

MATH 18.152 COURSE NOTES - CLASS MEETING # 13

18.152 Introduction to PDEs, Fall 2011

Professor: Jared Speck

Class Meeting # 13: Geometric Energy Estimates

1. \square_m , THE ENERGY-MOMENTUM TENSOR, AND COMPATIBLE CURRENTS

The following shorthand notation is often used for the “linear wave operator associated to m ”:

$$(1.0.1) \quad \square_m \stackrel{\text{def}}{=} (m^{-1})^{\alpha\beta} \partial_\alpha \partial_\beta.$$

Using this notation, the wave equation $-\partial_t^2 \phi + \Delta \phi = 0$ can be expressed as

$$(1.0.2) \quad \square_m \phi = 0.$$

We now introduce a very important object called the *energy-momentum tensor*. As we will see, it encodes some very important conservation laws associated to solutions of (1.0.2).

Definition 1.0.1. The energy-momentum tensor associated to equation (1.0.2) is

$$(1.0.3) \quad T_{\mu\nu} \stackrel{\text{def}}{=} \partial_\mu \phi \partial_\nu \phi - \frac{1}{2} m_{\mu\nu} (m^{-1})^{\alpha\beta} \partial_\alpha \phi \partial_\beta \phi.$$

Later in the course, we will hopefully have time to motivate its derivation in the larger context of *variational methods*. For now, we will simply study/use its useful properties.

Note that $T_{\mu\nu}$ is symmetric:

$$(1.0.4) \quad T_{\mu\nu} = T_{\nu\mu}.$$

In your homework, you will prove the following very important positivity property of T , which is called the *dominant energy condition*.

Lemma 1.0.1 (Dominant Energy Condition for $T_{\mu\nu}$).

$$(1.0.5)$$

$T(X, Y) \stackrel{\text{def}}{=} T_{\alpha\beta} X^\alpha Y^\beta \geq 0$ if X, Y are both timelike and future-directed or timelike and past-directed.

Since causal vectors are the limit of timelike vectors, we have the following consequence of (1.0.5):

$$(1.0.6)$$

$T(X, Y) \stackrel{\text{def}}{=} T_{\alpha\beta} X^\alpha Y^\beta \geq 0$ if X, Y are future-directed and causal or past-directed and causal.

As before, we can raise the indices of T :

$$(1.0.7) \quad T^{\mu\nu} = (m^{-1})^{\mu\alpha} (m^{-1})^{\nu\beta} T_{\alpha\beta}.$$

A very special case of Lemma 1.0.1 is the following, which corresponds to $X^\mu = Y^\mu = \delta_0^\mu = (1, 0, 0, \dots, 0)$ in the lemma:

$$(1.0.8) \quad T_{00} = T^{00} = \frac{1}{2} \sum_{\mu=0}^3 (\partial_\mu \phi)^2 = \frac{1}{2} |\nabla_{t,x} \phi|^2.$$

The derivation of (1.0.8) is a simple computation that you should do for yourself. Note that T_{00} is positive definite in all of the derivatives of ϕ . This fact will play an important role in Theorem 2.1 below.

The next lemma shows that $T^{\mu\nu}$ is divergence-free whenever ϕ verifies the wave equation. This fact is intimately connected to the derivation of conservation laws, which are fundamental ingredients in the study of hyperbolic PDEs.

Lemma 1.0.2 (The divergence of $T^{\mu\nu}$). *Let $T_{\mu\nu}$ be the energy-momentum tensor defined in (1.0.3). Then*

$$(1.0.9) \quad \partial_\mu T^{\mu\nu} = (\square_m \phi) (m^{-1})^{\nu\alpha} \partial_\alpha \phi.$$

