4. **Cauchy – Kovalevskaya Theorem**

As a warm up we will start with the corresponding result for ordinary differential equations.

**Theorem 4.1** (ODE Version of Cauchy – Kovalevskaya, I.)  *Suppose* $a > 0$ *and* $f: (-a, a) \to \mathbb{R}$ *is real analytic near* $0$ *and* $u(t)$ *is the unique solution to the ODE*

$$
\dot{u}(t) = f(u(t)) \text{ with } u(0) = 0.
$$

*Then* $u$ *is also real analytic near* $0$.

We will give four proofs. However it is the last proof that the reader should focus on for understanding the PDE version of Theorem 4.1.

**Proof.** (First Proof.) If $f(0) = 0$, then $u(t) = 0$ for all $t$ is the unique solution to Eq. (4.1) which is clearly analytic. So we may now assume that $f(0) \neq 0$. Let $G(z) := \int_0^z \frac{1}{f(\xi)} d\xi$, another real analytic function near $0$.

Then as usual we have

$$
\frac{d}{dt} G(u(t)) = \frac{1}{f(u(t))} \dot{u}(t) = 1
$$

and hence $G(u(t)) = t$. We then have $u(t) = G^{-1}(t)$ which is real analytic near $t = 0$ since $G'(0) = \frac{1}{f(0)} \neq 0$.

**Proof.** (Second Proof.) For $z \in \mathbb{C}$ let $u_z(t)$ denote the solution to the ODE

$$
\dot{u}_z(t) = zf(u_z(t)) \text{ with } u_z(0) = 0.
$$

Notice that if $u(t)$ is analytic, then $t \to u(tz)$ satisfies the same equation as $u_z$. Since $G(z, u) = zf(u)$ is holomorphic in $z$ and $u$, it follows that $u_z$ in Eq. (4.2) depends holomorphically on $z$ as can be seen by showing $\partial_z u_z = 0$, i.e. showing $z \to u_z$ satisfies the Cauchy Riemann equations. Therefore if $\epsilon > 0$ is chosen small enough such that Eq. (4.2) has a solution for $|t| < \epsilon$ and $|z| < 2$, then

$$
\dot{u}(t) = u_{1}(t) = \sum_{n=0}^{\infty} \frac{1}{n!} \partial^n u(t)|_{z=0}.
$$

Now when $z \in \mathbb{R}$, $u_z(t) = u(tz)$ and therefore

$$
\partial^n u_z(t)|_{z=0} = \partial^n u(tz)|_{z=0} = u^{(n)}(0)t^n.
$$

Putting this back in Eq. (4.3) shows

$$
u(t) = \sum_{n=0}^{\infty} \frac{1}{n!} u^{(n)}(0)t^n
$$

which shows $u(t)$ is analytic for $t$ near 0.

**Proof.** (Third Proof.) Go back to the original proof of existence of solutions, but now replace $t$ by $z \in \mathbb{C}$ and $\int_0^t f(u(\tau)) d\tau$ by $\int_0^z f(u(\xi)) d\xi = \int_0^1 f(u(tz)) z dt$. Then the usual Picard iterates proof work in the class of holomorphic functions to give a holomorphic function $u(z)$ solving Eq. (4.1).

**Proof.** (Fourth Proof: Method of Majorants) Suppose for the moment we have an analytic solution to Eq. (4.1). Then by repeatedly differentiating Eq. (4.1) we
The function is convergent and in particular, is a well defined analytic function for all non-negative integer coefficients. The first few polynomials are $p_1(x) = x$, $p_2(x, y) = xy$, $p_3(x, y, z) = x^2 z + xy^2$. Notice that these polynomials are universal, i.e. are independent of the function $f$ and

$$\left| u^{(n)}(0) \right| = \left| p_n \left( f(0), \ldots, f^{(n-1)}(0) \right) \right| \leq p_n \left( |f(0)|, \ldots, |f^{(n-1)}(0)| \right) \leq p_n \left( g(0), \ldots, g^{(n-1)}(0) \right)$$

where $g$ is any analytic function such that $|f^{(k)}(0)| \leq g^{(k)}(0)$ for all $k \in \mathbb{Z}_+$. (We will abbreviate this last condition as $f \ll g$.) Now suppose that $v(t)$ is a solution to

$$v(t) = g(v(t)) \text{ with } v(0) = 0,$$

then we know from above that

$$v^{(n)}(0) = p_n \left( g(0), \ldots, g^{(n-1)}(0) \right) \geq |u^{(n)}(0)|$$

for all $n$.

Hence if knew that $v$ were analytic with radius of convergence larger that some $\rho > 0$, then by comparison we would find

$$\sum_{n=0}^{\infty} \frac{1}{n!} \left| u^{(n)}(0) \right| \rho^n \leq \sum_{n=0}^{\infty} \frac{1}{n!} v^{(n)}(0) \rho^n < \infty$$

and this would show

$$u(t) := \sum_{n=0}^{\infty} \frac{1}{n!} p_n \left( f(0), \ldots, f^{(n-1)}(0) \right) t^n$$

is a well defined analytic function for $|t| < \rho$.

