Chapter 1

The Wave Equation and Special Relativity

1.1 The wave equation from Maxwell's equations

In this course we are going to take a PDE based approach to relativity. We will begin with exploring special relativity in the context of the wave equation:

$$-\frac{\partial^2 u}{\partial t^2} + c^2 \Delta u = 0.$$

This provides a convenient proxy for the study of Maxwell's equations. We could instead study Maxwell's equations directly, but since these are a system of PDEs for 6 components of the electromagnetic field they can be a bit unwieldy.

In fact, Maxwell's equations are closely related to the wave equation, as our first result will establish:

Lemma 1.1. *Let us denote* $S_T := \mathbb{R}^3 \times (-T, T)$. *Suppose* $\boldsymbol{E}, \boldsymbol{B} \in C^2(S_T)$ *satisfy Maxwell's equations. Then each component of* $\boldsymbol{E}, \boldsymbol{B}$ *satisfies the wave equation:*

$$\begin{aligned} -\frac{\partial^2 E_i}{\partial t^2} + c^2 \Delta E_i &= 0 \\ -\frac{\partial^2 B_i}{\partial t^2} + c^2 \Delta B_i &= 0 \end{aligned}$$

in S_T, *where* $c = (\mu_0 \epsilon_0)^{-\frac{1}{2}}$.

Proof. We start with (2) and differentiate in time to find

$$-\frac{\partial^2 \boldsymbol{B}}{\partial t^2} - \boldsymbol{\nabla} \times \frac{\partial \boldsymbol{E}}{\partial t} = 0.$$

Using (4) to replace the term involving $\dot{\boldsymbol{E}}$, we have

$$-\frac{\partial^2 \boldsymbol{B}}{\partial t^2} - \frac{1}{\mu_0 \epsilon_0} \boldsymbol{\nabla} \times (\boldsymbol{\nabla} \times \boldsymbol{B}) = 0.$$

Now, a standard vector calculus identity tells us that $\boldsymbol{\nabla} \times (\boldsymbol{\nabla} \times \boldsymbol{A}) = \boldsymbol{\nabla} (\boldsymbol{\nabla} \cdot \boldsymbol{A}) - \Delta \boldsymbol{A}$ for any C^2 vector field $\boldsymbol{A}$. Making use of this, together with (3), we deduce the result. □

Exercise($*$). Complete the proof by showing that the components of $\boldsymbol{E}$ obey the wave equation.

Later on in the course we shall be able to show a converse to this result, namely that we can find a solution of Maxwell's equations by solving the wave equation for each component.

At this stage, it's useful to assume that we are using units in which $c = 1$. For example, we can take the second as our unit of time and the light-second as our unit of length. This is convenient as it saves us carrying around a constant in our formulae. If you want to replace c it's always possible to do so by thinking about what units various quantities ought to carry.

1.2 The Cauchy problem for the wave equation

For the rest of this chapter, we shall be discussing solutions of the wave equation. We will start by considering the Cauchy problem. In general, the Cauchy problem for a PDE consists of specifying some data on a given surface and then trying to find a unique solution of the PDE in a neighbourhood of that surface. We will discuss this much more thoroughly when we come to the Cauchy-Kovalevskaya theorem. For now, let us consider what data we might expect to have to specify on the surface $\Sigma_0 := \mathbb{R}^3 \times \{0\}$ in order to find a unique solution of the wave equation:

$$-\frac{\partial^2 u}{\partial t^2} + \Delta u = 0 \tag{1.1}$$

in $S_T := \mathbb{R}^3 \times (-T, T)$. In order to see what data might be necessary to solve this problem, we can try and write u as a formal[1] power series in t about $t = 0$. That is, to try and find coefficients u_n so that the formal power series:

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

solves (1.1). Putting this series into the equation, we find that we can cancel the terms order by order if:

$$-u_{n+2}(x) + \Delta u_n(x) = 0, \qquad n = 0, 1, \ldots$$

In other words, once we have specified $u_0(x) = u(x,0)$ and $u_1(x) = u_t(x,0)$, at least in principle we can find the rest of the terms in the formal power series by repeated differentiation. In other words, we expect that the correct *Cauchy data* for the wave equation on Σ_0 are the values of $u|_{\Sigma_0}$ and $u_t|_{\Sigma_0}$.

The formal power series argument above, while suggestive, is not terribly useful. For the solutions we shall ultimately be interested in, the series expansion above will not converge, so any arguments based on manipulating these series are rather suspect. We need a better tool before we can understand solutions of the wave equation. In particular, we require some *a priori* estimates for the solutions. Recall from the Theory of PDE

[1] in this context, 'formal' means that we will ignore issues of convergence.

course that an a priori estimate is a an estimate that we can deduce directly from the equation, *without having to write down a solution*, i.e. they are estimates that must hold for all solutions. In the Theory of PDE course, the main a priori estimates we used were the Maximum Principle for Laplace's equation and for the heat equation. For the wave equation *no maximum principle exists*, instead we make use of energy estimates.

Theorem 1.1 (Basic energy estimate for the wave equation)**.** *Let* $S_T := \mathbb{R}^3 \times (-T, T)$ *and* $\Sigma_t = \mathbb{R}^3 \times \{t\}$*. Suppose that* $u \in C^2(S_T)$ *solves (1.1) in* S_T*, and that there exists* R *such that* $u(x,t) = 0$ *for* $|x| > R$*. Then if we define*

$$E[u](t) := \frac{1}{2}\int_{\Sigma_t} \left(u_t^2 + |\boldsymbol{\nabla} u|^2\right) d\sigma = \frac{1}{2}\int_{\mathbb{R}^3} \left(u_t(x,t)^2 + |\boldsymbol{\nabla} u(x,t)|^2\right) dx,$$

we have

$$\frac{dE[u]}{dt} = 0.$$

Proof. The assumptions on $u(x,t)$ imply that we can restrict the range of the integration in E to $B_{2R}(0)$ and that that we can differentiate $E(t)$ with respect to time and pass the time derivative under the integral[2] to obtain

$$\begin{aligned} \frac{dE[u]}{dt} &= \int_{B_{2R}(0)} (u_t u_{tt} + \boldsymbol{\nabla} u \cdot \boldsymbol{\nabla} u_t)\, dx \\ &= \int_{B_{2R}(0)} (u_t u_{tt} - \Delta u u_t)\, dx + \int_{\partial B_{2R}(0)} u_t \boldsymbol{\nabla} u \cdot \boldsymbol{n}\, d\sigma \end{aligned}$$

Here we have used the vector calculus identity $\boldsymbol{\nabla} \cdot (f\boldsymbol{\nabla} g) = \boldsymbol{\nabla} f \cdot \boldsymbol{\nabla} g + f\Delta g$ and then applied the divergence theorem. The first integral vanishes, because the integrand is proportional to $u_{tt} - \Delta u$ which is zero since u obeys the wave equation. The second integral vanishes because $u = 0$ for $|x| > R$. The result follows. □

Later on, we shall show that the assumption that solutions vanish outside a sufficiently large ball is in fact a reasonable one, because solutions to the wave equation exhibit *finite speed of propagation.* In the context of relativity, this is an analogue of the statement that no signal may travel faster than the speed of light.

From this result, we can immediately deduce that a solution of the wave equation is indeed uniquely determined by our proposed Cauchy data, namely $u|_{\Sigma_0}$ and $u_t|_{\Sigma_0}$.

Corollary 1.2. *Suppose* $u, v \in C^2(S_T)$ *solve (1.1) in* S_T*, and that there exists* R *such that* $u(x,t) = v(x,t) = 0$ *for* $|x| > R$*. Suppose further that:*

$$u|_{\Sigma_0} = v|_{\Sigma_0}, \qquad u_t|_{\Sigma_0} = v_t|_{\Sigma_0}.$$

Then $u = v$ *in* S_T*.*

---

[2] Check that you understand why this is true. You may wish to look at the appendix of the notes for the course MA3G1.

