

PDE 1: Symmetric Hyperbolic Systems

Scribed by Ethan Lu

December 20, 2021

Recall the following PDEs:

1. **The transport/reaction equation:** $\partial_t u + a \cdot \nabla u + bu = f$.
2. **The wave equation:** $\partial_t^2 u - \Delta u = f$.
3. **The general wave equation:** $\alpha \partial_t^2 u + \kappa \partial_t u - A : D^2 u + b \cdot \nabla u + cu = f$, where the operator associated to A, b, c is assumed to be elliptic.
4. **Maxwell's equations:**

$$\begin{cases} \operatorname{div} B = 0 \\ \operatorname{div} E = \rho \\ \partial_t E + J = \operatorname{curl} B \\ \partial_t B = -\operatorname{curl} E \end{cases}$$

where $E, B : \mathbb{R}^3 \times \mathbb{R} \rightarrow \mathbb{R}^3$ are vector fields describing the electric/magnetic field and J is the electric current density.

There are all actually part of the same class of PDEs! Here are some clues to why this might be true:

1. In 1 dimension, $(\partial_t^2 - \partial_x^2) = (\partial_t - \partial_x)(\partial_t + \partial_x)$, which establishes a connection between the transport and wave equations.
2. In Maxwell's equations, applying the divergence operator to both sides of the third equation yields

$$\partial_t \operatorname{div} E + \operatorname{div} J = \partial_t \rho + \operatorname{div} J = 0$$

which is a balance law that essentially says that the current is the flux vector associated to charge. Now recall that $\Delta = -\operatorname{curl}^2 + \nabla \operatorname{div}$ for \mathbb{R}^3 valued vector fields.

Applying the curl also yields

$$= \partial_t \operatorname{curl} E + \operatorname{curl} J = \operatorname{curl}^2 B = \operatorname{curl}^2 B - \nabla \operatorname{div} B = -\Delta B$$

which implies

$$\partial_t^2 B - \Delta B = \operatorname{curl} J$$

and similarly,

$$\partial_t^2 E - \Delta E = -\partial_t J - \nabla \rho.$$

Now write $\text{curl } X = \sum M_i \partial_i X$ where M_i are the appropriately chosen antisymmetric 3x3 matrices. Now note

$$\partial_t \begin{pmatrix} E \\ B \end{pmatrix} = \partial_t \begin{pmatrix} \text{curl } B \\ -\text{curl } E \end{pmatrix} + \begin{pmatrix} -J \\ 0 \end{pmatrix}$$

and also that

$$\begin{pmatrix} \text{curl } B \\ \text{curl } E \end{pmatrix} = \sum_j \begin{pmatrix} 0 & M_j \\ -M_j & 0 \end{pmatrix} \partial_j \begin{pmatrix} E \\ B \end{pmatrix}$$

hence the dynamic part of Maxwell's equations are just

$$\partial_t \begin{pmatrix} E \\ B \end{pmatrix} - \sum_j A_j \partial_j \begin{pmatrix} E \\ B \end{pmatrix} = \begin{pmatrix} -J \\ 0 \end{pmatrix}$$

which is just a vector-valued transport equation.

3. In the generalized wave equation, consider defining the function $U : \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}^{n+2}$ via

$$U = \begin{pmatrix} u \\ \partial_t u \\ \nabla u \end{pmatrix} \implies \partial_t U = \begin{pmatrix} \partial_t u \\ \partial_t^2 u \\ \nabla \partial_t u \end{pmatrix}, \partial_j U = \begin{pmatrix} \partial_j u \\ \partial_t \partial_j u \\ \nabla \partial_j u \end{pmatrix}.$$

Then U satisfies

$$\begin{pmatrix} 1 & 0 & | & 0 \\ 0 & \alpha & | & A \\ 0 & 0 & | & A \end{pmatrix} \partial_t U + \sum_{j \in [n]} \begin{pmatrix} 0 & | & 0 \\ 0 & -a_{1j} & \cdots & -a_{nj} \\ -a_{1j} & & & \\ \cdots & & \mathbf{0} & \\ -a_{nj} & & & \end{pmatrix} \partial_j U + \begin{pmatrix} 0 & -1 & | & 0 & \cdots & 0 \\ c & \kappa & | & b_1 & \cdots & b_n \\ 0 & & | & 0 & & \end{pmatrix} \partial_t U = \begin{pmatrix} 0 \\ f \\ 0 \end{pmatrix}$$

which can be rewritten as

$$\sum_{i=0}^n A^i \partial_i U + BU = F,$$

where ∂_0 is understood to mean ∂_t , and $A^j \in \mathbb{R}_{sym}^{n \times n}$, $B \in \mathbb{R}^{n \times n}$. Note also that

$$A^0 \begin{pmatrix} \omega \\ \psi \\ \xi \end{pmatrix} \cdot \begin{pmatrix} \omega \\ \psi \\ \xi \end{pmatrix} = \omega^2 + \alpha \psi^2 + A \xi \cdot \xi \geq \min(1, \alpha, \theta) \left\| \begin{pmatrix} \omega \\ \psi \\ \xi \end{pmatrix} \right\|^2$$

so if $\alpha > 0$, A^0 is positive-definite, which is an assumption we'll be making from now on.

