ENERGY METHOD FOR WAVE EQUATIONS

*Willie Wong**

Based on commit 5dfb7e5 of 2017-11-06 13:29

Abstract We give an elementary discussion of the energy method (and particularly the vector field method) in the context of linear wave equations on $\mathbb{R} \times \mathbb{R}^d$.

The goal of this short note is to describe the modern approach to the energy method for, generally speaking, nonlinear and variable-coefficient wave equations, by exposing the theory in the context of the linear wave equations on $\mathbb{R} \times \mathbb{R}^d$.

Our main object of study will be solutions $\phi \in C^2(\mathbb{R} \times \mathbb{R}^d; \mathbb{R})$ to the equation

$$\begin{aligned} \Box \phi &= 0 \\ \phi(0, x) &= g(x) \\ \partial_t \phi(0, x) &= h(x) \end{aligned} \tag{0.1}$$

where the couple of real-valued functions (g, h) on $\mathbb{R}^d$ is referred to as the *initial conditions*. The $\Box$ operator, as a reminder, is $-\partial^2_{tt} + \triangle$ where $\triangle$ is the Laplacian on $\mathbb{R}^d$.

The energy method is an L^2-based method; in other words, we study integral quantities related to the square norm of the solution. In particular, we will be particularly concerned with integrals over certain hypersurfaces in $\mathbb{R} \times \mathbb{R}^d$; for this their geometries play important roles. Most modern discussions of the energy method in wave equations in fact proceed in the settings Lorentzian geometry, which is the natural setting for these discussions. However, as a first introduction, we will present explicit formulae in these notes and forego the high-powered geometric techniques in favor of direct computations. This hopefully makes the gist of the arguments more approachable.

1. THE ENERGY MOMENTUM TENSOR

For convenience, we will identify $\mathbb{R} \times \mathbb{R}^d$ with $\mathbb{R}^{d+1}$, and the t coordinate with the x_0 coordinate (taking the coordinate system of $\mathbb{R}^{d+1}$ to run from $(x_0, x_1, \dots, x_d)$). All displayed matrices will have indices running from $0, \dots, d$ and the rows and

*Michigan State University, East Lansing, USA; wongwwy@member.ams.org

column aligned accordingly. We also use the short hand $\partial_i = \partial_{x_i}$. We will let m denote the diagonal matrix corresponding to the *Minkowski metric*:

$$m := \begin{pmatrix} -1 & & & & \\ & 1 & & & \\ & & 1 & & \\ & & & \ddots & \\ & & & & 1 \end{pmatrix} \tag{1.1}$$

Note that with matrix multiplication, $m \cdot m = \delta$ (we use the Kronecker δ symbol for the matrix with diagonal entries "1", i.e. the identity matrix).

Notice that with the matrix m we can write the $\Box$ operator as

$$\Box\phi = \sum_{i,j=0}^{d} \partial_i (m_{ij} \partial_j \phi). \tag{1.2}$$

Of principle importance in this discussion is the following object:

1.3 Definition

Let $\phi \in C^2(\mathbb{R} \times \mathbb{R}^d; \mathbb{R})$ be arbitrary; its *energy momentum tensor* Q is the matrix-valued function given by

$$Q_{ij} = \partial_i \phi \; \partial_j \phi - \frac{1}{2} m_{ij} \Big(\sum_{k,\ell} m_{k\ell} \partial_k \phi \partial_\ell \phi \Big), \qquad i,j \in \{0, \dots, d\}.$$

■

As we saw on Homework 4 of this course, the energy momentum tensor enjoys the following divergence property, which relates it to the wave equation (0.1).

1.4 Lemma

For any $\phi \in C^2(\mathbb{R} \times \mathbb{R}^d; \mathbb{R})$:

$$\sum_{i,j} \partial_i (m_{ij} Q_j k) = \Box\phi \; \partial_k \phi.$$

■

Now, taking $X = (X_0, X_1, \dots, X_d)$ an arbitrary (smooth) vector field, we can define the vector field with components

$${}^{(X)}\mathcal{J}_i := \sum_{j,k} m_{ij} Q_{jk} X_k. \tag{1.5}$$

The vector field ${}^{(X)}\mathcal{J}$ is called the *energy-momentum current* associated to X. Its (space-time) divergence can be computed to be

$$\begin{aligned}\sum_{i=0}^{d} \partial_i {}^{(X)}\mathcal{J}_i &= \sum_{i,j,k} \partial_i (m_{ij} Q_{jk} X_k) \\ &= \sum_{i,j,k} \partial_i (m_{ij} Q_{jk}) X_k + \sum_{i,j,k} Q_{jk} m_{ji} \partial_i X_k\end{aligned}$$

