MATH517: PDE SOME NOTES ON CAUCHY-KOWALEVSKAYA THEOTEM

Cauchy-Kowalevskaya theorem gives local existence of analytic solutions to a non-characteristic Cauchy problem of a partial differential equation (or system) that is analytic in its arguments. Its predecessor is Cauchy' theorem on the local existence of analytic solutions to the initial value problem for ordinary differential equations

$$\frac{du}{dt} = f(u(t), t), \quad u(0) = u_0,$$

when $f(u, t)$ is assumed to be analytic in a neighborhood of $(u_0, 0)$. A natural approach is to look for an analytic solution of the form $u(t) = \sum_{j=1}^{\infty} a_j t^j$ and determine the coefficients a_j through the initial condition and repeatedly differentiating the equation. Cauchy was able to show the convergence of the obtained series through his *method of majorants*. This theorem was extended by Cauchy, and later by Kowalevskaya, to the initial value problem for partial differential equations for the form:

$$\frac{\partial u(x,t)}{\partial t} = f(\partial_x u(x,t), u(x,t), x, t), \quad u(x,0) = g(x), \tag{1}$$

for (x, t) near $(x_0, 0)$, where $\partial_x u(x, t)$ stands for the gradient vector of $u(x, t)$ in the x-variables, and $f(p, u, x, t)$ is analytic in (p, u, x, t) near $(\partial_x g(x_0), g(x_0), x_0, 0)$. For the initial value problem for higher order partial differential equations, Cauchy discussed a procedure to reduce the problem to a (larger) system of first order partial differential equations of the form above. Kowalevskaya clarified the type of equations for which the method initiated by Cauchy would work (Kowalevskaya in fact was not aware of Cauchy's work). Kowalevskaya pointed out that although a formal power series solution can be determined for the initial value problem of the heat equation

$$\partial_t u(x,t) = \partial_x^2 u(x,t), \quad u(x,0) = g(x),$$

the power series does not need to converge. In fact, for $g(x) = \frac{1}{1-x}$, the formal power series for a solution $u(x, t)$ near $(0, 0)$ would be

$$\sum_{j=0}^{\infty} \frac{(2j)!}{j!(1-x)^{2j+1}} t^j,$$

which is not convergent for any $t \neq 0$! Thus Cauchy's theorem is not valid if one tries to allow terms of the kind $\partial_x^\alpha u(x, t)$ for $|\alpha| > 1$ in the right hand side of (1).

In the following we will formulate several versions of the Cauchy-Kowalevskaya's theorem so that we can conclude the existence of local analytic solutions without having to go through the reduction process to check whether the given problem can be reduced to one of the form (1). We will explain Cauchy-Kowalevskaya's theorem first in the context of the initial value problem for linear partial differential equations and the initial value is prescribed with respect to a distinguished variable t; then we will discuss how to formulate the Cauchy-Kowalevskaya's theorem when the initial surface is a general *non-characteristic* surface; and finally we describe the theorem for nonlinear partial differential equations.

Let's first examine the case of a linear differential operator in the form

$$P = \partial_t^m + \sum_{j<m,\, j+|\alpha|\le m} c_{j,\alpha}(x,t)\partial_t^j \partial_x^\alpha,$$

where the coefficients $c_{j,\alpha}(x,t)$ are analytic around a point $(x_0, 0)$ on the initial surface $t = 0$. We seek to solve

$$\begin{cases} Pu = f(x,t), \quad \text{near } (x_0,0), \\ \partial_t^j u(x,0) = g_j(x),\ j = 0, \cdots, m-1, \quad \text{near } x_0, \end{cases} \tag{2}$$

where $f(x,t)$, and $g_j(x)$ are anlytic functions around $(x_0, 0)$ and x_0 respectively.

Theorem 1 (Linear case with special non-characteristic initial surface)**.** *Suppose $c_{j,\alpha}(x,t)$ are analytic in a neighborhood V around $(x_0, 0)$. Then there is a neighborhood $U \subset V$ of $(x_0, 0)$, such that for any $f(x,t)$ analytic in U_1 around $(x_0, 0)$, and any $g_j(x)$ analytic in W around x_0, there is a unique analytic solution to* (2) *in $U \cap U_1 \cap (W \times \mathbb{R})$.*

Remark 1. *This existence of analytic solution does not imply the wellposedness of the Cauchy problem in the usual sense. For example, the above theorem applies to both $P_1 = \partial_t^2 - \partial_x^2$ and $P_2 = \partial_t^2 + \partial_x^2$ with $\{t = 0\}$ as initial surface, yet the Cauchy problem (with respect to t) is wellposed for P_1, but not for P_2.*

Remark 2. *There is no general local existence result when the analyticity assumptions are dropped. In 1956 H. Lewy constructed the first example of a linear differential equation that has no solution* anywhere.