With this additional structure, we'll now define the problem we're interested in studying.

Definition. A *symmetric hyperbolic system* is the PDE

$$\begin{cases} \sum_{i=0}^n A^i \partial_i u + Bu = f \\ u(t=0) = g \end{cases}$$

for some unknown $u : \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}^m$, given data f, g , symmetric matrices $A^j : \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}^{m \times m}$, $B \in \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}^{m \times m}$, and where we assume there exists $\theta > 0$ with $A^0 \xi \cdot \xi \geq \theta \|\xi\|^2$ for all $\xi \in \mathbb{R}^n$.

Remark. When $m = 1$ this is just the transport/reaction equation, and can be attacked with the method of characteristics. If $m \geq 2$, the same approach works for the special case of when each A^j is just a multiple of the identity, but otherwise fails. As we're about to see, it turns out that the key to progress will be quantitative estimates.

We begin first with some formal a priori estimates (using Einstein notation for convenience). Suppose first that u is a solution to the problem stated above. Then taking a dot product and integrating, we have that

$$\int_{\mathbb{R}^n} A^0 \partial_t u \cdot u + A^j \partial_j u \cdot u + \int_{\mathbb{R}^n} Bu \cdot u = \int_{\mathbb{R}^n} f \cdot u.$$

Now note that

$$\partial_j \left(A^j \frac{u \cdot u}{2} \right) = \partial_j A^j \frac{u \cdot u}{2} + A^j \partial_j u \cdot u$$

so assuming sufficiently nice decay of $\partial_j \left(A^j \frac{u \cdot u}{2} \right)$ at infinity, we have

$$\int_{\mathbb{R}^n} A^j \partial_j u \cdot u = - \int_{\mathbb{R}^n} \partial_j A^j \frac{u \cdot u}{2}$$

Thus, we have that

$$\begin{aligned} \partial_t \int_{\mathbb{R}^n} A^0 \frac{u \cdot u}{2} &\leq \int_{\mathbb{R}^n} |f \cdot u| + |Bu \cdot u| + \left| \frac{\partial_j A^j u \cdot u}{2} \right| \\ &\leq \frac{2 \|B\|_{L^\infty} + \|\partial_j A^j\|_{L^\infty} + 1}{\theta} \int_{\mathbb{R}^n} A^0 \frac{u \cdot u}{2} + \int_{\mathbb{R}^n} \frac{|f|^2}{2} \\ &= C \int_{\mathbb{R}^n} A^0 \frac{u \cdot u}{2} + \int_{\mathbb{R}^n} \frac{|f|^2}{2} \end{aligned}$$

This in turn implies that

$$\int_{\mathbb{R}^n} A^0 \frac{u \cdot u}{2}(t) \leq e^{Ct} \int_{\mathbb{R}^n} A^0 \frac{u \cdot u}{2}(0) + \int_0^t e^{C(t-s)} \int_{\mathbb{R}^n} \frac{|f(s)|^2}{2}$$

so we get the formal a priori estimate that

$$u \in L^\infty([0, T]; L^2(\mathbb{R}^n; \mathbb{R}^m))$$

provided that $u(t=0) \in L^2$, $f \in L^2((0, T); L^2(\mathbb{R}^n; \mathbb{R}^m))$, $A^j \in C^{0,1}$, $B \in L^\infty$.

Recall now that we're trying to solve the system

$$\begin{cases} A^0 \partial_t u + A^j \partial_j u + Bu = f \\ u(t=0) = g \end{cases}$$

for $u : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^m$. Last time, we showed that we can derive L^2 -type estimates for solutions to this system over finite time intervals given sufficient niceness of the underlying conditions. Going forward then, we'll be working under the following assumptions on the data/coefficients of our system, which should be able to rigorously justify our previous calculations:

1. $T \in (0, \infty)$ is the "time horizon" we're interested in.
2. $A^i, B \in C_b^{0,1}(\mathbb{R}^n \times [0, T], \mathbb{R}^{m \times n})$, with A^i symmetric over space time.
3. A^0 is uniformly coercive; that is, there exists $\theta > 0$ such that $A^0(x, t) \geq \theta I$ for all $x \in \mathbb{R}^n, t \in [0, T]$.
4. $g \in H^1(\mathbb{R}^n; \mathbb{R}^m), f \in L^\infty([0, T], L^2(\mathbb{R}^n; \mathbb{R}^m)) \cap L^2([0, T], H^1(\mathbb{R}^n; \mathbb{R}^m))$.

