2

The Laplace Equation and Wave Equation

Introduction

In this chapter we introduce the central linear partial differential equations of the second order, the Laplace equation

$$\Delta u = f \tag{0.1}$$

and the wave equation

$$\left(\frac{\partial^2}{\partial t^2} - \Delta\right) u = f. \tag{0.2}$$

For flat Euclidean space $\mathbb{R}^n$, the Laplace operator is defined by

$$\Delta u = \frac{\partial^2 u}{\partial x_1^2} + \cdots + \frac{\partial^2 u}{\partial x_n^2}. \tag{0.3}$$

The wave equation arose early in the history of continuum mechanics, in a mathematical description of the motion of vibrating strings and membranes. We discuss this in §1. The analysis, based on an appropriate version of Hamilton's stationary action principle, generally produces nonlinear partial differential equations, of a sort that will be studied more in Chaps. 14–16. The wave equation described by (0.2), which is linear, arises as a "linearized" PDE, describing such vibratory motion, as will be seen in §1.

In this chapter we consider the Laplace operator on a general Riemannian manifold and emphasize concepts defined in a coordinate-independent fashion. Also, more generally than the wave equation (0.2) on the Cartesian product of a spatial region with the time axis, we consider natural generalizations defined on a manifold endowed with a Lorentz metric.

Before defining the Laplace operator on Riemannian manifolds, we devote two sections to some first-order operators. In §2 we discuss the divergence operator applied to vector fields, and in §3 we generalize the operations of covariant derivative

M.E. Taylor, *Partial Differential Equations I: Basic Theory*,
Applied Mathematical Sciences 115, DOI 10.1007/978-1-4419-7055-8_2,

and divergence from vector fields to tensor fields. These concepts play important roles in the study of the Laplace and wave equations.

In §4 we define the Laplace operator acting on real- (or complex-) valued functions on a Riemannian manifold M, and in §5 we write down the wave equation for functions on $\mathbb{R} \times M$ and discuss energy conservation. In §6 we extend energy identities in a way that leads to proofs of results on finite propagation speed for solutions to such a wave equation.

In §7 we extend the notion of the wave equation from $\mathbb{R} \times M$ to a general Lorentz manifold. We extend the notion of energy conservation. To a solution of the wave equation is associated a second-order tensor field, the "stress-energy tensor," and the law of conservation of energy can be expressed as the vanishing of the divergence of this field, as is shown in §7. One can pass from such a "local" conservation law to an integral conservation law via the divergence theorem, for a certain class of Lorentz manifolds, namely those with a timelike Killing field. We derive the phenomenon of "finite propagaton speed" for solutions to the wave equation as a consequence of such a conservation law.

In §8 we consider a more general class of hyperbolic equations. To solutions we can still associate a tensor with some of the properties of a stress-energy tensor, but the energy conservation law may not hold, and instead we look for "energy estimates."

The Stokes formula used in §2 to derive the divergence theorem is a special case of a more general Stokes-type formula, which we discuss in §9. This more general formula is used in §10 to produce a variant of Green's formula for the Laplace operator acting on differential forms. In these sections we also make use of the notion of the "principal symbol" of a differential operator, as an invariantly defined function on the cotangent bundle.

In §11 we look at Maxwell's equations for the electromagnetic field. We show how they can be manipulated to yield the wave equation. This mathematical fact will be further exploited in Chap. 6. We deal with Maxwell's equations in the framework of relativity and work with the electromagnetic field on a general Lorentz 4-manifold.

Though we discuss some qualitative properties of solutions to the Laplace equation and the wave equation, such as Green's identities and finite propagation speed (in the case of the wave equation), we do not tackle the question of existence of solutions in this chapter, except for the very simplest case, namely the $n = 1$ case of (0.2), treated in §1. In the case of such equations on flat Euclidean space, Fourier analysis provides an adequate tool to construct and analyze solutions, and this will be developed in the next chapter. Then functional analytical methods, centered on the theory of Sobolev spaces, will be developed in Chap. 4 and applied in subsequent chapters. As we will see in Chap. 6, energy estimates, such as those derived in §8 of this chapter, in concert with Sobolev space theory, form the principal tools for existence theorems for linear hyperbolic equations. Existence of solutions to nonlinear hyperbolic equations, which requires somewhat more subtle analysis, will be studied in Chap. 16.

1. Vibrating strings and membranes

The problem of describing the motion of a vibrating string was one of the earliest problems of continuum mechanics, producing a partial differential equation. Such a PDE can be derived by a procedure similar to that described in §12 of Chap. 1, using a stationary action principle. To carry this out, we need formulas for the kinetic energy and the potential energy of a vibrating string.

Suppose our string is vibrating in $\mathbb{R}^k$; say its ends are tied down at two points, the origin 0 and a vector $Le_1 \in \mathbb{R}^k$, of length L. We suppose the string is uniform, of mass density m (i.e., total mass mL). The motion of the string is described by a function $u = u(t, x),\ t \in \mathbb{R},\ x \in [0, L]$, taking values in $\mathbb{R}^k$ and satisfying $u(t, 0) = 0,\ u(t, L) = Le_1$ for all t. Then the kinetic energy at time t is given by

$$T(t) = \frac{m}{2} \int_0^L |u_t(t, x)|^2 \, dx, \tag{1.1}$$

and the integral $\int_{t_0}^{t_1} T(t)\, dt$ is given by

$$J_0(u) = \frac{m}{2} \iint\limits_{I \times \Omega} |u_t(t, x)|^2 \, dx \, dt, \tag{1.2}$$

where $I = (t_0, t_1),\ \Omega = (0, L)$.

As for the potential energy at a given time t, we will use the law that the potential energy in a small piece of string is a function of the degree that the string has been stretched, namely,

$$V(t) = \int_0^L f\big(u_x(t, x)\big) \, dx \tag{1.3}$$

for a function

$$f : \mathbb{R}^k \longrightarrow \mathbb{R}. \tag{1.4}$$

This is known as Hooke's law. The case of an "ideal" string (where the force exerted by a small piece of string is proportional to the amount by which it has been stretched) is

$$f(y) = \sigma(|y| - a)^2, \tag{1.5}$$

where the unstretched string has length $aL < L$ and $\sigma > 0$ is a given constant. The term accompanying (1.2) in the expression for the action is

$$J_1(u) = \iint\limits_{I \times \Omega} f\big(u_x(t, x)\big) \, dx \, dt. \tag{1.6}$$

The stationary condition according to Hamilton's principle is

$$\frac{d}{ds}(J_0 - J_1)(u+sv)\big|_{s=0} = 0, \tag{1.7}$$

for all $v \in C_0^\infty(I \times \Omega, \mathbb{R}^k)$. A simple computation gives

$$\begin{aligned} \frac{d}{ds} J_0(u+sv)\big|_{s=0} &= \iint\limits_{I\times\Omega} m u_t v_t \, dx \, dt \\ &= -\iint m v u_{tt} \, dx \, dt, \end{aligned} \tag{1.8}$$

where the last identity is obtained by integration by parts. Furthermore, also integrating by parts, we have

$$\begin{aligned} \frac{d}{ds} J_1(u+sv)\big|_{s=0} &= \iint f'\big(u_x(t,x)\big)\cdot v_x(t,x)\, dx\, dt \\ &= -\iint \Big\{\frac{\partial}{\partial x} f'\big(u_x(t,x)\big)\Big\} \cdot v(t,x)\, dx\, dt. \end{aligned} \tag{1.9}$$

Note that

$$\frac{\partial}{\partial x} f'(u_x(t,x)) = f''(u_x)u_{xx}, \tag{1.10}$$

where $f''(y)$ is the $k \times k$ matrix valued function of second-order partial derivatives of $f : \mathbb{R}^k \to \mathbb{R}$, and u_{xx} takes values in $\mathbb{R}^k$. In other words,

$$\frac{d}{ds} J_1(u+sv)\big|_{s=0} = -\iint\limits_{I\times\Omega} f''(u_x)u_{xx}\cdot v\, dx\, dt. \tag{1.11}$$

Combining (1.8) and (1.11), we see that the stationary condition (1.7) is equivalent to the partial differential equation

$$m u_{tt} - f''(u_x)u_{xx} = 0. \tag{1.12}$$

If $f(y)$ is a second-order polynomial in y, that is, of the form

$$f(y) = a + b\cdot y + Ay\cdot y, \tag{1.13}$$

where $a \in \mathbb{R}, b \in \mathbb{R}^k$, and A is a real, symmetric, $k\times k$ matrix, then $f''(y) = 2A$, and the PDE (1.12) becomes

$$m u_{tt} - 2A u_{xx} = 0. \tag{1.14}$$

The example (1.5) does not satisfy this condition, and the resulting PDE is not linear. Let us rewrite this PDE, setting

$$u(t,x) = xe_1 + w(t,x), \tag{1.15}$$

so that $w(t,0) = 0$ and $w(t,L) = 0$ in $\mathbb{R}^k$. Then

$$J_1(u) = K_1(w) = \iint \varphi(w_x)\, dx\, dt, \tag{1.16}$$

where $\varphi : \mathbb{R}^k \to \mathbb{R}$ is given by

$$\varphi(y) = f(e_1 + y), \tag{1.17}$$

and the corresponding PDE for w is

$$mw_{tt} - \varphi''(w_x)w_{xx} = 0. \tag{1.18}$$

The *linearization* of this equation is, by definition, obtained by replacing $\varphi(y)$ by its quadratic part, that is, by the terms of order ≤ 2 in its power series about $y = 0$:

$$\varphi_0(y) = a_0 + b_0 \cdot y + \frac{1}{2} A_0 y \cdot y, \tag{1.19}$$

where $a_0 = \varphi(0) = f(e_1), b_0 = \varphi'(0) = f'(e_1)$, and $A_0 = \varphi''(0) = f''(e_1)$. For one reason why the term "linearization" is appropriate, see Exercise 4 at the end of this section. If φ is replaced by φ_0 in (1.16), the stationary condition yields the linear PDE

$$mw_{tt} - A_0 w_{xx} = 0 \quad \big(A_0 = \varphi''(0)\big). \tag{1.20}$$

In the case of an ideal string (1.5), this linearized PDE is readily computed to be

$$mw_{tt} - 2\sigma(I - aP)w_{xx} = 0, \tag{1.21}$$

where P is the orthogonal projection of $\mathbb{R}^k$ onto the orthogonal complement of e_1. (Compare the calculations (1.43)–(1.47) and (1.51)–(1.55) below.) Recall that we are assuming $0 < a < 1$.

For this linear equation, we can write $w = w^b + w^\#$, where w^b is parallel to e_1 and $w^\#$ is orthogonal to e_1. The equation (1.21) decouples, and we have

$$mw^b_{tt} - 2\sigma w^b_{xx} = 0 \tag{1.22}$$

as the equation for the *longitudinal* wave w^b and

$$mw^{\#}_{tt} - 2\sigma(1-a)w^{\#}_{xx} = 0, \tag{1.23}$$

as the equation for the *transverse* wave $w^{\#}$. Both of these equations are cases (with different values of c) of the wave equation

$$v_{tt} - c^2 v_{xx} = 0. \tag{1.24}$$

Here c is identified with the *propagation speed* for solutions to (1.24), for the following reason. Namely, for any C^2-functions f_j of one variable,

$$v(t,x) = f_1(x+ct) + f_2(x-ct) \tag{1.25}$$

is a solution to (1.24). Conversely, the general solution to (1.24) on $(t,x) \in \mathbb{R}\times\mathbb{R}$, satisfying the *initial conditions*

$$v(0,x) = g(x), \quad v_t(0,x) = h(x), \tag{1.26}$$

can be expressed in the form (1.25). Indeed, a solution to (1.24) in the form (1.25) satisfies these initial conditions if and only if

$$f_1(x) + f_2(x) = g(x) \ \text{ and } \ cf_1'(x) - cf_2'(x) = h(x). \tag{1.27}$$

This implies $f_1'(x) + f_2'(x) = g'(x)$, so we can solve algebraically for f_1' and f_2'; thus we can set

$$\begin{aligned} f_1(x) &= \frac{1}{2}g(x) + \frac{1}{2c}\int_0^x h(s)\,ds, \\ f_2(x) &= \frac{1}{2}g(x) - \frac{1}{2c}\int_0^x h(s)\,ds. \end{aligned} \tag{1.28}$$

That the solution (1.25) so produced is the only solution to (1.24) satisfying the initial conditions (1.26) is a special case of a uniqueness result proved in §5.

One can arrange that the *boundary condition*

$$v(t,0) = v(t,L) = 0 \tag{1.29}$$

be satisfied by taking g and h that satisfy

$$g(s) = g(s+2L) = -g(-s), \quad h(s) = h(s+2L) = -h(-s). \tag{1.30}$$

This is a special case of the *method of images*, discussed further in Chap. 3, §7.

Whenever one has the linear equation (1.14), if A is a positive-definite matrix, one can diagonalize A and construct solutions as above. Constructing solutions for the equation (1.12), or (1.18) in the nonlinear case, is much more difficult; Chap. 16 gives some results for this problem.

Now we look at the higher-dimensional case, of a vibrating membrane. Let Ω be some open region in $\mathbb{R}^n$. We consider vibrations of Ω in $\mathbb{R}^k$, with $k \geq n$. Define the inclusion $j : \mathbb{R}^n \hookrightarrow \mathbb{R}^k$ by

$$j(x_1, \dots, x_n) = (x_1, \dots, x_n, 0, \dots, 0).$$

This time suppose the boundary of Ω is tied down. The motion of the membrane is described by a function $u = u(t, x), t \in \mathbb{R}, x \in \overline{\Omega}$, taking values in $\mathbb{R}^k$ and satisfying $u(t, x) = j(x)$ for $x \in \partial\Omega$. We suppose the membrane is of a uniform substance, with mass density m. The kinetic energy at a given time t is then

$$T(t) = \frac{m}{2} \int_\Omega |u_t(t, x)|^2 \, dx, \tag{1.31}$$

parallel to (1.1), and the integral $\int_{t_0}^{t_1} T(t)\, dt = J_0(u)$ is again given by (1.2), with Ω now an n-dimensional domain. As for the potential energy, we will again work under the hypothesis that it is a function of the "stretching" of the membrane, of the form

$$V(t) = \int_\Omega f\big(u_x(t, x)\big)\, dx, \tag{1.32}$$

where, for each $(t, x) \in \mathbb{R} \times \Omega$,

$$u_x(t, x) \in \mathcal{L}(T_x\Omega, \mathbb{R}^k) \approx \mathcal{L}(\mathbb{R}^n, \mathbb{R}^k) \tag{1.33}$$

is the x-derivative, and

$$f : \mathcal{L}(\mathbb{R}^n, \mathbb{R}^k) \longrightarrow \mathbb{R} \tag{1.34}$$

is a given smooth function. Again $\int_{t_0}^{t_1} V(t)\, dt = J_1(u)$ is given by (1.6), the stationary action principle takes the form (1.7), and the variation of $J_0(u)$ is given by (1.8). The variation of $J_1(u)$ is also given by a formula of the form (1.11). More precisely, if we set

$$f = f(y), \quad y = \big(y_{\nu j}\big) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^k), \tag{1.35}$$

then (1.11) holds, with the interpretation

$$f''(u_x)u_{xx} \cdot v = \sum_{\mu,\nu=1}^{k} \sum_{i,j=1}^{n} \frac{\partial^2 f(u_x)}{\partial y_{\mu i}\, \partial y_{\nu j}}\, u^\mu_{x_i x_j} v^\nu, \tag{1.36}$$

where $u = (u^1, \dots, u^k)$, $v = (v^1, \dots, v^k) \in \mathbb{R}^k$. With this notation, the PDE obtained for u is again of the form (1.12).