The first ingredient of Cauchy's method is the (formal) dertermination of the Taylor series of a possible analytic solution $u(x,t)$ around $(x_0, 0)$. This is relatively straight forward in this set up: first note that any derivatives of order $m-1$ or less along $t = 0$ can be determined by the initial data alone: $\partial_x^\alpha u(x,0) = \partial_x^\alpha g_0(x)$ (in fact for all α), and for any $1 \le j < m$,

$$\partial_t^j \partial_x^\alpha u(x,0) = \partial_x^\alpha g_j(x).$$

Also

$$\begin{aligned} &\partial_t^m \partial_x^\beta u(x,0) = \partial_x^\beta \partial_t^m u(x,0) \\ =&\partial_x^\beta \left(- \sum_{j<m,\, j+|\alpha|\le m} c_{j,\alpha}(x,0)\partial_t^j \partial_x^\alpha u(x,0) + f(x,0) \right) \\ =&\partial_x^\beta \left(- \sum_{j<m,\, j+|\alpha|\le m} c_{j,\alpha}(x,0)\partial_x^\alpha g_j(x) + f(x,0) \right). \end{aligned}$$

Next we simply differentiate the equation and initial conditions inductively to represent all $\partial_t^j \partial_x^\alpha u(x_0, 0)$ in terms of given Cauchy data, the right hand side, and lower order derivatives.

The second ingredient is to prove the convergence of the constructed power series. This was done by Cauchy and Kowaleveskaya by the so called majorant method for power series.

Here we will not provide a full proof as given by Cauchy-Kowaleveskaya. Instead, we will describe an often used reduction procedure. In order to make a proof manageble, one typically first reduces the problem to an equivalent Cauchy problem of a system of first order differential equations. This is done by introducing new variables and use the compatability conditions as new equations: set $U = (u, \partial_t^j \partial_x^\alpha u | j + |\alpha| < m)$. Then for any $j + |\alpha| < m - 1$,

$$\partial_t U_{j,\alpha} = \partial_t^{j+1} \partial_x^\alpha u = U_{j+1,\alpha},$$

and for any $j + |\alpha| = m - 1$ with $\alpha \neq 0$,

$$\partial_t U_{j,\alpha} = \partial_t^{1+j} \partial_x^\alpha u = \partial_{x_{\alpha_1}} U_{1+j,\alpha-\alpha_1},$$

where α_1 is the first component of α that is not zero. Finally

$$\partial_t U_{m-1,0} = \partial_t^m u = - \sum_{j<m,\, j+|\alpha|\leq m} c_{j,\alpha}(x,t) \partial_t^j \partial_x^\alpha u(x,t) + f(x,t).$$

where terms with $j + |\alpha| \leq m - 1$ are linear combinations of $U_{j,\alpha}$, and terms with $j + |\alpha| = m$ and $j < m$ can be written as linear combinations of $\partial_{x_{\alpha_1}} U_{j,\alpha-\alpha_1}$. So we are led to studying a system of the form

$$\begin{cases} \partial_t u_i = \sum_{j=1}^{N} \sum_{\alpha=1}^{n} c_{ij}^\alpha(x,t) \partial_\alpha u_j(x,t) + \sum_{j=1}^{N} d_{ij}(x,t) u_j(x,t) + f_i(x,t), & \text{for } i = 0, \cdots, N \\ u_i(x,0) = g_i(x), & \text{for } i = 0, \cdots, N, \end{cases}$$

or using vector notation $u = (u_0, u_1, \cdots, u_N)$,

$$\begin{cases} \partial_t u = \sum_{\alpha=1}^{n} C^\alpha(x,t) \partial_\alpha u(x,t) + D(x,t) u(x,t) + F(x,t), \\ u(x,0) = G(x). \end{cases} \tag{3}$$

Next we formulate the notion of a non-characteristic initial surface with respect to a linear differential operator $P = \sum_{|\alpha|\leq m} c_\alpha(x) \partial_x^\alpha$ and formulate the Cauchy-Kowalevskaya theorem with respect to such an initial surface.