Our previous approaches to solving PDEs involved pulling out a derivative using integration by parts, then using this computation to find an associated bilinear form that we could apply functional analytic tools to. It turns out in this setting, we'll be unable to find such a bilinear form or reason about coercivity in the same way, so we'll need different tools to approach this problem. A first observation to be made is then the fact that any solution satisfies

$$\partial_t u + Lu := \partial_t u + (A^0)^{-1} [A^j \partial_j u + Bu] = (A^0)^{-1} f$$

which looks formally like an ODE. The issue, however, is that the operator L is unbounded from any Banach spaces to itself. While this issue could somewhat be circumvented via an appeal to Fréchet spaces or unbounded operators, for our discussions, we'll be handling this issue through mollification. In particular, we'll be modifying the differential operator so that it becomes bounded on Sobolev spaces. (Aside: this exact problem is actually the original motivation behind Friedrichs' invention of mollification.)

Definition. Let $\eta \in C_c^\infty$ be a positive radial approximate identity, and for $\varepsilon > 0$ write $\eta_\varepsilon(x) = \varepsilon^{-n} \eta(x/\varepsilon)$, which remains a positive radial approximate identity. Now define $K_\varepsilon := u \mapsto \eta_\varepsilon * u$, noting that this defines a bounded linear operator from any Sobolev space to any other Sobolev space, implying that

$$u \in L^2(\mathbb{R}^n; \mathbb{R}^m) \implies K_\varepsilon u = \bigcap_{k \in \mathbb{R}} H^k(\mathbb{R}^n; \mathbb{R}^m) \subseteq C_0^\infty(\mathbb{R}^n; \mathbb{R}^m)$$

Definition. For $\varepsilon > 0$, we say u_ε is an **approximate solution** to the problem if it solves

$$\begin{cases} A^0 \partial_t u_\varepsilon + K_\varepsilon [A^j \partial_j K_\varepsilon u_\varepsilon] + Bu_\varepsilon = K_\varepsilon f \\ u_\varepsilon(t=0) = g \end{cases}$$

Remark. The pre-mollification terms inside the scheme above are necessary to preserve the structure of the estimates we've calculated previously.

As we'll show now, modifying the equations in this way then makes reasoning about solutions much easier. Multiplying the first equation above by the inverse of A^0 , we have that

$$\partial_t u_\varepsilon + M_\varepsilon(t) u_\varepsilon = (A^0)^{-1} K_\varepsilon f$$

for some $M_\varepsilon(t) \in \mathcal{L}(H^1(\mathbb{R}^n; \mathbb{R}^m))$, which reduces our problem to the H^1 -valued ODE

$$\dot{u}_\varepsilon + M(t)u_\varepsilon = F(t)$$

for which we can use the following theorem:

Theorem. *Let $T > 0$, X be a real Banach space, and suppose $L \in L^\infty([0, T]; \mathcal{L}(X))$, $x_0 \in X$. Then there exists a unique $x \in C_b^{0,1}([0, T]; X) \cap W^{1,\infty}((0, T); X)$ such that*

$$\begin{cases} \dot{x}(t) + L(t)x(t) = F(t) & \text{a.e. in time} \\ x(0) = x_0 \end{cases}$$

Remark. *A theorem in functional analysis shows that if X is separable, then $W^{1,\infty}((0, T); X) = C_b^{0,1}([0, T]; X)$.*

Proof. For $\kappa > 0$ define $Y_\kappa = C_b^0([0, T]; X)$ with the norm $\|x\|_{Y_\kappa} = \max_{t \in [0, T]} e^{-\kappa t} \|x(t)\|_X$, which is equivalent to the usual norm. Given $v \in Y_\kappa$, define $Rv : [0, T] \rightarrow X$ via

$$Rv(t) := x_0 + \int_0^t F(s) - L(s)v(s) ds$$

which is clearly in $Y_\kappa \cap C_b^{0,1}([0, T]; X)$. We're looking for a fixed point, so now we calculate

$$\begin{aligned} \|Rv(t) - Ru(t)\|_X &= \left\| \int_0^t L(s)(v(s) - u(s)) ds \right\|_X \\ &\leq \|L\|_{L^\infty} \int_0^t \|u(s) - v(s)\|_X ds \\ &= e^{\kappa t} \|L\|_{L^\infty} \int_0^t e^{-\kappa(t-s)} e^{-\kappa s} \|u(s) - v(s)\|_X ds \\ &= e^{\kappa t} \|L\|_{L^\infty} \|u - v\|_{Y_\kappa} \frac{e^{-\kappa t}}{\kappa} (e^{\kappa t} - 1) \\ &= e^{\kappa t} \frac{\|L\|_{L^\infty}}{\kappa} \|u - v\|_{Y_\kappa} \end{aligned}$$

This implies that

$$\|Ru - Rv\|_{Y_\kappa} \leq \frac{\|L\|_{L^\infty}}{\kappa} \|u - v\|_{Y_\kappa}$$

so choosing $\kappa > 2\|L\| + 1$, we see that R is a contraction and hence admits a unique fixed point. \square

Applying the theorem, we immediately see that there always exists a unique approximate solution to the problem we're interested in solving; that is, for all $\varepsilon > 0$, there exists $u_\varepsilon \in C_b^{0,1}([0, T]; H^1) \cap W^{1,\infty}((0, T); H^1)$ such that

$$\begin{cases} A^0 \partial_t u_\varepsilon + K_\varepsilon [A^j \partial_j K_\varepsilon u_\varepsilon] + Bu = f \\ u_\varepsilon(t = 0) = g \end{cases}.$$

Now, in order to solve our original equation, we want to “send $\varepsilon \rightarrow 0$ ” and argue that the limit of such a sequence does the job. To do this, we'll need to formalize the a priori estimates done before. It turns out that, using the structure of an approximate solution, this can now be done rigorously.

