Causality, local well-posedness, and all that

Marcelo M. Disconzi*,

Vanderbilt University

The many faces of relativistic fluid dynamics

KITP, Santa Barbara, CA

June 2023

* Supported by NSF grant DMS-2107701.

Contents

Our presentation will differ from a typical seminar talk in that we will not be presenting new research result. Instead, we will carry out an informal discussion of some mathematical tools and concepts that have been proven very useful for the study of the evolution problem in relativistic hydrodynamics.

We will not make technical statements, focusing instead on intuitive explanations and illustrations via examples. We will attempt to be "sufficiently precise," meaning that we hope to present our main claims in a clear and precise fashion, while avoiding fully technical statements. This will be the case including for statements of theorems.

We encourage the audience to interrupt and ask questions. In particular, although we will assume familiarity with relativity, some conventions and notation used in the math literature are different than that of the physics literature, so do ask if anything looks out of place.

For simplicity, all functions etc. will be real, but everything generalize to complex-valued maps.

1. Solving a PDE

Consider the following linear partial differential equation (PDE):

$$M^\alpha \partial_\alpha \psi + B\psi = 0 \tag{$*$}$$

where $\psi = (\psi_1, \dots, \psi_N)$ and M^α, B are matrices. Given data ψ on $\Sigma_0 := \{x^0 = 0\}$, how (when) can we solve $(*)$?

Writing

$$M^0 \partial_0 \psi\big|_{\Sigma_0} = - M^i \partial_i \psi\big|_{\Sigma_0} - B\psi|_{\Sigma_0}$$

the RHS is entirely determined by ψ_0. Hence we know $\partial_0\psi|_{\Sigma_0}$ provided $(M^0)^{-1}$ exists, i.e., $\det M^0 \neq 0$. Taking ∂_t of $(*)$ and restricting to Σ_0,

$$M^0 \partial_0^2 \psi\big|_{\Sigma_0} = - M^i \partial_0 \partial_i \psi\big|_{\Sigma_0} - B\partial_0 \psi|_{\Sigma_0} - \partial_0 M^0 \partial_0 \psi\big|_{\Sigma_0} - \partial_0 M^i \partial_i \psi\big|_{\Sigma_0} - \partial_0 B \psi|_{\Sigma_0}\,.$$

For $\det M^0 \neq 0$, the RHS is determined by ψ_0 and thus so $\partial_0^2 \psi\big|_{\Sigma_0}$. Proceeding this way, we can determine

$$\partial_0^l \psi\Big|_{\Sigma_0}$$

for all $l \geq 0$ and define

$$\psi := \sum_{l=0}^{\infty} \frac{1}{l!}\, \partial_t^l \psi\Big|_{\Sigma_0} t^l \tag{$**$}$$

By construction, ψ will be a solution **provided the series converges.** It can be showed that **if the data ψ_0 is analytic, the series will converge** at least for small time (see below), and thus we **have a solution** (again, assuming that M^0 is invertible).

Remark 1.1. To say that ψ_0 is analytic means that ψ equals its convergent Taylor series. Thus, the statement is that if ψ_0 can be expressed as a Taylor series, then $(**)$ converges, at least for small t.

We want to better understand the structure of the above construction. In particular, we want to allow data given on other hypersurfaces than $\{x^0 = 0\}$.

Observe that Σ_0 is the level set $\{\phi = 0\}$ of the function $\phi(x) = x^0$. Taking the co-vector $\xi := d\phi = (1, 0, \ldots, 0)$,

$$\det M^0 = \det(M^\alpha \xi_\alpha),$$

which is a condition we will generalize.

Consider data ψ_0 given on a hypersurface Σ. Let us assume that Σ is given as the level set $\{\phi = 0\}$ of a function ϕ (this is always the case locally) and that Σ is regular, i.e., $d\phi \neq 0$ (as a vector, $d\phi$ is normal to Σ, thus the condition $d\phi \neq 0$ is stating that Σ has a well-define normal). Can we mimic the above argument and use $(*)$ to determine derivatives of ψ transverse to Σ in terms of ψ_0?

Since $d\phi \neq 0, \partial_\alpha \phi \neq 0$ for some $\alpha = 0, \ldots, n$. Let us assume for simplicity that $\partial_0 \phi \neq 0$. Introduce a change of variables:

$$\tilde{x}^\alpha = \begin{cases} \phi(x), & \alpha = 0 \\ x^\alpha, & \alpha \neq 0 \end{cases}$$

Then,

$$\frac{\partial \psi}{\partial x^\alpha} = \frac{\partial \psi}{\partial \tilde{x}^\beta} \frac{\partial \tilde{x}^\beta}{\partial x^\alpha} = \Lambda_\alpha^\beta \frac{\partial \psi}{\partial \tilde{x}^\beta},\ \Lambda_\alpha^\beta := \frac{\partial \tilde{x}^\beta}{\partial x^\alpha}$$

Thus,

$$M^\alpha \partial_\alpha \psi = M^\alpha \Lambda_\alpha^\beta \tilde{\partial}_\beta \psi = \tilde{M}^\beta \tilde{\partial}_\beta \psi,\ \tilde{\partial}_\beta = \frac{\partial}{\partial \tilde{x}^\beta},\ \tilde{M}^\beta = M^\alpha \Lambda_\alpha^\beta$$

and $(*)$ becomes (with $\tilde{\psi}(\tilde{x}) = \psi(x(\tilde{x}))$ and similarly for $\tilde{B}$)

$$\tilde{M}^\alpha \tilde{\partial}_\alpha \tilde{\psi} + \tilde{B} \tilde{\psi} = 0$$

with $\tilde{\psi}_0$ as data on Σ. The derivatives transverse to Σ are the $\tilde{x}^0$ derivatives, so we need, as above

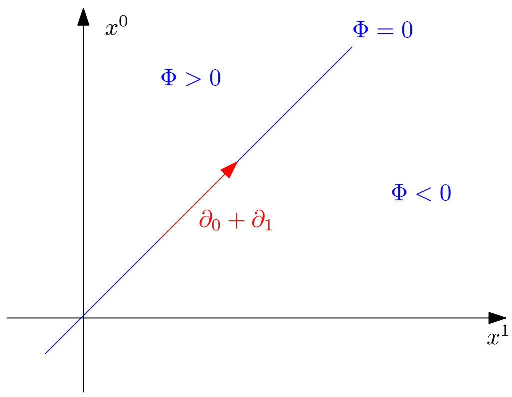
$$\det \tilde{M}^0 \neq 0$$

But

$$\tilde{M}^0 = M^\alpha \Lambda_\alpha^0 = M^\alpha \frac{\partial \tilde{x}^0}{\partial x^\alpha} = M^\alpha \frac{\partial \phi}{\partial x^\alpha}.$$

Thus, **derivatives of ψ transverse to Σ can be determined in terms of ψ_0 if $\det(M^\alpha \xi_\alpha) \neq 0$, $\xi = d\phi, \Sigma = \{\phi = 0\}$.**

Example 1.2. To illustrate the role of transverse as tangential derivatives, consider



$$
\begin{aligned}
\phi(x^0, x^1) &= x^0 - x^1,\ \tilde{x}^0 = x^0 - x^1,\ \tilde{x}^1 = x^1, \\
\tilde{\psi}(\tilde{x}^0, \tilde{x}^1) &= \psi(x^0, x^1) = \psi(\tilde{x}^0 + \tilde{x}^1, \tilde{x}^1) \\
\tilde{\psi}_0 &= \tilde{\psi}(0, \tilde{x}^1) = \psi_0(x^1, x^1)
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial}{\partial \tilde{x}^1} \tilde{\psi}(\tilde{x}^0, \tilde{x}^1) &= \frac{\partial}{\partial \tilde{x}^1} (\psi(\tilde{x}^0 + \tilde{x}^1, \tilde{x}^1)) \\
&= \partial_0 \psi(\tilde{x}^0 + \tilde{x}^1, \tilde{x}^1) \frac{\partial}{\partial \tilde{x}^1} (\tilde{x}^0 + \tilde{x}^1) + \partial_1 \psi(\tilde{x}^0 + \tilde{x}^1, \tilde{x}^1) \\
&= (\partial_0 + \partial_1)\psi|_{(\tilde{x}^0 + \tilde{x}^1, \tilde{x}^1)} \\
\frac{\partial}{\partial \tilde{x}^1} \tilde{\psi}\Big|_\Sigma &= (\partial_0 + \partial_1)\psi_0|_{(x^1, x^1)} \longrightarrow \text{ intrinsitc to } \Sigma \\
\frac{\partial}{\partial \tilde{x}^0} \tilde{\psi}(\tilde{x}^0, \tilde{x}^1) &= \frac{\partial}{\partial \tilde{x}^0} (\psi(\tilde{x}^0 + \tilde{x}^1, \tilde{x}^1)) \\
&= \partial_0 \psi(\tilde{x}^0 + \tilde{x}^1, \tilde{x}^1) \frac{\partial}{\partial \tilde{x}^0} (\tilde{x}^0 + \tilde{x}^1) + 0 \\
&= \partial_0 \psi|_{(\tilde{x}^0 + \tilde{x}^1, \tilde{x}^1)} \\
\frac{\partial}{\partial \tilde{x}^0} \tilde{\psi}\Big|_\Sigma &= \partial_0 \psi|_{(x^1, x^1)} \longrightarrow \text{ not intrinsitc to } \Sigma
\end{aligned}
$$

Conclusion 1.3. Given analytic data ψ_0 on an analytic hypersurface $\Sigma = \{\phi = 0\}$ (i.e. ϕ is analytic) then the PDE

$$M^\alpha \partial_\alpha \psi + B\psi = 0$$

admits a solution defined in a neighborhood of Σ and satisfying $\psi|_\Sigma = \psi_0$ provided that

$$\det(M^\alpha \xi_\alpha) \neq 0,\ \xi = d\phi$$

Definition 1.4. The hypersurface $\Sigma = \{\phi = 0\}$ such that

$$\det(M^\alpha \xi_\alpha) \neq 0,\ \xi = d\phi$$

are called characteristic surfaces/manifolds or simply **characteristics** for the PDE

$$M^\alpha \partial_\alpha \psi + B\phi = 0,$$

or more precisely for the differential operator $M^\alpha\partial_\alpha + B$. The matrix $M^\alpha \xi_\alpha$ is called the **characteristic matrix** and its determinant the **characteristic determinant**.