We represent an initial surface Σ as the level set of a defining funtions σ: $\Sigma = \{x : \sigma(x) = 0\}$, where $\nabla_x \sigma(x) \neq 0$ along Σ. One idea is to make a change of variables, locally, to "flatten" Σ. For instance, if $x_0 \in \Sigma$ is such that $\partial_{x_n} \sigma(x_0) \neq 0$, then in a neighborhood of x_0, Σ can be represented as a graph x_n in terms of $x_1, \cdots, x_{n-1}$. In fact, $x = (x', x_n) \mapsto (y', \tau)$, with $y' = x'$, and $\tau = \sigma(x)$ is a local diffeomorphism. If we adopt (y', τ) as new coordinates and set $v(y', \tau) = u(x', x_n)$, then Pu is expressed as a linear differential operator $\tilde{P} v$ of the same order m, and the coefficient of ∂_τ^m is given by $\sum_{|\alpha|=m} c_\alpha(x)(\nabla\sigma)^\alpha$. This is from the chain rule

$$\text{for } 1 \le j \le m-1, \partial_{x_j} = \partial_{y_j} + \sigma_{x_j}\partial_\tau, \quad \partial_{x_n} = \sigma_{x_n}\partial_\tau.$$

So

$\partial_x^\alpha u = (\nabla\sigma)^\alpha \partial_\tau^m v+$terms of order not higher than m and with $m-1$ or fewer derivatives in τ.

Definition. $\Sigma = \{x : \sigma(x) = 0\}$ *is called non-characteristic with respect to* P *at* $x_0 \in \Sigma$ *if* $\sum_{|\alpha|=m} c_\alpha(x_0)(\nabla\sigma(x_0))^\alpha \neq 0$. $\Sigma = \{x : \sigma(x) = 0\}$ *is called non-characteristic with respect to* P *if it is non-characteristic at every point on it. The level surfaces of* σ *are called characteristic with respect to* P *if* σ *is a non-trivial solution (*$\nabla\sigma \neq 0$*) to* $\sum_{|\alpha|=m} c_\alpha(x)(\nabla\sigma(x))^\alpha = 0$.

Definition. *A vector* $\xi \neq 0 \in \mathbb{R}^n$ *is called a characteristic direction for* P *at* x_0 *if*

$$\sum_{|\alpha|=m} c_\alpha(x_0)\xi^\alpha = 0.$$

P *is called elliptic at* x_0 *if it has no characteristic direction at* x_0, i.e., *for any* $\xi \in \mathbb{R}^n$,

$$\sum_{|\alpha|=m} c_\alpha(x_0)\xi^\alpha = 0 \Longrightarrow \xi = 0.$$

P *is called elliptic in a region if it is elliptic at every point in this region.*

Remark 3. *When dealing with a system, the coefficients* c_α *are interpreted as matrices, a vector* $\xi \neq 0 \in \mathbb{R}^n$ *is called a characteristic direction for* P *at* x_0 *if the matrix*

$$\sum_{|\alpha|=m} c_\alpha(x_0)\xi^\alpha$$

is singular. Thus level surfaces of σ *are called characteristic with respect to* P *if* σ *is a non-trivial solution (*$\nabla\sigma \neq 0$*) to*

$$\det\left(\sum_{|\alpha|=m} c_\alpha(x)(\nabla\sigma(x))^\alpha\right) = 0.$$

Example 1. *For a second order linear partial differential operator*

$$P = \sum_{i,j=0}^n a_{ij}(x)\partial^2_{x_i x_j} + \sum_{i=0}^n b_i(x)\partial_{x_i} + c(x),$$

level surfaces of σ *are characteristic with respect to* P *if* σ *is a non-trivial solution (*$\nabla\sigma \neq 0$*) to*

$$\sum_{i,j=0}^n a_{ij}(x)\partial_{x_i}\sigma(x)\partial_{x_j}\sigma(x) = 0.$$

For $P = \partial_t^2 - \Delta_x$*, this equation becomes* $|\partial_t\sigma(t,x)|^2 - |\nabla_x\sigma(t,x)|^2 = 0$. *If we further assume that* $\sigma(t,x)$ *has the form* $t - \phi(x)$*, then* $\phi(x)$ *must satisfy* $|\nabla_x\phi(x)| = 1$. *At the same time the surface* $t = \phi(x)$ *is non-characteristic, if* $|\nabla_x\phi(x)| \neq 1$ *on* $\{(t,x) : t = \phi(x)\}$.

Example 2. *For a first order linear partial differential operator*

$$P = a_0(t,x)\partial_t + \sum_{i=1}^{n} a_i(t,x)\partial_{x_i},$$

the initial surface $\sigma(t,x) = 0$ is characteristic if

$$a_0(t,x)\partial_t\sigma(t,x) + \sum_{i=1}^{n} a_i(t,x)\partial_{x_i}\sigma(t,x) = 0, \quad \text{for } (t,x) \text{ on } \sigma(t,x) = 0,$$

which means geometrically, when $\sigma(t,x) = 0$ is a non-degenerate surface, that the vector field

$$(t,x) \mapsto (a_0(t,x), a_1(t,x), \cdots, a_n(t,x))$$

is tangential to the surface $\sigma(t,x) = 0$, when (t,x) is on $\sigma(t,x) = 0$.