As in (1.15)–(1.17), we can concentrate on the deviation of u from the map $j : \Omega \to \mathbb{R}^k$. Set

$$u(t, x) = j(x) + w(t, x), \tag{1.37}$$

so the boundary condition becomes $w(t, x) = 0$ for $x \in \partial\Omega$; then the PDE for w is of the form (1.18), again interpreted as in (1.36), with

$$\varphi(y) = f(j + y), \tag{1.38}$$

for $y \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^k)$. As before, we have the linearized PDE

$$m w_{tt} - A w_{xx} = 0, \quad A = \varphi''(0), \tag{1.39}$$

where, for $w = (w^1, \dots, w^k)$,

$$\big(A w_{xx}\big)^\nu = \sum_{\mu=1}^k \sum_{i,j=1}^n \frac{\partial^2 \varphi(0)}{\partial y_{\mu i} \partial y_{\nu j}} \, w^\mu_{x_i x_j}. \tag{1.40}$$

We can regard A as defining a symmetric bilinear map

$$A : \mathcal{L}(\mathbb{R}^n, \mathbb{R}^k) \times \mathcal{L}(\mathbb{R}^n, \mathbb{R}^k) \longrightarrow \mathbb{R}. \tag{1.41}$$

There are a number of different forms the potential energy function $f(y)$ can take, depending on the physical properties of the membrane. In a number of models, one has $f(y) = \psi(y^* y)$, a function invariant under conjugating $y^* y$ by an orthogonal $n \times n$ matrix. These models have the form

$$f(y) = \Psi\big(\mathrm{Tr}\, g_1(y^* y), \dots, \mathrm{Tr}\, g_K(y^* y)\big), \tag{1.42}$$

where $g_\ell : \mathbb{R} \to \mathbb{R}$ is smooth and, for a self adjoint matrix $z = y^* y$, $g_\ell(z)$ is defined by the spectral representation; $g_\ell(z) v_j = g_\ell(\lambda_j) v_j$ for v_j in the λ_j-eigenspace of z. There is no loss in generality in assuming $g_\ell(1) = 0$.

To compute the linearized PDE when $f(y)$ is given by (1.42), start with

$$\begin{aligned} g_\ell\big((j^* + y^*)(j + y)\big) &= g_\ell(I + j^* y + y^* j + y^* y) \\ &= g_\ell(1) I + g_\ell'(1)(j^* y + y^* j + y^* y) \\ &\quad + \frac{1}{2} g_\ell''(1)(j^* y + y^* j)^2 + O(\|y\|^3). \end{aligned} \tag{1.43}$$

If $(1/2)\tau = \operatorname{Tr} j^*y = \operatorname{Tr} y^*j$, $\sigma = \operatorname{Tr} y^*y$, and $\gamma = \operatorname{Tr}(j^*y + y^*j)^2$, we obtain

$$\begin{aligned}\varphi(y) = f(j+y) &= \Psi(0) + \sum \partial_\ell \Psi(0)\big[g_\ell'(1)(\tau+\sigma) + \frac{1}{2} g_\ell''(1)\gamma\big] \\ &\quad + \sum \frac{1}{2}\partial_\ell\partial_m \Psi(0) g_\ell'(1) g_m'(1)\tau^2 + O(\|y\|^3).\end{aligned} \tag{1.44}$$

Thus the purely quadratic part, which yields the linearized PDE, is

$$\begin{aligned}\varphi_0(y) &= \sum_\ell \partial_\ell \Psi(0)\big[g_\ell'(1)\operatorname{Tr} y^*y + \frac{1}{2} g_\ell''(1)\operatorname{Tr}(j^*y+y^*j)^2\big] \\ &\quad + \sum_{\ell,m} \frac{1}{2}\partial_\ell\partial_m\Psi(0) g_\ell'(1) g_m'(1)\big[\operatorname{Tr}(j^*y+y^*j)\big]^2 \\ &= A \operatorname{Tr} y^*y + B \operatorname{Tr}(j^*y+y^*j)^2 + C\big(\operatorname{Tr}(j^*y+y^*j)\big)^2.\end{aligned} \tag{1.45}$$

As in the case of the linearized equations of the vibrating string, the resulting linear PDE decouples into an equation for the components of w orthogonal to the space $\mathbb{R}^n \subset \mathbb{R}^k$ in which Ω sits and an equation for the components of w parallel to this space. For the orthogonal component $w^\#$, since $j^*w^\# = 0$ in this case, we can replace $\varphi_0(y)$ by

$$\varphi^\#(y) = A \operatorname{Tr} y^*y, \quad y \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^{k-n}). \tag{1.46}$$

In this case, we have

$$\frac{\partial^2 \varphi^\#}{\partial y_{\mu i}\,\partial y_{\nu j}} = 2A\delta_{ij}\delta_{\mu\nu}. \tag{1.47}$$

Hence the linearized equation for the orthogonal (or transverse) wave is

$$m w_{tt}^\# - 2A\Delta w^\# = 0, \tag{1.48}$$

where Δ is the Laplace operator on $\mathbb{R}^n$:

$$\Delta v(x) = \frac{\partial^2 v}{\partial x_1^2} + \cdots + \frac{\partial^2 v}{\partial x_n^2}. \tag{1.49}$$

If $A > 0$, we can rewrite (1.48) in the form

$$v_{tt} - c^2 \Delta v = 0. \tag{1.50}$$

The equation (1.50) is typically called "the wave equation." As in (1.24), c is the propagation speed for waves satisfying (1.50); we will discuss this further in §6.

The construction of solutions to (1.50), satisfying initial conditions of the form (1.26), is not as elementary for $n > 1$ as the construction for $n = 1$ given by (1.25)–(1.28). In Chap. 3, we will give a construction, valid for $\Omega = \mathbb{R}^n$, using Fourier analysis. A symmetry trick similar to (1.30) will work if Ω is a rectangular solid in $\mathbb{R}^n$, though not for general bounded regions Ω. The existence and uniqueness of solutions to the wave equation (1.50) for such more general Ω are proven in Chap. 6.

The equation for the components of w parallel to the plane $\mathbb{R}^n$ of $\Omega \subset \mathbb{R}^k$, in this case, has a somewhat different form, as we now compute. Note that this case is the same as considering the entire linearized PDE for the case $k = n$. Then j is the identity map, so the linearization is of the form (1.39)–(1.40), with $\varphi(y)$ replaced by

$$\begin{aligned}\varphi^b(y) &= A \operatorname{Tr} y^* y + B \operatorname{Tr}(y + y^*)^2 + C\big(\operatorname{Tr}(y + y^*)\big)^2 \\ &= (A + 2B)\operatorname{Tr} y^* y + 2B \operatorname{Tr} y^2 + 4C\big(\operatorname{Tr} y\big)^2,\end{aligned} \tag{1.51}$$

since $\operatorname{Tr} y^* y = \operatorname{Tr} y y^*$ and $\operatorname{Tr} y^2 = \operatorname{Tr} (y^*)^2$, for a real $n \times n$ matrix y. If we denote the sum of the three terms on the last line in (1.51) by

$$\psi_0(y) + \psi_1(y) + \psi_2(y),$$

then, as in (1.47),

$$\frac{\partial^2 \psi_0}{\partial y_{\mu i}\, \partial y_{\nu j}} = (2A + 4B)\delta_{ij}\delta_{\mu\nu}. \tag{1.52}$$

Also, a brief computation gives

$$\frac{\partial^2 \psi_1}{\partial y_{\mu i}\, \partial y_{\nu j}} = 4B\delta_{\mu j}\delta_{\nu i} \tag{1.53}$$

and

$$\frac{\partial^2 \psi_2}{\partial y_{\mu i}\, \partial y_{\nu j}} = 8C\delta_{\mu i}\delta_{\nu j}. \tag{1.54}$$

Now, when φ is replaced by ψ_0, the differential operator of the form (1.40) is $(2A + 4B)\Delta$, similar to the computation giving (1.48). When φ is replaced by $\psi_1 + \psi_2$, the differential operator becomes

$$\begin{aligned}\big(\mathcal{L}w\big)^\nu &= 4B \sum_{\mu,i,j=1}^{n} \delta_{\mu j}\delta_{\nu i} w^\mu_{x_i x_j} + 8C \sum_{\mu,i,j=1}^{n} \delta_{\mu i}\delta_{\nu j} w^\mu_{x_i x_j} \\ &= (4B + 8C)\sum_j w^j_{x_\nu x_j}.\end{aligned} \tag{1.55}$$

We can write this as

$$\mathcal{L}w = (4B + 8C) \operatorname{grad} \operatorname{div} w, \tag{1.56}$$

where the *divergence* of the vector field $w = (w^1, \dots, w^n)$ is

$$\operatorname{div} w = \sum_j \frac{\partial w^j}{\partial x_j}, \tag{1.57}$$

and, as before, the gradient of a real-valued function on $\mathbb{R}^n$ is

$$\operatorname{grad} u = \Big(\frac{\partial u}{\partial x_1}, \dots, \frac{\partial u}{\partial x_n}\Big). \tag{1.58}$$

Thus the linearized PDE for vibration in the plane of Ω is

$$m w_{tt} - (2A + 4B)\Delta w - (4B + 8C) \operatorname{grad} \operatorname{div} w = 0. \tag{1.59}$$

The situation where $k = n$ represents a vibrating elastic solid, and the equation (1.59) is known as the equation of *linear elasticity*.

In linear elasticity it is common to linearize about an unstrained state. One writes (1.59) as

$$m w_{tt} - \mu\Delta w - (\lambda + \mu) \operatorname{grad} \operatorname{div} w = 0;$$

$\mu = 2A + 4B$ and $\lambda = 8C$ are called Lamé constants. For more on this, see [MH].

We will concentrate primarily on linear equations in this chapter, indeed, on scalar equations like (1.50). Methods of Chap. 16 will yield results on nonlinear equations of the form (1.12), in any number of x-variables, under a "hyperbolicity" assumption, which is that, for some $C > 0$,

$$\sum_{\mu,\nu=1}^{k} \sum_{i,j=1}^{n} \frac{\partial^2 f(y)}{\partial y_{\mu i}\, \partial y_{\nu j}}\, \xi_i \xi_j \tau_\mu \tau_\nu \ge C|\xi|^2|\tau|^2, \tag{1.60}$$

for $\xi \in \mathbb{R}^n,\ \tau \in \mathbb{R}^k$. A sufficient, though not necessary, condition for this to hold is that f be a strongly convex function of y. For example (in the case $k = n$), (1.60) holds for

$$f(y) = a \operatorname{Tr} y^* y + b \operatorname{Tr} y^2 \tag{1.61}$$

whenever $a > \max(0, -b)$, but such f is strongly convex only if $a > |b|$.

The notions of divergence, gradient, and Laplacian given above are for the case of Euclidean space $\mathbb{R}^n$. All these notions extend to more general Riemannian manifolds. The Laplacian will be defined in such a way as to generalize the identity

$$\int_{\mathbb{R}^n} (\Delta u)v \, dx = -\int_{\mathbb{R}^n} \text{grad } u \cdot \text{ grad } v \, dx, \tag{1.62}$$

for $u, v \in C_0^\infty(\mathbb{R}^n)$, which follows from the definition (1.49) by integration by parts. A further identity that generalizes to the case of Riemannian manifolds is

$$\Delta u = \text{ div grad } u, \tag{1.63}$$

which for a real-valued function on $\mathbb{R}^n$ follows immediately from the definitions of div, grad, and Δ given above.

We will discuss extensions of these concepts to Riemannian manifolds in the next few sections, starting with the notion of divergence in §2. Then we will derive a number of properties of solutions to wave equations, in §§5–8, and also discuss an extension of the wave equation (1.50) from the case $\mathbb{R} \times \mathbb{R}^n$ to Lorentz manifolds. The problem of proving existence of solutions will be tackled only in later chapters.

We will state here more precisely what the basic existence problem is. In the case of one of the wave equations produced above, say

$$\frac{\partial^2 u}{\partial t^2} - \Delta u = 0, \tag{1.64}$$

we desire to find u satisfying this PDE, given *initial conditions*

$$u(0, x) = f(x), \quad u_t(0, x) = g(x). \tag{1.65}$$

If $\partial\Omega \neq 0$, we also need to impose a boundary condition. There is in particular the *Dirichlet condition*

$$u(t, x) = 0, \ \text{ for } x \in \partial\Omega, \tag{1.66}$$

in the case of a membrane tied down along $\partial\Omega$, as discussed above. There are other boundary conditions that arise in other situations, such as the Neumann boundary condition described in §5, and others mentioned in subsequent chapters. We also can replace (1.64) and (1.66) by nonhomogeneous equations, that is, replace the zeros on the right by given functions.

In this section we have concentrated on evolution equations, involving motion with the passage of time. It is also of interest to study stationary problems, where there is no time dependence. In other words, one looks for stationary points for

$$J(u) = \int_{\Omega} f(u_x(x)) \, dx. \tag{1.67}$$

Thus one obtains a PDE of the form

$$f''(u_x)u_{xx} = 0, \tag{1.68}$$

interpreted via (1.36), as the stationary condition for $J(u)$. In the case $f(u_x) = |u_x|^2$, this becomes the Laplace equation

$$\Delta u = 0. \tag{1.69}$$

A typical boundary condition is the nonhomogeneous Dirichlet condition

$$u = \psi \text{ on } \partial\Omega. \tag{1.70}$$

The existence of a solution to this will follow from results of Chap. 5.