Proof. Consider $w = u - v$. This solves (1.1) in S_T and is in $C^2(S_T)$. Moreover, $w|_{\Sigma_0} = w_t|_{\Sigma_0} = 0$. Thus $E[w](0) = 0$. By Theorem 1.1 we have that $E[w](t) = 0$ for all $-T < t < T$. As a consequence, we have $|\boldsymbol{\nabla} w| = |w_t| = 0$ on S_T, which together with the fact that w vanishes for large x implies that $w = 0$ on S_T. □

Our energy estimate has shown that specifying u and u_t on a surface of constant time is enough to ensure the solution to the Cauchy problem, if it exists, is unique. In fact, the energy estimate gives us a statement of continuity: the L^2-norms of $u_t(\cdot, t)$ and $|\boldsymbol{\nabla} u|(\cdot, t)$ are controlled by the initial data. More precisely, we say that two solutions u, v as in Corollary 1.3 are $\epsilon-$close in the energy norm at time t if:

$$E[u - v](t) \leq \epsilon.$$

We clearly have

Corollary 1.3. *Suppose $u, v \in C^2(S_T)$ solve (1.1) in S_T, and that there exists R such that $u(x,t) = v(x,t) = 0$ for $|x| > R$. Suppose further that u and v are initially $\epsilon-$close in the energy norm. Then they remain $\epsilon-$close for all times $t \in (-T, T)$.*

We can improve the control over the solution to control higher derivatives:

Exercise 1.1. Let $S_T := \mathbb{R}^3 \times (-T, T)$ and $\Sigma_t = \mathbb{R}^3 \times \{t\}$. Fix $k \in \mathbb{N}$. Suppose that $u \in C^{2+k}(S_T)$, solves the wave equation (1.1) in S_T, and that there exists R such that $u(x,t) = 0$ for $|x| > R$. Define $u_0 := u|_{\Sigma_0}$ and $u_1 := u_t|_{\Sigma_0}$.

a) By deriving an equation for $\nabla_i u$ for $i = 1, 2, 3$ show that[3]

$$\frac{1}{2}\int_{\Sigma_t} \left(|\boldsymbol{\nabla} u_t|^2 + \left|\boldsymbol{\nabla}^2 u\right|^2 \right) d\sigma = \frac{1}{2}\int_{\mathbb{R}^3} \left(|\boldsymbol{\nabla} u_1|^2 + \left|\boldsymbol{\nabla}^2 u_0\right|^2 \right) dx$$

for $-T < t < T$.

b) Deduce that:

$$\frac{1}{2}\int_{\Sigma_t} \left(\left|\boldsymbol{\nabla}^k u_t\right|^2 + \left|\boldsymbol{\nabla}^{k+1} u\right|^2 \right) d\sigma = \frac{1}{2}\int_{\mathbb{R}^3} \left(\left|\boldsymbol{\nabla}^k u_1\right|^2 + \left|\boldsymbol{\nabla}^{k+1} u_0\right|^2 \right) dx.$$

for $-T < t < T$.

We have thus established that a solution, if it exists, to the Cauchy problem for the wave equation is unique and moreover that the solution depends continuously (in an appropriate sense) on the initial data. The final aspect of well posedness that remains to prove is that a solution does in fact exist. Rather than prove this now, we shall postpone our discussion to a later date, and just state a well posedness result.

[3] Here $\left|\boldsymbol{\nabla}^2 u\right|^2 = \sum_{i,j} \nabla_i \nabla_j u \nabla_i \nabla_j u$, etc.

Theorem 1.4 (Well posedness of the wave equation). *Suppose that $u_0, u_1 \in C_c^\infty(\mathbb{R}^3)$. Then there exists a unique solution $u \in C^\infty(\mathbb{R}^3 \times \mathbb{R})$ to the Cauchy problem:*

$$\begin{cases} -u_{tt} + \Delta u = 0 & \text{in } \mathbb{R}^3 \times \mathbb{R}, \\ u = u_0, \quad u_t = u_1 & \text{on } \mathbb{R}^3 \times \{0\}. \end{cases}$$

Moreover, if supp $u_i \subset B_R(0)$, for $i = 1, 2$, then supp $u(\cdot, t) \subset B_{R+|t|}(0)$.

Remarks: we have stated the theorem for smooth functions. Finite regularity versions of this theorem can be stated, however, if we work with spaces of C^k functions, then we have to assume more regularity on the initial data than we are able to recover for the solution. We say that the C^k-norms are not propagated by the wave equation. This is closely related to the absence of a maximum principle for the wave equation (see Exercise 1.2). The natural function spaces to consider are in fact the H^k spaces, whose norms *are* propagated by the equations[4].

Notice that the support of the function grows at most linearly in time. This is related to the finite speed of propagation. Finally, we have stated a result which is *global in* t, i.e. we do not restrict to a time interval $(-T, T)$. Obviously our result implies similar results on S_T.

Exercise 1.2. Let $\mathbb{R}^3_* := \mathbb{R}^3 \setminus \{0\}$, $S_{*,T} := \mathbb{R}^3_* \times (-T, T)$ and $|x| = r$. You may assume the result that if $u = u(r, t)$ is radial, we have

$$\Delta u(|x|, t) = \Delta u(r, t) = \frac{\partial^2 u}{\partial r^2}(r, t) + \frac{2}{r}\frac{\partial u}{\partial r}(r, t)$$

a) Suppose $u(x, t) = \frac{1}{r} v(r, t)$ for some function v. Show that u solves the wave equation on $\mathbb{R}^3_* \times (0, T)$ if and only if v satisfies the one-dimensional wave equation

$$-\frac{\partial^2 v}{\partial t^2} + \frac{\partial^2 v}{\partial r^2} = 0$$

on $(0, \infty) \times (-T, T)$.

b) Suppose $f, g \in C_c^2(\mathbb{R})$. Deduce that

$$u(x, t) = \frac{f(r + t)}{r} + \frac{g(r - t)}{r}$$

is a solution of the wave equation on $S_{*,T}$ which vanishes for large $|x|$.

c) Show that if $f \in C_c^3(\mathbb{R})$ is an odd function (i.e. $f(s) = -f(-s)$ for all s) then

$$u(x, t) = \frac{f(r + t) + f(r - t)}{2r}$$

extends as a C^2 function which solves the wave equation on S_T, with

$$u(0, t) = f'(t).$$

[4]You will have met these spaces if you took MA4A2: Advanced PDEs. Otherwise you may ignore this comment.

*d) By considering a suitable sequence of functions f, or otherwise, deduce that there exists no constant C independent of u such that the estimate

$$\sup_{S_T} (|u| + |u_t|) \leq C \sup_{\Sigma_0} (|u| + |u_t|)$$

holds for all solutions $u \in C^2(S_T)$ of the wave equation which vanish for large $|x|$.

1.3 Minkowski Spacetime

In the Theory of PDE course, we constructed solutions to the Cauchy problem for the heat equation on $\mathbb{R}^3$:

$$\begin{cases} u_t = \Delta u & \text{in } (0,T) \times \mathbb{R}^3, \\ u = u_0 & \text{on } \{0\} \times \mathbb{R}^3. \end{cases}$$

Physically, the function $u(t, \cdot)$ corresponds to the temperature at time t of an infinite uniform body whose initial temperature is given by the function u_0. In order to write the heat equation out in full, we pick an orthonormal basis for $\mathbb{R}^3$, say $\{\boldsymbol{e}_i\}_{i=1,2,3}$. Such a choice is called a frame. Once we have chosen a frame, we can define the partial derivatives in the $\boldsymbol{e}_i$ directions. For any function $f \in C^1(\mathbb{R}^3)$, we write:

$$\frac{\partial f}{\partial x^i}(\boldsymbol{x}) = \left.\frac{d}{ds} f(\boldsymbol{x} + s\boldsymbol{e}_i)\right|_{s=0}.$$

The Laplacian acting on a function u is then given by

$$\Delta f(\boldsymbol{x}) = g^{ij} \frac{\partial^2 f}{\partial x^i \partial x^j}(\boldsymbol{x}) = g^{ij} \frac{\partial}{\partial x^i} \left[\frac{\partial f}{\partial x^j} \right] (\boldsymbol{x}).$$

Here, I am using summation convention, so there is an implicit summation over $i, j = 1, 2, 3$. I have also introduced a new tensor, the Euclidean cometric tensor, whose components are given by:

$$g^{ij} = \begin{cases} 1 & i = j, \\ 0 & i \neq j. \end{cases}$$

As a result, when we perform the implicit summation, the only terms that survive are the diagonal ones and we have:

$$\Delta f(\boldsymbol{x}) = \frac{\partial^2 f}{\partial x^1 \partial x^1}(\boldsymbol{x}) + \frac{\partial^2 f}{\partial x^2 \partial x^2}(\boldsymbol{x}) + \frac{\partial^2 f}{\partial x^3 \partial x^3}(\boldsymbol{x}) \tag{1.2}$$

Now, in the physical explanation of what u represents, I made no reference to any particular frame. The heat equation (1.2), however, makes explicit reference to the coordinates x^i associated to the frame. To reconcile these two facts, we can show that if we change our orthonormal basis, we must find that the heat equation is unchanged.