I now claim that $u(t)$ solves Eq. (4.1). Indeed, both sides of Eq. (4.1) are analytic in $t$, so it suffices to show the derivatives of each side of Eq. (4.1) agree at $t = 0$. For example $u(0) = f(0)$, $u(0) = \frac{d}{dt} u(f(t))$, etc. However this is the case by the very definition of $u^{(n)}(0)$ for all $n$.

So to finish the proof, it suffices to find an analytic function $g$ such that $|f^{(k)}(0)| \leq g^{(k)}(0)$ for all $k \in \mathbb{Z}_+$ and for which we know the solution to Eq. (4.4) is analytic about $t = 0$. To this end, suppose that the power series expansion for $f(t)$ at $t = 0$ has radius of convergence larger than $r > 0$, then

$$\sum_{n=0}^{\infty} \frac{1}{n!} f^{(n)}(0) r^n$$

is convergent and in particular,

$$C := \max_n \left| \frac{1}{n!} f^{(n)}(0) r^n \right| < \infty$$

from which we conclude

$$\max_n \left| \frac{1}{n!} f^{(n)}(0) \right| \leq C r^{-n}.$$
Let

\[ g(u) := \sum_{n=0}^{\infty} C r^{-n} u^n = C \frac{1}{1 - u/r} = C \frac{r}{r - u}. \]

Then clearly \( f \ll g \). To conclude the proof, we will explicitly solve Eq. (4.4) with this function \( g(t) \),

\[ \dot{v}(t) = C \frac{r}{r - v(t)} \text{ with } v(0) = 0. \]

By the usual separation of variables methods we find

\[ rv(t) - \frac{1}{2} v^2(t) = Crt, \]

i.e.

\[ 2Crt - 2rv(t) + v^2(t) = 0 \]

which has solutions,

\[ v(t) = r - \sqrt{r^2 - 2Crt}. \]

We must take the negative sign to get the correct initial condition, so that

\[ (4.5) \quad v(t) = r - \sqrt{r^2 - 2Crt}. \]

which is real analytic for \( |t| < \rho := r/C \).

Let us now jazz up this theorem to that case of a system of ordinary differential equations. For this we will need the following lemma.

**Lemma 4.2.** Suppose \( h : (-a,a)^d \to \mathbb{R}^d \) is real analytic near \( 0 \in (-a,a)^d \), then

\[ h \ll \frac{Cr}{r - z_1 - \cdots - z_d} \]

for some constants \( C \) and \( r \).

**Proof.** By definition, there exists \( \rho > 0 \) such that

\[ h(z) = \sum_{\alpha} h_{\alpha} z^\alpha \text{ for } |z| < \rho \]

where \( h_{\alpha} = \frac{1}{\alpha!} \partial^\alpha h(0). \) Taking \( z = r(1,1,\ldots,1) \) with \( r < \rho \) implies there exists \( C < \infty \) such that \( |h_{\alpha}| r^{|\alpha|} \leq C \) for all \( \alpha \), i.e.

\[ |h_{\alpha}| \leq Cr^{-|\alpha|} \leq C \frac{|\alpha|!}{\alpha!} r^{-|\alpha|}. \]

This completes the proof since

\[ \sum_{\alpha} C \frac{|\alpha|!}{\alpha!} r^{-|\alpha|} z^\alpha = C \sum_{n=0}^{\infty} \sum_{|\alpha|=n} \frac{|\alpha|!}{\alpha!} \left( \frac{z}{r} \right)^\alpha = C \sum_{n=0}^{\infty} \left( \frac{z_1 + \cdots + z_d}{r} \right)^n \]

\[ = C \frac{1}{1 - \left( \frac{z_1 + \cdots + z_d}{r} \right)} = \frac{Cr}{r - z_1 - \cdots - z_d} \]

all of which is valid provided \( |z| := |z_1| + \cdots + |z_d| < r \).

**Theorem 4.3** (ODE Version of Cauchy – Kovalevskaya, II.). Suppose \( a > 0 \) and \( f : (-a,a)^d \to \mathbb{R}^d \) be real analytic near \( 0 \in (-a,a)^d \) and \( u(t) \) is the unique solution to the ODE

\[ (4.6) \quad \dot{u}(t) = f(u(t)) \text{ with } u(0) = 0. \]