Exercises

1. Compare the formulas (1.22) and (1.23) for longitudinal and transverse waves. For a piano wire, a is very close to 1. What does this imply about the relative propagation speeds of longitudinal and transverse waves along a piano wire? Which type of waves produce audible sounds?
2. For a function f appearing in (1.60), to be strongly convex means

$$\sum_{\mu,\nu}\sum_{i,j} \frac{\partial^2 f(y)}{\partial y_{\mu i}\partial y_{\nu j}} \lambda_{\mu i}\lambda_{\nu j} \geq C_0|\lambda|^2, \tag{1.71}$$

   where $|\lambda|^2 = \sum_{\mu,i} |\lambda_{\mu i}|^2$. Show that this estimate implies (1.60). Prove the statements made about $f(y) = a \operatorname{Tr} y^* y + b \operatorname{Tr} y^2$ after (1.61).
3. Suppose more generally that $f(y) = a \operatorname{Tr} y^* y + b \operatorname{Tr} y^2 + c(\operatorname{Tr} y)^2$. For what values of a, b, and c is f strongly convex? For what values of a, b, and c does one have the strong ellipticity condition (1.60)?
4. The following exercise relates to the choice of the word "linearization" in describing the relation between the (1.12) and (1.20). For $\Omega \subset \mathbb{R}^n$, bounded with smooth boundary, definev

$$F : C^2(\overline{\Omega}, \mathbb{C}^k) \to C(\overline{\Omega}, \mathbb{C}^k)$$

   by

$$F(u) = f''(u_x)u_{xx},$$

   the right side defined by (1.36). Assume f is C^∞. Show that F is differentiable, as a map between Banach spaces, and that

$$DF(j)w = Lw,$$

   where $Lw = Aw_{xx},\ A = f''(j)$, as defined by (1.40).
5. If $u = u(t,x)$ is a *real-valued* function on $\mathbb{R} \times \Omega$, show that the PDE for u giving the stationary condition for the function (1.67) can be written in the form

$$\operatorname{div} f_p(u_x) = 0, \tag{1.72}$$

where, if $f = f(p) = f(p_1, \dots, p_n)$, then $f_p(u_x)$ is the vector field with components $(\partial f/\partial p_j)(u_x)$. Compare (5.39).

2. The divergence of a vector field

Let M be an n-dimensional manifold, provided with a volume form $\omega \in \Lambda^n M$. Let X be a vector field on M. Then the divergence of X, denoted div X, is a function on M that measures the rate of change of the volume form under the flow generated by X. Thus it is defined by

$$\mathcal{L}_X \omega = (\operatorname{div} X)\omega. \tag{2.1}$$

Here, $\mathcal{L}_X$ denotes the Lie derivative. In view of the general formula $\mathcal{L}_X \alpha = d\alpha \rfloor X + d(\alpha \rfloor X)$, derived in Chap. 1, since $d\omega = 0$ for any n-form ω on M, we have

$$(\operatorname{div} X)\omega = d(\omega \rfloor X). \tag{2.2}$$

If $M = \mathbb{R}^n$, with the standard volume element

$$\omega = dx_1 \wedge \cdots \wedge dx_n, \tag{2.3}$$

and if

$$X = \sum X^j(x) \frac{\partial}{\partial x_j}, \tag{2.4}$$

then

$$\omega \rfloor X = \sum_{j=1}^{n} (-1)^{j-1} X^j(x)\, dx_1 \wedge \cdots \wedge \widehat{dx_j} \wedge \cdots \wedge dx_n. \tag{2.5}$$

Hence, in this case, (2.2) yields the formula used in (1.57) :

$$\operatorname{div} X = \sum_{j=1}^{n} \partial_j X^j, \tag{2.6}$$

where we use the notation

$$\partial_j f = \frac{\partial f}{\partial x_j}. \tag{2.7}$$

Suppose now that M is an oriented manifold endowed with a Riemannian metric $g_{jk}(x)$. Then M carries a natural volume element ω, determined by the condition that, if one has a coordinate system in which $g_{jk}(p_0) = \delta_{jk}$, then $\omega(p_0) = dx_1 \wedge \cdots \wedge dx_n$. This condition produces the following formula, in any oriented coordinate system:

$$\omega = \sqrt{g}\, dx_1 \wedge \cdots \wedge dx_n, \tag{2.8}$$

where

$$g = \det(g_{jk}). \tag{2.9}$$

In order to derive (2.8) , note that if coordinates y are related to x linearly, that is, $y_j = \sum A_{jk} x_k$, then

$$\sum dy_j^2 = \sum_{j,k,\ell} A_{jk} A_{j\ell}\, dx_k\, dx_\ell = \sum g_{k\ell}\, dx_k\, dx_\ell,$$

with

$$g_{k\ell} = \sum_j A_{\ell j} A_{jk},$$

provided $A = (A_{jk})$ is symmetric. Now construct A as the positive-definite square root of the positive-definite matrix $G = \big(g_{jk}(x_0)\big)$. In other words, if $\{v_j\}$ is an orthonormal basis of $\mathbb{R}^n$ with $Gv_j = c_j v_j$, set $Av_j = c_j^{1/2} v_j$. The transformation law for $\Lambda^n A$ on $\Lambda^n \mathbb{R}$ gives

$$\begin{aligned} dy_1 \wedge \cdots \wedge dy_n &= (\det A)\, dx_1 \wedge \cdots \wedge dx_n \\ &= \sqrt{g(x_0)}\, dx_1 \wedge \cdots \wedge dx_n, \end{aligned}$$

from which the formula (2.8) follows.

We now compute div X when the volume element on M is given by (2.8) . We have

$$\omega \rfloor X = \sum_j (-1)^{j-1} X^j \sqrt{g}\, dx_1 \wedge \cdots \wedge \widehat{dx_j} \wedge \cdots \wedge dx_n \tag{2.10}$$

and hence

$$d(\omega \rfloor X) = \partial_j(\sqrt{g} X^j)\, dx_1 \wedge \cdots \wedge dx_n. \tag{2.11}$$

Here, as below, we use the summation convention. Hence the formula (2.2) gives

$$\operatorname{div} X = g^{-1/2} \partial_j (g^{1/2} X^j). \tag{2.12}$$

We next derive a result known as the divergence theorem, as a consequence of Stokes' formula, proved in Chap. 1. Recall that Stokes' formula for differential forms is

$$\int_M d\alpha = \int_{\partial M} \alpha, \tag{2.13}$$

for an $(n-1)$-form on M, assumed to be a smooth, compact, oriented manifold with boundary. If $\alpha = \omega \rfloor X$, the formula(2.2) gives

$$\int_M (\text{div } X)\omega = \int_{\partial M} \omega \rfloor X. \tag{2.14}$$

This is one form of the divergence theorem. We will produce an alternative expression for the integrand on the right before stating the result formally.

Given that ω is the volume form for M determined by a Riemannian metric, we can write the interior product $\omega \rfloor X$ in terms of the volume element ω_∂ on ∂M, with its induced Riemannian metric, as follows. Pick normal coordinates on M, centered at $p_0 \in \partial M$, such that ∂M is tangent to the hyperplane $\{x_n = 0\}$ at $p_0 = 0$. Then it is clear that, at p_0,

$$j^*(\omega \rfloor X) = \langle X, \nu \rangle \omega_\partial, \tag{2.15}$$

where ν is the unit vector normal to ∂M, pointing out of M and $j : \partial M \hookrightarrow M$ is the natural inclusion. The two sides of (2.15), which are both defined in a coordinate-independent fashion, are hence equal on ∂M, and the identity (2.14) becomes

$$\int_M (\text{div } X)\omega = \int_{\partial M} \langle X, \nu \rangle \omega_\partial. \tag{2.16}$$

Finally, we adopt the following common notation: we denote the volume element on M by dV and that on ∂M by dS, obtaining the *divergence theorem*:

Theorem 2.1. *If M is a compact manifold with boundary, X a smooth vector field on M, then*

$$\int_M (\mathit{div}\ X)\, dV = \int_{\partial M} \langle X, \nu \rangle\, dS, \tag{2.17}$$

where ν is the unit outward-pointing normal to ∂M.

The only point left to mention here is that M need not be orientable. Indeed, we can treat dV and dS as measures and note that all objects in (2.17) are independent of a choice of orientation. To prove the general case, just use a partition of unity supported on orientable pieces.

The definition of the divergence of a vector field given by (2.1), in terms of how the flow generated by the vector field magnifies or diminishes volumes, is a good geometrical characterization, explaining the use of the term "divergence." There are other characterizations of the divergence operation, of a more analytical flavor, which are also quite useful. Here is one.

Proposition 2.2. *The divergence operation is the negative of the adjoint of the gradient operation on vector fields; if X is a vector field and u a function on M, one compactly supported on the interior of M, then*

$$(X, grad\, u)_{L^2(M)} = -(div\, X, u)_{L^2(M)}. \tag{2.18}$$

The asserted integral identity here is

$$\int\limits_M \langle X, \text{grad}\, u\rangle\, dV(x) = -\int\limits_M (\text{div}\, X)u\, dV(x),$$

provided either u or X has compact support in the interior of M. Note that

$$\langle X, \text{grad}\, u\rangle = \langle X, du\rangle = Xu.$$

In fact, we will use the divergence theorem to obtain a more general result, in which neither u or X is required to vanish on ∂M. We apply (2.17) with X replaced by uX. We have the following "derivation" identity:

$$\text{div}\, uX = u\,\text{div}\, X + \langle du, X\rangle = u\,\text{div}\, X + Xu, \tag{2.19}$$

which follows easily from the formula (2.12). The divergence theorem immediately gives the following result.

Proposition 2.3. *If M is a smooth, compact manifold with boundary, u a smooth function, X a smooth vector field on M, then*

$$\int\limits_M (div\, X)u\, dV + \int\limits_M Xu\, dV = \int\limits_{\partial M} \langle X, \nu\rangle u\, dS. \tag{2.20}$$

We can also express the adjoint of the differential operator X, defined by

$$\int\limits_M (X^*u)\overline{v}\, dV = \int\limits_M u(X\overline{v})\, dV, \tag{2.21}$$

for $v \in C_0^\infty(\mathring{M})$, using the divergence, as follows:

Proposition 2.4. *If X is a smooth vector field on M, then*

$$X^*u = -Xu - (div\, X)u. \tag{2.22}$$

This is equivalent to the statement that

$$\int_M \big[(Xu)v + u(Xv)\big]\,dV = -\int_M (\text{div}\,X)uv\,dV, \tag{2.23}$$

for $u, v \in C_0^\infty(\overset{\circ}{M})$. In fact, from (2.20) we can obtain the following more general result.

Proposition 2.5. *If u and v are smooth functions and X a smooth vector field on a compact manifold M with boundary, then*

$$\int_M \big[(Xu)v + u(Xv)\big]\,dV = -\int_M (\mathit{div}\,X)uv\,dV + \int_{\partial M} \langle X, \nu\rangle uv\,dS. \tag{2.24}$$

Proof. Replace u by uv in (2.20) and use the derivation identity $X(uv) = (Xu)v + u(Xv)$.

Exercises

1. Given a Hamiltonian vector field

$$H_f = \sum_{j=1}^n \Big[\frac{\partial f}{\partial \xi_j}\frac{\partial}{\partial x_j} - \frac{\partial f}{\partial x_j}\frac{\partial}{\partial \xi_j}\Big],$$

calculate div H_f directly from (2.6).

2. If M is a smooth domain in $\mathbb{R}^2$, apply the divergence theorem (2.17) to the vector field $X = g\partial/\partial x - f\partial/\partial y$ to deduce Green's formula:

$$\int_{\partial M} f\,dx + g\,dy = \iint_M \Big(\frac{\partial g}{\partial x} - \frac{\partial f}{\partial y}\Big)\,dx\,dy.$$

3. Show that the identity (2.19) for div (uX) follows from (2.2) and

$$du \wedge (\omega \rfloor X) = (Xu)\omega.$$

Prove this identity, for any n-form ω on M^n. What happens if ω is replaced by a k-form, $k < n$?

4. Relate Exercise 3 to the calculations

$$\mathcal{L}_{uX}\alpha = u\mathcal{L}_X\alpha + du \wedge (\iota_X\alpha) \tag{2.25}$$

and

$$du \wedge (\iota_X\alpha) = -\iota_X(du \wedge \alpha) + (Xu)\alpha, \tag{2.26}$$

valid for any k-form α. The last identity follows from (13.37) of Chap. 1; compare with formula (10.27) of this chapter.

5. Show that

$$\operatorname{div}[X,Y] = X(\operatorname{div} Y) - Y(\operatorname{div} X).$$

3. The covariant derivative and divergence of tensor fields

The covariant derivative of a vector field on a Riemannian manifold was introduced in Chap. 1, §11, in connection with the study of geodesics. We will briefly recall this concept here and relate the divergence of a vector field to the covariant derivative, before generalizing these notions to apply to more general tensor fields. A still more general setting for covariant derivatives is discussed in Appendix C.

If X and Y are vector fields on a Riemannian manifold M, then $\nabla_X Y$ is a vector field on M, the covariant derivative of Y with respect to X. We have the properties

$$\nabla_{(fX)} Y = f \nabla_X Y \tag{3.1}$$

and

$$\nabla_X (fY) = f \nabla_X Y + (Xf)Y, \tag{3.2}$$

the latter being the *derivation property*. Also, ∇ is related to the metric on M by

$$Z\langle X,Y\rangle = \langle \nabla_Z X, Y\rangle + \langle X, \nabla_Z Y\rangle, \tag{3.3}$$

where $\langle X,Y\rangle = g_{jk} X^j Y^k$ is the inner product on tangent vectors. The Levi–Civita connection on M is uniquely specified by (3.1) –(3.3) and the torsion free property:

$$\nabla_X Y - \nabla_Y X = [X,Y]. \tag{3.4}$$

There is the explicit defining formula (derived already in (11.22) of Chap. 1)

$$\begin{aligned} 2\langle \nabla_X Y, Z\rangle &= X\langle Y,Z\rangle + Y\langle X,Z\rangle - Z\langle X,Y\rangle \\ &\quad + \langle [X,Y],Z\rangle - \langle [X,Z],Y\rangle - \langle [Y,Z],X\rangle, \end{aligned} \tag{3.5}$$

which follows from cyclically permuting X, Y, and Z in (3.3) and combining the results, exploiting (3.4) to cancel out all covariant derivatives but one. Another way of writing this is the following. If

$$X = X^k D_k, \quad D_k = \frac{\partial}{\partial x_k} \text{ (summation convention)}, \tag{3.6}$$

then

$$\nabla_{D_j} X = X^k{}_{;j}\, D_k, \tag{3.7}$$

with

$$X^k{}_{;j} = \partial_j X^k + \sum_\ell \Gamma^k{}_{\ell j} X^\ell, \tag{3.8}$$

where the "connection coefficients" are given by the formula

$$\Gamma^\ell{}_{jk} = \frac{1}{2} g^{\ell\mu} \Big[\frac{\partial g_{j\mu}}{\partial x_k} + \frac{\partial g_{k\mu}}{\partial x_j} - \frac{\partial g_{jk}}{\partial x_\mu} \Big], \tag{3.9}$$

equivalent to (3.5). We also recall that $\partial g_{k\mu}/\partial x_j$ can be recovered from $\Gamma^\ell{}_{jk}$:

$$\frac{\partial g_{k\mu}}{\partial x_j} = g_{\ell\mu} \Gamma^\ell{}_{jk} + g_{\ell k} \Gamma^\ell{}_{j\mu}. \tag{3.10}$$

The divergence of a vector field has an important expression in terms of the covariant derivative.