The above result thus states that **we can solve the PDE for data prescribed on a non-characteristic surface.**

Note that **B plays an role in determining the characteristics**. The characteristics depend only on the principal part of the operator, which is the part containing the highest number of derivatives (i.e. $M^\alpha\partial_\alpha$ here). Thus, different operators with same principal part have the same characteristics.

Observe that the computation of $\det(M^\alpha\xi_\alpha)$ can be done without a priori fixing a surface Σ. In other words, given the PDE, we can ask which surfaces $\Sigma = \{\phi = 0\}$ are characteristic for our operator.

Example 1.5. Consider in $1+1$ dimensions

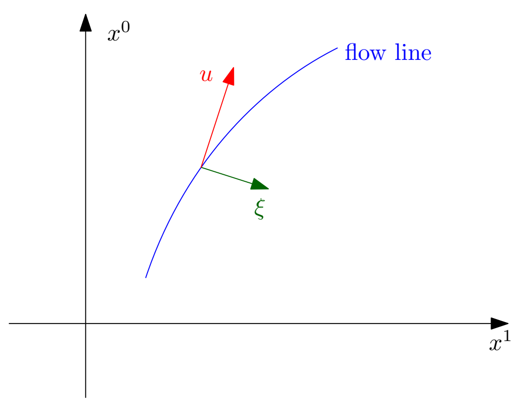
$$u^\mu\partial_\mu\psi + b\psi = 0,\ u = (u^0, u^1) = \text{vector}$$

Then the characteristics are given by

$$u^\mu \xi_\mu = 0$$

These are the flow lines (aka integral curves) of u. Indeed, u is tangent to the flow lines and ξ is orthogonal to the characteristics.

In higher dimensions the characteristics of $u^\mu\partial_\mu$ are still given by the flow lines of u, although in this case the flow lines are not hypersurfaces (think of this as a "degenerate" case).



Let us now consider second order PDE,

$$M^{\alpha\beta}\partial_\alpha\partial_\beta\psi + B^\alpha\partial_\alpha\psi + C\psi = 0$$

Reasoning as above, given $\psi_0 = \psi|_\Sigma$ and $\psi_1 = \partial_t\psi|_\Sigma$ on $\Sigma = \{x^0 = 0\}$, we can determine $\partial_t^\alpha\psi|_\Sigma$ and higher derivatives in terms of the data if

$$\det(M^{00}) \neq 0$$

For a general $\Sigma = \{\phi = 0\}$, a change of variables as above shows that transverse derivative of ψ are determined by data on Σ provided Σ is non-characteristic, i.e.,

$$\det(M^{\alpha\beta}\xi_\alpha\xi_\beta) \neq 0,$$

with the characteristics defined for the second order operator via

$$\det(M^{\alpha\beta}\xi_\alpha\xi_\beta) = 0$$

Again, the characteristics depend only on the principal part.

Example 1.6. For the wave operator $g^{\mu\nu}\partial_\mu\partial_\nu$, the characteristics are $g^{\mu\nu}\xi_\mu\xi_v = 0$, which are the light cones. Specifically for the Minkowski metric we have

$$-\xi_0^2 + |\vec{\xi}|^2 = 0,\ \xi_0 = \pm|\vec{\xi}|$$

The above reasoning generalizes to higher-order operators

$$M^{\alpha_1 \ldots \alpha_m}\partial_{\alpha_1}\ldots\partial_{\alpha_m}\psi + \underbrace{L.O.T}_{\text{lower order terms}}$$

The characteristics are given by

$$\det(M^{\alpha_1 \ldots \alpha_m}\xi_{\alpha_1}\ldots\xi_{\alpha_m}) = 0$$

(depending, once again, only on the principal part).

Example 1.7. For the third-order operator

$$u^\mu g^{\alpha\beta}\partial_\mu\partial_\alpha\partial_\beta$$

the characteristics are given by

$$u^\mu g^{\alpha\beta}\partial_\mu\partial_\alpha\partial_\beta = (u^\mu\xi_\mu)(g^{\alpha\beta}\xi_\alpha\xi_\beta) = 0$$

so the characteristics are the flow lines of u and the light cones. (Note that this illustrates that an operator can have multiple characteristics).

Example 1.8. For the Laplacian operator in $\mathbb{R}^n$

$$\triangle := \partial_1^2 + \cdots + \partial_n^2$$

The characteristics are given by

$$\xi_1^2 + \cdots + \xi_n^2 = 0 \Rightarrow \xi = 0$$

But recall that we required $\xi = d\phi \neq 0$, so the Laplacian has no characteristics.

Let us now consider nonlinear problems. Consider an operator of order m whose matrix of the principal part depends on at most $m-1$ of ϕ, i.e.,

$$\begin{aligned} P\psi &= M^{\alpha_1,\ldots,\alpha_m}(\psi, \partial\psi, \ldots, \partial^{m-1}\psi)\partial_{\alpha_1}\ldots\partial_{\alpha_m}\psi + F(\psi, \partial\psi, \ldots, \partial^{m-1}\psi) \\ &= M^{\alpha_1,\ldots,\alpha_m}(\partial^{m-1}\psi)\partial_{\alpha_1}\ldots\partial_{\alpha_m}\psi + F(\partial^{m-1}\psi) \end{aligned}$$

Operators of this form are called **quasilinar** and we have used the following:

Notation 1.9. Here and in what follows, $\partial^l\psi$ is abbreviation for the set of all derivatives of ψ of order l. When writing expressions like $M = M(\partial^{m-1}\psi)$ we mean that M depends on at most $m-1$ derivatives of ψ, i.e., we abbreviate $M(\psi, \partial\psi, \ldots, \partial^{m-1}\psi)$ as $M(\partial^{m-1}\psi)$.

Then, given ψ and its derivatives up to order not on Σ, i.e., (see the Cauchy problem below for further discussion)

$$\psi|_{\Sigma_0}, \partial\psi|_{\Sigma_0}, \ldots, \partial^{m-1}\psi\big|_{\Sigma}$$

we can, as above, solve for $\partial^m\psi|_\Sigma$ in terms of the data if Σ is non-characteristic, i.e., if

$$\det(M^{\alpha_1,\ldots,\alpha_m}(\partial^{m-1}\psi)\xi_{\alpha_1}\ldots\xi_{\alpha_m}) \neq 0$$

on $\Sigma = \{\phi = 0\}, \xi = d\phi$. This is well-defined because by assumption $M^{\alpha_1,\ldots,\alpha_m}(\partial^{m-1}\psi)\big|_\Sigma$ is determined by the data.

As in the linear case, the L.O.T $F(\partial^{m-1}\psi)$ plays no role in determining the characteristics $F(\partial^{m-1}\psi)\big|_\Sigma$ is determined by the data) and, as in the linear case, we do not need to fix Σ a priori, but can instead ask which Σ's are characteristic by computing the characteristic by computing the characteristic determinant.

Example 1.10. Consider the relativistic Euler equations

$$\begin{aligned}(p+\varrho)u^\alpha\nabla_\alpha u^\beta + \frac{\partial p}{\partial\varrho}\Pi^{\alpha\beta}\nabla_\alpha\varrho + \frac{\partial p}{\partial s}\Pi^{\alpha\beta}\nabla_\alpha s = 0\\ u^\alpha\nabla_\alpha\varrho + (p+\varrho)\nabla_\mu u^\mu = 0\\ u^\alpha\nabla_\alpha s = 0\end{aligned}$$

Or equivalently $A^\alpha\nabla_\alpha\Phi = 0$ when $\Phi = (u^\lambda, \varrho, s)$ and

$$A^\alpha = \begin{bmatrix}(p+\varrho)u^\alpha\delta^\beta_\lambda & \Pi^{\beta\alpha}\frac{\partial p}{\partial\varrho} & \Pi^{\beta\alpha}\frac{\partial p}{\partial s}\\ (p+\varrho)\delta^\alpha_\lambda & u^\alpha & 0\\ 0 & 0 & u^\alpha\end{bmatrix}$$

(Note that α indexes the matrices and not their entries.) The characteristic determinant is

$$\begin{aligned}\det(A^\alpha\xi_\alpha) &= u^\alpha\xi_\alpha \det\begin{bmatrix}(p+\varrho)u^\alpha\xi_\alpha\delta^\beta_\lambda & \Pi^{\beta\alpha}\xi_\alpha\frac{\partial p}{\partial\varrho}\\ (p+\varrho)\xi_\lambda & u^\alpha\xi_\alpha\end{bmatrix}\\ &= (p+\varrho)^4(u^\alpha\xi_\alpha)^4[(u^\alpha\xi_\alpha)^2 - \frac{\partial p}{\partial\varrho}\Pi^{\mu\nu}\xi_\mu\xi_\nu]\end{aligned}$$

Thus (for $p+\varrho\neq 0$) the characteristics are given by

$$u^\alpha\xi_\alpha = 0$$

which are the flow lines of u, and

$$(u^\mu\xi_\mu)^2 - \frac{\partial P}{\partial S}\Pi^{\mu\nu}\xi_\mu\xi_\nu = 0$$

which are the **sound cones**. The terminology sound cones can be understood as follows.

Introduce a frame $\{e_A\}_{A=0}^3$ with $e_0 = u$ and $\{e_1, e_2, e_3\}$ orthonormal and orthogonal to u. We also introduce the dual frame $\{e^A\}_{A=0}^3$ given by $(e^A)_\alpha := m^{AB}(e_B)_\alpha$ (where m is g expressed in this frame which then takes the form of the Minhowski metric), so that $e^A(e_B) = \delta^A_B$.

Decomposing ξ with respect to the dual frame $\xi_A = e^\mu_A\xi_\mu$ we have $\xi_{A=0} = -\xi^{A=0} = u^\mu\xi_\mu$ and $\xi_{A=i} = v_{A=i}$, where $v^\mu = \Pi^{\mu\alpha}\xi_\alpha$ and $v_A = e^\mu_A v_\mu$.