Suppose $\boldsymbol{e}'_i$ is another basis for $\mathbb{R}^3$. We must have that

$$\boldsymbol{e}_i = \boldsymbol{e}'_j \Lambda^j{}_i$$

for some real numbers $\Lambda^j{}_i$. Now, let us consider how changing basis affects the partial derivatives of a function $f : \mathbb{R}^3 \to \mathbb{R}$. Starting from the definition of the partial derivative, we have

$$\begin{aligned} \frac{\partial f}{\partial x^i}(\boldsymbol{x}) &= \frac{d}{ds} f(\boldsymbol{x} + s\boldsymbol{e}_i)\bigg|_{s=0} \\ &= \frac{d}{ds} f\left(\boldsymbol{x} + s\boldsymbol{e}'_j \Lambda^j{}_i\right)\bigg|_{s=0} \\ &= \Lambda^j{}_i \frac{d}{ds} f\left(\boldsymbol{x} + t\boldsymbol{e}'_j\right)\bigg|_{s=0} \\ &= \Lambda^j{}_i \frac{\partial f}{\partial x'^j}(\boldsymbol{x}) \end{aligned} \tag{1.3}$$

where we have used the chain rule to go from the second to the third line. Since $\Lambda^i{}_j$ don't depend on $\boldsymbol{x}$ we can differentiate again and we find

$$\frac{\partial^2 f}{\partial x^i \partial x^j}(\boldsymbol{x}) = \Lambda^k{}_i \Lambda^l{}_j \frac{\partial^2 f}{\partial x'^k \partial x'^l}(\boldsymbol{x})$$

We see then that the heat equation has the same form with respect to the basis $\{\boldsymbol{e}'_i\}$ as it did for the basis $\{\boldsymbol{e}_i\}$ if and only if

$$\Lambda^k{}_i g^{ij} \Lambda^l{}_j = g^{kl}. \tag{1.4}$$

Thinking of $\Lambda^i{}_j$ and g^{ij} as the components of two matrices, this is equivalent to the matrix equation:

$$\Lambda \Lambda^T = I,$$

which we recognise as the condition that $\Lambda \in O(3)$, i.e. Λ is orthogonal. In other words, the heat equation takes the same form with respect to any orthonormal frame for $\mathbb{R}^3$. This reflects a physical symmetry of the underlying problem: the invariance under rotations and reflections of a uniform body filling all of space.

We could in fact have taken the form of the heat equation to *define* the orthonormal frames. In this approach, we would say that a basis for $\mathbb{R}^3$ is orthonormal if and only if the heat equation takes the form

$$\frac{\partial u}{\partial t}(\boldsymbol{x}) = \frac{\partial^2 u}{\partial x^1 \partial x^1}(\boldsymbol{x}) + \frac{\partial^2 u}{\partial x^2 \partial x^2}(\boldsymbol{x}) + \frac{\partial^2 u}{\partial x^3 \partial x^3}(\boldsymbol{x})$$

with respect to this basis. In principle, we could hope to make physical measurements to determine the form of the heat equation, so this connects a mathematical construction (an orthonormal frame) to the physics of heat propagation.

The wave equation and the Lorentz group

Now we come to the equation we are actually interested in, the wave equation. The aim is, in a similar way to our brief discussion above, to look at the symmetries of the wave equation. The first thing we do is to drop the distinction between time and space.

Trivially, we can consider a point $(t, \boldsymbol{x}) \in \mathbb{R} \times \mathbb{R}^3$ as belonging to $\mathbb{R}^4$, which we wish to think of as the space-time manifold (what this means will become clearer as the course progresses). We introduce a basis for spacetime:

$$\vec{e}_0 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \vec{e}_1 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad \vec{e}_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad \vec{e}_3 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

Defining $x^0 := t$, we can write any point in $\mathbb{R}^4$ as

$$\vec{X} := (x^0, \boldsymbol{x}) = x^\mu \vec{e}_\mu$$

where we take the convention that greek indices μ, ν, etc. are summed over $0, \ldots, 3$. In order to write the wave equation concisely, we introduce a $(0, 2)-$tensor with components $\eta_{\mu\nu}$ called the metric tensor, given by:

$$\eta_{\mu\nu} = \begin{cases} -1 & \mu = \nu = 0 \\ 1 & \mu = \nu = 1, 2, \text{ or } 3, \\ 0 & \mu \neq \nu. \end{cases} \tag{1.5}$$

Notice that $\{\vec{e}_\mu\}$ is an orthonormal basis for η. The metric tensor has an inverse, a $(2, 0)-$tensor with components $\eta^{\mu\nu}$ satisfying

$$\eta^{\mu\sigma}\eta_{\sigma\nu} = \delta^\mu{}_\nu.$$

Here $\delta^\mu{}_\nu$ is the usual Kronecker delta, defined by

$$\delta^\mu{}_\nu = \begin{cases} 1 & \mu = \nu \\ 0 & \mu \neq \nu. \end{cases}$$

We can then write the wave equation very concisely as

$$-u_{tt} + \Delta u = \eta^{\mu\nu} \frac{\partial^2 u}{\partial x^\mu \partial x^\nu} = 0.$$

We'll sometimes use the notation

$$\Box u = \eta^{\mu\nu} \frac{\partial^2 u}{\partial x^\mu \partial x^\nu},$$

in the same way that Δ is used for the Laplacian. The operator $\Box$ is known as the *d'Alembertian* or *wave operator*.

Definition 1. We define the Minkowski spacetime, $\mathbb{E}^{1,3}$, to be $\mathbb{R}^4$ equipped with the metic tensor $\eta \in T^0{}_2(\mathbb{R}^4)$.

1.3.1 Lorentz transformations

As we did for the heat equation above, it is instructive to consider what happens when we choose another basis (or frame) for $\mathbb{R}^4$. Suppose that $\{\vec{e}_\mu{}'\}$ is a new basis for $\mathbb{R}^4$. We of course have

$$\vec{e}_\mu = \vec{e}_\nu{}' \Lambda^\nu{}_\mu,$$

for a set of real numbers $\Lambda^\nu{}_\mu$. The same calculation as (1.3) shows that for any f which is differentiable in a neighbourhood of $X \in \mathbb{E}^{1,3}$ we have

$$\frac{\partial f}{\partial x^\mu}(\vec{X}) = \Lambda^\nu{}_\mu \frac{\partial f}{\partial x'^\nu}(\vec{X}).$$

This means that the numbers

$$\frac{\partial f}{\partial x^\mu}(\vec{X})$$

transform like a $(0,1)-$tensor under a change of basis. It's convenient to write this as

$$\nabla_\mu f := \frac{\partial f}{\partial x^\mu},$$

so that the wave equation can be written

$$\nabla^\mu \nabla_\mu u = \eta^{\mu\nu} \nabla_\nu \nabla_\mu u = \nabla_\nu \eta^{\mu\nu} \nabla_\mu u = \nabla_\mu \nabla^\mu u = 0.$$

It is standard to suppress the argument to keep the notation clean, but it can be replaced if necessary.