Proposition 3.1. *Given a vector field X with components X^k as in (3.6),*

$$\mathit{div}\; X = X^j{}_{;j}. \tag{3.11}$$

Proof. This can be deduced from our previous formula for div X,

$$\begin{aligned} \operatorname{div} X &= g^{-1/2} \partial_j (g^{1/2} X^j) \\ &= \partial_j X^j + (\partial_j \log g^{1/2}) X^j. \end{aligned} \tag{3.12}$$

One way to see this is the following. We can think of ∇X as defining a tensor field of type $(1, 1)$:

$$(\nabla X)(Y) = \nabla_Y X. \tag{3.13}$$

Then the right side of (3.11) is the trace of such a tensor field:

$$X^j{}_{;j} = \operatorname{Tr} \nabla X. \tag{3.14}$$

This is clearly defined independently of any choice of coordinate system. If we choose an exponential coordinate system centered at a point $p \in M$, then $g_{jk}(p) = \delta_{jk}$ and $\partial g_{jk}/\partial x_\ell = 0$ at p, so (3.12) gives div $X = \partial_j X^j$ at p, in this coordinate system, while the right side of (3.11) is equal to $\partial_j X^j + \Gamma^j{}_{\ell j} X^\ell = \partial_j X^j$ at p. This proves the identity (3.11).

The covariant derivative can be applied to forms, and other tensors, by requiring ∇ to be a derivation. On scalar functions, set

$$\nabla_X u = Xu. \tag{3.15}$$

For a 1-form α, $\nabla_X \alpha$ is characterized by the identity

$$\langle Y, \nabla_X \alpha\rangle = X\langle Y, \alpha\rangle - \langle \nabla_X Y, \alpha\rangle. \tag{3.16}$$

Denote by $\mathfrak{X}(M)$ the space of smooth vector fields on M, and by $\Lambda^1(M)$ the space of smooth 1-forms; each of these is a module over $C^\infty(M)$. Generally, a tensor field of type (k, j) defines a map (with j factors of $\mathfrak{X}(M)$ and k of $\Lambda^1(M)$)

$$F : \mathfrak{X}(M) \times \cdots \times \mathfrak{X}(M) \times \Lambda^1(M) \times \cdots \times \Lambda^1(M) \longrightarrow C^\infty(M), \tag{3.17}$$

which is linear in each factor, over the ring $C^\infty(M)$. A vector field is of type $(1, 0)$ and a 1-form is of type $(0, 1)$. The covariant derivative $\nabla_X F$ is a tensor of the same type, defined by

(3.18)
$$\begin{aligned}(\nabla_X F)(Y_1, \dots, Y_j, \alpha_1, \dots, \alpha_k) = X \cdot \big(F(Y_1, \dots, Y_j, \alpha_1, \dots, \alpha_k)\big) \\ - \sum_{\ell=1}^{j} F(Y_1, \dots, \nabla_X Y_\ell, \dots, Y_j, \alpha_1, \dots, \alpha_k) \\ - \sum_{\ell=1}^{k} F(Y_1, \dots, Y_j, \alpha_1, \dots, \nabla_X \alpha_\ell, \dots, \alpha_k),\end{aligned}$$

where $\nabla_X \alpha_\ell$ is uniquely defined by (3.16). We can naturally consider ∇F as a tensor field of type $(k, j + 1)$:

$$(\nabla F)(X, Y_1, \dots, Y_j, \alpha_1, \dots, \alpha_k) = (\nabla_X F)(Y_1, \dots, Y_j, \alpha_1, \dots, \alpha_k). \tag{3.19}$$

For example, if Z is a vector field, ∇Z is a vector field of type $(1, 1)$, as already anticipated in (3.13). Hence it makes sense to consider the tensor field $\nabla(\nabla Z)$, of type $(1, 2)$. For vector fields X and Y, we define the *Hessian* $\nabla^2_{(X,Y)} Z$ to be the vector field characterized by

$$\langle \nabla^2_{(X,Y)} Z, \alpha\rangle = (\nabla\nabla Z)(X, Y, \alpha). \tag{3.20}$$

Since, by (3.19), if $F = \nabla Z$, we have

$$F(Y, \alpha) = \langle \nabla_Y Z, \alpha\rangle, \tag{3.21}$$

and, by (3.18),

$$(\nabla_X F)(Y,\alpha) = X \cdot \big(F(Y,\alpha)\big) - F(\nabla_X Y,\alpha) - F(Y,\nabla_X \alpha), \tag{3.22}$$

it follows by substituting (3.21) into (3.22) and using (3.16) that

$$\nabla^2_{(X,Y)} Z = \nabla_X \nabla_Y Z - \nabla_{(\nabla_X Y)} Z; \tag{3.23}$$

this is a useful formula for the Hessian of a vector field.

More generally, for any tensor field F, of type (j,k), the Hessian $\nabla^2_{(X,Y)} F$, also of type (j,k), is defined in terms of the tensor field $\nabla^2 F = \nabla(\nabla F)$, of type $(j,k+2)$, by the same type of formula as (3.20), and we have

$$\nabla^2_{(X,Y)} F = \nabla_X(\nabla_Y F) - \nabla_{(\nabla_X Y)} F, \tag{3.24}$$

by an argument similar to that for (3.23).

The metric tensor g is of type $(0,2)$, and the identity (3.3) is equivalent to

$$\nabla_X g = 0 \tag{3.25}$$

for all vector fields X (i.e., to $\nabla g = 0$). In index notation, this means

$$g_{jk;\ell} = 0 \;\text{ or, equivalently, }\; g^{jk}{}_{;\ell} = 0. \tag{3.26}$$

We also note that the zero torsion condition (3.4) implies

$$u_{;j;k} = u_{;k;j} \tag{3.27}$$

when u is a smooth scalar function, with second covariant derivative $\nabla\nabla u$, a tensor field of type $(0,2)$. It turns out that analogous second-order derivatives of a vector field differ by a term arising from the curvature tensor; this point is discussed in Appendix C, Connections and Curvature.

We have seen an expression for the divergence of a vector field in terms of the covariant derivative. We can use this latter characterization to provide a general notion of divergence of a tensor field. If T is a tensor field of type (k,j), with components

$$T_\alpha{}^\beta = T_{\alpha_1\cdots\alpha_j}{}^{\beta_1\cdots\beta_k} \tag{3.28}$$

in a given coordinate system, then div T is a tensor field of type $(k-1,j)$, with components

$$T_{\alpha_1\cdots\alpha_j}{}^{\beta_1\cdots\beta_{k-1}\ell}{}_{;\ell}. \tag{3.29}$$

In view of the special role played by the last index, the divergence of a tensor field T is mainly interesting when T has some symmetry property. In §7 we will introduce the stress-energy tensor, a symmetric second-order covariant tensor; raising indices produces a symmetric second-order tensor field of type $(2,0)$, whose divergence is an important object.

In view of (3.11), we know that a vector field X generates a volume-preserving flow if and only if $X^j{}_{;j} = 0$. Complementing this, we investigate the condition that the flow generated by X consists of isometries, that is, the flow leaves the metric g invariant, or equivalently

$$\mathcal{L}_X g = 0. \tag{3.30}$$

For vector fields U and V, we have

$$\begin{aligned}(\mathcal{L}_X g)(U,V) &= -\langle \mathcal{L}_X U, V\rangle - \langle U, \mathcal{L}_X V\rangle + X\langle U,V\rangle \\ &= \langle \nabla_X U - \mathcal{L}_X U, V\rangle + \langle U, \nabla_X V - \mathcal{L}_X V\rangle \\ &= \langle \nabla_U X, V\rangle + \langle U, \nabla_V X\rangle,\end{aligned} \tag{3.31}$$

where the first identity follows from the derivation property of $\mathcal{L}_X$, the second from the metric property (3.3) expressing $X\langle U,V\rangle$ in terms of covariant derivatives, and the third from the zero torsion condition (3.4). If U and V are coordinate vector fields $D_j = \partial/\partial x_j$, we can write this identity as

$$(\mathcal{L}_X g)(D_j, D_k) = g_{k\ell} X^\ell{}_{;j} + g_{j\ell} X^\ell{}_{;k}. \tag{3.32}$$

Thus X generates a group of isometries (one says X is a *Killing field*) if and only if

$$g_{k\ell} X^\ell{}_{;j} + g_{j\ell} X^\ell{}_{;k} = 0. \tag{3.33}$$

This takes a slightly shorter form for the covariant field

$$X_j = g_{jk} X^k. \tag{3.34}$$

We state formally the consequence, which follows immediately from (3.33) and the vanishing of the covariant derivatives of the metric tensor.

Proposition 3.2. *X is a Killing vector field if and only if*

$$X_{k;j} + X_{j;k} = 0. \tag{3.35}$$

Generally, half the quantity on the left side of (3.35) is called the *deformation tensor* of X. If we denote by ξ the 1-form $\xi = \sum X_j \, dx_j$, the deformation tensor is the *symmetric part* of $\nabla\xi$, a tensor field of type $(0,2)$. It is also useful to identify the antisymmetric part, which is naturally regarded as a 2-form.

Proposition 3.3. *We have*

$$d\xi = \frac{1}{2} \sum_{j,k} (X_{j;k} - X_{k;j})\, dx_k \wedge dx_j. \tag{3.36}$$

Proof. By definition,

$$d\xi = \frac{1}{2} \sum_{j,k} (\partial_k X_j - \partial_j X_k)\, dx_k \wedge dx_j, \tag{3.37}$$

and the identity with the right side of (3.36) follows from the symmetry $\Gamma^{\ell}{}_{jk} = \Gamma^{\ell}{}_{kj}$.

There is a useful generalization of the concept of a Killing field, namely a *conformal* Killing field, which is a vector field X whose flow consists of conformal diffeomorphisms of M, that is, preserves the metric tensor up to a scalar factor:

$$\mathcal{F}_X^{t*} g = \alpha(t,x) g \Longleftrightarrow \mathcal{L}_X g = \lambda(x) g. \tag{3.38}$$

Note that the trace of $\mathcal{L}_X g$ is $2 \operatorname{div} X$, by (3.32), so the last identity in (3.38) is equivalent to $\mathcal{L}_X g = (2/n)(\operatorname{div} X) g$ or, with $(1/2)\mathcal{L}_X g = \operatorname{Def} X$,

$$\operatorname{Def} X - \frac{1}{n} (\operatorname{div} X) g = 0 \tag{3.39}$$

is the equation of a conformal Killing field.

To end this section, and prepare for subsequent material, we note that concepts developed so far for Riemannian manifolds, that is, manifolds with positive-definite metric tensors, have extensions to indefinite metric tensors, including *Lorentz* metrics.

A Riemannian metric tensor produces a symmetric isomorphism

$$G : T_x M \longrightarrow T_x^* M, \tag{3.40}$$

which is positive. More generally, a symmetric isomorphism (3.40) corresponds to a nondegenerate metric tensor. Such a tensor has a well defined signature (j,k), $j + k = n = \dim M$; at each $x \in M$, $T_x M$ has a basis $\{e_1, \dots, e_n\}$ of mutually orthogonal vectors such that $\langle e_1, e_1 \rangle = \cdots = \langle e_j, e_j \rangle = 1$, while $\langle e_{j+1}, e_{j+1} \rangle = \cdots = \langle e_n, e_n \rangle = -1$. If $j = 1$ (or $k = 1$), we say M has a Lorentz metric.

The concepts discussed in this section in the Riemannian case, such as the covariant derivative, all extend with little change to the general nondegenerate case. We will see this in use, in the Lorentz case, in §7.

Exercises

1. Let φ be a tensor field of type $(0, k)$ on a Riemannian manifold, endowed with its Levi–Civita connection. Show that

$$(\mathcal{L}_X \varphi - \nabla_X \varphi)(U_1, \ldots, U_k) = \sum_j \varphi(U_1, \ldots, \nabla_{U_j} X, \ldots, U_k).$$

How does this generalize (3.31)?

2. Recall the formula (13.56) of Chap. 1, when ω is a k-form:

$$(d\omega)(X_0, \ldots, X_k) = \sum_{j=0}^{k} (-1)^j X_j \cdot \omega(X_0, \ldots, \widehat{X}_j, \ldots, X_k) + \sum_{0 \le \ell < j \le k} (-1)^{j+\ell}$$
$$\times\, \omega([X_\ell, X_j], X_0, \ldots, \widehat{X}_\ell, \ldots, \widehat{X}_j, \ldots, X_k).$$

Show that the last double sum can be replaced by

$$-\sum_{\ell<j} (-1)^j \omega(X_0, \ldots, \nabla_{X_j} X_\ell, \ldots, \widehat{X}_j, \ldots, X_k)$$
$$-\sum_{\ell>j} (-1)^j \omega(X_0, \ldots, \widehat{X}_j, \ldots, \nabla_{X_j} X_\ell, \ldots, X_k).$$

3. Using Exercise 2 and the expansion of $(\nabla_{X_j} \omega)(X_0, \ldots, \widehat{X}_j, \ldots, X_k)$ via the derivation property, show that

$$(d\omega)(X_0, \ldots, X_k) = \sum_{j=0}^{k} (-1)^j (\nabla_{X_j} \omega)(X_0, \ldots, \widehat{X}_j, \ldots, X_k). \tag{3.41}$$

Note that this generalizes Proposition 3.3.

4. Prove the identity
$$\frac{\partial \log \sqrt{g}}{\partial x_j} = \sum_\ell \Gamma^\ell{}_{\ell j}.$$

Use either the identity (3.11), involving the divergence, or the formula (3.9) for $\Gamma^\ell{}_{jk}$. Which is easier?

5. Show that the characterization (3.17) of a tensor field of type (k, j) is equivalent to the condition that F be a section of the vector bundle $\left(\otimes^j T^*\right) \otimes \left(\otimes^k T\right)$ or, equivalently, of the bundle Hom $(\otimes^j T, \otimes^k T)$. Think of other variants.

6. The operation $X_j = g_{jk} X^k$ is called *lowering indices*. It produces a 1-form (section of T^*M) from a vector field (section of TM), implementing the isomorphism (3.38). Similarly, one can raise indices:

$$Y^j = g^{jk} Y_k,$$

producing a vector field from a 1-form, that is, implementing the inverse isomorphism. Define more general operations raising and lowering indices, passing from tensor fields of type (j, k) to other tensor fields, of type (ℓ, m), with $\ell + m = j + k$. One says that these tensor fields are associated to each other via the metric tensor.