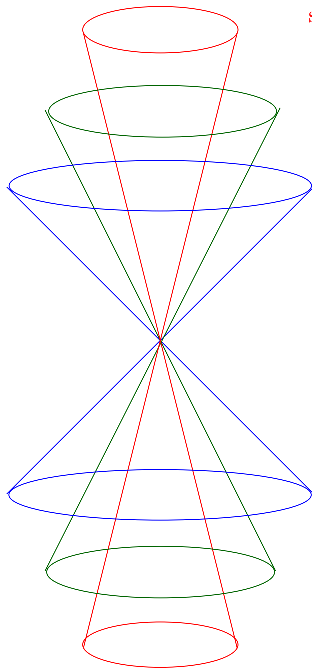
Therefore,

$$\xi^2_{A=0} - \frac{\partial p}{\partial\varrho}\sum_{i=1}^3 \xi^2_{A=i} = 0$$

If $\frac{\partial p}{\partial\varrho} < 0$, there are no real solutions, if $\frac{\partial p}{\partial\varrho} > 1$, the evolution will be acausal (we will see this later when we study causality). Thus, setting the sound speed as $c_s^2 = \frac{\partial p}{\partial\varrho}$,

$$\xi_A = \pm c_s(\sum_{i=1}^3 \xi^2_{A=i})^{\frac{1}{2}}$$

which has a cone structure very much the light cones. We will tie the sound cones to propagation of information at speed c_s when we discuss causality.



Remark 1.11. The characteristics are invariant under coordinate changes. Thus, it is legitimate to choose a convenient frame, as we did, to compute them.

Remark 1.12. The sound cones motivate the following definition. The **acoustical metric** is the Lorentzian metric defined by

$$\begin{aligned} G_{\alpha\beta} &:= c_s^2 g_{\alpha\beta} + (c_s^{-2} - 1) u_\alpha u_\beta \\ &= -u_\alpha u_\beta + c_s^{-2} \Pi_{\alpha\beta} \end{aligned}$$

One can verify that, for $0 < c_s \leq 1$ and $g_{\alpha\beta} u^\alpha u^\beta = -1$, G is indeed a Lorentzian metric, whose inverse is

$$\begin{aligned} (G^{-1})^{\alpha\beta} &= c_s^2 g^{\alpha\beta} + (c_s^2 - 1) u_\alpha u_\beta \\ &= -u_\alpha u_\beta + c_s^2 \Pi_{\alpha\beta} \end{aligned}$$

where we write G^{-1} because indices are raised with g but $G^{\alpha\beta} = g^{\alpha\mu} g^{\beta\nu} G_{\mu\nu} \neq (G^{-1})^{\alpha\beta}$. The sound cones are then

$$G^{\alpha\beta} \xi_\alpha \xi_\beta = 0$$

The acoustical metric reveals that the ideal fluid carries its own geometry, called the **acoustic geometry,** which is different from the spacetime geometry. When the ideal fluid is couple to Einstein's equations the acoustic and spacetime geometries interact with each other.

When $c_s \to 0, G$ is no longer defined. G^{-1} remains defined but it is no longer as invertible quadratic form,

$$(G^{-1})^{\alpha\beta} \xi_\alpha \xi_\beta \xrightarrow{c_s \to 0} (u^\alpha \xi_2)^2,$$

hence the sound cones degenerate to the flow lines, when $c_s \to 0$. This is a sense in which the flow lines can be though as a degenerate hypersurface, as alluded before.

2. The Cauchy problem

Definition 2.1. Consider the quasilinear m^{th}-order operator

$$P\psi := M^{\alpha_1,\dots,\alpha_m}(\partial^{m-1}\psi)\partial_{\alpha_1}\dots\partial_{\alpha_m}\psi + F(\partial^{m-1}\psi)$$

The **Cauchy problem** for P consists of the following. Given a hypersurface Σ and data on ψ consisting of ψ and its derivatives up to order $(m-1)$, i.e.,

$$\psi\big|_\Sigma , \partial\psi\big|_\Sigma , \dots , \partial^{m-1}\psi\big|_\Sigma ,$$

find a solution ψ to $P\psi = 0$ taking the prescribed data on Σ.

Remark 2.2. Typically we prescribe only transverse derivatives, e.g., $\partial_t\psi\big|_\Sigma , \dots , \partial_t^{m-1}\psi\big|_\Sigma$ if $\Sigma = \{x^0 = 0\}$. But if $\psi|_\Sigma$ is given (and is sufficiently differentiable) then its tangential derivatives up to order $(m-1)$ can be computed, e.g., $\partial_{x^1}^{m-1}\psi\big|_\Sigma$. Thus, there is no loss of generality in prescribing all derivatives up to order $(m-1)$.

By the foregoing discussion, we have the following result:

Theorem 2.3 (Cauchy-Kovalevskaya)**.** *The Cauchy problem for P is locally uniquely solvable in the analytic class provided Σ is non-characteristic.*

The theorem can be restated as follows. Assume that $M^{\alpha_1,\dots,\alpha_m}$ and F are analytic functions of their arguments. Given an analytic hypersurface $\Sigma = \{\phi = 0\}$ and analytic data

$$\psi\big|_\Sigma , \dots , \partial^{m-1}\psi\big|_\Sigma ,$$

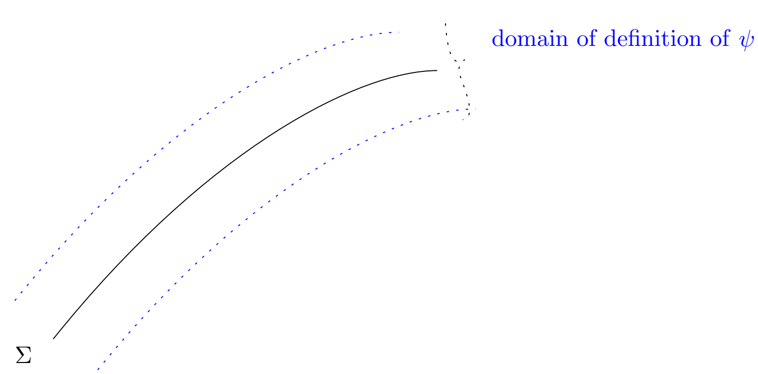
if

$$\det(M^{\alpha_1,\dots,\alpha_m}(\psi\big|_\Sigma , \dots , \partial^{m-1}\psi\big|_\Sigma)\xi_{\alpha_1}, \dots , \xi_{\alpha_m}) \neq 0$$

then there exists a unique analytic ψ, which is defined in a neighborhood of Σ, solving $P\psi = 0$ and taking the given data on Σ.

Remark 2.4. We note:

- The theorem guarantees that ψ is defined only in a neighborhood of Σ. E.g., if $\Sigma = \{x^0 = 0\}$ the solution will be defined for $|t| < \varepsilon$. This is what is meant by local existence (and uniqueness): the solution is guaranteed only for some (typically small) $|t| < \varepsilon$ or, more generally, in some neighborhood of Σ. We sometimes emphasize this with the terminology local solutions.
- The uniquness is the theorem is uniqueness of analytic solutions. The theorem leaves open the possibility that different, non-analytic solutions exist.



Example 2.5. Taking $\Sigma = \{x^0 = 0\}$, the Cauchy-Kovalevskaya theorem can be applied to

- $u^\mu\partial_\mu\psi + b\psi = 0, u$ not tangent to Σ

- $g^{\mu\nu}\partial_\mu\partial_\nu\psi = 0$
- $u^\mu g^{\alpha\beta}\partial_\mu\partial_\alpha\partial_\beta\psi = 0, u$ not tangent to Σ
- The relativistic Euler equations
- $\Delta\psi = 0$

2.1. **Difficulties and limitations.** A natural questions is what happens if we consider the Cauchy problem with data prescribed on a characteristic surface. As seen, if Σ is characteristic, then the PDE cannot be used to determine $\partial^m\psi|_\Sigma$ in terms of $\psi|_\Sigma, \dots, \partial^{m-1}\psi\big|_\Sigma$.

This means that the **PDE imposes non-trivial relations among the data** $\psi|_\Sigma, \dots, \partial^{m-1}\psi\big|_\Sigma$, **so in particular the data cannot be arbitrary, otherwise a solution will not exist.**

As an example, consider

$$\begin{aligned} -a(t,x)\partial_t^2\psi + \partial_x^2\psi = 0,\ a \geq 0 \\ \psi|_{t=0} = \psi_0, \\ \partial_t\psi|_{t=0} = \psi_1. \end{aligned}$$

The characteristics are

$$-a(t,x)\xi_0^2 + \xi_1^2 = 0,$$

giving the cones $\xi_0 = \pm\frac{|\xi_1|}{a}$ if $a \neq 0$.

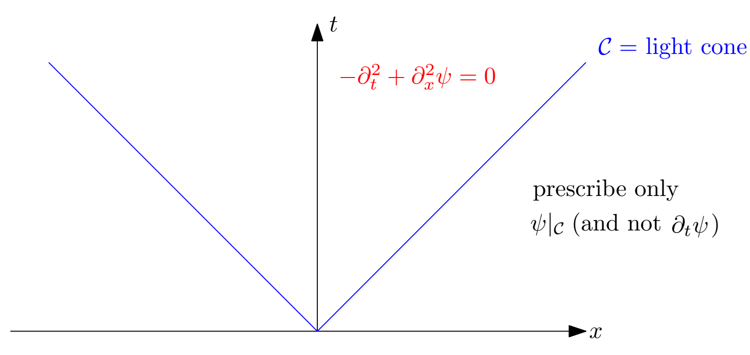
However, if $a(0,x) = 0$ then any $\xi = (\xi_0, 0)$ solves $-a(0,x)\xi_0^2 + \xi_1^2 = 0$, which means that the surface $\Sigma = \{t = 0\}$ is characteristic (i.e., $\phi = x^0, \xi = (1,0)$). At $t = 0$ the equation gives

$$0 = (-a(t,x)\partial_t^2\psi + \partial_x^2\psi)\big|_\Sigma = 0 + \partial_x^2\psi\big|_\Sigma = \partial_x^2\psi_0,$$

so the initial date cannot be arbitrary, having to satisfy the **constraint** $\partial_x^2\psi_0 = 0$.

There are important situations in physics where the data is not arbitrary, having to satisfy constrains. Such situations typically come about due to established principles, e.g., div $\mathrm{E} = 0 =$ div B in the (vacuum) Maxwell equations, $u^\alpha u_\alpha = -1$ for the relativistic Euler equations, or the Hamiltonian and momentum constraints for Einstein's equations. In such cases, one shows that the constraints are propagated by the flow.

There are also cases where we want to prescribe data on a characteristic surface, in which case we expect constraints or prescribe less data then stipulated by the Cauchy problem. For example, in the so called characteristic Cauchy problem for the wave equation, we prescribe data on a light cone, but only one function instead of two as in the usual Cauchy problem.



For an example of application of the characteristic Cauchy problem for the wave equation, see Hawking's original paper on Hawking radiation, where data is prescribed on past null infinity.