Theorem. *Assuming the minimal hypothesis, we have that*

$$\sup_{t \in [0, T]} \|u_\varepsilon(\cdot, t)\|_{H^1}^2 \leq C \left[\|g\|_{H^1}^2 + \|f\|_{L^2([0, T]; H^1)}^2 \right]$$

and that

$$\sup_{t \in [0, T]} \|\partial_t u_\varepsilon(\cdot, t)\|_{L^2}^2 \leq C \left[\|g\|_{H^1}^2 + \|f\|_{L^2([0, T]; H^1)}^2 + \|f\|_{L^\infty([0, T]; L^2)}^2 \right]$$

for all $\varepsilon > 0$ for some constant depending on n, m, A^0, A^j, B , and T (but not ε).

Proof. First note that

$$[0, T] \ni t \mapsto \int_{\mathbb{R}^n} A^0(\cdot, t) u_\varepsilon(\cdot, t) \cdot u_\varepsilon(\cdot, t)$$

is absolutely continuous, with

$$\partial_t \int_{\mathbb{R}^n} A^0(\cdot, t) u_\varepsilon(\cdot, t) \cdot u_\varepsilon(\cdot, t) = \int_{\mathbb{R}^n} \partial_t A^0 u_\varepsilon \cdot u_\varepsilon + 2A^0 \partial_t u_\varepsilon \cdot u_\varepsilon.$$

Thus, we see

$$\partial_t \int_{\mathbb{R}^n} \frac{1}{2} A^0 u_\varepsilon \cdot u_\varepsilon = \int_{\mathbb{R}^n} A^0 \partial_t u_\varepsilon \cdot u_\varepsilon + \partial_t A^0 u_\varepsilon \cdot u_\varepsilon / 2.$$

but by assumption, the first term on the RHS is

$$\int \left(K_\varepsilon f - B u_\varepsilon - K_\varepsilon [A^j \partial_j K_\varepsilon u_\varepsilon] \right) \cdot u_\varepsilon.$$

Using the structure of the convolution operator, we see that K_ε is self-adjoint and hence

$$\int \left(-K_\varepsilon [A^j \partial_j K_\varepsilon u_\varepsilon] \right) \cdot u_\varepsilon = - \int \left([A^j \partial_j K_\varepsilon u_\varepsilon] \right) \cdot K_\varepsilon u_\varepsilon$$

so integrating by parts, which may be done since $K_\varepsilon u_\varepsilon$ is in $\cap H^s$, this is equal to

$$\frac{1}{2} \int -\partial_j (A^j K_\varepsilon u_\varepsilon \cdot K_\varepsilon u_\varepsilon) + \partial_j A^j K_\varepsilon u_\varepsilon \cdot K_\varepsilon u_\varepsilon = \frac{1}{2} \int \partial_j A^j K_\varepsilon u_\varepsilon \cdot K_\varepsilon u_\varepsilon$$

Thus,

$$\partial_t \int \frac{A^0 u_\varepsilon \cdot u_\varepsilon}{2} \leq C \left[1 + \|B\|_{L^\infty} + \|\partial_t A^0\|_{L^\infty} + \|\partial_j A^j\|_{L^\infty} \right] \int \frac{1}{2} A^0 u_\varepsilon \cdot u_\varepsilon + \frac{1}{2} \|K_\varepsilon f\|_{L^2}^2.$$

which is exactly the type of estimate we can apply Gronwall to. In particular, since we have that

$$Z(t) \leq e^{Ct} Z(0) + \int_0^t e^{C(t-s)} F(s) ds$$

we get the free uniform estimate

$$\sup_{t \in [0, T]} Z(t) \leq e^{Ct} \left[Z(0) + \int_0^T F(s) ds \right]$$

so we now find that

$$\sup_{t \in [0, T]} \int_{\mathbb{R}^n} \frac{A^0 u_\varepsilon \cdot u_\varepsilon}{2} \leq e^{CT} \left[\int \frac{1}{2} A_0(t=0) g \cdot g + \int_0^T \int_{\mathbb{R}^n} |f|^2 \right]$$

and using coercivity again yields

$$\sup_{t \in [0, T]} \|u_\varepsilon(\cdot, t)\|_{L^2}^2 \lesssim \|g\|_{L^2}^2 + \int_0^T \|f(\cdot, t)\|_{L^2}^2 dt$$