The first term we can simplify using Lemma 1.4. For the second term we observe that Q is a symmetric matrix, so

$$\sum_{i,j,k} Q_{jk} m_{ji} \partial_i X_k = \sum_{i,j,k} Q_{jk} \cdot \frac{1}{2} \Big(m_{ji} \partial_i X_k + m_{ki} \partial_i X_j \Big).$$

1.6 Definition

Given the vector field X, its *deformation tensor* is the quantity

$$^{(X,0)}\pi_{jk} := \sum_{i=0}^{d} \Big(m_{ji} \partial_i X_k + m_{ki} \partial_i X_j \Big).$$

■

In this notation we conclude

$$\sum_{i=0}^{d} \partial_i {}^{(X)}\mathcal{J}_i = \Box\phi (X \cdot \partial\phi) + \frac{1}{2} \sum_{j,k} Q_{jk} {}^{(X,0)}\pi_{jk}. \tag{1.7}$$

This immediately implies

1.8 Lemma

Let ϕ be a solution to the wave equation. Then for any vector field X with vanishing deformation tensor, the current ${}^{(X)}\mathcal{J}$ is divergence free. ■

1.9 Corollary

If Ω is an open (space-time) region with piecewise C^1 boundaries, ϕ solves the wave equation, and X has vanishing deformation tensor, then

$$\int_{\partial\Omega} {}^{(X)}\mathcal{J} \cdot \vec{n} \, \mathrm{d}S = 0.$$

Here $\vec{n}$ refers to the unit outward normal vector field along $\partial\Omega \subset \mathbb{R} \times \mathbb{R}^d$. ■

1.10 Remark
Corollary 1.9 is in fact an instance of *Noether's theorem*.

Noether's theorem states, roughly, that associated to each symmetry of a system is a corresponding conservation law. In classical mechanics, time-translation symmetry corresponds to conservation of energy, spatial translation symmetry correspond to conservation of momentum, and rotational symmetry corresponds to conservation of angular momentum.

Here, the statement that ${}^{(X,0)}\pi \equiv 0$ is the assertion that X is a symmetry of the wave equations. The conclusion that a hypersurface integral vanishes is the conservation law. We will see better how this translates later when we discuss applications. ■

For our immediate applications, it suffices to observe that *any vector field with constant coefficients* automatically has vanishing deformation tensor; one should keep in mind that in further developments of the theory other vector fields can also come into play.

2. Space-time geometry

We make a few basic definitions; these should be familiar to readers with some experience in special relativity.

2.1 Definition
A vector X is said to be

- *time-like* if $X^T mX < 0$.
- *space-like* if $X^T mX > 0$.
- *light-like* or *null* if $X^T mX = 0$. ■

2.2 Example
The vector $(1,0,\dots,0)$ is time-like, but the vectors $(0,0,\dots,0,1,0,\dots,0)$ is space-like. The vector $(1,1,0,0,\dots,0)$ is null. ■

We can similarly characterize hypersurfaces by their unit normals.

2.3 Definition
A hypersurface $\Sigma \subset \mathbb{R} \times \mathbb{R}^d$ is said to be

- *space-like* (or sometimes, *co-time-like*) if $\vec{n}$ is time-like everywhere along Σ.
- *time-like* (or sometimes, *co-space-like*) if $\vec{n}$ is space-like everywhere along Σ.
- *light-like* or *null* if $\vec{n}$ is null.

Here $\vec{n}$ is a unit normal vector field along Σ. ■

2.4 Example

The hypersurface $\{x_0 = 0\}$ is space-like, the hypersurface $\{x_i = 0\}$ for any $i \in \{1, \ldots, d\}$ is time-like, and the hypersurface $\{x_0 + x_1 = 0\}$ is null. ■

A key proposition that gives the usefullness of the energy method is the following algebraic property (which you are asked to prove on HW5):

2.5 Proposition

If X, Y are two time-like vectors, such that $X \cdot Y > 0$. Then

$$^{(X)}\mathcal{J} \cdot Y \leq 0$$

with equality at a point (t, x) if and only if (t, x) is a critical point of ϕ. ■

We illustrate the proposition with a specific example:

2.6 Example

Let $X = Y = \partial_t$ be the vector corresponding to $(1, 0, 0, \ldots, 0)$. Clearly $X^T m X = m_{00} = -1 < 0$ so X is time-like. $X \cdot Y = 1 > 0$. We can compute

$$\begin{aligned}
^{(X)}\mathcal{J} \cdot Y &= \sum_{i,j,k} m_{ij} Q_{jk} X_k Y_i \\
&= \sum_j m_{0j} Q_{j0} \\
&= m_{00} Q_{00} \\
&= -\left[\partial_0\phi\partial_0\phi - \frac{1}{2} m_{00} \left(\sum_{i=1}^{d} (\partial_i\phi)^2 - (\partial_0\phi)^2 \right)\right] \\
&= -\frac{1}{2} \sum_{i=0}^{d} (\partial_i\phi)^2.
\end{aligned}$$

The expression is the negative of a sum of squares, and hence is non-positive, and vanishes precisely when $\partial_i\phi = 0$ for every i. ■

The following is the main technical result of these notes.

2.7 Theorem

Let Ω be an open space-time region with piecewise C^1 boundary. Assume that $\partial\Omega = \Sigma_+ \cup \Sigma_-$ where

- Both Σ_+ and Σ_- are space-like.
- On Σ_+ we have $\vec{n} \cdot (\partial_t) > 0$;

- and on Σ_- we have $\vec{n}\cdot(\partial_t) < 0$.

If $\phi \in C^2(\overline{\Omega};\mathbb{R})$ solves $\Box\phi = 0$ and $\nabla^{(t,x)}\phi \equiv 0$ along Σ_+ (resp. Σ_-), then $\nabla^{(t,x)}\phi \equiv 0$ along Σ_- (resp. Σ_+). ■

Proof From Corollary 1.9 we have that

$$\int_{\Sigma_+} {}^{(\partial_t)}\mathcal{J}\cdot\vec{n}\,\mathrm{d}S + \int_{\Sigma_-} {}^{(\partial_t)}\mathcal{J}\cdot\vec{n}\,\mathrm{d}s = 0.$$

By assumption the first term vanishes as a consequence of Proposition 2.5, the proposition also implying that the second term in the sum has non-negative integrand. That the total integral vanishes implies that the integrand must also vanish, which using the proposition again implies the desired result. □

3. Local uniqueness for wave equation

We now return to the initial value problem for the wave equation (0.1). One implication of Theorem 2.7 is the following.

3.1 Theorem

Let ϕ be a solution of (0.1). Suppose the initial data are such that $g \equiv 0 \equiv h$ on the set $\overline{B(0,r)}$. Then $\phi \equiv 0$ on the set

$$\{(t,x) \mid |t| + |x| \leq r\}.$$

■

Proof Define $\Sigma_0 = \{0\}\times\overline{B(0,r)}$. For $\tau \in (0,r)$, we let

$$\Sigma_\tau = \left[\{\tau\}\times B(0,\frac{r-\tau}{2})\right] \cup \left\{(t,x) \,\middle|\, r - \frac{\tau+r}{2\tau}t = |x|,\ t\in[0,\tau]\right\}.$$

Let Ω be the frustum bounded below by Σ_0 and above by Σ_τ. Quite clearly Σ_0 is a space-like hypersurface with $\vec{n}\cdot\partial_t < 0$. For Σ_τ, we observe that the top of the frustum is a portion of the hypersurface $\{t=\tau\}$ and so is space-like with $\vec{n}\cdot\partial_t > 0$. For the sides of the frustum the unit outward normal can be found to be the unit vector in the direction of

$$\left(\frac{\tau+r}{2\tau}, \frac{x_1}{|x|}, \frac{x_2}{|x|}, \ldots, \frac{x_d}{|x|}\right).$$

By virtue of the fact that $\tau < r$ we have $\tau + r > 2\tau$, this implies that $\vec{n}$ is time-like, and $\vec{n}\cdot\partial_t > 0$.

Applying Theorem 2.7 with Σ_- being Σ_0 and Σ_+ being Σ_τ for any of the fixed $\tau\in(0,r)$, we conclude that along Σ_τ, $\nabla^{(t,x)}\phi \equiv 0$.

Finally, observe that

$$\overline{\bigcup_{\tau\in(0,r)} \Sigma_\tau} = \{(t,x) \mid t \in [0,r], t + |x| \le r\}$$

we get that $\nabla\phi$ vanishes identically on this set, and since ϕ is therefore constant it must vanish by the initial conditions. We have the desired conclusion after treating the region with $t < 0$ with a time reflection. □

We can extend the previous theorem to initial data that vanish on sets of general shapes. We introduce the following notations.