In the case $n = 1$, the initial surface is simply a curve. It is characteristic iff it is an integral curve of the vector field $(t,x) \mapsto (a_0(t,x), a_1(t,x))$.

Example 3. *For the Maxwell system of electro-magnetism,*

$$\begin{cases} \partial_t \vec{E} = \nabla \times \vec{B}, \\ \partial_t \vec{B} = -\nabla \times \vec{E}, \end{cases}$$

the equation for a characteristic surface $\sigma(t,x) =$ const. can be found by following the procedure described above. Alternatively, if we recall that the objective is to check that, in flattening the surface, $(t,x) \mapsto (\tau = \sigma(t,x), x)$, whether we can solve the ∂_τ terms from the equations, and that the ∂_τ term in $\partial_t\vec{E}$ is $\partial_t\sigma(t,x)\partial_\tau\vec{E}$, and the ∂_τ term in $\nabla \times \vec{B}$ is $\nabla_x\sigma(t,x) \times \partial_\tau\vec{B}$, etc., we would be working with

$$\begin{cases} \partial_t\sigma(t,x)\partial_\tau\vec{E} = \nabla_x\sigma(t,x) \times \partial_\tau\vec{B} + \text{tangential derivatives of } \vec{B}, \\ \partial_t\sigma(t,x)\partial_\tau\vec{B} = -\nabla_x\sigma(t,x) \times \partial_\tau\vec{E} + \text{tangential derivatives of } \vec{E}. \end{cases}$$

From this system, we find

$$|\partial_t\sigma(t,x)|^2\partial_\tau\vec{E} = \nabla_x\sigma(t,x) \times \left(-\nabla_x\sigma(t,x) \times \partial_\tau\vec{E}\right) + \text{tangential derivatives of } \vec{E} \text{ and } \vec{B}.$$

Using

$$\nabla_x\sigma(t,x) \times \left(\nabla_x\sigma(t,x) \times \partial_\tau\vec{E}\right) = \left(\nabla_x\sigma(t,x) \cdot \partial_\tau\vec{E}\right)\nabla_x\sigma(t,x) - |\nabla_x\sigma(t,x)|^2\partial_\tau\vec{E}$$

we arrive at

$$\left(|\partial_t\sigma(t,x)|^2 - |\nabla_x\sigma(t,x)|^2\right)\partial_\tau\vec{E} + \left(\nabla_x\sigma(t,x) \cdot \partial_\tau\vec{E}\right)\nabla_x\sigma(t,x) = \text{tangential derivatives of } \vec{E} \text{ and } \vec{B}.$$

If $\nabla_x\sigma(t,x) \neq 0$, then let $\nu(t,x)$ denote the unit vector in the direction of $\nabla_x\sigma(t,x)$, and decompose $\partial_\tau\vec{E}$ as the vector sum of its component in the direction of $\nu(t,x)$ and an orthogonal component:

$$\partial_\tau\vec{E} = \left(\nu(t,x) \cdot \partial_\tau\vec{E}\right)\nu(t,x) + \left(\partial_\tau\vec{E}\right)^\perp,$$

the left hand side of the above equation becomes

$$\left(|\partial_t\sigma(t,x)|^2 - |\nabla_x\sigma(t,x)|^2\right)\left(\partial_\tau\vec{E}\right)^{\perp} + |\partial_t\sigma(t,x)|^2\left(\nu(t,x)\cdot\partial_\tau\vec{E}\right)\nu(t,x).$$

Thus $\sigma(t,x)=\tau$ *is non-characteristic iff* $\partial_t\sigma(t,x)\neq 0$ *and* $|\partial_t\sigma(t,x)|^2 - |\nabla_x\sigma(t,x)|^2 \neq 0$.

In setting up an initial value problem along a general surface Σ, one can prescribe the initial data in two ways. One way is to prescribe $u(x)$ restricted to Σ as $g_0(x)$, and normal derivatives of $u(x)$ along Σ up to order $m-1$:

$$\frac{\partial^j u(x)}{\partial\nu^j} = g_j(x), \qquad \text{for } x\in\Sigma \text{ and } j=1,\ldots,m-1.$$

Given $u(x)=g_0(x)$ for $x\in\Sigma$, tangential derivatives along Σ of $u(x)$ of any order can be computed through those of $g_0(x)$. Together with the prescribed normal derivatives of $u(x)$ along Σ up to order $m-1$, one can determine all partial derivatives of $u(x)$ of order $m-1$ or lower along Σ. These partial derivatives satisfy the compatibility conditions of mixed derivatives along Σ. Σ is non-characteristic for P iff all partial derivatives of $u(x)$ of order m (and therefore higher) restricted to Σ can be determined through the equation and the initial data. Another way to prescribe the initial data is to prescribe all partial derivatives of $u(x)$ of order $m-1$ or lower along Σ, subject to the natural compatibility conditions. An easy way to do this is to give a function $g(x)$, C^m or analytic near Σ, such that $\partial_\beta u = \partial_\beta g(x)$, along Σ for all $|\beta|\leq m-1$.