Now for any $k \in [n]$, we can apply ∂_k to the approximate equation to find that

$$A^0 \partial_t \partial_k u_\varepsilon + K_\varepsilon [A^j \partial_j K_\varepsilon \partial_k u_\varepsilon] + B \partial_k u_\varepsilon = K_\varepsilon \partial_k f - \partial_k A^0 \partial_t u_\varepsilon - K_\varepsilon [\partial_k A^j \partial_j K_\varepsilon u_\varepsilon] - \partial_k B u_\varepsilon$$

Performing estimates similarly to before, we thus find that

$$\partial_t \int_{\mathbb{R}^n} \frac{A^0}{2} \partial_k u_\varepsilon \cdot \partial_k u_\varepsilon \leq C \int_{\mathbb{R}^n} \frac{A_0}{2} \partial_k u_\varepsilon \cdot \partial_k u_\varepsilon + \frac{1}{2} \int_{\mathbb{R}^n} |\partial_k f|^2 + \int_{\mathbb{R}^n} \left[\underbrace{-\partial_k A^0 \partial_t u_\varepsilon}_{\text{I}} - \underbrace{K_\varepsilon [\partial_k A^j \partial_j K_\varepsilon u_\varepsilon]}_{\text{II}} - \underbrace{\partial_k B u_\varepsilon}_{\text{III}} \right] \cdot \partial_k u_\varepsilon.$$

Now we consider each of the labeled terms in turn. The first term is controlled by

$$\begin{aligned} |I| &= \left| \int -\partial_k A^0 \partial_t u_\varepsilon \cdot \partial_k u_\varepsilon \right| \lesssim \int |\partial_t u_\varepsilon|^2 + \partial_k u_\varepsilon \partial_k u_\varepsilon \\ &\lesssim \int \left| K_\varepsilon f - B u_\varepsilon - K_\varepsilon (A^j \partial_j K_\varepsilon u_\varepsilon) \right|^2 + \frac{A^0 \partial_k u_\varepsilon \cdot \partial_k u_\varepsilon}{2} \\ &\lesssim \int \frac{|f|^2}{2} + \frac{A^0 u_\varepsilon \cdot u_\varepsilon}{2} + \frac{A^0 \partial_k u_\varepsilon \cdot \partial_k u_\varepsilon}{2} \end{aligned}$$

Again using the fact that K_ε is self-adjoint, we have

$$|II| = \left| \int -\partial_k A^j \partial_j K_\varepsilon u_\varepsilon \cdot \partial_k K_\varepsilon u_\varepsilon \right| \lesssim \sup_{j \in [n]} \|\nabla A^j\|_{L^\infty} \int |K_\varepsilon \nabla u_\varepsilon|^2 \lesssim \int |\nabla u_\varepsilon|^2 \lesssim \int \frac{A^0 \partial_k u_\varepsilon \cdot \partial_k u_\varepsilon}{2}$$

Finally, we can handle the third term by noticing that

$$|III| \leq \|\nabla B\|_{L^\infty} \int_{\mathbb{R}^n} \frac{|u_\varepsilon|^2}{2} + \frac{\partial_k u_\varepsilon \cdot \partial_k u_\varepsilon}{2} \leq C \int_{\mathbb{R}^n} \frac{A^0 u_\varepsilon \cdot u_\varepsilon}{2} + \frac{A^0 \partial_k u_\varepsilon \cdot \partial_k u_\varepsilon}{2}.$$

Combining this with the previous estimates then yields

$$\partial_t \left(\sum_{|\alpha| \leq 1} \int_{\mathbb{R}^n} \frac{A^0 \partial^\alpha u_\varepsilon \cdot \partial^\alpha u_\varepsilon}{2} \right) \leq C \sum_{|\alpha| \leq 1} \int \frac{A^0 \partial^\alpha u_\varepsilon \cdot \partial^\alpha u_\varepsilon}{2} + C \|f\|_{H^1}^2$$

Gronwall then yields exactly the same type of estimate:

$$\sup_{t \in [0, T]} \|u_\varepsilon(\cdot, t)\|_{H^1}^2 \lesssim \|g\|_{H^1}^2 + \int_0^T \|f(\cdot, t)\|_{H^1}^2 dt.$$

Finally, to bound $\sup_{t \in [0, T]} \|\partial_t u_\varepsilon(\cdot, t)\|_{L^2}^2$ we solve for $\partial_t u_\varepsilon$ to see that

$$\partial_t u_\varepsilon = (A^0)^{-1} [K_\varepsilon f - Bu_\varepsilon - K_\varepsilon [A^j \partial_j K_\varepsilon u_\varepsilon]]$$

which implies that

$$\sup_{t \in [0, T]} \|\partial_t u_\varepsilon(\cdot, t)\|_{L^2}^2 \lesssim \sup_{t \in [0, T]} \|\partial_t u_\varepsilon(\cdot, t)\|_{H^1}^2 + \text{terms we don't care about.}$$

□

With these estimates out of the way we can now move onto the following theorem.