7. Using (3.16), show that if $\alpha = a_k(x)\,dx_k$ (summation convention), then $\nabla_{D_j}\alpha = a_{k;j}\,dx_k$, with
$$a_{k;j} = \partial_j a_k - \sum_\ell \Gamma^\ell{}_{kj} a_\ell.$$
Compare with (3.8). Use this to verify that (3.36)and (3.37) are equal. Work out a corresponding formula for $\nabla_{D_\ell} T$ when T is a tensor field of type (j,k), as in (3.28)
8. Using the formula (3.23) for the Hessian, show that, for vector fields X, Y, Z on M,
$$\big(\nabla^2_{(X,Y)} - \nabla^2_{(Y,X)}\big)Z = \big([\nabla_X, \nabla_Y] - \nabla_{[X,Y]}\big)Z.$$
Denoting this by $R(X,Y,Z)$, show that it is linear in each of its three arguments over the ring $C^\infty(M)$, for example, $R(X,Y,fZ) = f\ R(X,Y,Z)$ for $f \in C^\infty(M)$. Discussion of $R(X,Y,Z)$ as the *curvature tensor* is given in Appendix C, Connections and Curvature.
9. Verify (3.24). For a function u, to show that $\nabla^2_{(X,Y)}u = \nabla^2_{(Y,X)}u$, use the special case
$$\nabla^2_{(X,Y)}u = XYu - (\nabla_X Y)\cdot u$$
of (3.24). Note that this is an invariant formulation of (3.27). Show that
$$\nabla^2_{(X,Y)}u = \frac{1}{2}(\mathcal{L}_V g)(X,Y), \quad V = \text{ grad } u.$$
10. Let ω be the *volume form* of an oriented Riemannian manifold M. Show that $\nabla_X\omega = 0$ for all vector fields X.
11. Let X be a vector field on a Riemannian manifold M. Show that the formal adjoint of ∇_X, acting on vector fields, is
$$\nabla^*_X Y = -\nabla_X Y - (\text{div } X)Y. \tag{3.42}$$
12. Show that the formal adjoint of $\mathcal{L}_X$, acting on vector fields, is
$$\mathcal{L}^*_X Y = -\mathcal{L}_X Y - (\text{div } X)Y - 2\,\text{Def}(X)Y, \tag{3.43}$$
where $\text{Def}(X)$ is a tensor field of type $(1,1)$, given by
$$\frac{1}{2}(\mathcal{L}_X g)(Z,Y) = g(Z, \text{Def}(X)Y), \tag{3.44}$$
g being the metric tensor.
13. With div defined by (3.29) for tensor fields, show that
$$\text{div}\,(X \otimes Y) = (\text{div } Y)X + \nabla_Y X. \tag{3.45}$$
14. If X, Y, and Z have compact support, show that
$$(Z, \text{div}\,(X \otimes Y))_{L^2} = -(\nabla_Y Z, X)_{L^2}.$$
15. If $\gamma(s)$ is a unit-speed geodesic on a Riemannian manifold M, $\gamma'(s) = T(s)$, and X is a vector field on M, show that
$$\frac{d}{ds}\big\langle T(s), X(\gamma(s))\big\rangle = \frac{1}{2}\big(\mathcal{L}_X g\big)(T,T). \tag{3.46}$$

Deduce that if X is a Killing field, then $\langle T, X\rangle$ is constant on γ. Relate this to the conservation law for geodesic flow on a surface of revolution, discussed in Chap. 1, §16. (*Hint*: Show that the left side of (3.46) is equal to $\langle T, \nabla_T X\rangle$.)

16. If we define Def: $C^\infty(M, T) \to C^\infty(M, S^2T^*)$ by $\operatorname{Def}(X) = (1/2)\mathcal{L}_X g$, show that

$$\operatorname{Def}^* u = -\operatorname{div} u,$$

where $(\operatorname{div} u)^j = u^{jk}{}_{;k}$, as in (3.29).

4. The Laplace operator on a Riemannian manifold

We define the Laplace operator on a Riemannian manifold M, with metric g_{jk}, in a way that naturally generalizes the characterizations of the Laplace operator on Euclidean space, given by (1.49), (1.62), and (1.63). Taking (1.62) as fundamental, we define the Laplace operator Δ on M to be the second-order differential operator satisfying

$$-(\Delta u, v) = (du, dv) = (\operatorname{grad} u, \operatorname{grad} v), \tag{4.1}$$

for $u, v \in C_0^\infty(M)$. Here the left side is

$$-\int_M (\Delta u)\overline{v}\, dV, \tag{4.2}$$

where dV is the natural volume element, given in local coordinates by $\sqrt{g}dx_1 \cdots dx_n$. The right side of (4.1), for u and v supported in a coordinate patch, is

$$\begin{aligned}\int \langle du, dv\rangle\, dV &= \int g^{jk}(\partial_j u)(\partial_k \overline{v})\sqrt{g}\, dx \\ &= -\int \overline{v}\partial_k\big(g^{1/2}g^{jk}\,\partial_j u\big)g^{-1/2}g^{1/2}\, dx,\end{aligned} \tag{4.3}$$

integrating by parts, so we see that Δ is given in local coordinates by

$$\Delta u = g^{-1/2}\,\partial_j\big(g^{jk}g^{1/2}\,\partial_k u\big). \tag{4.4}$$

Soon we will see how to modify (4.1) when u and v do not vanish on ∂M, in case M is a compact Riemannian manifold with boundary.

We now show that (1.63)generalizes, that is, we have

$$\Delta u = \operatorname{div}\operatorname{grad} u. \tag{4.5}$$

In fact, in view of the formula

$$\operatorname{div} X = g^{-1/2}\,\partial_j(g^{1/2}X^j)$$

derived in (2.12), together with

$$X^j = g^{jk}\partial_k u, \quad \text{for } X = \text{ grad } u,$$

we see that (4.5) follows directly from the local coordinate formula (4.4). Note that the identity

$$(X, \text{grad } v)_{L^2} = -(\text{div } X, v)_{L^2}, \tag{4.6}$$

proved in (2.18), when applied to $X = \text{grad } u$, also gives (4.5) directly.

Applying the refinement (2.20) of (4.6) gives us important identities due to Green. Let us use the notation

$$\frac{\partial u}{\partial \nu} = \langle \text{grad } u, \nu \rangle \tag{4.7}$$

for the normal component of grad u; $\partial u/\partial \nu$ is called the *normal derivative* of u. If we exploit (2.20) with $X = \text{grad}\overline{v}$, we get the identity (4.8) below; if we interchange u and $\overline{v}$ and subtract the resulting expression from (4.8), we obtain (4.9). This provides a proof of *Green's identities*:

Proposition 4.1. *If M is a compact Riemannian manifold with boundary, then for $u, v \in C^\infty(M)$, we have*

$$-(u, \Delta v)_{L^2} = (du, dv) - \int_{\partial M} u\Big(\frac{\partial \overline{v}}{\partial \nu}\Big)\, dS \tag{4.8}$$

and

$$(\Delta u, v) - (u, \Delta v) = \int_{\partial M} \Big[\Big(\frac{\partial u}{\partial \nu}\Big)\overline{v} - u\Big(\frac{\partial \overline{v}}{\partial \nu}\Big)\Big]\, dS. \tag{4.9}$$

Next we express the Laplace operator in terms of covariant derivatives. As we have seen,

$$\text{div } X = X^j{}_{;j}.$$

If we set $X = \text{grad } u$, we obtain

$$\Delta u = g^{jk} u_{;j;k}, \tag{4.10}$$

using the fact that $g^{jk}{}_{;\ell} = 0$. Here, $\sum u_{;j;k} dx_k \otimes dx_j$ is a tensor field of type $(0,2)$, which is the same as $\nabla^2 u$. Recall that $\nabla^2 F$ is a tensor field of type $(j, k+2)$ whenever F is a tensor field of type (j,k). The formula (4.10) can be rewritten as

$$\Delta u = \text{ Tr}_g \nabla^2 u, \tag{4.11}$$

where Tr_g denotes the trace of $\nabla^2 u(x)$, as a quadratic form on $T_x M$, in terms of the quadratic form given by the metric tensor g. In other words, we can define a tensor field $H(u)$, of type $(1, 1)$, by

$$\langle H(u)X, Y\rangle = (\nabla^2 u)(X, Y), \tag{4.12}$$

and $\mathrm{Tr}_g \nabla^2 u = \ \mathrm{Tr}\ H(u)$.

Since the Laplace operator is defined in a coordinate-independent manner on a Riemannian manifold, it is clear that if $F : M \to M$ is a diffeomorphism and $F^* : C^\infty(M) \to C^\infty(M)$ is defined by $F^* u(x) = u(F(x))$, then F^* commutes with the Laplace operator provided F is an isometry. Thus, if X is a vector field on M, X commutes with Δ provided the flow $\mathcal{F}_X^t$ generated by X consists of isometries. This result has a converse.

Proposition 4.2. *A vector field X commutes with Δ if and only if X generates a group of isometries.*

The proof rests on a computation of independent interest. In fact, a manipulation of (4.10), which we leave to the reader, yields the general identity

$$\begin{aligned} [\Delta, X]u &= (X^{j;k} + X^{k;j})u_{;j;k} + (X^{j;k} + X^{k;j})_{;j}\, u_{;k} \\ &= g^{-1/2}\, \partial_j \big(g^{1/2}(X^{j;k} + X^{k;j})\, \partial_k u\big). \end{aligned} \tag{4.13}$$

Thus $[\Delta, X] = 0$ if and only if $X^{j;k} + X^{k;j} = 0$, which is equivalent to the condition (3.35) for a Killing field.

Exercises

1. If $u \in C^\infty(M)$, $X = \mathrm{grad}\ u$, the condition that X generates a volume-preserving flow is that $\Delta u = 0$. What PDE on u is equivalent to the statement that X is a Killing field?
2. Verify formula (4.13) for $[\Delta, X]$. Show that it has the invariant formulation

$$\frac{1}{2}[\Delta, X]u = \langle \mathrm{Def}(X), \nabla^2 u\rangle + \langle \mathrm{div}\ \mathrm{Def}(X), du\rangle = \ \mathrm{div}\big(\mathrm{Def}(X)\cdot du\big), \tag{4.14}$$

   in terms of the deformation tensor $\mathrm{Def}(X)$, with components $(1/2)(X^{j;k} + X^{k;j})$, that is, the type $(2, 0)$ analogue of the tensor field of type $(1, 1)$ given by (3.42), or the tensor field of type $(0, 2)$ equal to half of (3.35).
3. Show that the Laplace operator $\Delta = \partial^2/\partial x_1^2 + \cdots + \partial^2/\partial x_n^2$ on $\mathbb{R}^n$ has the following expressions in various coordinate systems:
   (a) Polar coordinates on $\mathbb{R}^2$: $x_1 = r\ \cos\theta,\ x_2 = r\ \sin\theta$.

$$\Delta = \frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\frac{\partial^2}{\partial \theta^2}. \tag{4.15}$$

(b) Spherical polar coordinates on $\mathbb{R}^3$: $x_1 = \rho\ \sin\varphi\ \sin\theta,\ x_2 = \rho\ \sin\varphi\ \cos\theta,$ $x_3 = \rho\ \cos\varphi.$

$$\Delta = \frac{\partial^2}{\partial\rho^2} + \frac{2}{\rho}\frac{\partial}{\partial\rho} + \frac{1}{\rho^2 \sin\varphi}\Big(\frac{\partial^2}{\partial\theta^2} + \sin\varphi\frac{\partial^2}{\partial\varphi^2} + \cos\varphi\frac{\partial}{\partial\varphi}\Big). \tag{4.16}$$

(c) Spherical polar coordinates on $\mathbb{R}^n$: $x = r\omega,\ \omega \in S^{n-1}$.

$$\Delta = \frac{\partial^2}{\partial r^2} + \frac{n-1}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\Delta_S, \tag{4.17}$$

where Δ_S is the Laplace operator on the unit sphere S^{n-1}. (Compare (4.19) below.) (*Hint*: Express the Euclidean metric tensor $ds^2 = dx_1^2 + \cdots + dx_n^2$ in these coordinates.)

4. Let N be a Riemannian manifold, of dimension $n-1$. Denote by $C(N)$ the cone with base N, that is, the space $\mathbb{R}^+ \times N$, with Riemannian metric

$$g = dr^2 + r^2 g_N. \tag{4.18}$$

Show that the Laplace operator on $C(N)$ is of the form

$$\Delta = \frac{\partial^2}{\partial r^2} + \frac{n-1}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\Delta_N, \tag{4.19}$$

where Δ_N is the Laplace operator on the base N. Apply this to the expression of the Laplace operator Δ on $\mathbb{R}^n$, in polar coordinates, with $N = S^{n-1}$.

5. Show that, in local coordinates,

$$\Delta u = g^{jk}\,\partial_j\partial_k u - g^{jk}\Gamma^\ell{}_{jk}\,\partial_\ell u.$$

5. The wave equation on a product manifold and energy conservation

The analysis of vibrating membranes in Euclidean space has important extensions to studies of vibrating manifolds. We will start with a fairly general situation, specializing quickly to models that give rise to "the wave equation"

$$\frac{\partial^2 u}{\partial t^2} - \Delta u = 0, \tag{5.1}$$

for $u = u(t,x)$, a scalar function on $\mathbb{R} \times M$, where Δ is the Laplace operator on M defined in §4.

We consider vibrations of one manifold M within another, N. Suppose these manifolds are endowed with Riemannian metric tensors g and h, respectively. The vibration is described by a map

$$u : \mathbb{R} \times M \longrightarrow N. \tag{5.2}$$

In §1 we dealt with the special case where M is a bounded region in $\mathbb{R}^n$ and $N = \mathbb{R}^k$. Now we allow M to be a compact manifold with boundary. We again use a stationary action principle to produce equations governing the vibration. The appropriate expression for "kinetic energy" is

$$T(t) = \frac{1}{2}\int_M m(x)|u_t(t,x)|^2 \, dV, \tag{5.3}$$

where dV is the natural volume element on M and $m(x) > 0$ is a given "mass density." The velocity $u_t(t,x)$ takes values in T_yN, with $y = u(t,x)$, and the square-norm in the integrand in (5.3) is given by the metric tensor h;

$$|u_t|^2 = h(u, u_t, u_t) \tag{5.4}$$

if $h(y, v, w)$ denotes the inner product of v and w in T_yN.