Theorem 2 (Linear case with general non-characteristic initial surface)**.** *Suppose there exists a neighborhood* V *of* $x_0\in\Sigma$ *such that* $c_\alpha(x)$ *are analytic in* V*, that* Σ *is analytic and non-characteristic with respect to* P *in* V*. Then there is a neighborhood* U *of* x_0*, such that for any analytic functions* $f(x)$ *and* $g(x)$ *in a neighbood* U_1 *around* x_0*, there is a unique analytic solution in* $U\cap U_1$ *to*

$$\begin{cases} Pu = f(x), & \textit{near } x_0,\\ \partial_\beta u = \partial_\beta g(x), & \textit{along } \Sigma \textit{ for all } |\beta|\leq m-1.\end{cases}$$

Remark 4. *The notion of non-characteristic initial surface implies the following consequence: Suppose* P *is an* m*-th order linear differential operator, and* u*,* v *are two* C^m *functions in a neighborhood* U *such that* $Pu\equiv Pv$ *in* U*. If* Σ *is a hypersurface which is non-characteristic with respect to* P *and* $\partial^\alpha u = \partial^\alpha v$ *in* $U\cap\Sigma$*, for all* $|\alpha|\leq m-1$*, then* $\partial^\alpha u=\partial^\alpha v$ *in* $U\cap\Sigma$*, for all* $|\alpha|=m$*; and if the* u *and* v *here were assumed to be* C^∞ *to begin with, then* $\partial^\alpha u=\partial^\alpha v$ *in* $U\cap\Sigma$*, for all* α*.*

Remark 5. *When the initial surface is characteristic, one may still be able to determine the formal power series expansion of a potential solution from the equation and initial data, the convergence of this constructed series is not gauranteed, as pointed out earlier for the heat operator* $\partial_t-\Delta$ *with the characteristic surfact* $\{t=0\}$ *as initial surface—a surface of the form* $\sigma(t,x)=0$ *is characterisc with respect to the heat operator* $\partial_t-\Delta$ *iff* $|\nabla_x\sigma(t,x)|^2=0$ *on it.*

Example 4. *For* $P_1 = \partial_t + i\partial_x$ *in* $\mathbb{R}^2$*, a characteristic direction* $\xi = (\zeta, \eta) \in \mathbb{R}^2$ *would have to satisfy* $\zeta + i\eta = 0$*, which would force* $\zeta = \eta = 0$*. Thus any regular curve in* $\mathbb{R}^2$ *in non-characteristic for* P_1*, and for any such analytic curve* γ *and analytic initial data* g *along* γ*, Cauchy-Kowalevskaya theorem can be applied. For* $P_2 = \partial_t + \partial_x$ *in* $\mathbb{R}^2$*, a characteristic direction* $\xi = (\zeta, \eta) \in \mathbb{R}^2$ *would have to satisfy* $\zeta + \eta = 0$*. Thus a regular curve of the form* $t - \phi(x) = 0$ *is non-characteristic for* P_2 *iff* $1 - \phi'(x) \neq 0$*; while* $\{(t, x) : t = x\}$ *would be a characteristic curve for* P_2*. One can not apply the Cauchy-Kowalevskaya theorem to*

$$\begin{cases} (\partial_t + \partial_x)\, u(t, x) = 0, \\ \qquad\qquad u(x, x) = g(x) \end{cases}$$

For one thing, if a solution existed, it would have forced g *to be a constant; secondly, no derivative of* u *in the direction transversal to the initial curve* $\{(t, x) : t = x\}$ *can be determined from the equation and the initial condition.*

Example 5. *The initial curve* $\{(t, x) : t = x^3\}$ *is non-characteristic with respect to the operator* $P = \partial_x$ *everywhere except at* $(0, 0)$*. The initial value problem*

$$\begin{cases} \partial_x u(t, x) = 0, \\ u(x^3, x) = x \end{cases}$$

has no analytic solution near $(0, 0)$*, for, a solution would have to be a function of* t *which takes value* $x = t^{1/3}$ *along* $\{(t, x) : t = x^3\}$*, thus would have to equal to* $t^{1/3}$*.*

The reduction process described above works almost identically for quasilinear operator

$$P = \sum_{|\alpha|=m} c_\alpha(x, \partial_x^\beta u)\partial_x^\alpha u(x) + D(x, \partial_x^\beta u),$$

where $\partial_x^\beta u$ denotes generic terms of differention of order $|\beta| \leq m - 1$. By this we mean the "flattening" process and the determination of the power series expansion based on the Cauchy data (of order $m - 1$ or less) and the equation. We may add that the process of reducing a higher order equation to a system of first order equation also works almost identically for higher order quasilinear operator, with one difference: the reduced system is quasilinear, instead of linear. Here the notion of non-characteristic initial hypersurface depends not only on the operator but also on the prescribed initial data.