Theorem. *There exists $u \in L_T^\infty(H^1)$ such that $\partial_t u \in L_T^\infty(L^2)$ solving*

$$\begin{cases} A^0 \partial_t u + A^j \partial_j u + Bu = f & \text{a.e. in } \mathbb{R}^n \times (0, T) \\ u(t = 0) = g & \text{in } \mathbb{R}^n \end{cases}$$

Further, u satisfies

$$\|u\|_{L^\infty(H^1)} + \|\partial_t u\|_{L^\infty(L^2)} \leq \text{the things in the previous theorem}$$

Proof. The estimates above give uniform bound for u_ε in $L_T^\infty(H^1)$ and $L_T^2(H^1)$ and $\partial_t u_\varepsilon$ in $L_T^\infty(L^2)$ and $L_T^2(L^2)$. Using weak-star convergence, we can find a sequence $\{\varepsilon_n\}$ with $\varepsilon_n \rightarrow 0$ such that

$$\begin{cases} u_\varepsilon \xrightarrow{*} u \text{ in } L_T^\infty(H^1) \\ u_\varepsilon \rightharpoonup u \text{ in } L_T^2(H^1) \\ \partial_t u_\varepsilon \xrightarrow{*} \partial_t u \text{ in } L_T^\infty(L^2) \\ \partial_t u_\varepsilon \rightharpoonup \partial_t u \text{ in } L_T^2(L^2) \end{cases}$$

where $\partial_t u$ is understood in the sense of distributions.

Lower semicontinuity in the weak and weak-star topologies provides the specified estimate for u , so it suffices to show that u actually solves the desired system.

Towards doing so, let $\varphi \in C_c^\infty((0, T))$, $v \in L^2(\mathbb{R}^n; \mathbb{R}^n)$. Then for all $t \in (0, T)$, we have

$$(A^0 \partial_t u_\varepsilon, v) + (A^j \partial_j K_\varepsilon u_\varepsilon, K_\varepsilon v) + (Bu_\varepsilon, v) = (K_\varepsilon f, v)$$

where (\cdot, \cdot) is the inner product on (spatial) L^2 . Multiplying by φ and integrating on $(0, T)$, we have the equality

$$\int_0^T \varphi (A^0 \partial_t u_\varepsilon, v) + \int_0^T \varphi (A^j \partial_j K_\varepsilon u_\varepsilon, K_\varepsilon v) + \int_0^T \varphi (Bu_\varepsilon, v) = \int_0^T \varphi (K_\varepsilon f, v)$$

which implies

$$\int_0^T (\partial_t u_\varepsilon, \varphi A^0 v) + \int_0^T (\partial_j u_\varepsilon, \varphi K_\varepsilon (A^j K_\varepsilon v)) + \int_0^T (Bu_\varepsilon, \varphi B^T v) = \int_0^T \varphi (K_\varepsilon f, v)$$

Since $K_\varepsilon h \xrightarrow{\varepsilon \rightarrow 0} h$ strongly in L^2 , we can send $\varepsilon \rightarrow 0$ in the identity above to find that

$$\int_0^T (\partial_t u, \varphi A^0 v) + \int_0^T (\partial_j u, \varphi(A^j v)) + \int_0^T (Bu, \varphi B^T v) = \int_0^T \varphi(f, v)$$

which then implies

$$\int_0^T \varphi \left[(A^0 \partial_t u + A^j \partial_j u + Bu - f, v) \right] = 0$$

for all φ, v , which implies equality almost everywhere in space and in time, implying the desired condition in the PDE.

Finally, note that

$$\{v \in L_T^\infty(H^1) \mid \partial_t v \in L_T^\infty(L^2) \text{ and } v(\cdot, 0) = g\}$$

is a convex and closed set with respect to the norm

$$v \mapsto \|v\|_{L^\infty(H^1)} + \|\partial_t v\|_{L_T^\infty(L^2)}$$

which can be thought of as the space $W_T^{1,\infty}(L^2) \hookrightarrow C^0([0, T]; L^2)$. Thus, this set is weakly closed, which shows that the weak limit u is also in this set. \square

Now to prove uniqueness, we need something more interesting.

Theorem (Finite speed of propagation). *Let u solve*

$$A^0 \partial_t u + A^j \partial_j u + Bu = 0$$

where u is in the two spaces from the last theorem. For $\alpha > 0, t_0 > 0, x_0 \in \mathbb{R}^n$, let

$$C_\alpha(x_0, t_0) := \{(x, t) \mid 0 < t < t_0, x \in B(x_0, \alpha(t_0 - t))\}$$

Then there exists $\alpha > 0$ such that if $u = 0$ in $B(x_0, \alpha t_0)$, then $u = 0$ in $C_\alpha(x_0, t_0)$.