Another example is the heat equation, considering

$$\partial_t \psi - \partial_x^2 \psi = 0,\ t > 0,$$
$$\psi|_{t=0} = \psi_0.$$

The heat equation is second order, so for its Cauchy problem we should specify $\psi|_{t=0}$ and $\partial\psi|_{t=0}$, but this will clearly lead to constraints: If $\psi|_{t=0} = \psi_0$ and $\partial_t\psi|_{t=0} = \psi_1$, then at $t = 0$ the equation gives $\psi_1 - \partial_x^2\psi_0 = 0$. (Note also that the characteristics for the heat operator are the surfaces t = constant, since we get $\xi_1^2 = 0, \xi_0 =$ free). The problem where we prescribe only $\psi|_{t=0}$ for the heat equation is sometimes called the Cauchy problem for the heat equation, but this terminology is very confusing since the Cauchy problem for an equation of order m is almost always defined as prescribing $\psi|_\Sigma, \ldots, \partial^{m-1}\psi\big|_\Sigma$. The terminology initial value problem for the heat equation would be more appropriate. In any case, for the heat equation we are taking clear advantage of a preferred time coordinate in order to prescribe $\psi|_{t=0}$ only, something undesirable in a relativistic setting.

We also note that solutions to heat equations will in general not be analytic even if $\psi|_{t=0}$ is. A counter example, due to Kovalevskaya, is

$$\psi|_{t=0} = \frac{1}{1+x^2},$$

which is analytic near zero, but the corresponding power series in t, constructed by using the equation to solve for $\partial_t\psi$ and higher derivatives at $t = 0$, does not converge for any $t > 0$ near the origin.

(As an aside on constraints, recall that in hydro, we have the constraint $u^\alpha u_\alpha = -1$, but this is by design and typically one uses it to decompose the equations in the directions parallel and perpendicular to u, treating all components of u in the same footing; the constraint $u^\alpha u_\alpha = -1$ is then showed to be propagated by the flow. In this sense, we can "forget" about the constraint.)

The upshot of the above discussion is that understanding the characteristics of a PDE is key to investigate the solvability of its Cauchy problem.

In particular, **if the characteristic determinant vanishes identically we cannot solve the Cauchy problem for any Σ, even in the nicest setting one could hope for, namely, when everything is analytic**. One then has to deal with constraints or a modification of the Cauchy problem (as in the characteristic Cauchy problem for the wave equation). Whether these difficulties come from some legitimate physical principle or are a true flaw of a theory can one be decided on case-by-case. For example, from a PDE perspective, the constraints for Einstein's equation can be understood as a consequence of the fact that its characteristic determinant vanishes identically if one does not fix a gauge. But this is understood as a consequence of the diffeomorphism invariance of the equations. Taking this into account allows us to properly formulate the Cauchy problem for Einstein's equations. On the other hand, certain formulations of chiral hydro have identically zero characteristic determinant, and in such cases there does not seem to exist a deeper principle, like in Einstein's equations, that would explain this, so such theories do seem fundamentally flawed (see my paper with Enrico, Fabio, and Jorge).

Even when we can apply the Cauchy-Kovalevskaya theorem, unfortunately, the outcome is not good enough for many applications of interest.

First, allowing only for analytic functions is **too restrictive**. For example, many situations in physics naturally lead us to consider bump functions that vanish outside a bounded set and these cannot be analytic (an analytic function vanishing on an open set vanishes everywhere). One often also has to deal with functions with limited differentiability, with concerns, etc. (not to mention distributions).

Second, even within the analytic class, we can get **undesirable results**, as the following example, by Hadamard shows. Consider the following Cauchy problem for the Laplacian in 2D:

$$\begin{cases} \Delta\psi &= 0, \\ \psi &= 0 \text{ on } \{x^2 = 0\}, \\ \frac{\partial\psi}{\partial x^2} &= \frac{1}{k}\sin(kx^1) \text{ on } \{x^2 = 0\}. \end{cases}$$

A solution is given by

$$\psi(x^1, x^2) = \frac{1}{k^2}\sin(kx^1)\sinh(kx^2)$$

which is analytic so is the solution given by Cauchy-Kovalevskaya (by uniqueness). When $k \to \infty$, the problem becomes

$$\begin{cases} \Delta\psi &= 0, \\ \psi &= 0 \text{ on } \{x^2 = 0\}, \\ \frac{\partial\psi}{\partial x^2} &= 0 \text{ on } \{x^2 = 0\} \end{cases}$$

which admits zero as solution, but

$$\psi(x^1, x^2) = \frac{1}{h^2}\sin(kx^1)\sinh(kx^2)$$

blows up when $h \to \infty$. Thus, the data converges to zero but the solution does not (we would have to be a bit more specific about how we are measuring convergence, but the statement will be true for any natural way of defining convergence of a sequence of functions. In particular, for any points $x' = \frac{\pi}{2k}, x^2 \neq 0, \psi(x^1, x^2) \to \infty$).

Another way to view the chore is as follows. Start with

$$\begin{cases} \Delta\psi &= 0, \\ \psi &= 0 \text{ on } \{x^2 = 0\}, \\ \frac{\partial\psi}{\partial x^2} &= 0 \text{ on } \{x^2 = 0\}. \end{cases}$$

Then consider

$$\begin{cases} \Delta\psi &= 0 \\ \psi &= 0 \text{ on } \{x^2 = 0\} \\ \frac{\partial\psi}{\partial x^2} &= \frac{1}{k}\sin(kx^1) \text{ on } \{x^2 = 0\} \end{cases}$$

Since $|\frac{\partial\psi}{\partial x^2}(x^1, 0)| = \frac{1}{k}|\sin(kx^1)| \leq \frac{1}{k}$ for every $(x^1, 0)$, this data is a perturbation of the zero data for large k. But the solution

$$\psi(x^1, x^2) = \frac{1}{k^2}\sin(kx^1)\sinh(kx^2)$$

is not a perturbation, for large h, of the zero solution (again, we would have to define how to measure perturbation, but this solution cannot be close to zero for every (x_1, x_2)). (The above example is due to Hadamard.)

This means that solutions do not vary continuously with the data; small changes on the data do not necessarily lead to small changes in the solution, a type of strong "instability" (we are using the word "instability" loosely) that we would like to avoid in physical systems.

In sum, **understanding the characteristic of a PDE is necessary to understand its solvability properties; it tells us whether we can solve it in at least the simplest case of analytic functions. But analyticity is not yet enough to provide us with a good existence and uniqueness theory.**

We will seek a more robust framework to investigate the solvability of PDE. In doing so, we will have to specify further conditions on the kinds of PDE, we will consider.

3. Local well-posedness

Definition 3.1. The Cauchy problem for a PDE is said to be **locally well-posed (LWP)** if:

a. A solution exists (meaning, a **local** solution that solves the equation and takes the given data). (Existence.)

b. There is at most one solution (meaning, given two local solutions taking the data, they coincide). (Uniqueness).

c. Small variations in the data produce smell variations in solutions (if a sequence of data converges to a given data, the corresponding solutions converge, where convergence has to be properly defined, see below). (Continuous dependence on the data.)

From now on we will say simply LWP for LWP of the Cauchy problem. We will also say solution for local solution, unless stated otherwise.

We need to make a few qualifications:

- It is possible that a PDE is LWP for a certain class of functions X (e.g., analytic functions) but not LWP with respect to another class of functions Y (e.g., smooth functions). (i.e., given data in X, LWP holds, but given data in Y, LWP does not hold). Thus, one needs to talk about **LWP in X, where X is some space of functions** (we will give examples of relevant function spaces soon).
- When we talk about LWP in X, we mean that given data in X a solution exists and the solution also belongs to X. For example, we saw that for the heat equation it is possible to prescribe analytic data such that the corresponding solution will not be analytic (and in this case we are not talking about the Cauchy problem, as seen, but this illustrates the point).
- To talk about continuous dependence on the data, we need to be more specific about how to quantify "small variations" or "convergence". This will depend on the function space X we work on. For example, if X has a norm then we can consider convergence in that norm.

Example 3.2. The wave equation is LWP in analytic spaces. The Laplacian is not LWP in analytic spaces, as seen. This is why, for the Laplacian, one does not consider the Cauchy problem, prescribing $\psi|_\Sigma$ and $\partial\psi|_\Sigma$ but rather the Dirichlet problem (prescribing only $\psi|_\Sigma$) or the Neumann problem (prescribing only $\partial\psi|_\Sigma$). (One can formulate a notion of LWP for the Dirichlet / Neumann problems for the Laplacian and show that they are LWP in most spaces of interest).

4. Function spaces

As discussed, we could like to consider more general classes of functions than analytic functions. Consider a domain $\Omega \subset \mathbb{R}^n$ (possibly $\Omega = \mathbb{R}^n$).

We denote (these are all vector spaces):

- $C^\omega(\Omega)$ = space of analytic functions on Ω.

- $G^s(\Omega)$ = space of Gevrey functions of class s, s a real number ≥ 1. This is a generalization of analytic functions that retain many good properties of analytic functions. Move precisely, $\psi : \Omega \to \mathbb{R}$ belongs to $G^s(\Omega)$ if it is smooth and for every compact set $K \subset \Omega$ there exists a constant C such that for all $k = 0, 1, 2, \ldots$ and all $x \in \Omega$.
$$|\partial^k \psi(x)| \leq C^{k+1}(k!)^s.$$
What all this means is that ψ can be expressed as a convergent (on some domain) series (here, in 1D for simplicity)
$$\psi(x) = \sum_{k=0}^{\infty} \frac{\partial^k \psi(0)}{(k!)^s} x^k$$

We see that $G^1(\Omega) = C^\omega(\Omega)$. The point of using $G^s(\Omega)$ is that Gevrey function admit convergent series as above, which is a powerful tool, but they are not as rigid as analytic functions. E.g., $G^s(\Omega), s > 1$, admits bump functions which vanish outside a compact set.