Since the $\Lambda^\nu{}_\mu$ are constant, we see that the form of the wave equation is preserved if

$$\eta^{\mu\nu} = \Lambda^\mu{}_\sigma \eta^{\sigma\tau} \Lambda^\nu{}_\tau$$

Thinking of $\Lambda^\mu{}_\nu$ and $\eta_{\mu\nu}$ as matrices, we can write this condition as:

$$\Lambda \eta^{-1} \Lambda^T = \eta^{-1}. \tag{1.6}$$

Definition 2. 1. A basis with respect to which η takes the form (1.5) is called an *inertial frame*. Equivalently, an inertial frame is one for which the wave equation takes the form

$$-\frac{\partial^2 u}{\partial x^0 \partial x^0}(\vec{X}) + \frac{\partial^2 u}{\partial x^1 \partial x^1}(\vec{X}) + \frac{\partial^2 u}{\partial x^2 \partial x^2}(\vec{X}) + \frac{\partial^2 u}{\partial x^3 \partial x^3}(\vec{X}) = 0$$

2. A matrix Λ satisfying (1.6) is said to be a *Lorentz transformation*, and represents a transformation between inertial frames. The set of all Lorentz transformations $O(1,3)$ forms a group under matrix multiplication, called the *Lorentz group*.

The connection between an inertial frame and the wave equation allows us, in principle, to determine whether a frame is inertial or not by making physical measurements of some quantity obeying the wave equation.

Lemma 1.2. *The following conditions are equivalent to (1.6):*

$$\Lambda^{-1} = \eta^{-1}\Lambda^T\eta.$$

and

$$\Lambda^T\eta\Lambda = \eta.$$

Proof. Multiplying (1.6) on the right by η, we have

$$\Lambda\eta^{-1}\Lambda^T = \eta^{-1} \iff \Lambda\left(\eta^{-1}\Lambda^T\eta\right) = I \iff \Lambda^{-1} = \eta^{-1}\Lambda^T\eta.$$

since Λ is a square matrix. For the second condition, multiply the first on the right by Λ and on the left by η to see

$$\Lambda^{-1} = \eta^{-1}\Lambda^T\eta \iff \eta = \eta\eta^{-1}\Lambda^T\eta\Lambda = \Lambda^T\eta\Lambda.$$

□

Since the three conditions are equivalent, we can take any of them to define the Lorentz group.

Example 1. The Lorentz group contains the orthogonal group $O(3)$ as a subgroup. To see this, consider the set of matrices of the form

$$\Lambda = \begin{pmatrix} 1 & 0_{1\times 4} \\ 0_{4\times 1} & R_{3\times 3} \end{pmatrix}$$

where $R_{3\times 3} \in O(3)$. It is straightforward to verify that $\Lambda \in O(1,3)$.

The Lorentz group is a *Lie group*[5]. We can think of the Lorentz group as a Lie subgroup of $GL(4)$ (since by Lemma 1.2 Lorentz transformations are invertible).

Theorem 1.5. *The Lorentz group is a Lie subgroup of $GL(4)$, and its Lie algebra, denoted by $\mathfrak{o}(1,3)$ is the set of matrices ℓ satisfying*

$$\ell\eta^{-1} + \eta^{-1}\ell^T = 0.$$

Proof. Consider the map $f : GL(4) \to Sym(4\times 4)$ given by

$$f(A) = A\eta^{-1}A^T.$$

Clearly $O(1,3)$ is the preimage of η^{-1}, so it will be enough to show that η^{-1} is a regular point of the map f. To see this, we calculate

$$\begin{aligned} df|_A &= (dA)\eta^{-1}A^T + A\eta^{-1}(dA)^T \\ &= (dA)A^{-1}\eta^{-1} + \eta^{-1}\left[(dA)A^{-1}\right]^T \end{aligned}$$

[5]Don't worry if you've not met these before. All it means is that in addition to having a group structure, $O(1,3)$ can be made into a manifold, in such a way that the group operations are continuous.

for $A \in f^{-1}(\eta^{-1})$. To show that the derivative map is surjective, we must show that for any point $A \in f^{-1}(\eta^{-1})$ and any symmetric matrix $S \in Sym(4 \times 4)$ we can find a matrix B such that

$$BA^{-1}\eta^{-1} + \eta^{-1}\left[BA^{-1}\right]^T = S$$

For this, we can simply take $B = \frac{1}{2}S\eta A$. Thus η^{-1} is a regular point of f, and we deduce that $O(1,3)$ is a closed subgroup of $GL(4)$ and hence is a Lie subgroup. To complete the proof, we recall that the Lie algebra of a matrix Lie group is simply the tangent space at the identity, i.e. $\mathfrak{o}(1,3) = T_I O(1,3)$. From the regular value theorem we know that $T_I O(1,3) = \text{Ker } df|_I$, which corresponds precisely to those matrices ℓ satisfying

$$\ell\eta^{-1} + \eta^{-1}\ell^T = 0.$$

□

The group $O(1,3)$ has four connected components. The connected component containing the identity is a subgroup of $O(1,3)$ called the *proper orthochronous Lorentz group* and is denoted $SO^+(1,3)$. There are two important Lorentz transformations which do not belong to $SO^+(1,3)$. These have matrices[6]:

$$\mathbb{T} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}; \qquad \mathbb{P} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}.$$

Using T, P we can write $O(1,3)$ as the disjoint union of connected cosets:

$$O(1,3) = SO^+(1,3) \cup \mathbb{T}\left(SO^+(1,3)\right) \cup \mathbb{P}\left(SO^+(1,3)\right) \cup \mathbb{PT}\left(SO^+(1,3)\right).$$

We've thus reduced the problem of understanding $O(1,3)$ to the problem of understanding $SO^+(1,3)$. Any element of $SO^+(1,3)$ may be written as

$$\Lambda = e^\ell = \sum_{n=0}^{\infty} \frac{\ell^n}{n!}$$

for some $\ell \in \mathfrak{o}(1,3)$, where the exponential here is the matrix exponential.

Example 2. We can take

$$\ell = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

[6] Check that these satisfy (1.6)

We check that

$$\ell\eta^{-1} + \eta^{-1}\ell^T = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} + \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$
$$= \begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = 0.$$

Clearly if $\ell \in \mathfrak{o}(1,3)$, then so is $-s\ell$. The matrix exponential of $-s\ell$ is straightforward to calculate using the fact that

$$\ell^2 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

and $\ell^3 = \ell$. We find

$$\begin{aligned} e^{-s\ell} &= \sum_{n=0}^{\infty} \frac{(-s)^n \ell^n}{n!} = I - \ell \sum_{n=0}^{\infty} \frac{s^{2n+1}}{(2n+1)!} + \ell^2 \sum_{n=1}^{\infty} \frac{s^{2n}}{(2n)!} \\ &= I - \ell \sinh s + \ell^2(\cosh s - 1) \\ &= \begin{pmatrix} \cosh s & -\sinh s & 0 & 0 \\ -\sinh s & \cosh s & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \end{aligned}$$

This transformation is called a boost in the x direction of rapidity s.

Let us see how the boost of rapidity s changes our coordinates. Recall that our transformation relates the old basis to the new basis by

$$\vec{e}_\mu = \vec{e}_\nu{}' \Lambda^\nu{}_\mu,$$

Suppose we have a point $\vec{X} \in \mathbb{R}^4$ with coordinates x^μ relative to the old basis and x'^μ relative to the new basis. We have:

$$\vec{X} = x^\mu \vec{e}_\mu = x^\mu \Lambda^\nu{}_\mu \vec{e}_\nu{}' = x'^\nu \vec{e}_\nu{}',$$

so that

$$x'^\nu = \Lambda^\nu{}_\mu x^\mu.$$

Thus we have:

$$\begin{aligned} x'^0 &= x^0 \cosh s - x^1 \sinh s, \\ x'^1 &= -x^0 \sinh s + x^1 \cosh s, \\ x'^2 &= x^2, \\ x'^3 &= x^2. \end{aligned}$$

We can take this to mean that the frame defined by $\vec{e}_\mu{}'$ is moving at a speed $v = \tanh s$ along the positive x-direction, relative to the frame defined by $\vec{e}_\mu$. Notice that $|v| < 1$. Thus two inertial frames related by a boost of this kind cannot have a relative speed faster than the speed of light.