3.2 Definition

Given any open $\Omega \in \mathbb{R}^d$, its *domain of dependence* is the set

$$\mathscr{D}(\Omega) = \overline{\bigcup_{(y,r)} \left\{(t,x) \in \mathbb{R}\times\mathbb{R}^d \;\middle|\; |t| + \left|x-y\right| \le r\right\}}$$

where the union is taken over all $(y,r) \in \Omega \times \mathbb{R}_+$ such that $B(y,r) \subset \Omega$.

Given any open $\Omega \in \mathbb{R}^d$, its *domain of influence* is the set

$$\mathscr{I}(\Omega) = (\mathbb{R}\times\mathbb{R}^d) \setminus \mathscr{D}(\mathbb{R}^d \setminus \overline{\Omega}).$$ ■

3.3 Proposition

A point $(t,x) \in \mathbb{R}\times\mathbb{R}^d$ is in $\mathscr{I}(\Omega)$ if and only if there exists a point $y \in \Omega$ such that the vector $(t, x-y)$ is time-like. ■

Proof Observe that the condition $(t,x-y)$ is time-like $\iff y \in B(x,|t|)$. We work by proving the contrapositive: since $B(x,|t|)$ is open, we have that $B(x,|t|)\cap\Omega = \emptyset$ if and only if $B(x,|t|) \subset \mathbb{R}^d \setminus \overline{\Omega}$, which is in trun equivalent to $(t,x) \in \mathscr{D}(\mathbb{R}^d \setminus \overline{\Omega})$ by definition. □

3.4 Theorem

Let ϕ be a solution of (0.1). Suppose the initial data are such that there exists some open $\Omega \subset \mathbb{R}^d$ such that $g = h = 0$ on Ω. Then $\phi \equiv 0$ on $\mathscr{D}(\Omega)$. ■

Proof Immediate consequence of Theorem 3.1. □

This is precisely *finite speed of propagation property*: indeed if $\operatorname{supp} g \cup \operatorname{supp} h \subset \Omega$ for some Ω open, then the above theorem implies $\operatorname{supp}\phi \subset \mathscr{I}(\Omega)$, which only contains points (t,x) such that $\operatorname{dist}(x,\Omega) < |t|$.

A uniqueness statement is also available for solutions on bounded domains.

3.5 Theorem
Let $\Omega \subset \mathbb{R}^d$ be open, bounded, with C^1 boundary. If $\phi \in C^2([0,T]\times\overline{\Omega})$ solves the initial-boundary-value problem

$$\begin{gathered} \Box\phi = 0 \\ \phi(0,x) = \partial_t\phi(0,x) = 0, \quad x \in \Omega \\ \phi(t,y) = 0, \quad t \in [0,T], y \in \partial\Omega \end{gathered}$$

Then $\phi \equiv 0$. ■

Proof Let $\Omega_\tau = (0,\tau)\times\Omega$. Apply Corollary 1.9 to Ω_τ with $X = \partial_t$. The boundary $\partial\Omega_\tau$ has three pieces: $\{\tau\}\times\Omega$, $\{0\}\times\Omega$, and $[0,\tau]\times\partial\Omega$. On the very last portion, note that $\vec{n}$ is orthogonal to ∂_t, and so

$$^{(\partial_t)}\mathcal{J}\cdot\vec{n} = \sum_i Q_{i0}\vec{n}_i = \partial_t\phi\partial_n\phi.$$

This quantity vanishes as along $[0,\tau]\times\partial\Omega$, by assumption $\phi \equiv 0$ and hence $\partial_t\phi \equiv 0$. Therefore we are left with

$$\int_{\{0\}\times\Omega} {}^{(\partial_t)}\mathcal{J}\cdot\vec{n}\,\mathrm{d}S + \int_{\{\tau\}\times\Omega} {}^{(\partial_t)}\mathcal{J}\cdot\vec{n}\,\mathrm{d}S = 0.$$

The first term vanishes by the initial condition. The integrand in the second term is signed by Proposition 2.5, and hence must identically vanish as the integral is zero. By the proposition again we have that $\nabla^{(t,x)}\phi \equiv 0$ on $\{\tau\}\times\Omega$. As τ is arbitrary, this means that $\nabla^{(t,x)}\phi \equiv 0$ on $(0,T)\times\Omega$, and hence is constant, and hence vanishes. □

4. Propagation of regularity

This final section addresses something mentioned when we discussed the fundamental solution of the wave equation, that the solution seems to be *less differentiatble* than the given initial data, with the situation getting worse and worse in higher and higher dimensions. It turns out that this can be ameliorated if one thinks of regularity not in the pointwise differentiability sense, but in the sense of energy integrals.