In particular, if ϕ is a solution to (1.0.2), then

$$(1.0.10) \quad \partial_\mu T^{\mu\nu} = 0.$$

Proof. The proof is a computation that uses the symmetry property $(m^{-1})^{\mu\nu} = (m^{-1})^{\nu\mu}$ and the fact that we are allowed to interchange the order of partial derivatives (if ϕ is sufficiently smooth):

$$(1.0.11) \quad \begin{aligned} \partial_\mu T^{\mu\nu} &= \partial_\mu \left((m^{-1})^{\mu\alpha} (m^{-1})^{\nu\beta} \partial_\alpha \phi \partial_\beta \phi - \frac{1}{2} (m^{-1})^{\mu\nu} (m^{-1})^{\alpha\beta} \partial_\alpha \phi \partial_\beta \phi \right) \\ &= (\square_m \phi) (m^{-1})^{\nu\beta} \partial_\beta \phi + (m^{-1})^{\mu\alpha} (m^{-1})^{\nu\beta} (\partial_\alpha \phi) \partial_\mu \partial_\beta \phi \\ &\quad - \frac{1}{2} (m^{-1})^{\mu\nu} (m^{-1})^{\alpha\beta} (\partial_\mu \partial_\alpha \phi) \partial_\beta \phi - \frac{1}{2} (m^{-1})^{\mu\nu} (m^{-1})^{\alpha\beta} (\partial_\alpha \phi) \partial_\mu \partial_\beta \phi \\ &= (\square_m \phi) (m^{-1})^{\nu\beta} \partial_\beta \phi, \end{aligned}$$

where the last three terms have canceled each other. □

As we will soon see, the energy-momentum tensor provides an amazingly convenient way of bookkeeping in the divergence theorem. However, in order to apply the divergence theorem, we need to find a **useful** vectorfield to take the divergence of. By useful, we mean a vectorfield that can be used to control a solution ϕ to the wave equation. One way of constructing a useful vectorfield is to start with an auxiliary vectorfield X and then to contract it with the energy momentum tensor to form a new vectorfield J . The next definition shows how to do this.

Definition 1.0.2. Given any vectorfield X , we associate to it the following *compatible current*, which is itself a vectorfield:

$$(1.0.12) \quad (X)J^\mu \stackrel{\text{def}}{=} T^{\mu\alpha} X_\alpha.$$

So which vectors X are the useful ones? It turns out that the answer is *causal vectors*. This fact is closely connected to the dominant energy condition (1.0.5). This will become more clear in our proof of theorem 2.1 below; Note that by Lemma 1.0.1, $J^\mu Y_\mu \stackrel{\text{def}}{=} T^{\mu\alpha} X_\alpha Y_\mu = T_{\alpha\beta} X^\alpha Y^\beta \stackrel{\text{def}}{=} T(X, Y) \geq 0$ if X, Y are both timelike and future-directed (i.e., $X^0, Y^0 > 0$) or past-directed (i.e., $X^0, Y^0 < 0$).

In order to apply the divergence theorem to ${}^{(X)}J^\mu$, we of course need to know its divergence. We carry out this computation in the next corollary.

Corollary 1.0.3. *Using (1.0.4) and (1.0.10), we have that*

$$(1.0.13) \quad \partial_\mu \left({}^{(X)}J^\mu \right) = T^{\alpha\beta(X)} \pi_{\alpha\beta},$$

where

$$(1.0.14) \quad {}^{(X)}\pi_{\mu\nu} \stackrel{\text{def}}{=} \frac{1}{2} \left(\partial_\mu X_\nu + \partial_\nu X_\mu \right)$$

is called the deformation tensor of X .

We now state a version of the divergence theorem that is tailored to our study of the linear wave equation.

Theorem 1.1 (Divergence Theorem). *Let ϕ be a solution to the linear wave equation $\square_m \phi = 0$. Let X be any vectorfield, and let ${}^{(X)}J$ be the compatible current define in Definition 1.0.2. Let $\Omega \subset \mathbb{R}^{1+n}$ be a domain with boundary $\partial\Omega$. Then the following integral identity holds:*

$$(1.0.15) \quad \int_{\partial\Omega} \hat{N}_\alpha {}^{(X)}J^\alpha[\phi(\sigma)] d\sigma = \int \partial_\mu \left({}^{(X)}J^\mu[\phi(t, x)] \right) dt d^n x.$$

2. ENERGY ESTIMATES AND UNIQUENESS

We will now use the results of the previous section to derive some extremely important *energy estimates* for solutions to $\square_m \phi = 0$. The results we derive are a geometry version of integration by parts + the divergence theorem. They could alternatively be derived by multiplying both sides of the wave equation by a suitable quantity and then integrating by parts over a suitable hypersurfaces, but there is a substantial gain in geometric insight that accompanies our use of compatible currents.