Classifying vectors

The metric η allows us to define an inner product for vectors in Minkowski space. We define

$$\eta\left(\vec{X}, \vec{Y}\right) = x^\mu y^\mu \eta_{\mu\nu}.$$

By construction, this is invariant under a Lorentz transformation to a new frame.

Exercise 1.3. Show that if $\Lambda^\nu{}_\mu$ is the matrix of a Lorentz transformation and

$$x'^\nu = \Lambda^\nu{}_\mu x^\mu, \qquad y'^\nu = \Lambda^\nu{}_\mu y^\mu.$$

then

$$x'^\mu y'^\mu \eta_{\mu\nu} = x^\mu y^\mu \eta_{\mu\nu}.$$

In particular, for any vector $\vec{X}$ in Minkowski space, we have an invariant quantity $\eta\left(\vec{X}, \vec{X}\right)$ which is the same no matter which inertial frame we calculate it in. Unlike a traditional inner product, this quantity does not need to be positive. We classify non-zero vectors according to the sign.

Definition 3. A non-zero vector $\vec{X} \in \mathbb{E}^{1,3}$ is

i) *Timelike* if $\eta\left(\vec{X}, \vec{X}\right) < 0$,

ii) *Null*, or *lightlike* if $\eta\left(\vec{X}, \vec{X}\right) = 0$,

iii) *Spacelike* if $\eta\left(\vec{X}, \vec{X}\right) > 0$.

A vector is *causal* if it is timelike or null.

We say that a C^1–curve $\gamma : (a, b) \to \mathbb{E}^{1,3}$ is timelike/null/spacelike/causal if its tangent vector $\dot{\gamma}(s)$ is timelike/null/spacelike/causal for *every* $s \in (a, b)$.

You should think of a future directed timelike vector as representing the instantaneous velocity of a particle travelling at less than the speed of light, while a future directed null vector represents the instantaneous velocity of a particle travelling *at* the speed of light. A timelike curve should be thought of as an allowable trajectory for a massive particle.

Since a Lorentz transformation does not change the quantity $\eta\left(\vec{X}, \vec{X}\right)$, we see that if a vector is timelike/null/spacelike with respect to one frame it is timelike/null/spacelike with respect to *every* frame. Similarly, the condition of being past/future directed is independent of frame. In Figure 1.1 we show the surfaces $\eta\left(\vec{X}, \vec{X}\right) = \pm 1, 0$, which consist of vectors of constant Minkowski norm.

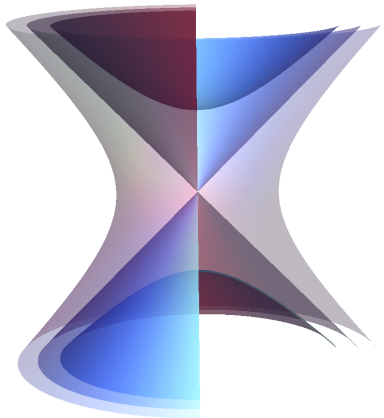

Figure 1.1 The surfaces $\eta\left(\vec{X},\vec{X}\right)=1$ (one sheeted hyperboloid) $\eta\left(\vec{X},\vec{X}\right)=-1$ (two sheeted hyperboloid) and $\eta\left(\vec{X},\vec{X}\right)=0$ (cone)

Exercise 1.4. Using $\mathbb{P},\mathbb{T}$ and the transformations of Examples 1, 2 or otherwise:

a) Suppose that $\vec{X}\in\mathbb{E}^{1,3}$ is a unit timelike vector, i.e. $\eta\left(\vec{X},\vec{X}\right)=-1$. Show that there exists an inertial frame $\{\vec{e}_\mu\}_{\mu=0,\ldots,3}$, such that writing $\vec{X}=x^\mu\vec{e}_\mu$, we have

$$x^\mu=(1,0,0,0).$$

Deduce that if $\vec{X}$ is timelike and $\vec{Y}\neq 0$ satisfies $\eta\left(\vec{X},\vec{Y}\right)=0$, then $\vec{Y}$ is spacelike.

b) Suppose that $\vec{X}\in\mathbb{E}^{1,3}$ is a null vector, i.e. $\eta\left(\vec{X},\vec{X}\right)=0$. Show that there exists an inertial frame $\{\vec{e}_\mu\}_{\mu=0,\ldots,3}$ such that writing $\vec{X}=x^\mu\vec{e}_\mu$, we have

$$x^\mu=\lambda(1,1,0,0).$$

for some $\lambda>0$. Deduce that if $\vec{X}$ is null and $\vec{Y}\neq 0$ satisfies $\eta\left(\vec{X},\vec{Y}\right)=0$, then either $\vec{Y}$ is either spacelike or parallel to $\vec{X}$.

c) Suppose that $X\in\mathbb{E}^{1,3}$ is a unit spacelike vector, i.e. $\eta\left(\vec{X},\vec{X}\right)=1$. Show that there exists an inertial frame $\{\vec{e}_\mu\}_{\mu=0,\ldots,3}$ compatible with the time orientation such that writing $\vec{X}=x^\mu\boldsymbol{e}_\mu$, we have

$$x^\mu=(0,1,0,0).$$

Deduce that if $\vec{X}$ is spacelike and $\vec{Y}\neq 0$ satisfies $\eta\left(\vec{X},\vec{Y}\right)=0$, then $\vec{Y}$ can be timelike, null or spacelike.

It is also useful to classify some surfaces[7] according to their causal properties. We define

[7] for us a surface has three dimensions

Definition 4. We say that a surface $\Sigma \subset \mathbb{E}^{1,3}$ is

i.) *Timelike* if at every point in Σ there is a timelike tangent vector.

ii.) *Null* if at every point in Σ there is a null tangent vector, but no timelike tangent vector.

iii.) *Spacelike* if every vector tangent to Σ is spacelike.

We have the following useful result, which can be proven by choosing a suitable orthonormal basis.

Lemma 1.3. *i.) If Σ is timelike, then locally there exists a spacelike vector field which is normal to the tangent plane (with respect to η).*

ii.) If Σ is spacelike, then locally there exists a timelike vector field which is normal to the tangent plane (with respect to η).

1.3.2 Causal geometry

We can define a relation between timelike vectors $\vec{T}_1$, $\vec{T}_2$ by:

$$\vec{T}_1 \sim \vec{T}_2 \qquad \Longleftrightarrow \qquad \eta(\vec{T}_1, \vec{T}_2) < 0. \tag{1.7}$$

Exercise 1.5. Show that the relation $\sim$ between timelike vectors defined in (1.7) is an equivalence relation. [Hint: the final part of Exercise 1.4 *a*) may be useful]

Definition 5. A *time orientation* for the Minkowski spacetime is an equivalence class of timelike vectors $[\vec{T}]_\sim$. We say that a causal vector, X, is *future directed* if $\eta(\vec{X}, \vec{T}) < 0$ and *past directed* if $\eta(\vec{X}, \vec{T}) > 0$ for any $\vec{T} \in [\vec{T}]_\sim$. We say that an inertial frame $\{\vec{e}_\mu\}_{\mu=0,\ldots,3}$ is compatible with the time orientation $\vec{T}$ if $\vec{e}_0$ is future directed.

Suppose an inertial frame $\{\vec{e}_\mu\}_{\mu=0,\ldots,3}$ is compatible with the time orientation $\vec{T}$, and that $\{\vec{e}_\mu{}'\}_{\mu=0,\ldots,3}$ is related to $\{\vec{e}_\mu\}_{\mu=0,\ldots,3}$ by a Lorentz transformation Λ. We say that Λ preserves the time orientation if $\{\vec{e}_\mu{}'\}_{\mu=0,\ldots,3}$ is compatible with the time orientation. We define the *orthochronous Lorentz group*, $O^+(1,3)$, to be the subgroup of $O(1,3)$ consisting of all of the Lorentz transformations which preserve the time orientation.