- $C^\infty(\Omega)$ = space of smooth functions on Ω
- $C^k(\Omega)$ = space of k-times continuously differentiable functions on Ω. We can also consider the subspace $C^k_B(\Omega)$ of $C^k(\Omega)$ of functions with all derives up to order k uniformly bounded. $C^k_B(\Omega)$ is a normed space with norm
$$\|\psi\|_{C^k_B(\Omega)} := \max_{l=0,\dots,k} \sup_{x\in\Omega} |\partial^l \psi(x)|.$$
- $H^s(\Omega)$ = space of functions on Ω whose derivatives up to order s are square integrable. $H^s(\Omega)$ is a normed space with norm
$$\|\psi\|_{H^s(\Omega)} := [\sum_{l=0}^s \int_\Omega |\partial^l \psi(x)|^2 \, dx]^{\frac{1}{2}}.$$

(Recall that $\partial^l\psi$ is abbreviation for the set of all derivatives of ψ of order l. So, e.g.,
$$|\partial^2\psi|^2 = |\frac{\partial^2\psi}{\partial(x^0)^2}|^2 + |\frac{\partial^2\psi}{\partial x^0 \partial x^1}|^2 + \dots + |\frac{\partial^2\psi}{\partial(x^1)^2}|^2 + \dots)$$
When $s = 0$, we have
$$\|\psi\|_{H^0(\Omega)} = (\int_\Omega |\psi(x)|^2 \, dx)^{\frac{1}{2}}$$
which is the space of square integrable functions on Ω that we know from quantum mechanics. Thus, $H^s(\Omega)$ is simply an extension where we now require the derivatives to be square integrable.

We note that $H^s(\Omega)$ is a Hilbert space with inner product
$$\langle \psi, \phi\rangle_{H^s(\Omega)} = [\sum_{l=0}^s \int_\Omega \partial^l\psi(x)\partial^l\phi(x) \, dx]^{\frac{1}{2}}$$
(replace ϕ by $\bar\phi$ for complex-valued functions). When $\Omega = \mathbb{R}^n$, we can equivalently write
$$\|\psi\|_{H^s(\mathbb{R}^n)} = [\int_{\mathbb{R}^n} (1 + |\xi|^2)^s |\hat\xi(\psi)|^2 \, d\xi]^{\frac{1}{2}}$$
where $\hat\psi$ is the Fourier transform of ψ.

(Technical note: the derivatives $\partial^l\psi$ in the definition of H^s are distributional derivatives.)

Morally, we should think of the above spaces in increasing generality, i.e.,
$$C^\omega(\Omega) \subset C^s(\Omega) \subset C^\infty(\Omega) \subset C^k_B(\Omega) \subset H^s(\Omega) \subset L^2(\Omega)$$
We say "morally" because these inclusions are not true in general. For example,
$$1 \in C^\omega(\mathbb{R}) \text{ but } 1 \notin L^2(\mathbb{R})$$
or
$$\tan x \in C^\infty((\frac{-\pi}{2}, \frac{\pi}{2})) \text{ but } \tan x \notin C^k_B((\frac{-\pi}{2}, \frac{\pi}{2}))$$
However, as we will discuss, the class of PDEs we will ultimate consider will enjoy causality, which allows us to localize the problem within closed and bounded sets for which the above inclusions do hold. For example, for any $[a, b] \in \mathbb{R}$,
$$1 \in C^\omega([a, b]) \text{ and } 1 \in L^2([a, b])$$

or, for any $[x_0+\epsilon, x_0-\epsilon] \subset (\frac{-\pi}{2}, \frac{\pi}{2})$,

$$\tan x \in C^\omega([x_0+\epsilon, x_0-\epsilon]) \text{ and } \tan x \in C^k_B([x_0+\epsilon, x_0-\epsilon])$$

Some of the above inclusions also work in a **density** sense. To illustrate what we mean, consider $C^\infty(\mathbb{R})$ and $H^s(\mathbb{R})$. $C^\infty(\mathbb{R}) \not\subset H^s(\mathbb{R})$ since $1 \in C^\infty(\mathbb{R})$ but $1 \notin H^s(\mathbb{R})$. But given any $\psi \in H^s(\mathbb{R})$ we can find a sequence $\{\psi_k\}_{k=1}^\infty \subset C^\infty(\mathbb{R})$ that converges to ψ in the $H^s(\mathbb{R})$ norm, i.e.,

$$\lim_{k\to\infty} \|\psi_k - \psi\|_{H^s(\Omega)} = 0$$

I.e., we can approximate any $\psi \in H^s(\Omega)$ by smooth functions (with approximation quantified in the H^s norm). We say, for this approximation property, that "$C^\infty(\mathbb{R})$ is dense in $H^s(\Omega)$." I.e., although,

$$C^\infty(\mathbb{R}) \not\subset H^s(\mathbb{R})$$

we do have,

$$C^\infty(\mathbb{R}) \cap H^s(\Omega) \subset_{dense} H^s(\Omega)$$

(Of course, the inclusion $C^\infty(\mathbb{R}) \cap H^s(\Omega) \subset H^s(\Omega)$ is trivial: it is the "dense" part that is non-trivial.) Similar density holds for many general domains Ω.

(Technical note: It turns out that for many PDE techniques we need density of $C^\infty(\Omega)$ in $H^s(\Omega)$ and not $C^\infty(\Omega) \subset H^s(\Omega)$)

Example 4.1. To illustrate the importance of choice of function space, consider the heat equation

$$-\partial_t\psi + \partial_x^2\psi = 0,\ t > 0,$$
$$\psi|_{t=0} = 0.$$

$\psi \equiv 0$ is a solution. Consider

$$\phi(t,x) = \sum_{h=0}^{\infty} \frac{f^k(t)}{(2k)!} x^{2k},$$

where

$$f(t) = \begin{cases} e^{-t^{-l}} & ,t > 0 \\ 0 & ,t \le 0 \end{cases}$$

and $l > 1$ is a constant. One can show that ϕ is a smooth function (there is a constant C_l such that $|f^{(k)}(t)| \le \frac{k!}{(C_l t)^k} e^{\frac{-1}{2}t^{-l}}$) that solves the heat equation, $\phi(0,x) = 0$, but for $t > 0$

$$\phi(t,0) = f(t) \ne 0.$$

Thus, solutions to the initial-value problem for the heat equation are not unique in the C^∞ class. (This example is by Tychonoff.) However, if we consider the problem in $H^s(\mathbb{R}), s \ge 2$, then uniqueness holds.

4.1. **Hyperbolic equations.** We will investigate classes of PDE, for which LWP can be established. More precisely, we will focus on C^∞ and H^s, which are standard spaces for investigation of LWP. C^∞ is a natural space to consider; we can differentiate at will but C^∞ is significantly more general than C^ω or G^s. H^s is sufficiently general to include many interesting functions, while still having many good properties (e.g. it is a Hilbert space).

We will also consider only the case when data is given on a non-characteristic hypersurface Σ. We have seen that if Σ is characteristic we cannot solve the Cauchy problem in the class of analytic functions; if we cannot solve it in the most restrictive function space, there is no hope of solving it in more general spaces (recall our inclusions of spaces). Finally, we will assume throughout that the operators, i.e. the matrices $M^{\alpha_1,\dots,\alpha_m}(\partial^{m-1}\psi)$ etc. are smooth functions of their arguments.

Definition 4.2. Consider the operator

$$P\psi = M^{\alpha_1,\dots,\alpha_m}(\partial^{m-1}\psi)\partial_{\alpha_1}, \partial_{\alpha_n}\psi + F(\partial^{m-1}\psi)$$

where $M^{\alpha_1,\dots,\alpha_m}$ is $N \times N$. P is called **strictly hyperbolic** if

$$\det(M^{\alpha_1,\dots,\alpha_m}(\partial^{m-1}\psi))\xi_{\alpha_1},\dots,\xi_{\alpha_m} = 0$$

admits mN real distinct solutions $\xi_0 = \xi_0(\xi_{\alpha_1},\dots,\xi_{\alpha_m})$ for each $\xi \neq 0$. (We are omitting certain technical aspects of the definition.)

Remark 4.3. We note:

- We are making use of coordinates in singling out ξ_0. It is possible to give an invariant definition of strictly hyperbolic operators, but we do not do it here for simplicity.
- The definition allows for dependence of $M^{\alpha_1,\dots,\alpha_m}$ on $\partial^{m-1}\psi$. When we want to solve the Cauchy problem, we do not yet have ψ, but as in the case of the Cauchy-Kovalevskaya, we have $\psi|_\Sigma,\dots,\partial^{m-1}\psi\big|_\Sigma$, which allows ut to determine $M^{\alpha_1,\dots,\alpha_m}(\partial^{m-1}\psi\big|_\Sigma)$ and hence calculate the roots.
- $F(\partial^{m-1}\psi)$ does not enter in the definition of strict hyperbolicity.
- There are other notions of hyperbolicity (strong hyperbolic, symmetric hyperbolic, weak hyperbolic). They are share some important properties. We will make comments on other definitions of hyperbolicity later.

Example 4.4. Consider the below exapmles:

- $u^\mu\partial_\mu$ is strictly hyperbolic if $u^0 \neq 0$
- $g^{\mu\nu}\partial_\mu\partial_\nu$ is strictly hyperbolic
- $u^\mu g^{\alpha\beta}\partial_\mu\partial_\nu\partial_\beta$ is strictly hyperbolic u is timelike.
- $\Delta, \partial_t - \Delta$ are not strictly hyperbolic
- The relativistic Euler equations and Einstein's equation in wave coordinates are not strictly hyperbolic, but they will satisfy other definitions of hyperbolicity introduced below.

Theorem 4.5 (Leray)**.** *The Cauchy problem for a strictly hyperbolic operator with data given on a non-characteristic hypersurface is locally well-posed in C^∞ and in H^s for sufficiently large s.*

Remark 4.6. How large s has to be depends on details of $M^{\alpha_1,\dots,\alpha_m}$.

Example 4.7. aqui We obtain LWP for the strictly hyperbolic operators of the previous example.

Consider now operators with diagonal principle part, e.g., Einstein's equations in wave coordinates

$$\begin{bmatrix} g^{\mu\nu}\partial_{\mu\nu}g_{00} & 0 \\ \dots & \dots \\ 0 & g^{\mu\nu}\partial_{\mu\nu}g_{33} \end{bmatrix} = F(\partial g)$$

The characteristic determinant is the degree 20 polynomial $(g^{\mu\nu}\xi_\mu\xi_\nu)^{10}$ but the corresponding roots are repeated thus Leray's theorem does not apply. Such a repetition of the roots, however, is "spurious" in the sense that we are multiplying the same $g^{\mu\nu}\xi_\mu\xi_\nu$, whereas the diagonal structure of the equation suggests that we should treat each equation in the system separately. I.e., each

$$g^{\mu\nu}\partial_\mu\partial_\nu g_{\alpha\beta} = F_{\alpha\beta}(\partial g)$$

has a strictly hyperbolic principal part and the coupling among the equations is only at "lower order" (think from the point of view of the Cauchy-Kovalevskaya theorem: for each fixed $\alpha\beta$, we can determine $\partial_t^\alpha g_{\alpha\beta}|_\Sigma$ without appealing to other equations, i.e. we do not have to invert the a full matrix with all $\partial_t^2 g_{\alpha\beta}$, $\alpha, \beta = 0, \ldots, 3$).