Definition. $\Sigma = \{x : \sigma(x) = 0\}$ *is called non-characteristic with respect to* P *at* $x_0 \in \Sigma$ on the initial data g *if* $\sum_{|\alpha|=m} c_\alpha(x_0, \partial_x^\beta g(x_0))(\nabla\sigma(x_0))^\alpha \neq 0$*.* $\Sigma = \{x : \sigma(x) = 0\}$ *is called non-characteristic with respect to* P *if it is non-characteristic at every point on it.*

Theorem 3 (Quasilinear case with general non-characteristic initial surface)**.** *Suppose that* $c_\alpha(x, \partial_x^\beta u)$ *are analytic in its arguments around* $(x_0, \partial_x^\beta g(x_0))$*, that the initial data* g *is analytic around* $x_0 \in \Sigma$*, and that* Σ *is analytic around* x_0 *and non-characteristic with respect to*

P on the initial data g. Then there is a neighborhood of x_0, with a unique analytic solution to

$$\begin{cases} Pu = 0, & \text{near } x_0, \\ \partial_\beta u = \partial_\beta g(x), & \text{along } \Sigma \text{ for all } |\beta| \le m-1. \end{cases}$$

Example 6. *Consider the quasilinear problem*

$$\begin{cases} u_t(t,x) + u(t,x)u_x(t,x) = 0, \\ \qquad\qquad u(0,x) = g(x). \end{cases}$$

The initial curve $\{(t,x) : t = 0\}$ is non-characteristic for any initial data g. However, for the quasilinear problem

$$\begin{cases} u_t(t,x) + u(t,x)u_x(t,x) = 0, \\ \qquad\qquad u(t,0) = h(t), \end{cases}$$

the initial curve $\{(t,x) : x = 0\}$ is non-characteristic at $(t,0)$ for the given data h iff $h(t) \neq 0$. Thus one can apply the Cauchy-Kowalevskaya Theorem to establish local existence of an analytic solution to the above problem near $(0,0)$ if $h(t) = 1+t$, but not if $h(t) = t$.

Example 7. *If we modify Examples 2 into a quasilinear problem, making $t = x_0$ for simplicity, and allowing the a_i to depend on $x = (x_0, \cdots, x_n)$ as well as on $u(x)$, then, for a given hypersurface Σ as the initial surface and a given function g as initial data of u on Σ, the initial value problem*

$$\begin{cases} \sum_{i=0}^{n} a_i(x, u(x))\partial_{x_i} u(x) = c(x, u(x)), \\ \qquad\qquad u(x) = g(x), \quad \text{for } x \in \Sigma, \end{cases} \tag{4}$$

is non-charateristic near $\bar{x} \in \Sigma$, iff

the vector $(a_0(\bar{x}, g(\bar{x})), \cdots, a_n(\bar{x}, g(\bar{x})))$ is transveral to Σ at $\bar{x}$.

This can also be seen directly as follows. Describe Σ near $\bar{x}$ parametrically by a map $\phi : \Omega \subset \mathbb{R}^{n-1} \mapsto \Sigma \subset \mathbb{R}^n$ with $0 \in \Omega$ and $\phi(0) = \bar{x}$. We require that the Jacobian of ϕ at 0

$$\begin{pmatrix} \partial_1\phi_1 & \partial_1\phi_2 & \cdots & \partial_1\phi_n \\ \vdots & \vdots & \vdots & \vdots \\ \partial_{n-1}\phi_1 & \partial_{n-1}\phi_2 & \cdots & \partial_{n-1}\phi_n \end{pmatrix} \tag{5}$$

has rank $n-1$ at $s = 0$. Then, from

$$u(\phi_1(s), \cdots, \phi_n(s)) = g(s),$$

if we differentiate in the s_j direction, $j = 1, \cdots, n-1$, we obtain

$$\sum_{i=1}^{n} u_{x_i}(\phi_1(s), \cdots, \phi_n(s))\partial_{s_j}\phi_i(s) = g_{s_j}, \quad j = 1, \cdots, n-1. \tag{6}$$

These $n-1$ linear equations are the compatability conditions for the n quantities $u_{x_i}(\phi(s))$. (4) *provides one more compatability condition. We can determine $u_{x_i}(\phi(s))$ along Σ near $\bar{x}$, if*

$$\begin{pmatrix} \partial_1\phi_1 & \partial_1\phi_2 & \cdots & \partial_1\phi_n \\ \vdots & \vdots & \vdots & \vdots \\ \partial_{n-1}\phi_1 & \partial_{n-1}\phi_2 & \cdots & \partial_{n-1}\phi_n \\ a_1(\bar{x}, g(\bar{x})) & a_2(\bar{x}, g(\bar{x})) & \cdots & a_n(\bar{x}, g(\bar{x})) \end{pmatrix} \quad \text{is non-degenerate at } \bar{x}. \tag{7}$$

This arrives at the same non-charateristic condition as given in the beginning of this example.