Theorem 2.1 (Energy estimates in a cone). *Let $\phi(t, x)$ be a C^2 solution to the $1+n$ dimensional global Cauchy problem for the linear wave equation*

$$(2.0.16) \quad \square_m \phi = 0,$$

$$(2.0.17) \quad \phi(0, x) = f(x), \quad x \in \mathbb{R}^n,$$

$$(2.0.18) \quad \partial_t \phi(0, x) = g(x), \quad x \in \mathbb{R}^n.$$

Let $R \in [0, \infty]$, let X be the past-directed timelike vector defined by $X^\mu = -\delta_0^\mu$, and let ${}^{(X)}J^\mu[\phi(t, y)]$ be the compatible current (1.0.12) associated to X . Note that by (1.0.8), ${}^{(X)}J^\mu[\phi(t, y)] = |\nabla_{t,y} \phi(t, y)|^2 = \sum_{\mu=0}^n (\partial_\mu \phi)^2 = (\partial_t \phi)^2 + \sum_{i=1}^n (\partial_i \phi)^2$. Define the square of the energy $E[\phi](t)$ by

$$(2.0.19) \quad E^2[\phi](t) \stackrel{\text{def}}{=} \int_{B_{R-t}(p)} \hat{N}_\mu {}^{(X)}J^\mu[\phi(t, y)] d^n y = \frac{1}{2} \int_{B_{R-t}(p)} |\nabla_{t,y} \phi(t, y)|^2 d^n y,$$

where $\hat{N}_\mu = \delta_\mu^0$ (and therefore $\hat{N}^\mu = -\delta_0^\mu$) is the past-pointing unit normal covector to $\{t\} \times B_{R-t}(p) \subset \mathbb{R}^4$, and $B_R(p) \subset \mathbb{R}^3$ denotes the solid Euclidean ball of radius R centered at p . Then

$$(2.0.20) \quad E[\phi](t) \leq E[\phi](0).$$

Proof. The goal is to apply Theorem 1.1 to the solid truncated backwards light cone $\mathcal{C}_{t,p;R} \stackrel{\text{def}}{=} \{(\tau, y) \in [0, \infty) \times \mathbb{R}^n \mid |y - p| \leq R - \tau\}$ and to make use of the dominant energy condition. It is easy to see that $\partial\mathcal{C}_{t,p;R} = \mathcal{B} \cup \mathcal{M}_{t,p;R} \cup \mathcal{T}$, where $\mathcal{B} \stackrel{\text{def}}{=} \{0\} \times B_R(p)$ is the flat base of the truncated cone, $\mathcal{T} \stackrel{\text{def}}{=} \{t\} \times B_{R-t}(p)$ is the flat top of the truncated cone, and $\mathcal{M}_{t,p;R} \stackrel{\text{def}}{=} \{(\tau, y) \in [0, \infty) \times \mathbb{R}^n \mid |y - p| = R - \tau\}$ is the mantle of the truncated cone.

By Theorem 1.1, we have that

$$(2.0.21) \quad E[\phi](t) - E[\phi](0) + F[\phi] = \int_{\mathcal{C}_{t,p}} \partial_\mu \left({}^{(X)}J^\mu[\phi(\tau, y)] \right) d\tau d^n y,$$

where

$$(2.0.22) \quad F[\phi] \stackrel{\text{def}}{=} \int_{\mathcal{M}_{t,p;R}} \hat{N}_\alpha {}^{(X)}J^\alpha[\phi(\sigma)] d\sigma$$

is the “flux” associated to $\mathcal{M}_{t,p;R}$. Since ϕ solves the wave equation (2.0.16), and since ${}^{(X)}\pi_{\mu\nu} = 0$, the identity (1.0.13) implies that the right-hand side of (2.0.21) is 0. Therefore,