Operators of this type are called **strictly hyperbolic operators in diagonal form**. We will not give the precise definition here because it is a bit cumbersome to define what coupling "only at lower order" means, but it essentially means that $M^{\alpha_1,\ldots,\alpha_m}(\partial^{m-1}\psi)$ is diagonal with each diagonal entry being strictlt hyperbolic.

We now state:

Theorem 4.8 (Leray)**.** *The Cauchy problem for a strictly hyperbolic operator in diagonal from with data given on a non-characteristic hypersurface is locally well-posed in C^∞ and in H^s for sufficiently large s.*

It follows that Einstein's equation in wave coordinates, are LWP in C^∞ and in H^s. From this we obtain a LWP result for the full (i.e., coordinate free) Einstein's equations. (In view of the diffeomorphism invariance of Einstein's equations, one needs to precisely formulate what is mean by the Cauchy problem for the full Einstein equations, including dealing with the constraints. This was done by Choquet-Bruhat and Choquet-Bruhat and Geroch.)

Definition 4.9. Consider the first-order operator

$$P\psi = A^\alpha(\psi)\partial_\alpha\psi + B(\psi)$$

P is called **symmetric hyperbolic** if the matrices $A^\alpha(\psi)$ are symmetric and $A^0(\psi)$ is positive definite. P is called **strongly hyperbolic** if $A^0(\psi)$ is invertible and for each $\vec{\xi} \neq 0$, the eigenvalue problem

$$((A^0)^{-1}A^i\xi_i - \lambda I)V = 0$$

has only real eigenvalues λ and a complete set of eigenvectors V.

Remark 4.10. Once again we are singling out ξ_0 but these definitions can be given invariant meaning.

Example 4.11. Assume that the sound speed c_s satisfies $0 < c_s \leq 1$. Then:

- The relativistic Euler equations are strongly hyperbolic.
- Upon multiplication by a suitable matrix, the relativistic Euler equations can be written as a first-order symmetric hyperbolic system.
- Moreover, upon introducing an evolution equation for the vorticity, one can rewrite the relativistic Euler equations as a strictly hyperbolic system in a diagonal form.

Example 4.12. We have:

- The BDNK equations, suitably rewritten as a first-order system, are strongly hyperbolic (under suitable conditions on the transport coefficients).
- The ideal MHD system (under suitable conditions, e.g., $0 < c_s \leq 1$) is strongly hyperbolic.
- The IS equations with only bulk viscosity, with $0 < c_s \leq 1$, and conditions on ζ and τ_Π, is a first-order symmetric hyperbolic system.
- The IS-like bulk theory by Jorge and Lorenzo is a first-order symmetric hyperbolic system with conditions on ζ and τ_Π.

Theorem 4.13 ((various authors, but see in particular Kato)**.** *The Cauchy problem for first-order symmetric hyperbolic operators and for strongly hyperbolic operators with data given on a non-characteristic surface is locally well-posed in C^∞ and in H^s (s large.)*

Considering the above results, we obtain LWP (in C^∞ and H^s) for:

- The relativistic Euler equations.
- The Einstein-Euler system.
- The BDNK equations.
- The BDNK-Einstein system.
- The ideal MHD system.
- The GRMHD (Einstein coupled to ideal MHD system).
- The IS equations with only bulk viscosity.
- The IS equations with only bulk viscosity coupled to Einstein.
- The IS-like theory by Jorge & Lorenzo.
- The IS-like theory by Jorge & Lorenzo coupled to Einstein.

(In all cases for data satisfying conditions indicated in the above examples.)

Finally, we mention that a **weakly hyperbolic** operator is, roughly, one that can be written as a product of strictly hyperbolic operators. The precise definition is a bit technical so we avoid it here, but state:

Theorem 4.14 (Leray, Ohya)**.** *Weakly hyperbolic operators are LWP in Gevrey spaces.*

Example 4.15. The DNMR equations (without current) are weakly hyperbolic under suitable conditions, hence LWP in Gevrey spaces (the result also holds for coupling with Einstein).

5. Causality

We are all familiar with the notion of causality in relativity, which says that no information propagates faster than the speed of light. We would like to formulate this notion in PDE terms and investigate PDEs for which it holds.

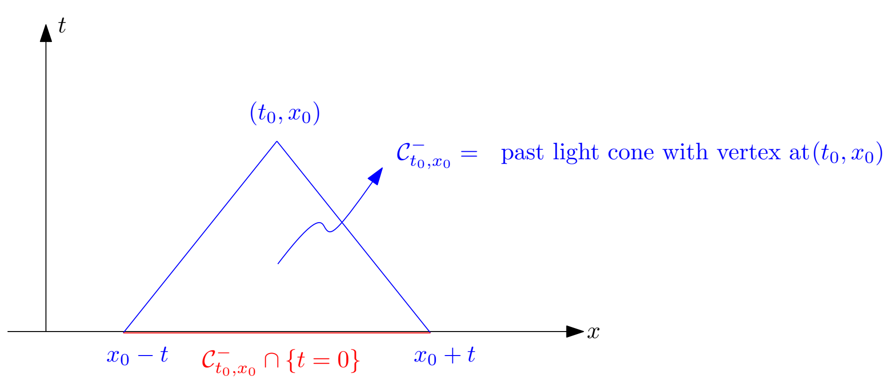
Consider the 1D wave equation

$$\begin{aligned} -\partial_t^2\psi + \psi_x^2\psi &= 0, \\ \psi|_{t=0} &= \psi_0, \\ \partial_t\psi|_{t=0} &= \psi_1. \end{aligned}$$

D'Alembert's formula give,

$$\psi(t,x) = \frac{\psi_0(x+t)+\psi_0(x-t)}{2} + \frac{1}{2}\int_{x-t}^{x+t}\psi_1(y)dy$$

We see that at (t,x), ψ is completely determined by the data on the past light cone with vertex at (t,x) intersected with $\{t=0\}$.

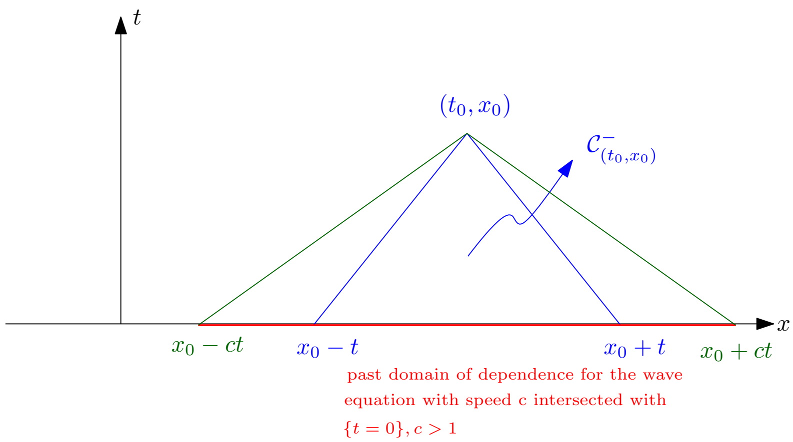


To say that ψ at (t,x) depends only on its values at $\mathcal{C}(t,x)^- \cap \{t=0\}$ is the statement that information propagates within the light cone.

For a wave equation with speed c, we have

$$\psi(t,x) = \frac{\psi_0(x+ct)+\psi_0(x-ct)}{2} + \frac{1}{2c}\int_{x-ct}^{x+ct} \psi_1(y)dy$$

and the same situation holds except that the light cone is replaced by the triangle with sides of slope $\pm\frac{1}{c}$ in the (t,x)-plane. We call such triangle the **past domain of dependence** of (t,x) (for the wave equation with speed c we see that the solution depends only on the past domain of dependence of (t,x) intersected with $\{t=0\}$). Thus, information propagates at finite speed a but causality will hold only when $c \leq 1$. Note that the lines $x \pm ct =$ constant are the characteristics of the wave equation.



We see that we need to distinguish between PDE, where information propagates at finite speed from those where information propagates causally. It is easier to first understand the former, since its requirements are less stringent, and then ask when the latter holds.

One of our main tasks is to understand how to generalize the above arguments without appealing to explicit formulas, since these are not available for general PDEs.

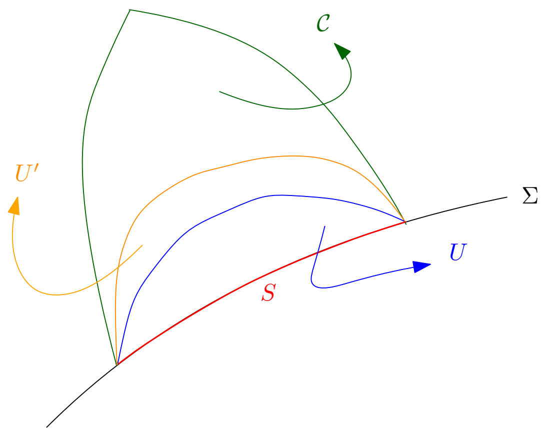
Remark 5.1. In investigating these questions, we will be asking about properties of solutions. The question of existence of solutions was investigated above. So, we can assume a given solution and ask when information propagates at finite speed. The resulting conclusions will be properties of the operators under investigation because they will hold for general solutions.

Unless otherwise specified, we will be working with operators of the form

$$P\psi = P(\partial^{m-1})\psi = M^{\alpha_1,\dots,\alpha_m}(\partial^{m-1}\psi)\partial_{\alpha_1},\dots,\partial_{\alpha_m}\psi + F(\partial^{m-1}\psi)$$

5.1. **Domains of dependence.** Consider a solution to $P\psi = 0$ in a neighborhood of a non-characteristic surface Σ. Suppose that we are in the analytic setting.

Assume that $\psi \equiv \cdots \equiv \partial^{m-1}\psi \equiv 0$ on $S \subset \Sigma$. Then, by the Cauchy-Kovalevskaya theorem, ψ vanishes on a small neighborhood U of S with "ends" on S. This notion can be made precise, but we illustrate it with a picture instead:



Now take the "upper boundary" of U. $\psi = 0$ there (technical note: we have to be a bit careful to justify this because in principle $\psi = 0$ in the open set that is the interior of u but this can be done by a continuity argument). If the upper boundary of U is not characteristic, we can repeat the argument, obtaining that ψ vaniehs on a larger region U' with ends on S. We continue this process until we reach a region $\mathcal{C}$ whose upper boundary is characteristic . We have $\psi = 0$ in $\mathcal{C}$. $\mathcal{C}$ plays a role analogous to the light cone with base on S. If $\psi|_S = 0$, then $\psi = 0$ in $\mathcal{C}$.