To undertand the fully nonlinear version better, it is instructive to first examine the case with special initial surfact $\{t = 0\}$. Given a fully nonlinear operator

$$F = F(x, t, \partial_t^j \partial_x^\alpha u(x,t) | \, j + |\alpha| \le m),$$

where F is analytic in its arguments. Also given is initial data in the form of

$$\partial_t^j u(x, 0) = g_j(x), \quad \text{for } 0 \le j \le m - 1.$$

Then all the terms $\partial_t^j \partial_x^\alpha u(x,0) | \, j + |\alpha| \le m$ in F along $(x, 0)$ are determined by $g_j(x)$ with the exception of one term: $\partial_t^m u(x, 0)$. We require that (i) algebraically we can solve for $\partial_t^m u(x,0) = \tilde{u}_m$ *at this one point* x_0, and (ii) we can solve for $\partial_t^m u(x,t)$, locally, from

$$F(x, t, \partial_t^j \partial_x^\alpha u(x,t) | \, j + |\alpha| \le m) = 0,$$

in terms of the other arguments. By the implicit function theorem, this can be done if we assume

$$\frac{\partial}{\partial \left(\partial_t^m u\right)} F(x_0, 0, \partial_x^\alpha g_j(x_0) | \, (j + |\alpha| \le m, j < m), \tilde{u}_m) \neq 0. \tag{8}$$

Theorem 4 (fully nonlinear case with special non-characteristic initial surface)**.** *Suppose that the inital data $g_j(x)$ are analytic around x_0, that $F(x_0, 0, \partial_x^\alpha g_j(x_0) | \, (j + |\alpha| \le m, j < m), \tilde{u}_m) = 0$ has a solution $\tilde{u}_m$, that $F(x, t, \partial_t^j \partial_x^\alpha u(x,t) | \, j + |\alpha| \le m)$ is analytic in its arguments around $(x_0, 0, \partial_x^\alpha g_j(x_0) | \, (j + |\alpha| \le m, j < m), \tilde{u}_m)$, and that* (8) *holds. Then there is a neighborhood of x_0, with a unique analytic solution to*

$$\begin{cases} F = 0, & \text{near } x_0, \\ \partial_\beta u = \partial_\beta g(x), & \text{along } \Sigma \text{ for all } |\beta| \le m - 1. \end{cases}$$

Example 8. *Consider the problem*

$$\begin{cases} u_t^2(t,x) + u_x^2(t,x) = 1, \\ \qquad\qquad u(0,x) = g(x). \end{cases}$$

If $|g'(0)| < 1$, then near $(0,0)$, we can recast the equation as

$$u_t(t,x) = \pm\sqrt{1 - u_x^2(t,x)}.$$

We can choose to work with either branch of the square root. This corresponds to two possible choices for $u_t(0,0) = \pm\sqrt{1-|g'(0)|^2}$. *In such cases, the initial curve* $\{(0,x)\}$ *is non-characteristic near* $(0,0)$.

With this version in hand, the fully nonlinear case with general non-characteristic initial surface can be formulated in the same spirit as we did in the case for linear/quasilinear case. In fact, the easiest approach is to solve the differentiated equation $\partial F = 0$ first. This is a quasilinear equation of order $m+1$, the coefficient in front of the highest order term $\partial_x^\gamma \partial u$ with $|\gamma| = m$ is

$$\frac{\partial}{\partial\left(\partial_x^\gamma u\right)} F(x, \partial_x^\alpha u(x) |\, |\alpha| \leq m).$$

There are some issues to be dealt with. First, from the given Cauchy data

$$\partial_x^\beta u(x) = \partial_x^\beta g(x), \quad \text{along } \Sigma \text{ for all } |\beta| \leq m-1,$$

we still need to determine the terms $\partial_x^\alpha u$ along Σ for $|\alpha| = m$. They have to satisfy some compatability conditions along Σ. One way to resolve this issue is to assume that *at the point* $x_0 \in \Sigma$, we can find $\partial_x^\alpha u(x_0) = \tilde{u}_\alpha$ for $|\alpha| = m$ satisfying the compatability conditions (will be illustrated later in a simple case) and $F(x_0, \partial_x^\alpha u(x_0)) = 0$. We may assume $\sigma_{x_n}(x_0) \neq 0$, so can take the ∂ above to be ∂_{x_n}. We have discussed that $\Sigma = \{\sigma = 0\}$ is non-characteristic with respect to the equation $\partial_{x_n} F = 0$ on the initial data g (and $\tilde{u}_\alpha$) is

$$\sum_{\gamma=m} \frac{\partial}{\partial\left(\partial_x^\gamma u\right)} F(x_0, \partial_x^\alpha u(x_0) |\, |\alpha| \leq m) \left(\nabla_x \sigma(x_0)\right)^\gamma \sigma_{x_n}(x_0) \neq 0,$$

which is now equivalent to

$$\sum_{\gamma=m} \frac{\partial}{\partial\left(\partial_x^\gamma u\right)} F(x_0, \partial_x^\alpha u(x_0) |\, |\alpha| \leq m) \left(\nabla_x \sigma(x_0)\right)^\gamma \neq 0. \tag{9}$$

The remaining issue to solve the Cauchy problem for the quasilinear system $\partial_{x_n} F = 0$ is the the appropriate determination of the Cauchy data along Σ—we have data up to order $m-1$ prescribed along Σ and have assumed the determination of data of order m at x_0. It turns out the this, together with (9), allows to extend the Cauchy data to a neighborhood of x_0 along Σ by the implicit function theorem.

Theorem 5 (fully nonlinear case with general non-characteristic initial surface)**.** *Suppose that* Σ *is anlytic around* x_0 *and that the inital data* $g_j(x)$ *are analytic around* x_0, *that there exist* $\tilde{u}_\alpha$ *for* $|\alpha| = m$ *such that, with* $\partial_x^\alpha u(x_0) = \tilde{u}_\alpha$ *for* $|\alpha| = m$, *and* $\partial_x^\alpha u(x_0) = \partial_x^\alpha g(x_0)$ *for* $|\alpha| < m$, *we have* $F(x_0, \partial_x^\alpha u(x_0)) = 0$, *that* $F(x, \partial_t^j \partial_x^\alpha u(x,t) |\, j + |\alpha| \leq m)$ *is analytic in its arguments around* $(x_0, \partial_x^\alpha u(x_0))$, *and that* (9) *holds. Then there is a neighborhood of* x_0, *with a unique analytic solution to*

$$\begin{cases} F = 0, & \textit{near } x_0, \\ \partial_\beta u = \partial_\beta g(x), & \textit{along } \Sigma \textit{ for all } |\beta| \leq m-1. \end{cases}$$

Example 9. *For the problem*

$$\begin{cases} u_t^2(t,x) + u_x^2(t,x) = 1, \\ u(\phi(x), x) = g(x), \end{cases} \tag{10}$$

the initial curve is $\gamma : t = \phi(x)$*. It follows that the compatibility equation*

$$u_t(\phi(x), x)\phi'(x) + u_x(\phi(x), x) = g'(x)$$

must hold. We must first determine $u_t(\phi(x), x)$ *and* $u_x(\phi(x), x)$ *from*

$$\begin{cases} u_t^2(t,x) + u_x^2(t,x) = 1, \\ u_t(\phi(x), x)\phi'(x) + u_x(\phi(x), x) = g'(x). \end{cases}$$

This system of algebraic equations in $u_t(\phi(x), x)$ *and* $u_x(\phi(x), x)$ *has solution iff*

$$|g'(x)| \leq \sqrt{1 + |\phi'(x)|^2}.$$

The initial curve $\gamma : t = \phi(x)$ *is characteristic at* $(\phi(x), x)$ *iff it is characteristic with respect to the differentiated problem:*

$$u_t(t,x)u_{tt}(t,x) + u_x(t,x)u_{xt}(t,x) = 0,$$

where the initial data $u_t(\phi(x), x)$ *and* $u_x(\phi(x), x)$ *are determined from the steps above. So the characteristic equation for* $\gamma : t = \phi(x)$ *becomes—by taking* $\sigma(t,x) = t - \phi(x)$ *in the characteristic equation*

$$\begin{cases} u_t(t,x) - u_x(t,x)\phi'(x) = 0, \\ u_t^2(t,x) + u_x^2(t,x) = 1, \\ u_t(\phi(x), x)\phi'(x) + u_x(\phi(x), x) = g'(x). \end{cases}$$

Substituting the first equation into the remaining two, we find

$$|g'(x)|^2 = 1 + |\phi'(x)|^2.$$

Thus, if g *and* ϕ *are analytic near* x_0 *and* $|g'(x_0))|^2 < 1 + |\phi'(x_0)|^2$*, then* (10) *has a unique local analytic solution.*