The form that we will consider for the potential energy is the following generalization of (1.3):

$$V(t) = \int_M f(x, u(t,x), u_x(t,x)) \, dV, \tag{5.5}$$

where

$$u_x(t,x) \in \mathcal{L}(T_xM, T_{u(t,x)}N), \tag{5.6}$$

and f is a smooth, real-valued function defined on the bundle $\mathcal{L}$ over $M \times N$ with fiber over (x, y) given by $\mathcal{L}(T_xM, T_yN)$:

$$f = f(x, y, A), \quad A \in \mathcal{L}(T_xM, T_yN). \tag{5.7}$$

In particular, one has examples analogous to (1.42), that is,

$$f(x, y, A) = \Psi\big(\mathrm{Tr}\, g_1(A^*A), \dots, \mathrm{Tr}\, g_K(A^*A)\big), \tag{5.8}$$

where $A^* \in \mathcal{L}(T_yN, T_xM)$ is the adjoint of A, defined using the inner products on T_xM and T_yN defined by their Riemannian metrics. The $g_\ell(A^*A)$ are defined as described below (1.42). Many interesting cases of this sort arise naturally, including

$$f(x, y, A) = \mathrm{Tr}\, A^*A. \tag{5.9}$$

Applying the stationary action principle will yield for u a second-order system of PDE of a form that generalizes (1.12). We look here at the details for a special case.

Namely, take $N = \mathbb{R}$, and suppose $f(x, y, A)$ is independent of $y \in \mathbb{R}$. In other words, we consider a potential energy of the form

$$V(t) = \int_M f(x, u_x(t,x))\, dV, \tag{5.10}$$

where $u_x(t,x) \in T_x^*M$ and $f = f(x,\xi)$ is a smooth, real-valued function defined on T^*M, or perhaps on some open subset. In that case, the stationary condition for $(J_0 - J_1)(u) = \int_{t_0}^{t_1} \big[T(t) - V(t)\big]\, dt$ is derived from the following calculations. First, as in (1.8),

$$\frac{d}{ds} J_0(u + sv)\big|_{s=0} = -\iint m u_{tt} v \, dV\, dt, \tag{5.11}$$

provided $v \in C_0^\infty(I \times \overset{\circ}{M})$, $I = (t_0, t_1)$. Here $\overset{\circ}{M}$ denotes the interior of M. Furthermore, for such v,

$$\frac{d}{ds} J_1(u + sv)\big|_{s=0} = \iint f_\xi(x, u_x) \cdot v_x \, dV\, dt, \tag{5.12}$$

where, in local coordinates,

$$f_\xi(x, u_x) \cdot v_x = \sum_j \frac{\partial f}{\partial \xi_j} \frac{\partial v}{\partial x_j}. \tag{5.13}$$

If v is supported in a coordinate patch, in which $dV = \sqrt{g}dx$, we can integrate by parts and write

$$\frac{d}{ds} J_1(u+sv)\big|_{s=0} = -\iint \sum_j g^{-1/2}\, \partial_{x_j}\big(g^{1/2} f_{\xi_j}(x, u_x)\big) v \sqrt{g}\, dx\, dt. \tag{5.14}$$

Thus we get the following PDE for u, in a local coordinate system:

$$m u_{tt} - g^{-1/2} \partial_{x_j}\big(g^{1/2} f_{\xi_j}(x, u_x)\big) = 0, \tag{5.15}$$

using the summation convention. Written out more fully, this is

(5.16)

$$m u_{tt} - \Big[f_{\xi_j \xi_k}(x, u_x) u_{x_j x_k} + f_{\xi_j x_j}(x, u_x) + \frac{1}{2} g^{-1} (\partial_{x_j} g) f_{\xi_j}(x, u_x) \Big] = 0.$$

An invariant formulation of this PDE is given in the exercises.

The choice of $f(x, \xi)$ that produces a wave equation of the form (5.1) is that of a constant times the Riemannian metric on covariant vectors:

$$f(x, \xi) = \sigma\, g(x, \xi, \xi) = \sigma\, g^{jk}\, \xi_j \xi_k, \tag{5.17}$$

with σ a positive constant. In that case, (5.15) becomes

$$m u_{tt} - 2\sigma \Delta u = 0 \tag{5.18}$$

in view of the local coordinate formula

$$\Delta u = g^{-1/2}\, \partial_j (g^{1/2} g^{jk}\, \partial_k u) \tag{5.19}$$

derived in §4. If m is a constant, this is of the form (5.1) provided $2\sigma = m$, which could be arranged by a rescaling of the t-variable.

Other choices of $f(x, \xi)$ arise naturally in the study of vibrating membranes, choices that lead to nonlinear PDE. We will return to this in Chap. 16, but for now we concentrate on the linear case (5.18), until the very end of this section where we make a few brief comments on nonlinear problems.

Let us redo the calculation of the variation of $J_1(u)$ in an invariant fashion, when $f(x, \xi)$ is given by (5.17), so

$$J_1(u) = \sigma \iint\limits_{I \times M} |d_x u|^2\, dV\, dt. \tag{5.20}$$

We have, for $v \in C_0^\infty(I \times \overset{\circ}{M})$,

$$\frac{d}{ds} J_1(u + sv)\big|_{s=0} = 2\sigma \iint \langle d_x u, d_x v\rangle\, dV\, dt, \tag{5.21}$$

and Green's formula (4.8) shows that this is equal to

$$-2\sigma \iint (\Delta u) v\, dV\, dt, \tag{5.22}$$

since the boundary integral vanishes in this case. Again the stationary condition for $(J_0 - J_1)(u)$ is seen to be the wave equation (5.18).

As in(1.26), it is typical to specify *initial conditions*, of the form

$$u(0, x) = f(x), \quad u_t(0, x) = g(x). \tag{5.23}$$

If $\partial M \neq \emptyset$, we also need to specify a *boundary condition* for u. One typical condition is

$$u(t, x) = 0, \quad \text{for } x \in \partial M. \tag{5.24}$$

This is known as the *Dirichlet* boundary condition for u. It models a vibrating drum head that is firmly attached to its boundary. Tying down the boundary provides a justification for considering only variations v that vanish on $I \times \partial M$ in the specification of the stationary condition above. Another natural physical problem is to describe vibrations of M when the boundary is allowed to move freely. Then we should allow any $v \in C^\infty(\overline{I} \times M)$ that vanishes at $t = t_0$ and $t = t_1$, as a variation. The formula (5.11) for the variation of $J_0(u)$ continues to hold, and so does (5.21), but an application of Green's formula to (5.21) now yields

$$(5.25)\quad \frac{d}{ds} J_1(u + sv)\big|_{s=0} = -2\sigma \iint\limits_{I \times M} (\Delta u) v \, dV \, dt + 2\sigma \iint\limits_{I \times \partial M} v \, \frac{\partial u}{\partial \nu} \, dS \, dt.$$

If we do apply this to the subclass of $v \in C_0^\infty(I \times \overset{\circ}{M})$, we see that the wave equation (5.18) must still be satisfied for u to be a stationary point. Now, granted that u satisfies (5.18), we hence have

$$(5.26)\qquad \frac{d}{ds} (J_0 - J_1)(u + sv)\big|_{s=0} = -2\sigma \iint\limits_{I \times \partial M} v \, \frac{\partial u}{\partial \nu} \, dS \, dt,$$

for all $v \in C^\infty(\overline{I} \times M)$ that vanish at $t = t_0$ and at $t = t_1$. This yields the following boundary condition for freely vibrating M:

$$(5.27)\qquad \frac{\partial u}{\partial \nu} = 0, \quad \text{for } x \in \partial M.$$

This is known as the *Neumann* boundary condition for u. Another situation it models is the propagation of small-amplitude sound waves in a region bounded by a hard wall.

Since we have introduced the kinetic energy and the potential energy, we should look at the total energy. In the case when (5.17) gives the potential energy, if we take $m = 1$ and $\sigma = 1/2$, the total energy is

$$(5.28)\qquad E(t) = \frac{1}{2} \int\limits_M \big[|u_t(t,x)|^2 + \langle d_x u, d_x u \rangle \big] \, dV(x).$$

We aim to establish the energy conservation law

$$(5.29)\qquad E(t) = \text{ const.}$$

whenever u is a sufficiently smooth solution to the wave equation (5.1), assuming that u satisfies either the Dirichlet condition (5.24) or the Neumann condition (5.27) on ∂M. In fact, we have

$$(5.30)\qquad \frac{dE}{dt} = \int\limits_M \big[u_t u_{tt} + \langle d_x u_t, d_x u \rangle \big] \, dV.$$

We want to factor u_t out of the integrand, so we integrate by parts the last term in (5.30), using Green's identity to get

$$\frac{dE}{dt} = \int_M u_t(u_{tt} - \Delta u)\, dV + \int_{\partial M} u_t \frac{\partial u}{\partial \nu}\, dS. \tag{5.31}$$

The right side of (5.31) vanishes provided u satisfies the wave equation and either the Dirichlet or Neumann boundary condition. This proves the energy conservation law (5.29), equivalent to

$$\int_M \big[|u_t(t,x)|^2 + \langle d_x u, d_x u\rangle\big]\, dV = \int_M \big[|g(x)|^2 + \langle d_x f, d_x f\rangle\big]\, dV, \tag{5.32}$$

given the initial conditions (5.23).

We continue briefly the discussion of stationary problems from the end of §1. These problems do not involve t-dependence, that is, they arise via describing critical points for a function

$$J(u) = \int_M f\big(x, u(x), u_x(x)\big)\, dV, \tag{5.33}$$

with

$$f = f(x, y, A), \quad A \in \mathcal{L}(T_x M, T_y N). \tag{5.34}$$

If $N = \mathbb{R}$ and $f(x, y, \xi) = f(x, \xi)$ is given by (5.17), then the PDE obtained as the stationary condition for $J(u)$ is

$$\Delta u = 0, \tag{5.35}$$

involving the Laplace operator (5.19). A typical boundary condition is the nonhomogeneous Dirichlet condition

$$u = \psi \ \text{ on } \partial M. \tag{5.36}$$

Another is the nonhomogeneous Neumann condition

$$\frac{\partial u}{\partial \nu} = \varphi \ \text{ on } \partial M. \tag{5.37}$$

These will be studied in Chap. 5.

There are also very important nonlinear problems arising from the problem of finding stationary points, particularly extrema, of (5.33). We mention in particular the choice (5.9) for $f(x, y, A)$, namely, Tr A^*A. Maps $u : M \to N$ critical for such $J(u)$ are called *harmonic maps*. In case $N = \mathbb{R}^k$, these are just functions whose components are harmonic in the sense of (5.35), but for a nonflat Riemannian manifold N, one gets a nonlinear problem. For example, as seen in Chap. 1, for $M = I \subset \mathbb{R}$, one gets the geodesic equation. Harmonic maps will be studied in Chap. 14, by variational methods, and in Chap. 15, via techniques involving nonlinear parabolic PDE.

Exercises

1. For $J_1(u) = \int_M f(x, u_x)dV$ as in (5.10), $f : T^*M \to \mathbb{R}$, demonstrate the invariant formula
$$\frac{d}{ds} J_1(u + sv)\big|_{s=0} = \int\limits_M \langle A_f(x, u_x), v_x\rangle \, dV,$$
where $A_f : T^*M \to TM$ is given by
$$A_f(x, \xi) = D\pi(x, \xi) H_f, \tag{5.38}$$
H_f being the Hamiltonian vector field of f, and $\pi : T^*M \to M$ the natural projection. For fixed t, $u_x = d_x u$ is a 1-form on M. Consequently, $A_f(x, u_x)$ is a vector field on M.
2. In the context of Exercise 1, show that the resulting PDE (5.15)has the invariant description
$$mu_{tt} - \operatorname{div} A_f(x, u_x) = 0. \tag{5.39}$$
Compare (1.72).
3. Show that (under an appropriate nondegeneracy hypothesis) maps of the form A_f invert Legendre transformations $\lambda : TM \to T^*M$, discussed in §12 of Chap. 1.
(*Hint*: Using (12.9)–(12.18), consider the Legendre transform associated to the function $F(x, v)$ on TM defined implicitly by
$$F\big(x, f_\xi(x, \xi)\big) = f(x, \xi) - \xi \cdot f_\xi(x, \xi)$$
or, in the notation used above,
$$F\big(A_f(x, \xi)\big) = f(x, \xi) - \langle A_f(x, \xi), \xi\rangle.)$$

6. Uniqueness and finite propagation speed

We study some properties of solutions to the wave equation on $\mathbb{R} \times M$:
$$u_{tt} - \Delta u = 0, \tag{6.1}$$
with initial conditions
$$u(0, x) = f(x), \quad u_t(0, x) = g(x), \tag{6.2}$$

and boundary condition either the Dirichlet condition or the Neumann condition, if $\partial M \neq \emptyset$. We leave aside for the present the issue of the existence of solutions, for arbitrarily given f and g. We examine the uniqueness; u is assumed sufficiently smooth. If u_1 and u_2 solve (6.1) with initial data $f_j,\ g_j$, then $u_1 - u_2$ solves (6.1) with initial data $f = f_1 - f_2,\ g = g_1 - g_2$. To establish uniqueness, it suffices to show that if $f = g = 0$, then the solution $u = 0$ for all t. But by energy conservation, we have, for all t,

$$\int_M \big[u_t^2 + \langle d_x u, d_x u\rangle\big]\, dV = \int_M \big[|g|^2 + \langle d_x f, d_x f\rangle\big]\, dV = 0. \tag{6.3}$$

Thus u is constant. Since $u(0,x) = 0$, we conclude that $u = 0$ everywhere. This establishes uniqueness.

A closer look at how Green's formula enters into this argument will produce both a generalization of the notion of energy conservation and a localization of this uniqueness theorem to a result implying finite propagation speed for solutions to the wave equation. Note that the identity (5.31) can be written as

$$E(t_2) - E(t_1) = \int_{t_1}^{t_2} \int_M u_t(u_{tt} - \Delta u)\, dV\, dt + \int_{t_1}^{t_2} \int_{\partial M} u_t \frac{\partial u}{\partial \nu}\, dS\, dt. \tag{6.4}$$

In particular, for u satisfying either the Dirichlet or Neumann condition on ∂M, with $\Omega = [t_1, t_2] \times M$, we have

$$\begin{aligned} &\int_\Omega u_t(u_{tt} - \Delta u)\, dV\, dt = \\ &\frac{1}{2} \int_{\{t=t_2\}} \big[|u_t|^2 + |d_x u|^2\big]\, dV - \frac{1}{2} \int_{\{t=t_1\}} \big[|u_t|^2 + |d_x u|^2\big]\, dV. \end{aligned} \tag{6.5}$$

Next we want to look at the left side of (6.5) when Ω is a more general sort of region in $\mathbb{R} \times M$ than a product region $[t_1, t_2] \times M$.