If we now consider the vertex" q of $\mathcal{C}$, which is the point whose characteristics emanating from q toward S are precisely the upper boundary of $\mathcal{C}$, then we can state that $\psi(q)$ **depends only its values on** $\mathcal{C} \cap \Sigma = S$, meaning that the operator P propagates information at finite speed.

In order to see that this is indeed the case, we need to understand what it means to say that $\psi(q)$ depends only on its value on S. It means that if we have two solutions ψ_1 and ψ_2 that agree, with their derivatives, i.e., the data induced on S satisfies

$$\psi_1|_S = \psi_2|_S, \dots, \partial^{m-1}\psi_1\big|_S = \partial^{m-1}\psi_2\big|_S$$

then

$$\psi_1(q) = \psi_2(q).$$

But this can be reduced to considering zero data on S upon taking the difference $\psi_1 - \psi_2$.

(Technical note: there are some points that need to considered more carefully. For example, since P is in principle quasilinear, we need to write $P(\partial^{m-1}\psi_1)\psi_1 - P(\partial^{m-1}\psi_2)\psi_2$ as a linear equation for $\psi_1 - \psi_2$, which can be done since ψ_1 and ψ_2 are assumed given. This difference will depend on $\partial^m\psi_1$ or $\partial^m\psi_2$, depending on how one writes it, with principal part

$$M^{\alpha_1,\dots,\alpha_m}(\partial^{m-1}\psi_1)\partial_{\alpha_1},\dots,\partial_{\alpha_m} \text{ or } M^{\alpha_1,\dots,\alpha_m}(\partial^{m-1}\psi_2)\partial_{\alpha_1},\dots,\partial_{\alpha_m},$$

again, depending on how one writes things. Both cases need to be considered and compared since we do not know yet that $\psi_1 = \psi_2$ in $\mathcal{C}$, so the characteristics for $M^{\alpha_1,\dots,\alpha_m}(\partial^{m-1}\psi_1)$ and $M^{\alpha_1,\dots,\alpha_m}(\partial^{m-1}\psi_2)$ at this stage cannot be assumed to coincide.)

The precise formulation of the above argument was given by Holmgren & John. We will not give it here because we will present a more general approach below, but we will make the following remarks:

Remark 5.2. We note that:

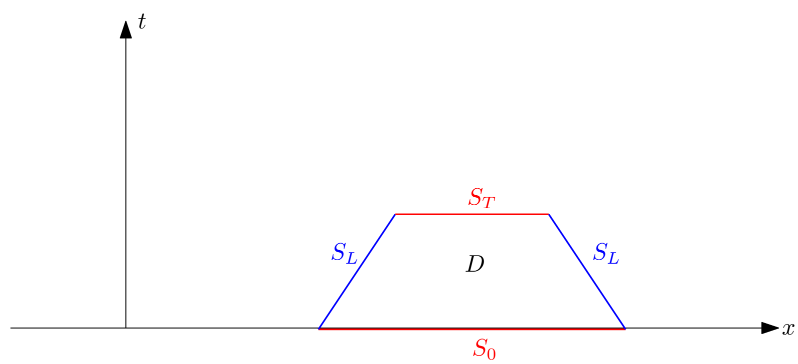
- Holmgren's and John's theorems are for linear equations, but as we assumed ψ known, $P(\partial^{m-1}\psi)\psi$ can be viewed as the linear operator $P(\partial^{m-1}\psi)$ acting on ψ.
- We invoked the Cauchy-Kovalevskaya theorem to conclude $\psi = 0$. Recall, however, that the Cauchy-Kovalevskaya theorem does exclude non-uniqueness outside the analytic class. Holmgren's and John's theorem establish uniqueness outside analytic functions as well, but still require some assumptions of analyticity.

We will give a criterion to identify when operators propagate information at finite speed. But the audience should keep in mind the above heuristic argument as a guide.

In order to gain intuition, let us once again appeal to the wave equation. Our goal is to devise an argument that does not depend on explicit formulas for the solution. Consider:

$$\begin{aligned} -\partial_t^2 + \partial_x^2\psi &= 0, \\ \psi|_{t=0} &= \psi_0, \\ \partial_t\psi|_{t=0} &= \psi_1. \end{aligned}$$

Multiply the equation by $\partial_t\psi$, integrate over a region $\mathcal{D}$ bounded by two slabs $S_T \subset \{t = T\}$ and $S_0 \subset \{t = 0\}$ and a lateral boundary S_L, and integrate by parts



$$\begin{aligned} \int_{\mathcal{D}} \partial_t\psi(-\partial_t^2\psi + \partial_x^2\psi) = &\int_{\mathcal{D}} (-\frac{1}{2}\partial_t(\partial_t\psi)^2 + \partial_t\psi\partial_x^2\psi) \\ &= -\frac{1}{2}\int_{S_T} \nu^0(\partial_t\psi)^2 - \frac{1}{2}\int_{S_0} \nu^0(\partial_t\psi)^2 - \frac{1}{2}\int_{S_L} \nu^0(\partial_t\psi)^2 \\ &\quad - \int_{\mathcal{D}} \underbrace{\partial_t\partial_x\psi\partial_x\psi}_{=\frac{1}{2}\partial_t(\partial_x\psi)^2} + \int_{S_T} \partial_t\psi\nu^1\partial_x\psi + \int_{S_0} \partial_t\psi\nu^1\partial_x\psi + \int_{S_L} \partial_t\psi\nu^1\partial_x\psi \\ &= -\frac{1}{2}\int_{S_T} \nu^0(\partial_t\psi)^2 - \frac{1}{2}\int_{S_0} \nu^0(\partial_t\psi)^2 - \frac{1}{2}\int_{S_L} \nu^0(\partial_t\psi)^2 \\ &\quad - \frac{1}{2}\int_{S_T} \nu^0(\partial_x\psi)^2 - \frac{1}{2}\int_{S_0} \nu^0(\partial_x\psi)^2 - \frac{1}{2}\int_{S_L} \nu^0(\partial_x\psi)^2 \\ &\quad + \int_{S_T} \partial_t\psi\nu^1\partial_x\psi + \int_{S_0} \partial_t\psi\nu^1\partial_x\psi + \int_{S_L} \partial_t\psi\nu^1\partial_x\psi \end{aligned}$$

where $\nu = (\nu^0, \nu^1)$ is the unit outer normal to $S_T \cup S_0 \cup S_L$. $\nu^0 = 1$ on S_L and $\nu^0 = -1$ on S_0, and $\nu^1 = 0$ on $S_T \cup S_0$ Thus:

$$\begin{aligned} \frac{1}{2}\int_{S_T} (\partial_t\psi)^2 + \frac{1}{2}\int_{S_T} (\partial_x\psi)^2 = &\frac{1}{2}\int_{S_0} (\partial_t\psi)^2 + \frac{1}{2}\int_{S_0} (\partial_x\psi)^2 \\ &- \frac{1}{2}\int_{S_L} \nu^0(\partial_t\psi)^2 - \frac{1}{2}\int_{S_L} \nu^0(\partial_x\psi)^2 + \int_{S_L} \partial_t\psi\nu^1\partial_x\psi \end{aligned}$$

Let us now assume that the boundary S_L is characteristic, i.e., a light cone boundary. Then $\nu = \frac{1}{\sqrt{2}}(1, \pm 1)$ and the Cauchy-Schwarz inequality gives

$$-\frac{1}{2}(\partial_t\psi)^2 - \frac{1}{2}(\partial_x\psi)^2 + \partial_t\psi\partial_x\psi \leq -\frac{1}{2}(\partial_t\psi)^2 - \frac{1}{2}(\partial_x\psi)^2 + \frac{1}{2}(\partial_t\psi)^2 + \frac{1}{2}(\partial_x\psi)^2 \leq 0$$

On the right part of S_L and similarly on the left part of S_L. Thus

$$\begin{aligned}\frac{1}{2}\int_{S_T}(\partial_t\psi)^2+\frac{1}{2}\int_{S_T}(\partial_x\psi)^2 &\leq \frac{1}{2}\int_{S_0}(\partial_t\psi)^2+\frac{1}{2}\int_{S_0}(\partial_x\psi)^2\\ &= \frac{1}{2}\int(\psi_0^2+\psi_1^2)\end{aligned}$$

If $\psi_0=\psi_1=0$ we find

$$\frac{1}{2}\int_{S_T}(\partial_t\psi)^2+\frac{1}{2}\int_{S_T}(\partial_x\psi)^2\leq 0.$$

Thus,

$$\partial_t\psi|_{S_0}=\partial_x\psi|_{S_T}=0\Rightarrow\psi\equiv 0 \text{ on } S_T.$$

This is another way of verifying causality for the wave equation. Passing to the limit when S_T converges to the vertex of light cone we recover the statement previously obtained with D'Alembert's formula.

Observe that this was obtained **without imposing boundary conditions on S_L. We have used only the structure of the wave operator and that S_L is characteristic, i.e., that S_L is null.** The same remains true if S_L is spacelike. However, if S_L is timelike, then boundary conditions would be needed (e.g., if $S_L=\{x=\text{constant}\}$ then the boundary integrals become $\pm\int_{S_L}\partial_t\psi\partial_x\psi$, which a priori does not have a sign, so we need Dirichlet or Neumann boundary conditions). We will not consider boundary conditions here.

We want to generalize the above argument, so the idea is to identify which properties an operator needs to produce a quadratic form like $(\partial_t\psi)^2+(\partial_x\psi)^2$ and boundary terms with a sign. We anticipate from the foregoing discussion that the "lateral" boundary will be given by characteristics.