$$(2.0.23) \quad E[\phi](t) - E[\phi](0) + F[\phi] = 0.$$

We claim that $F[\phi] \geq 0$. The energy inequality (2.0.20) would then follow from (2.0.23). The key observation for showing that $F[\phi] \geq 0$ is the following. Along the mantle $\mathcal{M}_{t,p;R}$, it is easy to see (draw the picture!) that $\hat{N}_\mu = \underline{L}_\mu$, where \underline{L} is a past-directed null vector. Therefore, the integrand in (2.0.22) is equal to $T_{\alpha\beta} X^\alpha \underline{L}^\beta$, and since X is a past-directed timelike vector, the dominant energy condition (1.0.6) implies that $T_{\alpha\beta} X^\alpha \underline{L}^\beta \geq 0$. Therefore, $F[\phi] \geq 0$ as desired. \square

Theorem 2.1 can easily be used to prove the following **local** uniqueness result for solutions to the linear wave equation.

Corollary 2.0.4 (Uniqueness). *Suppose that two C^2 solutions ϕ_1 and ϕ_2 to (2.0.16) have the same initial data on $B_R(p) \subset \{(\tau, y) \mid \tau = 0\}$. Then the two solutions agree on the “solid backwards light cone” $\mathcal{C}_{p;R} \stackrel{\text{def}}{=} \{(\tau, y) \mid 0 \leq \tau \leq R, 0 \leq |y - p| \leq R - \tau\}$.*

Proof. Define $\psi \stackrel{\text{def}}{=} \phi_1 - \phi_2$. Then ψ verifies (2.0.16) and furthermore, $E[\psi](0) = 0$. Thus, by Theorem 2.1, $E[\psi](t) = 0$ for $0 \leq t \leq R$. Therefore, from the definition of $E[\psi](t)$, it follows that $\nabla_{\tau, y} \psi(\tau, y) = 0$ for $(\tau, y) \in \mathcal{C}_{p;R}$. Thus, by elementary analysis, ψ is constant in $\mathcal{C}_{p;R}$. But $\psi(0, x) = 0$ for $(0, x) \in \mathcal{C}_{p;R}$. Thus, $\psi(\tau, y) = 0$ for all points $(\tau, y) \in \mathcal{C}_{p;R}$. \square

Corollary 2.0.4 is one illustration of the *finite speed of propagation* property associated to the linear wave equation. Another way to think about it is the following. Suppose you alter the initial conditions *outside* of $B_R(p)$, but not on $B_R(p)$ itself. Then this alteration has *no effect whatsoever* on the behavior of the solution in the spacetime region $\mathcal{C}_{p;R}$. Think about this claim yourself; it follows easily from the Corollary!

3. DEVELOPMENTS, DOMAIN OF DEPENDENCE, AND RANGE OF INFLUENCE

We will now develop a language for discussing the finite speed of propagation properties of the linear wave equation in more detail. If we had more time in this course, we could adopt a more geometric point of view that would apply to many other hyperbolic PDEs. This would involve fleshing out our discussion of Lorentzian geometry, and also developing a generalized version of geometry that applies to a large class of PDEs.

Warning 3.0.1. Some people permute or even severely alter the following definitions, which can be very confusing. The definitions below therefore indicate some of my biases.

Definition 3.0.3 (Development). Let $S \subset \{(t, x) \mid t = 0\}$ be a set. Assume that that we know the initial data $\phi(0, x) = f(x)$, $\partial_t \phi(0, x) = g(x)$ for the wave equation (1.0.2), but only for $x \in S$. Then a *future development* Ω of S is defined to be a “future” region of spacetime $\Omega \subset \mathbb{R}^{1+n} \cap \{(t, x) \mid t \geq 0\}$ on which the solution $\phi(t, x)$ to (1.0.2) is uniquely determined by the initial data on S . A *past development* $\mathcal{D}^-(S)$ can be analogously defined (replace $t \geq 0$ with $t \leq 0$ in the previous definition).

Example 3.0.1. If $B_R(p)$ and $\mathcal{C}_{p;R}$ are as in Corollary 2.0.4, then $\mathcal{C}_{p;R}$ is a development of $B_R(p)$. You can imagine that the solution knows how to “develop” in $\mathcal{C}_{p;R}$ from the initial conditions on its subset $B_R(p)$.