Then \( u \) is also real analytic near \( 0 \).
Proof. All but the first proof of Theorem 4.1 may be adapted to the cover this case. The only proof which perhaps needs a little more comment is the fourth proof. By Lemma 4.2, we can find \( C, r > 0 \) such that
\[
\frac{C r}{r - z_1 - \cdots - z_d}
\]
for all \( j \). Let \( v(t) \) denote the solution to the ODE,
\[
\dot{v}(t) = g(v(t)) = \frac{C r}{r - v_1(t) - \cdots - v_d(t)}(1, 1, \ldots, 1)
\]
with \( v(0) = 0 \). By symmetry, \( v_j(t) = v_j(1) =: w(t) \) for each \( j \) so Eq. (4.7) implies
\[
\dot{w}(t) = \frac{C r}{r - d w(t)} = \frac{C (r/d)}{(r/d) - w(t)} \quad \text{with} \quad w(0) = 0.
\]
We have already solved this equation (see Eq. (4.5) with \( r \) replaced by \( r/d \)) to find
\[
w(t) = r/d - \sqrt{r^2/d^2 - 2C r t/d} = r/d \left( 1 - \sqrt{1 - 2C d t/r} \right).
\]
Thus \( v(t) = w(t)(1, 1, \ldots, 1) \) is a real analytic function which is convergent for \( |t| < r/(2Cd) \).

Now suppose that \( u \) is a real analytic solution to Eq. (4.6). Then by repeatedly differentiating Eq. (4.6) we learn
\[
\begin{align*}
\dot{u}_j(t) &= \partial_i f_j(u(t)) \dot{u}_i(t) = \partial_i f_j(u(t)) f_i(u(t)) \\
u_j^{(3)}(t) &= \partial_k \partial_i f_j(u(t)) \dot{u}_k(t) \dot{u}_i(t) + \partial_i f_j(u(t)) \ddot{u}_i(t) \\
& \vdots \\
u_j^{(n)}(t) &= p_n \left( \{\partial^\alpha f_j(u(t))|_{|\alpha| < n}\}, \left\{u_j^{(k)}(t)\right\}_{k<n, 1 \leq i \leq d} \right)
\end{align*}
\]
where \( p_n \) is a polynomial with all non-negative integer coefficients. We now define \( u_j^{(n)}(0) \) inductively so that
\[
u_j^{(n)}(0) = p_n \left( \{\partial^\alpha f_j(u(0))|_{|\alpha| < n}\}, \left\{u_j^{(k)}(0)\right\}_{k<n, 1 \leq i \leq d} \right)
\]
for all \( n \) and \( j \) and we will attempt to define
\[
(4.10) \quad u(t) = \sum_{n=0}^\infty \frac{1}{n!} u_j^{(n)}(0) t^n.
\]
To see this sum is convergent we make use of the fact that the polynomials \( p_n \) are universal i.e. are independent of the function \( f_j \) and have non-negative coefficients so that by induction
\[
\left|u_j^{(n)}(0)\right| \leq p_n \left( \{\partial^\alpha f_j(u(0))|_{|\alpha| < n}\}, \left\{|\dot{u}_j^{(k)}(0)|\right\}_{k<n, 1 \leq i \leq d} \right)
\leq p_n \left( \{\partial^\alpha g_j(u(0))|_{|\alpha| < n}\}, \left\{|\dot{v}_j^{(k)}(0)|\right\}_{k<n, 1 \leq i \leq d} \right) = v_j^{(n)}(0).
\]
Notice the when \( n = 0 \) that \( |u_j(0)| = 0 = v_j(0) \). Thus we have shown \( u \ll v \) and so by comparison the sum in Eq. (4.10) is convergent for \( t \) near 0. As before \( u(t) \) solves Eq. (4.6) since both functions \( \dot{u}(t) \) and \( f(u(t)) \) are analytic functions of \( t \) which have common values for all derivatives in \( t \) at \( t = 0 \).

4.1. PDE Cauchy Kovalevskaya Theorem. In this section we will consider the following general quasi-linear system of partial differential equations

\[
\sum_{|\alpha|=k} a_\alpha(x, J^{k-1}u) \partial_\alpha u(x) + c(x, J^{k-1}u) = 0
\]

where

\[
J^l u(x) = (u(x), Du(x), D^2u(x), \ldots, D^l u(x))
\]

is the “\( l \)-jet” of \( u \). Here \( u : \mathbb{R}^n \to \mathbb{R}^m \) and \( a_\alpha(J^{k-1}u, x) \) is an \( m \times m \) matrix. As usual we will want to give boundary data on some hypersurface \( \Sigma \subset \mathbb{R}^n \). Let \( \nu \) denote a smooth vector field along \( \Sigma \) such that \( \nu(x) \notin T_x \Sigma \) \( (T_x \Sigma \text{ is the tangent space to } \Sigma \text{ at } x) \) for \( x \in \Sigma \). For example we might take \( \nu(x) \) to be orthogonal to \( T_x \Sigma \) for all \( x \in \Sigma \). To hope to get a unique solution to Eq. (4.11) we will further assume there are smooth functions \( g_l \) on \( \Sigma \) for \( l = 0, \ldots, k-1 \) and we will require

\[
D^l u(x)(\nu(x), \ldots, \nu(x)) = g_l(x) \text{ for } x \in \Sigma \text{ and } l = 0, \ldots, k-1.
\]

**Proposition 4.4.** Given a smooth function \( u \) on a neighborhood of \( \Sigma \) satisfying Eq. (4.12), we may calculate \( D^l u(x) \) for \( x \in \Sigma \) and \( l < k \) in terms of the functions \( g_l \) and there tangential derivatives.