Proof. WLOG we'll assume $T = t_0$. Let $E : [0, T] \rightarrow \mathbb{R}$ via

$$E(t) = \int_{B(x_0, \alpha(t_0 - t))} (A^0 u \cdot u)(t)$$

for some α to be chosen. We note that E is absolutely continuous with

$$\begin{aligned} \dot{E}(t) &= \int_{B(x_0, \alpha(t_0 - t))} \partial_t A^0 u \cdot u + 2A^0 \partial_t u \cdot u - \alpha \int_{\partial B(x_0, \alpha(t_0 - t))} A^0 u \cdot u \\ &= \int_{B(x_0, \alpha(t_0 - t))} \partial_t A^0 u \cdot u + 2[-A^j \partial_j u - Bu] \cdot u - \alpha \int_{\partial B(x_0, \alpha(t_0 - t))} A^0 u \cdot u \end{aligned}$$

We also know that

$$\int -2A^j \partial_j u \cdot u = \int -\partial_j(A^j u \cdot u) + \partial_j A^j u \cdot u$$

so substituting, \dot{E} is

$$\int_{B(x_0, \alpha(t_0-t))} \partial_t A^0 u \cdot u - \int_{B(x_0, \alpha(t_0-t))} -2Bu \cdot u + A^j \partial_j u \cdot u + \int_{\partial B(x_0, \alpha(t_0-t))} -\alpha A^0 u \cdot u - A^j v_j u \cdot u$$

which is controlled by

$$C \int_{B(x_0, \alpha(t_0-t))} A^0 u \cdot u + \left(\max_j \|A^j\|_{L^\infty} - \alpha \right) \int_{\partial B(x_0, \alpha(t_0-t))} A^0 u \cdot u$$

Picking $\alpha = \max_j \|A^j\|_{L^\infty}$ then shows that $\dot{E}(t) \leq CE(t) \implies E(t) \leq e^{Ct} E(0) = 0$ which concludes. \square

As a corollary, we see that u as constructed before is unique.

Now we proceed with some results relating spatial and temporal regularity of such solutions.

Theorem. Let $1 \leq k \in \mathbb{N}$ and suppose that $A^0, A^j, B \in C_b^{0,1}(\mathbb{R}^n \times [0, T]; \mathbb{R}^m) \cap L^\infty((0, T); C_b^{k-1,1}(\mathbb{R}^n; \mathbb{R}^{m \times m}))$, $g \in H^k(\mathbb{R}^n; \mathbb{R}^m)$ and $f \in L^2((0, T); H^k(\mathbb{R}^n; \mathbb{R}^m)) \cap L^\infty((0, T); H^{k-1}(\mathbb{R}^n; \mathbb{R}^m))$. Then

$$\sup_{t \in [0, T]} \|u(\cdot, t)\|_{H^k} \leq C \left(\|g\|_{H^k} + \|f\|_{L_T^2(H^k)} \right)$$

and

$$\sup_{t \in [0, T]} \|\partial_t u(\cdot, t)\|_{H^{k-1}} \leq C \left(\|g\|_{H^k} + \|f\|_{L_T^2(H^k)} + \|f\|_{L_T^\infty(H^{k-1})} \right)$$

Proof. Recall that the approximate problem reads

$$\begin{cases} \partial_t u_\varepsilon + \mathcal{M}_\varepsilon u_\varepsilon = (A^0)^{-1} K_\varepsilon f \in L^\infty(H^k) \\ u_\varepsilon(t=0) = g \in H^k \end{cases}$$

where

$$\mathcal{M}_\varepsilon v := (A^0)^{-1} K_\varepsilon [A^j \partial_j K_\varepsilon v] + (A^0)^{-1} B v.$$

One can show that the higher regularity of A^j, B implies that

$$\mathcal{M}_\varepsilon \in L^\infty([0, T]; \mathcal{L}(H^k)).$$

Then the ODE result shows that

$$u_\varepsilon \in C_b^{0,1}([0, T]; H^k) \cap W^{1,\infty}((0, \infty); H^k).$$

We can then apply ∂^α for $\alpha \in \mathbb{N}^n$ with $|\alpha| \leq k$ and argue as before to show that

$$\partial_t \left(\sum_{|\alpha| \leq k} \int_{\mathbb{R}^n} A^0 \partial^\alpha u_\varepsilon \cdot \partial^\alpha u_\varepsilon \right) \leq C \left(\sum_{|\alpha| \leq k} \int_{\mathbb{R}^n} A^0 \partial^\alpha u_\varepsilon \cdot \partial^\alpha u_\varepsilon \right) + \|f\|_{H^k}^2$$

Then solving for $\partial_t u_\varepsilon$ and arguing as before finishes. \square

Theorem (Higher regularity, temporal derivatives). *Now suppose that $k \geq 2$, $A^j, B \in C_b^{k-1,1}(\mathbb{R}^n \times [0, T])$, $g \in H^k$, $f \in L_T^2(H^k)$ with*

$$\partial_t^\ell f \in L_T^\infty(H^{k-1-\ell}) \text{ for } \ell = 0, \dots, k-1.$$