First, we assume for simplicity that Ω does not intersect $\mathbb{R} \times \partial M$. We suppose $\partial\Omega$ consists of two smooth surfaces, Σ_1 and Σ_2, as indicated in Fig. 6.1. We denote by Ω_t the intersection of Ω with $\{t\} \times M \subset \mathbb{R} \times M$. Now, making use of formula (2.19), we have

$$\begin{aligned} \int_\Omega u_t(u_{tt} - \Delta u)\, dV\, dt = &\int_\Omega \frac{\partial}{\partial t}\Big(\frac{1}{2} u_t^2\Big)\, dV\, dt + \int_\Omega \langle d_x u_t, d_x u\rangle\, dV\, dt \\ &- \int_\Omega \operatorname{div}_x(u_t \operatorname{grad}_x u)\, dV\, dt. \end{aligned} \tag{6.6}$$

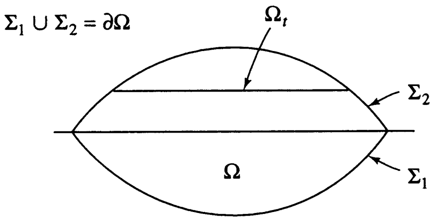

FIGURE 6.1 Spacelike Bounded Region

Note that

$$\langle d_x u_t, d_x u\rangle = \frac{1}{2}\,\frac{\partial}{\partial t}\langle d_x u, d_x u\rangle. \tag{6.7}$$

Applying the fundamental theorem of calculus to the first two integrals on the right side of (6.6), and the divergence theorem to the last integral, we get

$$\int_{\Omega} u_t(u_{tt}-\Delta u)\,dV\,dt = \frac{1}{2}\int_{\partial\Omega}\big[u_t^2+\langle d_x u, d_x u\rangle\big]\omega - \int\int_{\partial\Omega_t} u_t\frac{\partial u}{\partial \nu_x}\,dS_t\,dt. \tag{6.8}$$

Both integrals on the right side of (6.8) are integrals over $\partial\Omega$. Here ω is the volume form on M, thought of as an n-form on $\mathbb{R}\times M$, pulled back to $\partial\Omega$, and dS_t is the natural surface measure on $\partial\Omega_t$, thought of as a surface in M. We want to express both ω and $dS_t\;dt$ in terms of the natural surface measure on $\partial\Omega$, induced from the inclusion $\partial\Omega\subset\mathbb{R}\times M$, endowed with the natural product Riemannian metric. Indeed, we easily obtain

$$\omega = N_t\,dS, \quad dS_t\,dt = |N_x|\,dS, \tag{6.9}$$

where $N = (N_t, N_x)$ is the outward unit normal to $\partial\Omega\subset\mathbb{R}\times M$. Hence (6.8) becomes

$$\int_{\Omega} u_t(u_{tt}-\Delta u)\,dV\,dt = \frac{1}{2}\int_{\partial\Omega}\Big\{\big[u_t^2+|d_x u|^2\big]N_t - 2u_t\frac{\partial u}{\partial\nu_x}|N_x|\Big\}\,dS. \tag{6.10}$$

Thus, if u satisfies the wave equation in Ω, we see that

$$\begin{aligned}&\int_{\Sigma_2}\Big\{\big[u_t^2+|d_x u|^2\big]|N_t| - 2u_t\frac{\partial u}{\partial\nu_x}|N_x|\Big\}\,dS\\ &\quad=\int_{\Sigma_1}\Big\{\big[u_t^2+|d_x u|^2\big]|N_t| + 2u_t\frac{\partial u}{\partial\nu_x}|N_x|\Big\}\,dS.\end{aligned} \tag{6.11}$$

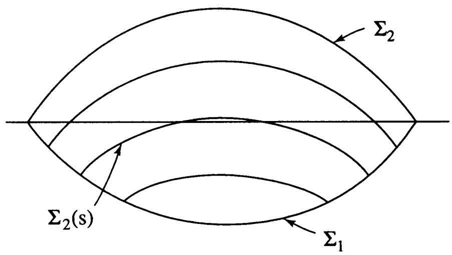

FIGURE 6.2 Spacelike Sweep

This is a useful "energy identity" provided the integrands are positive-definite quadratic forms in $du = (u_t, d_x u)$. Note that Cauchy's inequality implies

$$2\left|u_t \frac{\partial u}{\partial \nu_x}\right| \le u_t^2 + |d_x u|^2. \tag{6.12}$$

Thus the integrands have the desired property, provided

$$|N_x| < |N_t|. \tag{6.13}$$

Definition. *A surface $\Sigma \subset \mathbb{R} \times M$ is called spacelike provided its normal $N = (N_t, N_x)$ satisfies (6.13). A vector satisfying (6.13) is called timelike.*

Clearly any surface $t =$ const. is spacelike, as is a small perturbation of such a surface. Suppose $\Omega \subset \mathbb{R} \times M$ is bounded by spacelike surfaces Σ_1 and Σ_2 and furthermore is swept out by spacelike surfaces $\Sigma_2(s)$, as in Fig. 6.2. We call Ω a domain of influence for its lower boundary Σ_1.

Theorem 6.1. *Suppose $\Omega \subset \mathbb{R} \times M$ is a domain of influence for its lower boundary Σ_1. If u solves the wave equation $u_{tt} - \Delta u = 0$ on $\mathbb{R} \times M$, and if u and $du = (u_t, d_x u)$ vanish on Σ_1, then u vanishes throughout Ω.*

Proof. The energy identity implies that du vanishes on each $\Sigma_2(s)$; hence du vanishes on Ω, so u is constant on Ω. Since $u = 0$ on Σ_1, this constant is 0.

One interpretation of this theorem is that it shows that signals propagate at speed at most 1. In other words, in the special case $\Sigma_1 = \{t = 0\}$, if $u(0, x) = f(x)$ and $u_t(0, x) = g(x)$ vanish on some open set $\mathcal{O} \subset M$, then the solution to the wave equation vanishes on $\{(t, x) : x \in \mathcal{O},\ \mathrm{dist}(x, \partial\mathcal{O}) > |t|\}$.

A slight variation of the argument above treats the case when $\partial\Omega$ consists of three parts, Σ_1 and Σ_2, both spacelike as above, and a part in $\mathbb{R} \times \partial M$, provided the solution u to $u_{tt} - \Delta u = 0$ satisfies the Dirichlet or Neumann boundary condition.

Exercises

1. Use(1.24)–(1.28) to write out the explicit solution to the initial value problem (6.1)–(6.2) in case $\Delta = \partial^2/\partial x^2$ on $\mathbb{R}$, and explicitly observe finite propagation speed in this case.
2. Extend the finite propagation speed argument of Theorem 6.1 to the case where M has a boundary, on which either the Dirichlet or Neumann boundary condition is imposed.
3. Consider the equations of linear elasticity, derived in (1.59), $Lu = 0$, where

$$Lu = mu_{tt} - \mu\Delta u - (\lambda + \mu) \operatorname{grad} \operatorname{div} u.$$

Suppose $\mu > 0, \lambda + 2\mu > 0, m > 0$. For each $(t, x) \in \mathbb{R} \times M$, $u(t, x) \in T_x M$. Take $M = \mathbb{R}^n$. Let Ω be a region in $\mathbb{R} \times M$ of the form depicted in Fig. 6.1. Perform an integration by parts of

$$\int_{\Omega} u_t \cdot Lu \, dV \, dt,$$

along the lines of (6.6)–(6.10), to derive an identity similar to (6.11). What geometrical conditions should be placed on Σ_1 and Σ_2, replacing the "spacelike" condition (6.13), in order to ensure that the resulting integrands are positive-definite quadratic forms in $\nabla u = (u_t, \nabla_x u)$? Derive a finite propagation speed result.

7. Lorentz manifolds and stress-energy tensors

The analysis of the wave equation in the last section made strong use of the fact that we were working with $\partial^2/\partial t^2 - \Delta$ on a product $\mathbb{R} \times M$. We will take a deeper look at the notion of energy, which will produce concepts that are important in the study of the wave equation on more general Lorentz manifolds.

For starters, we will stick with the product case $\mathbb{R} \times M$, M a Riemannian manifold. This has a natural structure of a Lorentz manifold, with metric

$$h = -dt^2 + g. \tag{7.1}$$

Contrast this with the Riemannian metric $dt^2 + g$ on $\mathbb{R} \times M$ we considered in the last section. In coordinates, h_{jk} has the form

$$\left(h_{jk}\right) = \begin{pmatrix} -1 & 0 \\ 0 & g_{\mu\nu} \end{pmatrix}. \tag{7.2}$$

The *stress-energy tensor* T associated with u is supposed to be a symmetric, second order tensor such that, if Z is a unit timelike vector (representing the "world line" of an observer), then $T(Z, Z)$ gives the observed energy density. The energy density $(1/2)u_t^2 + (1/2)\langle d_x u, d_x u\rangle$ encountered before specifies

$$T_{00} = \frac{1}{2}u_t^2 + \frac{1}{2}\langle d_x u, d_x u\rangle = u_t^2 + \frac{1}{2}\langle du, du\rangle, \tag{7.3}$$

where

$$\langle du, du\rangle = h^{jk}\,\partial_j u\,\partial_k u \tag{7.4}$$

is the Lorentz square-length of du. If we expect that T is constructed in a "natural" manner from du and the metric tensor h, we are led to require

$$T(Z,Z) = \langle Z, du\rangle^2 + \frac{1}{2}\langle du, du\rangle \ \text{ whenever } \ \langle Z, Z\rangle = -1.$$

If $\langle Z, Z\rangle = -z^2$, this leads to $T(Z,Z) = \langle Z, du\rangle^2 - (1/2)\langle du, du\rangle\langle Z, Z\rangle$, and polarizing this identity gives

$$T(Z,W) = \langle Z, du\rangle\langle W, du\rangle - \frac{1}{2}\langle du, du\rangle\langle Z, W\rangle. \tag{7.5}$$

This should hold for all vectors Z, W. Equivalently, we write

$$T = du \otimes du - \frac{1}{2}\langle du, du\rangle h. \tag{7.6}$$

We call(7.6)the stress-energy tensor associated to a wave $u = u(t,x)$. See the exercises for more on the construction of T.

More generally, let Ω be any Lorentz manifold, with metric tensor, of signature $(n,1)$, denoted h. The "Laplacian" in this metric is defined by

$$\Box u = |h|^{-1/2}\,\partial_j(h^{jk}|h|^{1/2}\,\partial_k u) = h^{jk}u_{;j;k}, \tag{7.7}$$

in analogy with the formula for the Laplace operator on a Riemannian manifold. Here, $|h| = |\det(h_{jk})|$. The wave equation on a general Lorentz manifold is

$$\Box u = 0. \tag{7.8}$$

In this more general context, it is still meaningful to assign to u the tensor T, defined by (7.5) and (7.6). We continue to call T the stress-energy tensor. We have the following important result.

Proposition 7.1. *For a solution to (7.8) on a general Lorentz manifold Ω, the stress-energy tensor has vanishing divergence, that is,*

$$T^{jk}{}_{;k} = 0. \tag{7.9}$$

More generally, for any u,

$$T^{jk}{}_{;k} = u^{;j}\,\Box u. \tag{7.10}$$

Proof. This is a straightforward calculation. We have

$$T^{jk} = u^{;j}u^{;k} - \frac{1}{2}h^{jk}h^{\mu\nu}u_{;\mu}u_{;\nu}, \tag{7.11}$$

where $u^{;j} = h^{jk}u_{;k}$ denotes the gradient. Hence, using $h^{jk}{}_{;\ell} = 0$, we obtain

$$\begin{aligned} T^{jk}{}_{;k} &= u^{;j}{}_{;k}u^{;k} + u^{;j}u^{;k}{}_{;k} - \frac{1}{2}h^{jk}h^{\mu\nu}u_{;\mu;k}u_{;\nu} - \frac{1}{2}h^{jk}h^{\mu\nu}u_{;\mu}u_{;\nu;k} \\ &= u^{;j}\square u + u^{;j}{}_{;k}u^{;k} - h^{jk}h^{\mu\nu}u_{;\mu;k}u_{;\nu} \\ &= u^{;j}\square u + u^{;j}{}_{;k}u^{;k} - u_{;\mu}{}^{;j}u^{;\mu}. \end{aligned}$$

Since, as we have seen, $u_{;j;k} = u_{;k;j}$, we obtain (7.10), and the proposition follows.

We have seen that the divergence theorem applies to reduce the integral $\int_\Omega (\operatorname{div} X)\, dV$ to a boundary integral, when X is a vector field; in particular, when X is a divergence-free vector field, it yields that a certain boundary integral is zero or, equivalently, that integrals over two parts of $\partial\Omega$ are equal in magnitude. However, T is not a divergence-free vector field; it is a second-order tensor field. In general vanishing of div T will not lead to integral conservation laws. It will, however, in the following case.

Suppose a Lorentz manifold Ω has a timelike Killing field Z, that is, a timelike vector field whose flow preserves the metric tensor h. As derived in the Riemannian case, the condition for the metric to be preserved is

$$Z_{j;k} + Z_{k;j} = 0, \quad Z_j = h_{jk}Z^k. \tag{7.12}$$

Here, "timelike" means that $h(Z, Z) < 0$. This means Z lies inside the light cone determined by the Lorentz metric.

Lemma 7.2. *If T^{jk} is divergence free and Z^k is a Killing field, then*

$$X^j = T^{jk}Z_k \text{ is divergence free.} \tag{7.13}$$

Proof. We have

$$X^j{}_{;j} = T^{jk}{}_{;j}Z_k + T^{jk}Z_{k;j}.$$

Now the symmetry of T^{jk} implies $T^{jk}{}_{;j} = 0$ and

$$T^{jk}Z_{k;j} = \frac{1}{2}T^{jk}(Z_{k;j} + Z_{j;k}) = 0,$$

assuming (7.12) holds. This proves the lemma.

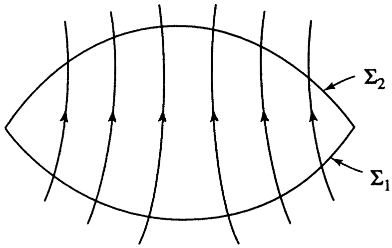

FIGURE 7.1 Timelike Curves

We denote the vector(7.13) by

$$X = \widetilde{T}Z. \tag{7.14}$$

Suppose $\mathcal{O}$ is a region in the Lorentz manifold Ω, bounded by two surfaces Σ_1 and Σ_2, as in Fig. 7.1.