Lemma 1.4. *We can decompose the orthochronous Lorentz group as:*

$$O^+(1,3) = SO^+(1,3) \cup \mathbb{P}(SO^+(1,3)).$$

Proof. We first show that transformations in $SO^+(1,3)$ preserve time orientation. Suppose for contradiction that $\{\vec{e}_\mu{}'\}_{\mu=0,\ldots,3}$ is not compatible with the time orientation $[\vec{T}]_\sim$, so that $\eta(\vec{e}_0{}', \vec{T}) \geq 0$ and that the bases are related by

$$\vec{e}_\mu = \vec{e}_\nu{}' \Lambda^\nu{}_\mu,$$

with $\Lambda \in SO^+(1,3)$. This implies that there exists a continuous map $\Gamma : [0,1] \to O(1,3)$ with $\Gamma(0) = I, \Gamma(1) = \Lambda$. Consider the map $f : [0,1] \to \mathbb{R}$ given by

$$f(s) = \eta\left(\vec{e}_\nu{}' \Gamma^\nu{}_0(s), \vec{T}\right).$$

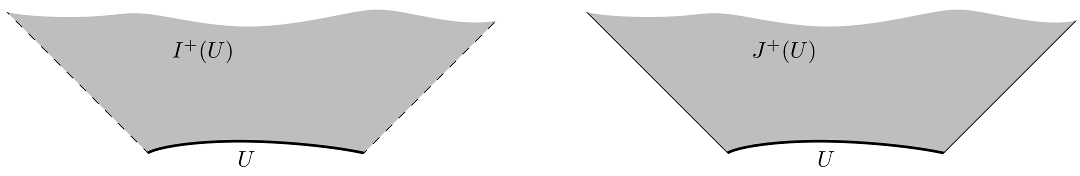

Figure 1.2 The chronological (left) and causal (right) future of a set U

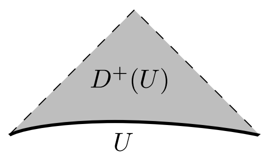

Figure 1.3 The future Cauchy development of a set U

We have that f is continuous, $f(0) < 0$ and $f(1) \geq 0$, so there exists $s_0 \in [0, 1]$ such that $f(s_0) = 0$. Since $\Gamma \in SO^+(1, 3)$, we know that $\vec{e}_\nu{}' \Gamma^\nu{}_0(s_0)$ is a unit timelike vector, however any non-zero vector orthogonal to the timelike vector $\vec{T}$ must be spacelike, which gives a contradiction.

Finally, we observe that $\mathbb{P}$ preserves the time orientation and $\mathbb{T}$ does not. Together with the decomposition of $O(1, 3)$, we obtain the result. □

Obviously, we can define a future directed timelike/causal curve by requiring that the tangent vector is future directed timelike/causal at all points on the curve. This allows us to talk about the future and past of a subset $U \subset \mathbb{E}^{1,3}$. We will define three important sets:

Definition 6. 1. The *chronological future* of $U \subset \mathbb{E}^{1,3}$, denoted $I^+(U)$, is the set of all points $p \in \mathbb{E}^{1,3}$ which can be reached from U by a future directed timelike curve.

2. The *causal future* of $U \subset \mathbb{E}^{1,3}$, denoted $J^+(U)$, is the set of all points $p \in \mathbb{E}^{1,3}$ which can be reached from U by a future directed causal curve.

3. The *future Cauchy development*, or *domain of dependence* of U, denoted $D^+(U)$ is the set of all points $p \in \mathbb{E}^{1,3}$ such that every past inextensible timelike curve through p intersects U.

1.4 Lorentz geometry and the wave equation

1.4.1 Doppler shifts

One simple calculation we can do with the machinery that we've built up in the last section allows us to explain the phenomenon of the Doppler shift. We know this best from the field of acoustics: the engine of a speeding car sounds higher pitched as it approaches us than it does as it drives away. The Doppler shift also occurs for light waves. Light coming from an object moving towards us is 'blue-shifted', meaning the frequency of the electromagnetic radiation increases. Light coming from an object moving away is 'red-shifted'. This fact is behind the dreaded radar guns used to enforce speed limits.

To see how the geometrical picture we've built up can help us understand this phenomenon, let us fix an inertial frame $\{\vec{e}_\mu\}_{\mu=0,\ldots,3}$ and consider the following function[8]

$$\begin{array}{rccl} u_{\vec{k}} & : & \mathbb{E}^{1,3} & \to \mathbb{C} \\ & & X & \mapsto e^{i\eta(\vec{X},\vec{k})} = e^{ix^\mu k_\mu} \end{array}$$

where $\vec{k}$ is a constant vector. We have that

$$\nabla_\mu u_{\vec{k}} = ik_\mu u_{\vec{k}}$$

so that

$$\Box u_{\vec{k}} = -k_\mu k^\mu u_{\vec{k}}.$$

If we choose $\vec{k}$ to be a null vector, so that $k_\mu k^\mu = 0$, then u_k is a solution of the wave equation. These are the plane wave solutions. In this frame, writing $x^\mu = (t, \boldsymbol{x})$, we see

$$u_k = e^{-ik^0 t + i\boldsymbol{k}\cdot\boldsymbol{x}}$$

so that solution in this frame is seen to be a wave with frequency k^0 propagating in the direction $\boldsymbol{n} = \frac{\boldsymbol{k}}{k^0}$.

Now, since our solution is written in a manifestly Lorentz invariant form, we know that $u_{\vec{k}}$ is given by the same expression, regardless of which inertial frame we choose. As a result, if we instead choose an inertial frame $\{\vec{e}_\mu{}'\}_{\mu=0,\ldots,3}$, then we would view the solution as a wave with frequency k'^0 propagating in the direction $\boldsymbol{n}' = \frac{\boldsymbol{k}}{k^0}$. Suppose that the frames are related by the boost transformation of Example 2. Then we have that

$$\begin{aligned} k'^0 &= k^0 \cosh s - k^1 \sinh s, \\ k'^1 &= -k^0 \sinh s + k^1 \cosh s, \\ k'^2 &= k^2, \\ k'^3 &= k^2. \end{aligned}$$

Thus we see that viewed in the two different frames, the frequency and the wavenumber in the x^1 direction differ. We can simplify the situation by assuming that $k^2 = k^3 = 0$ and $k^1 = k^0$, so that our boost is in the direction of propagation. Then we have $k'^2 = k'^3 = 0$ and

$$k'^0 = k'^1 = k^0 e^{-s}.$$

[8] If you prefer to work with real solutions of the wave equation, we are always free to take the real part.

Recalling that the rapidity s corresponds to a relative velocity between the two frames of $v = \tanh s$, we deduce that the frequency in the frame $\{\vec{e}_\mu{}'\}_{\mu=0,\ldots,3}$ is given in terms of the frequency in the frame $\{\vec{e}_\mu\}_{\mu=0,\ldots,3}$ by

$$k'^0 = k^0\sqrt{\frac{1-v}{1+v}}$$

Thus, if $v > 0$, an observer at rest in the primed frame will measure a lower frequency relative to an observer at rest in the unprimed frame. This is because the observer in the primed frame is moving in the positive x^0-direction relative to an observer in the unprimed frame, so sees a red shifted wave.

1.4.2 The energy-momentum tensor

Our motivation for introducing the Lorentz group was to understand symmetries of the wave equation. We found that the wave equation looks the same in any inertial frame. A very deep principle of theoretical physics says that symmetries give rise to conserved quantities. In the case of the wave equation, this fact is demonstrated by the existence of a tensor, the energy-momentum tensor, which has some very useful properties.