**Proof.** Let us begin by choosing a coordinate system \( y \) on \( \mathbb{R}^n \) such that \( \Sigma \cap D(y) = \{ y_n = 0 \} \) and let us extend \( \nu \) to a neighborhood of \( \Sigma \) by requiring \( \frac{\partial \nu}{\partial y_n} = 0 \). To complete the proof, we are going to show by induction on \( k \) that we may compute

\[
\left( \frac{\partial}{\partial y} \right)^\alpha u(x) \text{ for all } x \in \Sigma \text{ and } |\alpha| < k
\]

from Eq. (4.12).

The claim is clear when \( k = 1 \), since \( u = g_0 \) on \( \Sigma \). Now suppose that \( k = 2 \) and let \( \nu_i = \nu_i(y_1, \ldots, y_{n-1}) \) such that

\[
\nu = \sum_{i=1}^n \nu_i \frac{\partial}{\partial y_i} \text{ in a neighborhood of } \Sigma.
\]

Then

\[
g_1 = (Du) \nu = \nu u = \sum_{i=1}^n \nu_i \frac{\partial u}{\partial y_i} = \sum_{i<n} \nu_i \frac{\partial g_0}{\partial y_i} + \nu_n \frac{\partial u}{\partial y_n}.
\]

Since \( \nu \) is not tangential to \( \Sigma = \{ y_n = 0 \} \), it follows that \( \nu_n \neq 0 \) and hence

\[
\frac{\partial u}{\partial y_n} = \frac{1}{\nu_n} \left( g_1 - \sum_{i<n} \nu_i \frac{\partial g_0}{\partial y_i} \right) \text{ on } \Sigma.
\]

\[1\text{The argument shows that } v_j^{(n)}(0) \geq 0 \text{ for all } n. \text{ This is also easily seen directly by induction using Eq. (4.9) with } f \text{ replaced by } g \text{ and the fact that } \partial^a g_j(0) \geq 0 \text{ for all } \alpha.\]
For $k = 3$, first observe from the equality $u = g_0$ on $\Sigma$ and Eq. (4.13) we may compute all derivatives of $u$ of the form $\frac{\partial^\alpha u}{\partial y^\alpha}$ on $\Sigma$ provided $\alpha_n \leq 1$. From Eq. (4.12) for $l = 2$, we have

$$g_2 = (D^2 u)(v,v) = v^2 u + \text{l.o.t.s.} = \sum \nu_j \frac{\partial}{\partial y_j} \left( \nu_i \frac{\partial u}{\partial y_i} \right) + \text{l.o.t.s.} = \nu_n^2 \frac{\partial^2 u}{\partial y_n^2} + \text{l.o.t.s.}$$

where l.o.t.s. denotes terms involving $\frac{\partial^\alpha u}{\partial y^\alpha}$ with $\alpha_n \leq 1$. From this result, it follows that we may compute $\frac{\partial^2 u}{\partial y^2}$ in terms of derivatives of $g_0$, $g_1$ and $g_2$. The reader is asked to finish the full inductive argument of the proof. \[\Box\]

**Remark 4.5.** The above argument shows that from Eq. (4.12) we may compute $\frac{\partial^\alpha u}{\partial y^\alpha}$ for any $\alpha$ such that $\alpha_n < k$.

To study Eq. (4.11) in more detail, let us rewrite Eq. (4.11) in the $y$-coordinates. Using the product and the chain rule repeatedly Eq. (4.11) may be written as

$$\sum_{|\alpha|=k} b_\alpha(y,J^{k-1}u)\partial^\alpha y u(y) + c(y,J^{k-1}u) = 0$$

where

$$J^l u(y) = (u(y), Du(y), D^2 u(y), \ldots, D^l u(y)).$$

We will be especially concerned with the $b_{(0,0,\ldots,0,k)}$ coefficient which can be determined as follows:

$$\sum_{|\alpha|=k} a_\alpha \left( \frac{\partial}{\partial x} \right)^\alpha = \sum_{|\alpha|=k} a_\alpha \left( \sum_{j=1}^n \frac{\partial y_j}{\partial x} \frac{\partial}{\partial y_j} \right)^\alpha = \sum_{|\alpha|=k} a_\alpha \left( \frac{\partial y_n}{\partial x} \frac{\partial}{\partial y_n} \right)^\alpha + \text{l.o.t.s.} = \sum_{|\alpha|=k} a_\alpha \left( \frac{\partial n}{\partial x} \right)^\alpha + \text{l.o.t.s.}$$

where l.o.t.s. now denotes terms involving $\frac{\partial^\alpha u}{\partial y^\alpha}$ with $\alpha_n < k$. From this equation we learn that

$$b_{(0,0,\ldots,0,k)}(y,J^{k-1}u) = \sum_{|\alpha|=k} a_\alpha \left( \frac{\partial y_n}{\partial x} \right)^\alpha = \sum_{|\alpha|=k} a_\alpha \left( \frac{\partial y_n}{\partial x} \right)^\alpha = a_{(0,0,\ldots,0,k)}(y,J^{k-1}u).$$

**Definition 4.6.** We will say that boundary data $(\Sigma, g_0, \ldots, g_{k-1})$ is non-characteristic for Eq. (4.11) at $x \in \Sigma$ if

$$b_{(0,0,\ldots,0,k)}(y,J^{k-1}u) = \sum_{|\alpha|=k} a_\alpha(x,J^{k-1}u(x)) \left( \frac{\partial y_n}{\partial x} \right)^\alpha$$

is invertible at $x$.

Notice that this condition is independent of the choice of coordinate system $y$. To see this, for $\xi \in (\mathbb{R}^n)^*$ let

$$\sigma(\xi) = \sum_{|\alpha|=k} a_\alpha(x,J^{k-1}u(x)) \left( \xi \left( \frac{\partial}{\partial x} \right) \right)^\alpha$$
which is $k$–linear form on $(\mathbb{R}^n)^*$. This form is coordinate independent since if $f$ is a smooth function such that $f(x) = 0$ and $df_x = \xi$, then

$$
\sigma(\xi) = \frac{1}{k!} \sum_{|\alpha| = k} a_\alpha(x, J^{k-1}u(x)) \left( \frac{\partial}{\partial x} \right)^\alpha f^k|_x.
$$

Noting that

$$
b_{(0,0,\ldots,k)}(y, J^{k-1}u) = \sigma(dy_n)
$$

our non-characteristic condition becomes, $\sigma(dy_n)$ is invertible. Finally $dy_n$ is the unique element $\xi$ of $(\mathbb{R}^n)^* \setminus \{0\}$ up to scaling such that $\xi|_{T_x \Sigma} \equiv 0$. So the non-characteristic condition may be written invariantly as $\sigma(\xi)$ is invertible for all (or any) $\xi \in (\mathbb{R}^n)^* \setminus \{0\}$ such that $\xi|_{T_x \Sigma} \equiv 0$.

Assuming the given boundary data is non-characteristic, Eq. (4.11) may be put into “standard form,”

$$
\sum_{|\alpha| = k} b_\alpha(y, J^{k-1}u) \partial_y^\alpha u(y) + c(y, J^{k-1}u) = 0
$$

with

$$
\frac{\partial^l u}{\partial y_n^l} = g_l \text{ on } y_n = 0 \text{ for } l < k
$$

where $b_{(0,0,\ldots,k)}(y, J^{k-1}u) = Id$ - matrix and

$$
J^l u(y) = (u(y), Du(y), D^2 u(y), \ldots, D^l u(y)).
$$

By adding new dependent variables and possible a new independent variable for $y_n$ one may reduce the problem to solving the system in Eq. (4.20) below. The resulting theorem may be stated as follows.

**Theorem 4.7** (Cauchy Kovalevskaya). *Suppose all the coefficients in Eq. (4.11) are real analytic and the boundary data in Eq. (4.12) are also real analytic and non-characteristic near some point $a \in \Sigma$. Then there is a unique real analytic solution to Eqs. (4.11) and (4.12). (The boundary data in Eq. (4.12) is said to be real analytic if there exists coordinates $y$ as above which are real analytic and the functions $\nu$ and $g_l$ for $l = 0,\ldots,k - 1$ are real analytic functions in the $y$–coordinate system.)*

**Example 4.8.** Suppose $a, b, C, r$ are positive constants. We wish to show the solution to the quasi-linear PDE

$$
w_t = \frac{Cr}{r - y - aw} [bw_y + 1] \text{ with } w(0, y) = 0
$$

is real analytic near $(t, y) = (0, 0)$. To do this we will solve the equation using the method of characteristics. Let $g(y, z) := \frac{Cr}{r - y - az}$, then the characteristic equations are

$$
t' = 0 \text{ with } t(0) = 0,
y' = -bg(y, z) \text{ with } y(0) = y_0 \text{ and } z' = g(y, z) \text{ with } z(0) = 0.
$$
From these equations we see that we may identify $t$ with $s$ and that $y + bz = y_0$.
Thus $z(t) = w(t, y(t))$ satisfies
\[
\dot{z} = g(y_0 - bz, z) = \frac{Cr}{r - y_0 + bz - az} = \frac{Cr}{r - y_0 + (b - a)z}
\]
with $z(0) = 0$.

Integrating this equation gives
\[
Crt = \int_0^t (r - y_0 + (b - a)z(\tau)) \dot{z}(\tau) d\tau = (r - y_0)z - \frac{1}{2} (a - b) z^2
\]
\[
= (r - y - bz)z - \frac{1}{2} (a - b) z^2 = (r - y)z - \frac{1}{2} (a + b) z^2,
\]
i.e.
\[
\frac{1}{2} (a + b) z^2 - (r - y) z + Cr t = 0.
\]
The quadratic formula gives
\[
w(t, y) = \frac{1}{a + b} \left[ (r - y) \pm \sqrt{(r - y)^2 - 2(a + b)Cr} \right]
\]
and using $w(0, y) = 0$ we conclude
\[
(4.17) \quad w(t, y) = \frac{1}{a + b} \left[ (r - y) - \sqrt{(r - y)^2 - 2(a + b)Cr} \right].
\]

Notice the $w$ is real analytic for $(t, y)$ near $(0, 0)$.

In general we could use the method of characteristics and ODE properties (as in Example 4.8) to show
\[
u_t = a(x, u)u_x + b(x, u) \text{ with } u(0, x) = g(x)
\]
have local real analytic solutions if $a, b$ and $g$ are real analytic. The method would also work for the fully non-linear case as well. However, the method of characteristics fails for systems while the method we will present here works in this generality.

**Exercise 4.1.** Verify $w$ in Eq. (4.17) solves Eq. (4.16).

**Solution 1 (4.1).** Let $\rho := \sqrt{(r - y)^2 - 2(a + b)Cr}$, then
\[
w(t, y) = \frac{1}{a + b} \left[ r - y - \rho \right] = \frac{r - y}{a + b} - \frac{1}{a + b} \rho,
\]
\[
w_t = C r / \rho, \quad \rho = r - y - (a + b)w \text{ and}
\]
\[
bw_y + 1 = \frac{b}{a + b} \left[ -1 + (r - y) / \rho \right] + 1 = \frac{1}{a + b} \left[ a + b (r - y) / \rho \right].
\]
Hence
\[
\frac{bw_y + 1}{w_t} = \frac{1}{(a + b) Cr} \left[ \rho a + b (r - y) \right]
\]
\[
= \frac{1}{(a + b) Cr} \left[ (r - y - (a + b)w) a + b (r - y) \right]
\]
\[
= \frac{1}{Cr} \left[ r - y - aw \right]
\]
as desired.
**Example 4.9.** Now let us solve for

\[ v(t, x) = (v^1, \ldots, v^m) (t, x_1, \ldots, x_n) \]

where \( v \) satisfies

\[ v^j_t = \frac{Cr}{r-x_1-\cdots-x_n-x_j} + \sum_{k=1}^{m} \alpha_{j}^k v^k \]

with \( v(0, x) = 0 \).

By symmetry, \( v^j = v^1 = w(t, y) \) for all \( j \) where \( y = x_1 + \cdots + x_n \). Since \( \partial_i v^j = w_y \), the above equations all may be written as

\[ w_t = \frac{Cr}{r-y-mw} [mw_y + 1] \text{ with } w(0, y) = 0. \]

Therefore from Example 4.8 with \( a = m \) and \( b = mn \), we find

\[ w(t, y) = \frac{1}{m(n+1)} [(r-y) - \sqrt{(r-y)^2 - 2m(n+1)Crt}] \]

and hence that

\[ v(t, x) = w(t, x_1 + \cdots + x_n) (1, 1, 1, \ldots, 1) \in \mathbb{R}^m. \]

**4.2. Proof of Theorem 4.7.** As is outlined in Evans, Theorem 4.7 may be reduced to the following theorem.

**Theorem 4.10.** Let \((t, x, z) = (t, x_1, \ldots, x_n, z_1, \ldots, z_m) \in \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^m \) and assume \((t, x, z) \to B_j(t, x, z) \in \{m \times m \text{ matrices}\} \) (for \( j = 1, \ldots, n \) and \((t, x, z) \to c(t, x, z) \in \mathbb{R}^m \) are real analytic functions near \((0, 0, 0) \in \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^m \) and \( x \to f(x) \in \mathbb{R}^m \text{ is real analytic near } 0 \in \mathbb{R}^n \). Then there exists, in a neighborhood of \((t, x) = (0, 0) \in \mathbb{R} \times \mathbb{R}^n \), a unique real analytic solution \( u(t, x) \in \mathbb{R}^m \) to the quasi-linear system

\[ u_t(t, x) = \sum_{j=1}^{n} B_j(t, x, u(t, x))\partial_j u(t, x) + c(t, x, u(t, x)) \text{ with } u(0, x) = f(x). \]

**Proof.** (Sketch.)

**Step 0.** By replacing \( u(t, x) \) by \( u(t, x) - f(x) \), we may assume \( f \equiv 0 \). By letting \( u^{m+1}(t, x) = t \) if necessary, we may assume \( B_j \) and \( c \) do not depend on \( t \). With these reductions we are left to solve

\[ u_t(t, x) = \sum_{j=1}^{n} B_j(x, u(t, x))\partial_j u(t, x) + c(x, u(t, x)) \text{ with } u(0, x) = 0. \]

**Step 1.** Let

\[ g(x, z) := \frac{Cr}{r-x_1-\cdots-x_n-x_j-z_1-\cdots-z_m} \]

where \( C \) and \( r \) are positive constants such that

\[ (B_j)_{kl} \ll g \text{ and } c_k \ll g \]

for all \( k, l, j \). For this choice of \( C \) and \( r \), let \( v \) denote the solution constructed in Example 4.9 above.

**Step 2.** By repeatedly differentiating Eq. (4.20), show that if \( u \) solves Eq. (4.20) then \( \partial_x^\ell \partial_t^k u^j(0, 0) \) is a **universal** polynomial in the derivatives \( \{\partial_x^\ell \partial_t^k \} \alpha, \beta < k \).
of the entries of $B_j$ and $c$ and $u$ with all coefficients being non-negative. Use this fact and induction to conclude

$$|\partial^\alpha_x \partial^k_t w^j(0,0)| \leq \partial^\alpha_x \partial^k_t v^j(0,0)$$

for all $\alpha, k$ and $l$.

**Step 3.** Use the computation in Step 2. to define $\partial^\alpha_x \partial^k_t w^j(0,0)$ for all $\alpha$ and $k$ and then defined

$$u(t,x) := \sum_{\alpha,k} \frac{\partial^\alpha_x \partial^k_t u(0,0)}{\alpha!k!} t^k x^\alpha.$$ 

Because of step 2. and Example 4.9, this series is convergent for $(t,x)$ sufficiently close to zero.

**Step 4.** The function $u$ defined in Step 3. solves Eq. (4.20) because both

$$u_t(t,x)$$

and

$$\sum_{j=1}^n B_j(x, u(t,x)) \partial_j u(t,x) + c(x, u(t,x))$$

are both real analytic functions in $(t,x)$ each having, by construction, the same derivatives at $(0,0)$.

4.3. **Examples.**

**Corollary 4.11 (Isothermal Coordinates).** Suppose that we are given a metric $ds^2 = Edx^2 + 2F dxdy + Gdy^2$ on $\mathbb{R}^2$ such that $G/E$ and $F/E$ are real analytic near $(0,0)$. Then there exists a complex function $u$ and a positive function $\rho$ such that $Du(0,0)$ is invertible and $ds^2 = \rho |du|^2$ where $du = u_x dx + u_y dy$.

**Proof.** Working out $|du|^2$ gives

$$|du|^2 = |u_x|^2 dx^2 + 2 \Re(u_x \bar{u}_y) dxdy + |u_y|^2 dy^2.$$ 

Writing $u_y = \lambda u_x$, the previous equation becomes

$$|du|^2 = |u_x|^2 \left( dx^2 + 2 \Re(\lambda) dxdy + |\lambda|^2 dy^2 \right).$$

Hence we must have

$$E = \rho |u_x|^2, \quad F = \rho |u_x|^2 \Re \lambda$$

and

$$G = \rho |u_x|^2 |\lambda|^2$$

or equivalently

$$\frac{F}{E} = \Re \lambda \quad \text{and} \quad \frac{G}{E} = |\lambda|^2.$$ 

Writing $\lambda = a + ib$, we find $a = F/E$ and $a^2 + b^2 = G/E$ so that

$$\lambda = \frac{F}{E} \pm i \sqrt{G/E - (F/E)^2} = \frac{1}{E} \left( F \pm i \sqrt{GE - F^2} \right).$$

We make a choice of the sign above, then we are looking for $u(x,y) \in \mathbb{C}$ such that $u_y = \lambda u_x$. Letting $u = \alpha + i\beta$, the equation $u_y = \lambda u_x$ may be written as the system of real equations

$$\alpha_y = \Re [(a + ib) (\alpha_x + i\beta_x)] = a\alpha_x - b\beta_x$$

and

$$\beta_y = \Im [(a + ib) (\alpha_x + i\beta_x)] = a\beta_x + b\alpha_x.$$ 

which is equivalent to

$$\left( \begin{array}{c} \alpha \\ \beta \end{array} \right)_y = \left( \begin{array}{cc} a & -b \\ b & a \end{array} \right) \left( \begin{array}{c} \alpha \\ \beta \end{array} \right)_x.$$
So we may apply the Cauchy Kovalevskaya theorem 4.10 with \( t = y \) to find a real analytic solution to this equation with (say) \( u(x, 0) = x \), i.e. \( \alpha(x, 0) = x \) and \( \beta(x, 0) = 0 \). (We could take \( u(x, 0) = f(x) \) for any real analytic function \( f \) such that \( f'(0) \neq 0 \).) The only thing that remains to check is that \( Du(0,0) \) is invertible. But

\[
Du(0,0) = \begin{pmatrix} \Re u_x & \Re u_y \\ \Im u_x & \Im u_y \end{pmatrix} = \begin{pmatrix} \alpha_x & \alpha_y \\ \beta_x & \beta_y \end{pmatrix} = \begin{pmatrix} \alpha_x & \alpha \alpha_x - b \beta_x \\ \beta_x & \beta \beta_x + b \alpha_x \end{pmatrix}
\]

so that

\[
\det [Du] = b (\alpha_x^2 + \beta_y^2) = \Im \lambda |u_x|^2.
\]