By (2.14)), we have

$$\begin{aligned} 0 = \int_{\mathcal{O}} (\operatorname{div} \widetilde{T}Z)\, dV &= \int_{\Sigma_1 \cup \Sigma_2} \omega \rfloor (\widetilde{T}Z) \\ &= \int_{\Sigma_1} \langle \widetilde{T}Z, \nu_1 \rangle \, dS - \int_{\Sigma_2} \langle \widetilde{T}Z, \nu_2 \rangle \, dS, \end{aligned} \tag{7.15}$$

where ν_j is the unit vector, normal to Σ_j, with respect to the Lorentz metric h, pointing in the same "forward" direction as Z. The last identity in (7.15) holds in analogy with (2.15). We make the hypothesis as before, that Σ_1 and Σ_2 are spacelike (i.e., ν_j are timelike), so it makes sense to specify that they lie inside the forward light cone. Equation (7.15) is equivalent to

$$\int_{\Sigma_2} T(Z, \nu_2)\, dS = \int_{\Sigma_1} T(Z, \nu_1)\, dS. \tag{7.16}$$

The volume element dS on Σ_j is determined here by the Riemannian metric on Σ_j, induced by restricting the Lorentz metric h to tangent vectors to Σ_j.

Again we seek to guarantee that the integrand in (7.16), which is a quadratic form in du for T given by (7.5), is positive-definite. In order to check this at a point $p_0 \in \partial\mathcal{O}$, choose a coordinate system such that

$$\big(h_{jk}(p_0)\big) = \begin{pmatrix} -1 & 0 \\ 0 & I \end{pmatrix}, \quad \nu(p_0) = (1, 0, \dots, 0)^t \quad (\nu = \nu_1 \text{ or } \nu_2), \tag{7.17}$$

which is always possible. Suppose $Z(p_0) = (Z^0, Z^1, \dots, Z^n)$. The condition that $Z(p_0)$ belong to the forward light cone is

$$Z^0 > 0, \quad (Z^0)^2 > (Z^1)^2 + \cdots + (Z^n)^2. \tag{7.18}$$

Now, if we set $M = \widetilde{T}\nu$, then, at p_0,

$$M^0 = -\frac{1}{2}\big[(\partial_0 u)^2 + (\partial_1 u)^2 + \cdots + (\partial_n u)^2\big], \quad M^j = (\partial_0 u)(\partial_j u), \tag{7.19}$$

if T is given by (7.5). Consequently, at p_0,

$$\begin{aligned} T(Z,\nu) &= \langle Z, M\rangle = -Z^0 M^0 + \sum_{j=1}^n Z^j M^j \\ &= \frac{1}{2} Z^0\big[(\partial_0 u)^2 + \cdots + (\partial_n u)^2\big] + \sum_{j=1}^n Z^j(\partial_0 u)(\partial_j u). \end{aligned} \tag{7.20}$$

The positive definiteness of this quadratic form in $(\partial_0 u, \dots, \partial_n u)$ follows immediately from Cauchy's inequality, granted (7.18). This definiteness calculation does not use the hypothesis that Z is a Killing field, of course. For positive definiteness of $T(Z,\nu)$ in du, it suffices that Z and ν both be nonzero timelike vectors inside the forward light cone.

In order to emphasize that the dependence of $T(Z,\nu)$ on du has fundamental significance, we adopt the following notation. Set

$$\begin{aligned} E_{Z,\nu}(du) &= T(Z,\nu) \\ &= \Big(du \otimes du - \frac{1}{2}\langle du, du\rangle h\Big)(Z,\nu) \\ &= \langle du, Z\rangle\langle du,\nu\rangle - \tfrac{1}{2}\langle Z,\nu\rangle\langle du, du\rangle. \end{aligned} \tag{7.21}$$

The calculation above establishes the following result.

Lemma 7.3. *If Z and ν are nonzero timelike vectors pointing inside the forward light cone, then*

$$E_{Z,\nu}(du) \text{ is positive-definite in } du.$$

Note that the identity (7.16) is

$$\int_{\Sigma_2} E_{Z,\nu_2}(du)\, dS = \int_{\Sigma_1} E_{Z,\nu_1}(du)\, dS. \tag{7.22}$$

It follows that if $\mathcal{O}$, as in Fig. 7.1, is swept out by spacelike surfaces, as in Fig. 7.2, then the same argument as given in §6 leads to the uniqueness result: $\Box u = 0$ in $\mathcal{O}$, u and $du = 0$ on Σ_1 imply $u = 0$ in $\mathcal{O}$, provided Ω has a timelike Killing field Z. This gives finite propagation speed for solutions to the wave equation on such a Lorentz manifold.

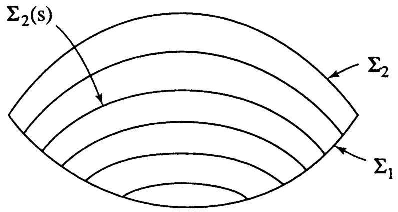

FIGURE 7.2 Spacelike Surfaces

If a Lorentz manifold Ω has no timelike Killing field, which is typical, then natural energy identities such as (7.22) do not arise. However, there are *inequalities* involving the stress-energy tensor, that are powerful enough to imply the local uniqueness (finite propagation speed) of solutions to the wave equation $\Box u = 0$ on a general Lorentz manifold. In the next section we will establish this as a special case of a more general result on hyperbolic equations.

Exercises

1. If M is a Lorentz manifold, $S \subset M$ a hypersurface (codimension 1), show that S is spacelike if and only if the metric tensor restricted to S is positive-definite. In the product case (7.1), show that the definitions of "spacelike" given in this section and the previous one are equivalent.
2. On $\mathbb{R}^{n+1}$, with coordinates $(x_0, \dots, x_n)$, place the Lorentz inner product

$$\langle u, v \rangle = -u_0 v_0 + u_1 v_1 + \cdots + u_n v_n.$$

Show that $A : \mathbb{R}^{n+1} \to \mathbb{R}^{n+1}$, defined by

$$A(u_0, u_1, u_2, \dots, u_n) = (u_1, u_0, u_2, \dots, u_n)$$

is *skew-adjoint* for the Lorentz metric (i.e., $\langle Au, v \rangle = -\langle u, Av \rangle$), and hence the group $\mathcal{F}(t) = e^{tA}$ preserves the Lorentz metric.
3. Consider the hyperboloids

$$M = M_s = \{x \in \mathbb{R}^{n+1} : \langle x, x \rangle = s\}.$$

Show that M_s is spacelike if and only if $s < 0$.
4. If $s > 0$ and M_s is as in Exercise 3, show that M_s gets a Lorentz metric, induced from $\mathbb{R}^{n+1}$. Show that the group $\mathcal{F}(t)$ of Exercise 2 leaves M_s invariant and its generator is a timelike Killing field on M_s.
5. We consider a general approach to constructing a second-order tensor of the form

$$T^{jk} = A^{jk\ell m} u_{;\ell} u_{;m},$$

where $A^{jk\ell m}$ is a tensor field of type $(4, 0)$, such that the conclusion (7.10) of Proposition 7.1 holds. Let us assume that $\nabla A = 0$. Show that

$$T^{jk}{}_{;k} = B^{jk\ell m} u_{;k;\ell} u_{;m},$$

where

$$B = P^{23} P^{34} A.$$

Here, $P^{\mu\nu}$ denotes the operation on tensors of type $(4,0)$ of symmetrizing with respect to the μth and νth indices, for example, $(P^{23}C)^{jk\ell m} = (1/2)[C^{jk\ell m} + C^{j\ell km}]$. Consequently, (8.10) holds provided

$$P^{23} P^{34} A = H, \quad H^{jk\ell m} = h^{jm} h^{k\ell}.$$

6. Show that $P^{\mu\nu}$ are all projections of the same rank and H belongs to the range of P^{23}. Show that Ker $P^{23} \cap R(P^{34}) = 0$ and hence

$$P^{23} : R(P^{34}) \longrightarrow R(P^{23}) \text{ is an isomorphism.}$$

(*Hint*: If $B \in$ Ker $P^{23} \cap R(P^{34})$, show that $B^{jk\ell m} = -B^{jmk\ell}$. $(k\ \ell\ m) \mapsto (m\ k\ \ell)$ is a cyclic permutation of order 3, so apply this transformation three times.)
7. Deduce that the equation $P^{23} P^{34} A = H$ has a solution A, given uniquely, mod Ker P^{34}, and hence that the tensor $T^{jk} = A^{jk\ell m} u_{;\ell} u_{;m}$ is uniquely determined by the conditions set in Exercise 5.
8. Show that, for general smooth scalar u, with T defined by (7.6), then

$$\operatorname{div} \widetilde{T} Z = (Zu)\Box u + \langle T, \operatorname{Def}(Z)\rangle, \tag{7.23}$$

where $\operatorname{Def}(Z)$ is the deformation tensor of Z, with components $(1/2)(Z_{j;k} + Z_{k;j})$ and $\langle T, V\rangle = T^{jk} V_{jk}$. This implies Lemma 7.2. Show that (7.23) follows from the general identity

$$\operatorname{div}(\widetilde{T} Z) = \langle Z, \operatorname{div} T\rangle + \langle T, \operatorname{Def} Z\rangle. \tag{7.24}$$

8. More general hyperbolic equations; energy estimates

In this section we derive estimates for a solution to a nonhomogeneous hyperbolic equation of the form

$$Lu = f \ \text{ in } \Omega, \tag{8.1}$$

where L is given in local coordinates by

$$Lu = h^{jk}\, \partial_j \partial_k u + b^j(x)\, \partial_j u + c(x) u. \tag{8.2}$$

By definition, to say L is hyperbolic is to say that (h^{jk}) is a symmetric matrix of signature $(n, 1)$, if $\dim \Omega = n + 1$. One can then use the inverse matrix (h_{jk}) to define a Lorentz metric on Ω, and in view of the formula (7.7), we can write (8.2) as

$$Lu = \Box u + Xu, \tag{8.3}$$

for some first-order differential operator X on Ω.

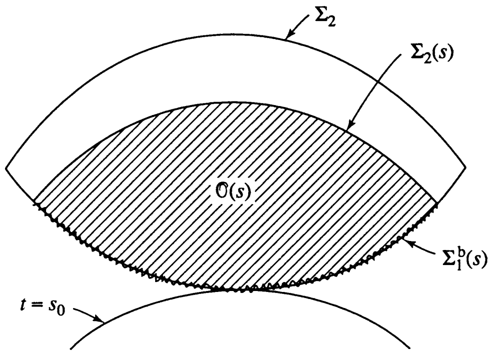

FIGURE 8.1 Spacelike Bounded Regions

Suppose $\mathcal{O} \subset \Omega$ is bounded by two surfaces Σ_1 and Σ_2, both spacelike. As at the end of §7, we suppose that $\mathcal{O}$ is swept out by spacelike surfaces. Specifically, we suppose that there is a smooth function on a neighborhood of $\overline{\mathcal{O}}$, which in fact we denote by t, such that dt is timelike, and set

$$\mathcal{O}(s) = \overline{\mathcal{O}} \cap \{t \le s\}, \quad \Sigma_2(s) = \overline{\mathcal{O}} \cap \{t = s\}.$$

We suppose $\mathcal{O}$ is swept out by $\Sigma_2(s)$, $s_0 \le s \le s_1$, as illustrated in Fig. 8.1, with $\Sigma_2 = \Sigma_2(s_1)$. Also set

$$\Sigma_1^b(s) = \Sigma_1 \cap \{t \le s\}.$$

As in (7.15), the divergence theorem implies

$$\int_{\Sigma_2(s)} E_{Z,\nu_2}(du)\, dS = \int_{\Sigma_1^b(s)} E_{Z,\nu_1}(du)\, dS - \int_{\mathcal{O}(s)} (\operatorname{div} \widetilde{T}Z)\, dV, \tag{8.4}$$

where $E_{Z,\nu}(du)$ is defined by (7.21) and T by (7.5), though at this point it is not physically meaningful in general to think of T as the stress-energy tensor. Here ν_1 is the forward-pointing unit normal to Σ_1, with respect to the Lorentz metric, and ν_2 is the normalization of grad t, the vector field obtained from dt via the Lorentz metric. Z is any timelike vector field; we will set $Z = \nu_2$. Note that Lemma 7.3 applies to the integrands $E_{Z,\nu_j}(du)$.

We no longer have $\operatorname{div} \widetilde{T}Z = 0$, but we can *estimate* this quantity, as follows. First,

$$\operatorname{div} \widetilde{T}Z = T^{jk}{}_{;k} h_{j\ell} Z^\ell + T^{jk} h_{j\ell} Z^\ell{}_{;k} = \langle \operatorname{div} T, Z\rangle + \langle T, \nabla Z\rangle. \tag{8.5}$$

The term $\langle T, \nabla Z\rangle$ is a quadratic form in du, and hence, by Lemma (7.7), we have an estimate

$$|\langle T, \nabla Z\rangle| \le K\, E_{Z,Z}(du). \tag{8.6}$$

As for the first term on the right side of (8.5), (7.10) implies

$$\operatorname{div} T = (\operatorname{grad} u)\Box u. \tag{8.7}$$

If u satisfies $Lu = f$, this implies

$$\operatorname{div} T = (\operatorname{grad} u)(f - Xu). \tag{8.8}$$

Cauchy's inequality together with Lemma 7.3 gives an estimate

$$|\langle \operatorname{div} T, Z\rangle| \le K\ E_{Z,Z}(du) + K|u|^2 + K|f|^2. \tag{8.9}$$

Consequently, (8.4) yields the estimate

$$\begin{aligned} &\int_{\Sigma_2(s)} E_{Z,Z}(du)\, dS \le \\ &\int_{\Sigma_1^b(s)} E_{Z,\nu_1}(du)\, dS + K \int_{\mathcal{O}(s)} \left[2E_{Z,Z}(du) + |u|^2 + |f|^2\right] dV. \end{aligned} \tag{8.10}$$

Suppose that u satisfies the following initial conditions on Σ_1:

$$u = g, \quad du = \omega \ \text{ on } \Sigma_1. \tag{8.11}$$

We want to estimate the left side of (8.10) in terms of f, g, and ω. Our first goal will be to deive a variant of (8.10) without the $|u|^2$ term. We can work on the term $\int_{\mathcal{O}(s)} |u|^2\, dV$ on the right side of (8.10) as follows. An easy consequence of the fundamental theorem of calculus, Cauchy's inequality, and Lemma 7.3 gives

$$\int_{\mathcal{O}(s)} |u|^2\, dV \le C \int_{\Sigma_1^b(s)} |g|^2\, dS + C \int_{\mathcal{O}(s)} E_{Z,Z}(du)\, dV, \tag{8.12}$$

which can be applied to (8.10).

At this point, it is convenient to set

$$E(s) = \int_{\mathcal{O}(s)} E_{Z,Z}(du)\, dV. \tag{8.13}$$

We will want to estimate the rate of change of $E(s)$. