}, \bar{y}, \bar{z})$ satisfies the initial value problem (3). By the uniqueness result for ODEs, $(\bar{x}, \bar{y}, \bar{z})$ must be identical to $(x, y, \vec{z})$ from the proof of existence. Thus,

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

which completes the proof of uniqueness. □

Fully Nonlinear PDE. We now turn our attention to the fully nonlinear PDE

$$F(x, y, u, \partial_x u, \partial_y u) = 0, \qquad u(f(r), g(r)) = h(r). \tag{8}$$

Here, the objects of interest are defined as follows:

- The initial data curve Γ is defined as before
$$\Gamma := \{(f(r), g(r)) \mid r \in I\}, \qquad f, g : I \to \mathbb{R},$$
where $I \subseteq \mathbb{R}$ is once again an open interval.
- The unknown is a function $u : \mathcal{U} \to \mathbb{R}$, with $\mathcal{U} \subseteq \mathbb{R}^2$ a neighbourhood of some point on Γ.
- The function F defining the PDE has the usual form
$$F : \mathbb{R}^2_{x,y} \times \mathbb{R}_z \times \mathbb{R}^2_{p,q} \to \mathbb{R}.$$
- The initial value for u is given by the function $h : I \to \mathbb{R}$.

The idea now is to reduce (8) to a system of quasilinear PDE of the form (1). This can be done via a standard trick in PDEs: even if (8) itself has no particular structure, by formally differentiating (8), we obtain quasilinear structure at the level of one derivative higher.

To see this, suppose u is a solution of (8), and consider the function

$$(x, y) \mapsto F(x, y, u(x, y), \partial_x u(x, y), \partial_y u(x, y)).$$

Taking partial derivatives of the above and abbreviating

$$\Phi(x, y) := (x, y, u(x, y), \partial_x u(x, y), \partial_y u(x, y))$$

yields the relations:

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

In particular, rearranging, we see that $\partial_x u$ and $\partial_y u$ satisfy a co-propagating system:

$$\begin{aligned} &\partial_p F(x, y, u, \partial_x u, \partial_y u) \cdot \partial_x(\partial_x u) + \partial_q F(x, y, u, \partial_x u, \partial_y u) \cdot \partial_y(\partial_x u) \\ &\quad = -\partial_x F(x, y, u, \partial_x u, \partial_y u) - \partial_z F(x, y, u, \partial_x u, \partial_y u) \cdot \partial_x u, \\ &\partial_p F(x, y, u, \partial_x u, \partial_y u) \cdot \partial_x(\partial_y u) + \partial_q F(x, y, u, \partial_x u, \partial_y u) \cdot \partial_y(\partial_y u) \\ &\quad = -\partial_y F(x, y, u, \partial_x u, \partial_y u) - \partial_z F(x, y, u, \partial_x u, \partial_y u) \cdot \partial_y u. \end{aligned} \tag{9}$$

The main point is to consider $\partial_x u$ and $\partial_y u$ as additional unknowns and to solve the resulting system. To be more precise, we define the vector-valued unknown $\vec{u} := (u, u_x, u_y)$, with the last two components representing "$\partial_x u$" and "$\partial_y u$". Combining (9) with the trivial relation

$$\begin{aligned} &\partial_p F(x, y, u, u_x, u_y) \cdot \partial_x u + \partial_q F(x, y, u, u_x, u_y) \cdot \partial_y u \\ &\quad = \partial_p F(x, y, u, u_x, u_y) \cdot u_x + \partial_q F(x, y, u, u_x, u_y) \cdot u_y, \end{aligned}$$

we obtain the system:

$$a(x, y, \vec{u}) \cdot \partial_x \vec{u} + b(x, y, \vec{u}) \cdot \partial_y \vec{u} = \vec{c}(x, y, \vec{u}), \qquad \vec{u}(f(r), g(r)) = (h(r), w(r), v(r)), \tag{10}$$

where:

- The unknown $\vec{u} = (u, u_x, u_y)$ is a $\mathbb{R}^3$-valued function on some neighbourhood Ω of Γ.
- The initial data are given by functions $h, w, v : I \to \mathbb{R}$.