Definition 7. Suppose $U \subset \mathbb{E}^{1,3}$ is open. Given $\psi \in C^2(U)$, we define a symmetric $(0,2)-$tensor, the *energy momentum tensor* with components given by

$$T_{\mu\nu}[\psi] := \nabla_\mu\psi\nabla_\nu\psi - \frac{1}{2}\eta_{\mu\nu}\nabla_\sigma\psi\nabla^\sigma\psi.$$

Theorem 1.6. *The energy momentum tensor has the following properties:*

1. *We have a formula for the divergence:*

$$\nabla_\mu T^\mu{}_\nu[\psi] = (\Box\psi)\,\nabla_\nu\psi$$

2. *Fix an inertial frame* $\{\vec{e}_\mu\}_{\mu=0,\ldots,3}$. *The* $00-$*component of* T *is the local energy density*

$$T_{00}[\psi] = \frac{1}{2}\left[(\nabla_0\psi)^2 + |\boldsymbol{\nabla}\psi|^2\right]$$

3. *Fix an inertial frame* $\{\vec{e}_\mu\}_{\mu=0,\ldots,3}$. *If* $\vec{V} = V^\mu\vec{e}_\mu$ *is any future directed unit timelike vector, then:*

$$V^0\left[(\nabla_0\psi)^2 + |\boldsymbol{\nabla}\psi|^2\right] \geq V^\mu T_{\mu 0}[\psi] \geq \frac{1}{4V^0}\left[(\nabla_0\psi)^2 + |\boldsymbol{\nabla}\psi|^2\right]$$

Proof. 1. We calculate

$$\begin{aligned}
\nabla_\mu T^\mu{}_\nu[\psi] &= \nabla_\mu\left(\nabla^\mu\psi\nabla_\nu\psi - \frac{1}{2}\delta^\mu{}_\nu\nabla_\sigma\psi\nabla^\sigma\psi\right)\\
&= (\nabla_\mu\nabla^\mu\psi)\,\nabla_\nu\psi + \nabla^\mu\psi\nabla_\mu\nabla_\nu\psi - \frac{1}{2}\nabla_\nu\left(\nabla_\mu\psi\nabla^\mu\psi\right)\\
&= (\nabla_\mu\nabla^\mu\psi)\,\nabla_\nu\psi + \nabla^\mu\psi\nabla_\mu\nabla_\nu\psi - (\nabla_\nu\nabla_\mu\psi)\,\nabla^\mu\psi\\
&= (\Box\psi)\,\nabla_\nu\psi.
\end{aligned}$$

2. Again, a straightforward calculation gives us:

$$\begin{aligned} T_{00}[\psi] &= \nabla_0\psi\nabla_0\psi + \frac{1}{2}\left(-\nabla_0\psi\nabla_0\psi + \nabla_i\psi\nabla^i\psi\right) \\ &= \frac{1}{2}\left[(\nabla_0\psi)^2 + |\boldsymbol{\nabla}\psi|^2\right] \end{aligned}$$

3. Since $\vec{V}$ is timelike, future directed, and unit, we have $V^0 > 0$ and

$$(V^0)^2 - |\boldsymbol{V}|^2 = 1.$$

We calculate, similarly to the previous part:

$$V^\mu T_{\mu 0}[\psi] = \nabla_0\psi(V^0\nabla_0\psi + V^i\nabla_i\psi) + \frac{V^0}{2}\left(-\nabla_0\psi\nabla_0\psi + \nabla_i\psi\nabla^i\psi\right).$$

Now, we use the Cauchy-Schwarz inequality and Young's inequality (or the AM-GM inequality) to deduce

$$\left|\nabla_0\psi V^i\nabla_i\psi\right| \leq |\nabla_0\psi|\,|\boldsymbol{V}|\,|\boldsymbol{\nabla}\psi| \leq \frac{|\boldsymbol{V}|}{2}\left[(\nabla_0\psi)^2 + |\boldsymbol{\nabla}\psi|^2\right]$$

We therefore have

$$\frac{V^0 + |\boldsymbol{V}|}{2}\left((\nabla_0\psi)^2 + |\boldsymbol{\nabla}\psi|^2\right) \geq V^\mu T_{\mu 0}[\psi] \geq \frac{V^0 - |\boldsymbol{V}|}{2}\left((\nabla_0\psi)^2 + |\boldsymbol{\nabla}\psi|^2\right)$$

Now, since $(V^0)^2 - |\boldsymbol{V}|^2 = 1$, we have $V^0 - |\boldsymbol{V}| = (V^0 + |\boldsymbol{V}|)^{-1}$ and $2V^0 > |\boldsymbol{V}| + V^0$, which completes the proof.

□

Remark: Property 2. is sometimes known as the *weak energy condition* and property 3. as the *dominant energy condition.*

A useful Corollary is the following

Corollary 1.7. *Suppose that $\psi \in C^2(U)$ solves the wave equation in U, and let $\vec{V}$ be a C^1 vector field*[9] *on U. Defining the vector field:*

$${}^{\vec{V}}\vec{J}[\psi] = V^\nu T_\nu{}^\mu[\psi]\vec{e}_\mu$$

we have

$$\nabla_\mu\left({}^{\vec{V}}J^\mu[\psi]\right) = {}^{V}\Pi_{\mu\nu}T^{\mu\nu}[\psi]$$

where

$${}^{\vec{V}}\Pi_{\mu\nu} = \nabla_{(\mu}V_{\nu)}$$

is the deformation tensor *of $\vec{V}$. ${}^{\vec{V}}\vec{J}$ is called the* energy current *associated to the vector field $\vec{V}$.*

[9] This means that with respect to a basis for $\mathbb{E}^{1,3}$ we may write $\vec{V} = V^\mu \boldsymbol{e}_\mu$, where $V^\mu \in C^1(U)$.

Proof. We can calculate, using part 1. of Theorem 1.6:

$$\begin{aligned}\nabla_\mu \left({}^{\vec{V}}J^\mu\right) &= \nabla_\mu \left(V^\nu T_\nu{}^\mu\right) = (\nabla_\mu V^\nu) T_\nu{}^\mu + V^\nu \nabla_\mu \left(T_\nu{}^\mu\right) \\ &= \nabla_\mu V_\nu T^{\mu\nu} + V^\nu (\Box\psi) \nabla_\nu \psi \\ &= \nabla_{(\mu} V_{\nu)} T^{\mu\nu} = {}^{\vec{V}}\Pi_{\mu\nu} T^{\mu\nu}.\end{aligned}$$

□

This corollary is particularly useful when the deformation tensor vanishes for $\vec{V}$. These vector fields are especially important:

Definition 8. A *Killing vector* is a vector field such that

$${}^{\vec{V}}\Pi = 0.$$

Killing vectors are closely related to symmetries of the spacetime. To explore a bit more fully the Killing vectors, let us pick an orthonormal basis $\{\vec{e}_\mu\}$ for $\mathbb{E}^{1,3}$ and use it to identify each point $X \in \mathbb{E}^{1,3}$ with its components x^μ with respect to this basis. First, let us note that the constant vector field

$$\vec{P}_0 = \vec{e}_0,$$

is a Killing field: in fact we have that $\nabla_\mu V_\nu = 0$ for all μ, ν. Similarly for the constant vector fields

$$\vec{P}_i = \vec{e}_i.$$

Now, suppose we have a matrix $\ell \in \mathfrak{o}(1,3)$, which can be written in components as $\ell^\mu{}_\nu$. Then we can define a vector field on $\mathbb{E}^{1,3}$ by:

$$\vec{V}_\ell = x^\nu \ell^\mu{}_\nu \vec{e}_\mu.$$

I claim this is a Killing field. To see this, first note that $(V_\ell)^\mu = x^\nu \ell^\mu_\nu$, so that

$$\nabla_\mu (V_\ell)_\nu = \eta_{\nu\sigma} \ell^\sigma{}_\mu$$

hence

$$\begin{aligned}{}^{\vec{V}_\ell}\Pi_{\mu\nu} &= \frac{1}{2}\left(\eta_{\nu\sigma}\ell^\sigma{}_\mu + \eta_{\mu\sigma}\ell^\sigma{}_\nu\right) \\ &= \left[\eta\ell + (\eta\ell)^T\right]_{\nu\mu} \\ &= \left[\eta\left(\ell\eta^{-1} + \eta^{-1}\ell^T\right)\eta\right]_{\nu\mu} = 0.\end{aligned}$$