Then

$$\sup_{t \in [0, T]} \sum_{\ell=0}^k \|\partial_t^\ell u(\cdot, t)\|_{H^{k-\ell}} \leq C \left[\|g\|_{H^k} + \|f\|_{L_T^2(H^k)} \right] + \sum_{\ell=0}^{k-1} \|\partial_t^\ell f\|_{L_T^\infty(H^{k-1-\ell})}.$$

Corollary. *If we pick k large enough in the previous theorem, then u is a classical solution; i.e., $u \in C_b^1(\mathbb{R}^n \times [0, T]; \mathbb{R}^m)$ and*

$$\begin{cases} A^0 \partial_t u + A^j \partial_j u + Bu = f \text{ everywhere in } \mathbb{R}^n \times [0, T] \\ u(t=0) = g \end{cases}$$

(See the Aubin-Lions-Simon space-time compactness lemma).

Now let's revisit the previous examples we had.

Example 1 (Wave Equations). *Recall that we had the system*

$$\begin{cases} \alpha \partial_t^2 u + \kappa \partial_t u - A : D^2 u + b \cdot \nabla u + cu = f \\ u(t=0) = u_0 \\ \partial_t u(t=0) = v_0 \end{cases} \quad (1)$$

Which led to the matrices

$$A^0 = \left(\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & \alpha & 0 \\ 0 & 0 & A \end{array} \right), A^j = \left(\begin{array}{c|ccc} 0 & & & \\ \hline 0 & 0 & -a_{1j} & \cdots & -a_{nj} \\ -a_{1j} & & & & \\ \cdots & & & \mathbf{0} & \\ -a_{nj} & & & & \end{array} \right), B = \left(\begin{array}{cc|ccc} 0 & -1 & 0 & \cdots & 0 \\ \hline c & \kappa & b_1 & \cdots & b_n \\ 0 & & 0 & & \end{array} \right), F = \begin{pmatrix} 0 \\ f \\ 0 \end{pmatrix}$$

We saw before that if u solves (1), then $U := \begin{pmatrix} u \\ \partial_t u \\ \nabla u \end{pmatrix}$ solves

$$\begin{cases} A^0 \partial_t U + A^j \partial_j U + BU = F \\ U(t=0) = \begin{pmatrix} u_0 \\ v_0 \\ \nabla u_0 \end{pmatrix} \end{cases} \quad (2)$$

Now suppose we go the other way around, and use our theory to solve (2). Then using the second condition, we can reverse engineer the system to see that

$$u = U_1, \partial_t u = U_2, \nabla u = \begin{pmatrix} U_3 \\ \vdots \\ U_{n+2} \end{pmatrix}$$

and that u solves the original formulation.

Example 2 (Maxwell's equations). Recall Maxwell's equations again, which read

$$\begin{cases} \operatorname{div} B = 0 \\ \operatorname{div} E = \rho \\ \partial_t E + J = \operatorname{curl} B \\ \partial_t B = -\operatorname{curl} E \end{cases}.$$

We saw previously that any solution to these equations satisfy

$$\begin{cases} \partial_t \rho + \operatorname{div} J = 0 \\ \partial_t \begin{pmatrix} E \\ B \end{pmatrix} + \sum_{j=1}^3 \begin{pmatrix} 0 & -M_j \\ M_j & 0 \end{pmatrix} \partial_j \begin{pmatrix} E \\ B \end{pmatrix} = \begin{pmatrix} -J \\ 0 \end{pmatrix} \end{cases} \quad (3)$$

for appropriately chosen antisymmetric $M_j \in \mathbb{R}^{3 \times 3}$. Again using our theory to solve the latter for $U \in \mathbb{R}^6$, and setting $\begin{pmatrix} E \\ B \end{pmatrix} = U$ exactly as in the previous example, we see again that

$$\begin{cases} \partial_t E + J = \operatorname{curl} B \\ \partial_t B = -\operatorname{curl} E \end{cases}.$$

Now suppose that ρ, J are given and satisfy

$$\partial_t \rho + \operatorname{div} J = 0.$$

Applying the divergence to our dynamics equations implies that

$$\begin{cases} \partial_t \operatorname{div} E + \operatorname{div} J = 0 \implies \partial_t (\operatorname{div} E - \rho) = 0 \implies \operatorname{div} E - \rho = \operatorname{div} E_0 - \rho_0 \\ \partial_t \operatorname{div} B = 0 \implies \operatorname{div} B = \operatorname{div} B_0 \end{cases},$$

Hence if our data satisfies $\begin{cases} \operatorname{div} B_0 = 0 \\ \operatorname{div} E_0 = \rho_0 \end{cases}$, then $\begin{cases} \operatorname{div} B = 0 \\ \operatorname{div} E = \rho \end{cases}$ in $\mathbb{R}^n \times [0, T]$.