Definition 3.0.4 (Maximal development). The *maximal future development* of S , which we denote by $\mathcal{D}^+(S)$, is defined to be the union of all future developments of S . The *maximal past development* $\mathcal{D}^-(S)$ can be analogously defined. The *maximal development* of S is defined to be $\mathcal{D}^+(S) \cup \mathcal{D}^-(S)$.

Example 3.0.2. Consider the plane $P \stackrel{\text{def}}{=} \{(t, x^1, x^2, x^3) \mid x^1 = 0\}$. Then using techniques from a more advanced course, one could show that $\mathcal{D}(P) = P$ for the wave equation (1.0.2). That is, knowing the conditions of a solution ϕ along P is not enough information to determine the solution anywhere else. This is closely connected to the fact that all smooth curves in P have tangent vectors that are timelike relative to the Minkowski metric.

Definition 3.0.5 (Domain of dependence). Let $\Omega \subset \mathbb{R}^{1+n}$. Assume that ϕ is a solution to the wave equation (1.0.2) in Ω . A *domain of dependence* for Ω is a set S such that ϕ is completely determined on Ω from only the data $\phi|_S$ and $\nabla_{t,x}\phi|_S$.

Remark 3.0.1. For general nonlinear hyperbolic PDEs, domains of dependence depend both on Ω and the solution ϕ itself. However, for the linear wave equation, domains of dependence do not depend on the solution. Roughly speaking, this is because the “geometry of the solution” is predetermined by the Minkowski metric m .

Example 3.0.3. In 1+1 dimensions, a domain of dependence for the spacetime point (t, x) (for the wave equation (1.0.2)) is the “initial data” interval $\{0\} \times [x-t, x+t]$. Another domain of dependence for this point is the interval $\{t/2\} \times [x-t/2, x+t/2]$. A trivial example is that (t, x) is a domain of dependence for itself.

Example 3.0.4. In 1+3 dimensions, a domain of dependence for the positive t axis $\{(t, x^1, x^2, x^3) \mid x^1 = x^2 = x^3 = 0, t \geq 0\}$ (for the wave equation (1.0.2)) is all of “space:” $\{(t, x^1, x^2, x^3) \mid t = 0\}$. Any subset of space is *not* a domain of dependence for the positive t axis.

The next definition is complementary to the notion of domain of dependence.

Definition 3.0.6 (Range of influence). Assume that ϕ is a solution to the wave equation (1.0.2) in \mathbb{R}^{1+n} . The *range of influence* \mathcal{R} for a set $S \subset \mathbb{R}^{1+n}$ is the set of all points $(t, x) \in \mathbb{R}^{1+n}$ such that $\phi(t, x)$ is affected by the initial data $\phi|_S$ and $\nabla_{t,x}\phi|_S$.

Example 3.0.5. In 1 + 1 dimensions, the (future) range of influence (for $t \geq 0$) of the interval $S = \{0\} \times [-1, 1]$ is $\mathcal{R} = \{(t, x) \mid -t - 1 \leq x \leq t + 1\}$.

Example 3.0.6. In 1 + 1 dimensions, the (future) range of influence (for $t \geq 0$) of the t axis $S = \{(t, 0) \mid t \geq 0\}$ is $\mathcal{R} = \{(t, x) \mid t \geq 0\}$.

Example 3.0.7. In 1 + 3 dimensions, the (future) range of influence (for $t \geq 0$) of $S = \{0\} \times \partial B_1(0)$ is $\mathcal{R} = \{(t, x) \mid 0 \leq t \leq 1, |x| = 1 - t\} \cup \{(t, x) \mid 0 \leq t < t, |x| = 1 + t\}$ where $|x| \stackrel{\text{def}}{=} \sqrt{(x^1)^2 + (x^2)^2 + (x^3)^2}$. This is a consequence of the Sharp Huygens' Principle.

MIT OpenCourseWare
<http://ocw.mit.edu>

18.152 Introduction to Partial Differential Equations.
Fall 2011

For information about citing these materials or our Terms of Use, visit: <http://ocw.mit.edu/terms>.