Consider a strictly hyperbolic operator:

$$P(\partial^{m-1}\psi)\psi=M^{\alpha_1,\ldots,\alpha_m}(\partial^{m-1}\psi)\partial_{\alpha_1},\ldots,\partial_{\alpha_m}\psi+F(\partial^{m-1}\psi)$$

and the characteristic determinant

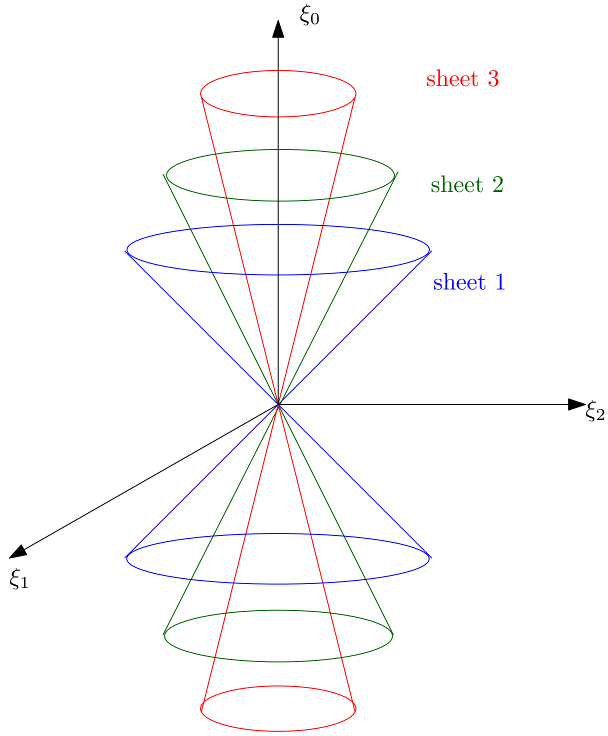
$$M^{\alpha_1,\ldots,\alpha_m}(\partial^{m-1}\psi)\xi_{\alpha_1}\ldots\xi_{\alpha_m}=0$$

Leray proved that (in spacetime dimensions at least 3), the roots $\xi_0=\xi_0(\vec{\xi})$ define a family of sheets $(\xi_0(\vec{\xi}),\vec{\xi})$ whose interiors have a common intersection that forms the interior of two opposite half-cones $\Gamma^{*,+},\Gamma^{*,-}$. (Once again, we can define these notions without singling out ξ_0.)

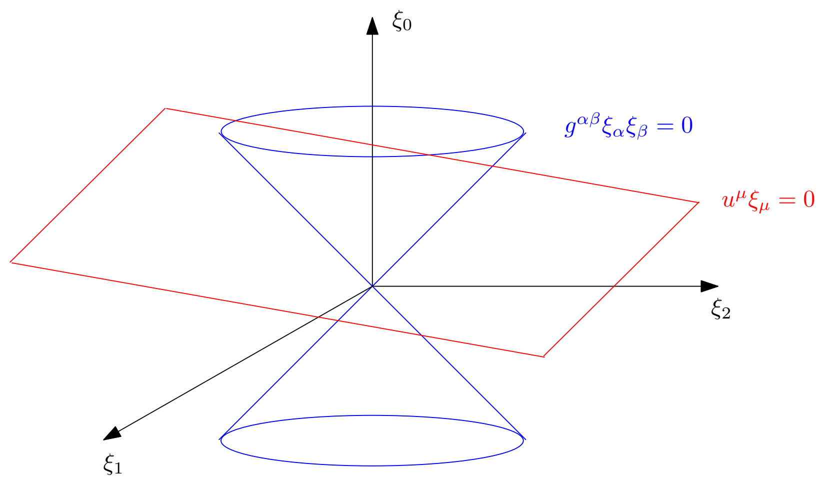
(The restriction on the dimension can be understood from the $1+1$ wave equation:

$$-\partial_t^2\psi+\partial_x^2\psi=0$$

allows us to switch t and x.)



Example 5.3. Consider $u^\mu g^{\alpha\beta}\partial_\mu\partial_\alpha\partial_\beta$, where u timelike. The common interior is the light cone. (The cone associated with $u^\mu \xi_\mu$ is composed of the two half spaces separated by $u^\mu\xi_\mu = 0$)



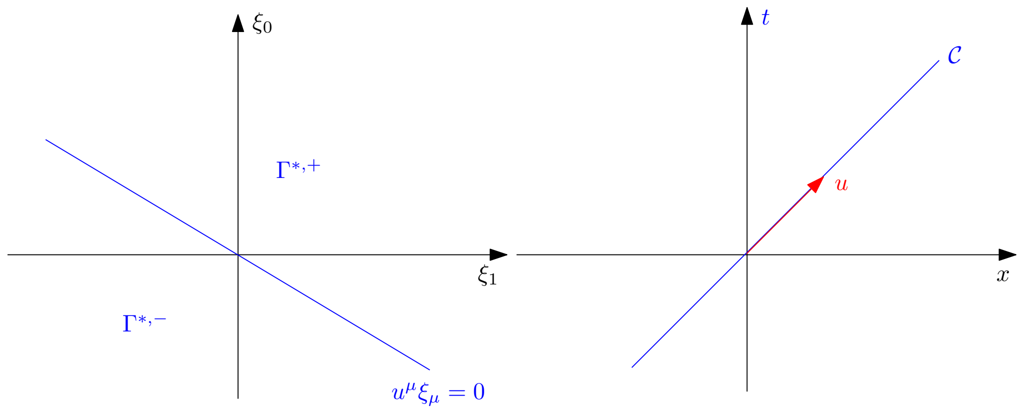
Leray also proved the following. Define $\mathcal{C}^\pm$ as the set of vectors v such that $\xi(v) \geq 0$ for all $\xi \in \Gamma^{*,\pm}$. Then, $\mathcal{C}^\pm$ also forms two opposite half cones. Set $\mathcal{C} = \mathcal{C}^+ \cup \mathcal{C}^-$. It follows that the **boundary of $\mathcal{C}$ is characteristic**.

Example 5.4. Consider $u^\mu\partial_\mu$. Then

$$u^\mu\xi_\mu = u^0\xi_0 + u^i\xi_i = 0,\ \xi_0 = -\frac{u^i\xi_i}{u^0}$$

$$\xi(v) = \xi_0 v^0 + \xi_i v^i = -\frac{v^i\xi_i}{u^0}v^0 + \xi_i v^i = (-\frac{u^i}{u^0}v^0 + v^i)\xi_i \geq 0$$

for all $\xi \in \Gamma^{*,\pm} \Rightarrow \frac{v^i}{v^0} = \frac{u^i}{u^0}$ i.e., v is parallel to u. Thus the cose $\mathcal{C}^{\pm}$ is the line through u (as in previous examples, this should be viewed as a degenerate cone). This is another way of seeing that the characteristics of $u^\mu \partial_\mu$ are the flow lines of u.



We define a vector v to be **timelike/null-like/spacelike with respect to the operator** P, or simply P**-timelike/null-like/spacelike**, if v belongs to the interior/boundary/exterior of $\mathcal{C}$.

A curve is said to be timelike/null-like/spacelike with respect to P, or simply P-timelike/null-like/spacelike, if its tangent vector is P-timelike/null-like/spacelike. A vector/curve is **causal with respect to** P**,** or simply P**-causal**, if it is P-timelike or P-causal.

With these notions, we can define a causal structure on spacetime, but this is a **causal structure determined by** P **without reference to a metric.** As we will see shortly, the operator P **will be causal in the relativity sense when the causal structure determined by** P **is compatible with the causal structure given by the spacetime metric.**

In particular, we can define the **past domain of dependence of a point** q **with respect to** P (or P-past domain of dependence), denoted $J_P^-(q)$, as the set of past causal curves from q. (It is legitimate to talk about past/future because the above notions also allow us to define a time orientation with respect to P.)

With these notions, we can now carry out an argument similar to that done for the wave equation, integrating $P\psi = 0$ multiplied by a suitable combination of derivatives of ψ over a spacetime region with lateral boundaries belonging to $J_P^-(q)$. It is possible to show that a similar result holds: one obtains the integral of a definite quadratic from plus boundary terms that have the correct sign to produce an inequality similar to that of the wave equation. We obtain:

Theorem 5.5 (Leray)**.** *For solutions to* $P(\partial^{m-1}\psi)\psi = 0$*,* $\psi(q)$ *depends only on* $J_P^-(q)$*. In particular, in the Cauchy problem,* $\psi(q)$ *is completely determined by the Cauchy data on* $J_P^-(q) \cap \Sigma$*, where* Σ *is a non-characteristic.*

This means that the operator P propagates information at finite speed (this is often referred to as saying that P has the domain of dependence property or finite speed of propagation property). In order to know that P is causal in the relativistic sense we need to know that the finite speed at which P propagates information is less or equal than the speed of light. **This will be the case if and only if the cones** $\mathcal{P}$ **associated with the operator** P **are inside (or coincide with) light cones.**

We would like a **simple criterion to determine whether** P **is causal.** The criterion is the following. Consider the roots $\xi_0 = \xi(\vec{\xi})$ of the characteristic determinant

$$\det(M^{\alpha_1,\ldots,\alpha_m}(\partial^{m-1}\psi)\xi_{\alpha_1}\ldots\xi_{\alpha_m} = 0$$

P will be causal if and only if the sheets $(\xi_0(\vec{\xi}), \vec{\xi})$ are outside or coincide with the light cone, i.e., **if and only if the roots** $\xi = (\xi_0(\vec{\xi}), \vec{\xi})$ **are spacelike or null-like.**

The requirement that the sheets be outside or coincide with the light cone, i.e., $\xi = (\xi_0(\vec{\xi}), \vec{\xi})$ be spacelike or null-like, is because we are looking at cones and sheets in the variable ξ (i.e., in momentum space). The duality $\xi(v) = 0$ that determines the cone $\mathcal{C}$ in physical space takes sheets outside the light cone in ξ space to sheets inside the light cone in **physical** space; see the above example of $u^\mu \partial_\mu$ for the extreme case where the half spaces $\Gamma^{*,\pm}$ are mapped onto a line. This property is not seen with the light cone itself because it is "self dual", i.e. the light cone in ξ is mapped to a light cone in physical space.

We finish mentioning that the above domains of dependence theorem and causality result also holds for strictly hyperbolic operators in diagonal form, weakly hyperbolic operators, strongly hyperbolic operators, and first-order symmetric hyperbolic operators.

Remark 5.6. We have not discussed mixed-order systems, i.e. systems where, say, some equations involve third-order derivatives whereas other equations involve only first-order derivatives. In fact, some of the examples mentioned above are mixed-order (e.g. Einstein-Euler involves second- and first-order PDEs). But the results we stated can be generalized to such cases, and we implicitly used this when stating results for mixed-order systems in the above examples. For more details, including how to correctly identify the principal part of mixed-order systems, see my paper "On the existence of solutions and causality for relativistic viscous conformal fluids."

6. Conclusion

In sum, the LWP and causility of systems of PDEs is intrinsically tied to its characteristic. There are, in fact, many other properties (e.g. global existence of solutions, formation of shocks) that are tied to the behavior of the characteristics of hyperbolic operators.