We illustrate this point by first writing down two "energy estimates".

4.1 Proposition
Suppose ϕ solves (0.1) with $g,h \in C_c^\infty(\mathbb{R}^d)$. Then for every $t \in \mathbb{R}$ we have that

$$\int_{\mathbb{R}^d} \left|\partial_t\phi(t,x)\right|^2 + \left|\nabla^{(x)}\phi(t,x)\right|^2\,\mathrm{d}x = \int_{\mathbb{R}^d} |h|^2 + |\nabla g|^2\,\mathrm{d}x.$$

■

Proof This follows from applying Corollary 1.9 to the slab $[0,t]\times\mathbb{R}^d$ and using that by finite speed of propagation $x\mapsto\phi(t,x)$ has compact support for every t. $\square$

4.2 Proposition

Suppose ϕ solves (0.1). Then for any $(t,x)\in\mathbb{R}\times\mathbb{R}^d$ and $r\in\mathbb{R}_+$, we have that

$$\int_{B(x,r)} \left|\partial_t\phi(t,y)\right|^2+\left|\nabla\phi(t,y)\right|^2\,\mathrm{d}y \le \int_{B(x,r+|t|)} |h|^2+|\nabla g|^2\,\mathrm{d}y.$$

■

Proof The claim follows from applying Corollary 1.9 to the frustum given by $\Omega=\mathscr{D}(B(x,r+|t|))\cap[0,t]\times\mathbb{R}^d$. Notice that by a direct computation, for the portion of $\partial\Omega$ that is the sloped sides of the frustum, we can check ${}^{(\partial_t)}\mathcal{J}\cdot\vec{n}\le 0$ (similar to the proof of Theorem 3.1). $\square$

Now, on homework 5 you are asked to show also the following result.

4.3 Lemma

If ϕ solves the wave equation, and if X is any vector field with ${}^{(X,0)}\pi\equiv 0$, then the function $\psi=\sum_{i=0}^d X_i\partial_i\phi$ also solves the wave equation. ■

So the energy estimates in Proposition 4.1 and 4.2 would also apply, provided that g,h are replaced by their correct counterparts. In particular, we have that for α any multi-index for $\mathbb{R}^d$,

$$\int_{\mathbb{R}^d}\left|\partial_t\partial^\alpha\phi(t,x)\right|^2+\left|\nabla^{(x)}\partial^\alpha\phi(t,x)\right|^2\,\mathrm{d}x=\int_{\mathbb{R}^d}|\partial^\alpha h|^2+|\nabla\partial^\alpha g|^2\,\mathrm{d}x \tag{4.4}$$

in the case g,h have compact support, or in the general case

$$\int_{B(x,r)} \left|\partial_t\partial^\alpha\phi(t,y)\right|^2+\left|\nabla\partial^\alpha\phi(t,y)\right|^2\,\mathrm{d}y \le \int_{B(x,r+|t|)} |\partial^\alpha h|^2+|\nabla\partial^\alpha g|^2\,\mathrm{d}y. \tag{4.5}$$

In this sense, we have the informal statement that

> A solution of the wave equation on $\mathscr{D}(\Omega)$ is as regular as its initial data in Ω, provided we measure regularity in the energy sense.

This is one of the reasons why in the second course in this qualifying exam sequence you will be learning about *Sobolev spaces*, which will allow you to make the above informal statement more precise.

References

[1] Demetrios Christodoulou and Sergiu Klainerman. "Asymptotic properties of linear field equations in Minkowski space". In: *Comm. Pure Appl. Math.* 43.2 (1990), pp. 137–199. DOI: 10.1002/cpa.3160430202. URL: http://dx.doi.org/10.1002/cpa.3160430202.

[2] S. Klainerman. "A commuting vectorfields approach to Strichartz-type inequalities and applications to quasi-linear wave equations". In: *Int. Math. Res. Not. IMRN* 2001.5 (2001), pp. 221–274. ISSN: 1073-7928. DOI: 10.1155/S1073792801000137. URL: http://dx.doi.org/10.1155/S1073792801000137.

[3] Sergiu Klainerman. "Lecture Notes for Introduction to Analysis". Princeton University, MAT 535, 2004 Fall semester Updated to 2008 version. 2004 / 2008.