In fact, one can show that any Killing field of the Minkowski spacetime is a linear combination of fields of the form $\vec{P}_\mu$ and $\vec{V}_\ell$. All of these vector fields are useful, but for us, the most useful will be the vector field $\vec{P}_0$, which is everywhere timelike. It owes its existence to the time translation symmetry of the Minkowski spacetime. We will make use of this vector field to prove the final result of this section:

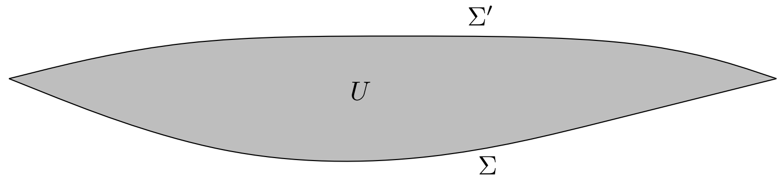

Figure 1.4 The geometry of Theorem 1.8

Theorem 1.8 (Finite speed of propagation). *Suppose that $U \subset \mathbb{E}^{1,3}$ is a open, bounded, set such that the boundary ∂U consists of two smooth compact components Σ, Σ' which are both spacelike, and such that Σ' lies in $J^+(\Sigma)$ (see Figure 1.4). Then if $\psi \in C^2(\overline{U})$ solves the wave equation in U we have the estimate*

$$\int_U \left((\nabla_0 \psi)^2 + |\boldsymbol{\nabla}\psi|^2 \right) dX \leq C \int_\Sigma \left((\nabla_0 \psi)^2 + |\boldsymbol{\nabla}\psi|^2 \right) d\sigma$$

for a constant C which depends only on the domain U. In particular if $\psi, \partial_t \psi$ vanish on Σ, then ψ vanishes throughout U.

Proof. Pick an inertial frame $\{\vec{e}_\mu\}$ for the wave equation, and choose coordinates $\vec{X} = x^\mu \vec{e}_\mu$. Consider the vector field

$$\vec{V} = e^{-x^0} \vec{e}_0$$

We calculate

$$\nabla_\mu V_\nu = \begin{cases} e^{-x^0} & \mu = \nu = 0, \\ 0 & \text{otherwise.} \end{cases}$$

We want to consider the current ${}^{\vec{V}}\vec{J}$. We have

$$\nabla_\mu \left({}^{\vec{V}} J^\mu[\psi] \right) = e^{-x^0} T_{00}$$

We will apply Corollary B.2 (the divergence theorem) to ${}^{\vec{V}}\vec{J}$. We find

$$\int_U \nabla_\mu \left({}^{\vec{V}} \vec{J}^\mu[\psi] \right) dx = \int_\Sigma {}^{\vec{V}} J^\mu[\psi] t_\mu d\sigma - \int_{\Sigma'} {}^{\vec{V}} J^\mu[\psi] t_\mu d\sigma$$

where $\vec{t}$ is the future directed unit normal[10], and $d\sigma$ is the (positive) surface measure. Inserting our expressions for $\nabla_\mu \left({}^{\vec{V}} J^\mu[\psi] \right)$ and ${}^{\vec{V}}\vec{J}$ we have:

$$\int_\Sigma e^{-x^0} T_{0\mu} t^\mu d\sigma = \int_{\Sigma'} e^{-x^0} T_{0\mu} t^\mu d\sigma + \int_U e^{-x^0} T_{00} dx \tag{1.8}$$

Now, since Σ, Σ' are smooth and compact and U is bounded, we have that there exists a constant $c > 1$ such that

$$0 < c^{-1} \leq \sup_{\Sigma \cup \Sigma'} t^0, \ \sup_U |x^0| \leq c < \infty.$$

[10] notice that the signs here are different from what one might expect.

As a result, we have that

$$\int_{\Sigma} e^{-x^0} T_{0\mu} t^\mu d\sigma \leq c e^c \int_{\Sigma} \left((\nabla_0 \psi)^2 + |\boldsymbol{\nabla} \psi|^2 \right) d\sigma$$

$$\int_{\Sigma'} e^{-x^0} T_{0\mu} t^\mu d\sigma \geq \frac{e^{-c}}{4c} \int_{\Sigma'} \left((\nabla_0 \psi)^2 + |\boldsymbol{\nabla} \psi|^2 \right) d\sigma \geq 0$$

and

$$\int_U e^{-x^0} T_{00} dX \geq e^{-c} \int_U \left((\nabla_0 \psi)^2 + |\boldsymbol{\nabla} \psi|^2 \right) dx dt$$

Putting these estimates together with (1.8), we have the result. □

Corollary 1.9. *Suppose that $u \in C^\infty(\mathbb{E}^{1,3})$ solves the wave equation, and that*

$$(supp \ \ u|_{x^0=0}) \bigcup (supp \ \ \nabla_0 u|_{x^0=0}) \subset B_R(0)$$

Then

$$(supp \ \ u|_{x^0=T}) \bigcup (supp \ \ \nabla_0 u|_{x^0=T}) \subset B_{R+|T|}(0).$$

This justifies some of the assumptions made in §1.2, where we considered solutions of the wave equation vanishing for large $|\boldsymbol{x}|$ at any fixed time x^0.

Exercise 2.1. Let $U \subset \mathbb{E}^{1,3}$ be open. Define an antisymmetric $(0,2)-$tensor field, F on U with components[11]

$$[F_{\mu\nu}] = \begin{pmatrix} 0 & -E_1 & -E_2 & -E_3 \\ E_1 & 0 & B_3 & -B_2 \\ E_2 & -B_3 & 0 & B_1 \\ E_3 & B_2 & -B_1 & 0 \end{pmatrix}.$$

Where $E_i, B_i \in C^1(U)$. Show that the vacuum Maxwell equations for $\boldsymbol{E}, \boldsymbol{B}$ (with units such that $c^2 = \epsilon_0 \mu_0 = 1$) hold in U if and only if F satisfies the equations

$$\nabla_\mu F^\mu{}_\nu = 0, \qquad \nabla_{[\mu} F_{\nu\sigma]} = 0,$$

where for any $(0,3)-$tensor $A_{\mu\nu\sigma}$ we define:

$$A_{[\mu\nu\sigma]} = \frac{1}{6} \left(A_{\mu\nu\sigma} + A_{\nu\sigma\mu} + A_{\sigma\mu\nu} - A_{\nu\mu\sigma} - A_{\mu\sigma\nu} - A_{\sigma\nu\mu} \right).$$

Exercise 2.2. Suppose that F is as in Exercise 2.1. Fix an inertial frame $\{\vec{e}_\mu\}$. Define

$$T_{\mu\nu}[F] = F_{\mu\sigma} F_\nu{}^\sigma - \frac{1}{4} \eta_{\mu\nu} F_{\sigma\tau} F^{\sigma\tau}.$$

Show that T has the following properties

---

[11] The convention is that the first index specifies the row, and the second index the column.

a) We have a formula for the divergence:

$$\nabla_\mu T^\mu{}_\nu[F] = \left(\nabla_\mu F^\mu{}_\sigma\right) F_\nu{}^\sigma + \frac{3}{2}\left(\nabla_{[\mu} F_{\nu\sigma]}\right) F^{\mu\sigma}$$

b) The 00−component of T is the local energy density

$$T_{00}[F] = \frac{1}{2}\left[|\boldsymbol{E}|^2 + |\boldsymbol{B}|^2\right]$$

c) If $\vec{V}$ is any future directed unit timelike vector, then:

$$V^0\left[|\boldsymbol{E}|^2 + |\boldsymbol{B}|^2\right] \geq V^\mu T_{\mu 0}[F] \geq \frac{1}{4V^0}\left[|\boldsymbol{E}|^2 + |\boldsymbol{B}|^2\right]$$

Hence, or otherwise, deduce that the electromagnetic field exhibits finite speed of propagation.