- The coefficients of (10) are given by:

$$a(x,y,\vec{u}) := \partial_p F(x,y,u,u_x,u_y), \tag{11}$$
$$b(x,y,\vec{u}) := \partial_q F(x,y,u,u_x,u_y),$$
$$\vec{c}(x,y,\vec{u}) := \begin{bmatrix} \partial_p F(x,y,u,u_x,u_y) \cdot u_x + \partial_q F(x,y,u,u_x,u_y) \cdot u_y \\ -\partial_x F(x,y,u,u_x,u_y) - \partial_z F(x,y,u,u_x,u_y) \cdot u_x \\ -\partial_y F(x,y,u,u_x,u_y) - \partial_z F(x,y,u,u_x,u_y) \cdot u_y \end{bmatrix},$$

The derivations we have done can now be interpreted as follows:

Proposition 1. *Let $f, g, h \in C^2(I)$, and let $w, v \in C^1(I)$. Let Ω be a neighbourhood of Γ, and suppose $u \in C^2(\Omega)$ is a solution of (8) that also satisfies*

$$\partial_x u(f(r), g(r)) = w(r), \qquad \partial_y u(f(r), g(r)) = v(r), \qquad r \in I.$$

Then, $\vec{u} := (u, \partial_x u, \partial_y u)$ is a C^1-solution of (10), (11) on Ω.

The bigger, and more relevant, challenge is to prove the converse: that a solution of (10), (11) yields a solution of (8). This is not entirely straightforward for the following reasons:

(1) In solving (10), we are solving the *derivative* of (8) and not (8). Thus, to show (8), we will need some sort of additional constraint on the initial conditions. This can be captured by imposing that (8) is initially satisfied on Γ.
(2) In addition, from the point of view of (10), there is no reason a priori that u_x should correspond to $\partial_x u$, and similarly for u_y and $\partial_y u$. In fact, this will also have to be imposed (indirectly) as a constraint on Γ and then propagated to the domain of u.

The precise statement of this is given below:

Proposition 2. *Let $f, g, h \in C^2(I)$, and let $w, v \in C^1(I)$. Also, let $\mathcal{U} \subseteq \mathbb{R}^2$ be a neighbourhood of $x_0 \in \Gamma$, and suppose $\vec{u} := (u, u_x, u_y)$ is the C^1-solution on $\mathcal{U}$ of (10) obtained via Theorem 1. Furthermore, suppose that for each $r \in I$ with $(f(r), g(r)) \in \Gamma \cap \mathcal{U}$, the following constraints hold:*

$$f'(r) \cdot w(r) + g'(r) \cdot v(r) = h'(r), \tag{12}$$
$$F(f(r), g(r), h(r), w(r), v(r)) = 0.$$

Then, $u \in C^2(\mathcal{U})$, and u solves (8), with the additional initial conditions

$$\partial_x u(f(r), g(r)) = w(r), \qquad \partial_y u(f(r), g(r)) = v(r), \qquad (f(r), g(r)) \in \Gamma \cap \mathcal{U}.$$

Remark. *Note that from the perspective of (8), one only imposes initial data for u, but not for ∇u. Indeed, one cannot impose data for ∇u freely, as its values on Γ are restricted precisely by the constraints (12). However, if one can find some $w(r_0)$ and $v(r_0)$ at a single point of Γ so that (12) holds, then one can generate appropriate $w(r)$ and $v(r)$ locally via the implicit function theorem; see Lemma 1 in the other notes [1] on the website.*

Proof sketch of Proposition 2. Let $(x, y, \vec{z}) := (x, y, z, p, q)$ denote the solution to the characteristic equations (3) corresponding to the problem (10). In particular, these equations are given by

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

where $\tilde{\gamma} := (x, y, z, p, q)$. Note these are the usual characteristic equations for fully nonlinear equations; see [1, Eq. (10)]. Recall also from the proof of Theorem 1 that

$$z(r,s) = u(x(r,s), y(r,s)), \qquad p(r,s) = u_x(x(r,s), y(r,s)), \qquad q(r,s) = u_y(x(r,s), y(r,s)). \tag{13}$$

To show that u solves (8), we must show that

$$\partial_x u = u_x, \qquad \partial_y u = u_y, \tag{14}$$

and we must show that u satisfies the PDE,

$$F(x, y, u(x, y), u_x(x, y), u_y(x, y)) = 0, \qquad (x, y) \in \mathcal{U}. \tag{15}$$

We omit the details of these derivations. However, we mention that the main step in this process is to obtain the following propagation of constraint conditions:

$$\begin{aligned} \partial_r x(r, s) \cdot p(r, s) + \partial_r y(r, s) \cdot q(r, s) &= \partial_r z(r, s), \\ F(x(r, s), y(r, s), z(r, s), p(r, s), q(r, s)) &= 0. \end{aligned} \tag{16}$$

The proof of this is essentially the same as that of [1, Lemma 3].

Note that (15) follows from the second relation in (16). For (14), the reader is referred to the proof of [1, Eq. (16)] (the hardest part of this is the first relation in (16)).

References

1. A. Shao, *The Cauchy problem via the method of characteristics*, `http://wwwf.imperial.ac.uk/~cshao/pde/char_fnl.pdf`.