Thus

\[
\det [Du(0,0)] = \Im \lambda (0,0) = \pm \sqrt{G/E - (F/E)^2}_{(0,0)} \neq 0.
\]

**Example 4.12.** Consider the linear PDE,

(4.23) \[ u_y = u_x \text{ with } u(x, 0) = f(x). \]

where \( f(x) = \sum_{m=0}^{\infty} a_m x^m \) as real analytic function near \( x = 0 \) with radius of convergence \( \rho \). (So for any \( r < \rho \), \( |a_m| \leq Cr^{-n}. \)) Formally the solution to Eq. (4.23) should be given by

\[
u(x, y) = \sum_{n=0}^{\infty} \frac{1}{n!} \frac{\partial^n u(x, y)|_{y=0}}{y^n}.
\]

Now using the PDE (4.23),

\[
\partial_y^n u(x, y)|_{y=0} = \partial_x^n u(x, 0) = f^{(n)}(x).
\]

Thus we get

(4.24) \[ u(x, y) = \sum_{n=0}^{\infty} \frac{1}{n!} f^{(n)}(x) y^n. \]

By the Cauchy estimates,

\[
\left| f^{(n)}(x) \right| \leq \frac{n! \rho}{(\rho - |x|)^{n+1}}
\]

and so

\[
\sum_{n=0}^{\infty} \frac{1}{n!} \left| f^{(n)}(x) y^n \right| \leq \rho \sum_{n=0}^{\infty} \frac{|y|^n}{(\rho - |x|)^{n+1}}
\]

which is finite provided \( |y| < \rho - |x| \), i.e. \( |x| + |y| < \rho \). This of course makes sense because we know the solution to Eq. (4.23) is given by

\[ u(x, y) = f(x + y). \]

Now we can expand Eq. (4.24) out to find

\[
u(x, y) = \sum_{n=0}^{\infty} \frac{1}{n!} \left( \sum_{m=n}^{\infty} m(m-1) \ldots (m-n+1) a_m x^m y^n \right)
\]

(4.25) \[ = \sum_{m \geq n \geq 0} \binom{m}{n} a_m x^{m-n} y^n. \]
Since
\[
\sum_{m \geq n \geq 0} \binom{m}{n} |a_m x^{m-n} y^n| \leq C \sum_{m \geq n \geq 0} \binom{m}{n} |r^{-m} x^{m-n} y^n| = C \sum_{m \geq 0} r^{-m} (|x| + |y|)^m < \infty
\]
provided \(|x| + |y| < r\). Since \(r < \rho\) was arbitrary, it follows that Eq. (4.25) is convergent for \(|x| + |y| < \rho\).

Let us redo this example. By the PDE in Eq. (4.23), \(\partial^m \partial^n x u(x, y) = \partial^{n+m} u(x, y)\) and hence
\[
\partial^m \partial^n u(0, 0) = f(m+n)(0).
\]
Written another way
\[
D^\alpha u(0, 0) = f(|\alpha|)(0)
\]
and so the power series expansion for \(u\) must be given by
\[
(4.26) \quad u(x, y) = \sum_{\alpha} \frac{f(|\alpha|)(0)}{\alpha!} (x, y)^\alpha.
\]
Using \(f^{(m)}(0)/m! \leq C r^{-m}\) we learn
\[
\sum_{\alpha} \left| \frac{f(|\alpha|)(0)}{\alpha!} (x, y)^\alpha \right| \leq C \sum_{\alpha} \left| \frac{f(|\alpha|)(0)}{\alpha!} \right| |x|^{|\alpha|} |y|^{|\alpha|} = C \sum_{m=0}^{\infty} \frac{f^{(m)}(0)}{m!} \sum_{|\alpha|=m} \frac{m!}{\alpha!} |x|^{|\alpha|} |y|^{|\alpha|} < \infty
\]
if \(|x| + |y| < r\). Since \(r < \rho\) was arbitrary, it follows that the series in Eq. (4.26) converges for \(|x| + |y| < \rho\).

Now it is easy to check directly that Eq. (4.26) solves the PDE. However this is necessary since by construction
\[
D^\alpha y u_y(0, 0) = D^\alpha u_x(0, 0)
\]
for all \(\alpha\). This implies, because \(u_y\) and \(u_x\) are both real analytic, that \(u_x = u_y\).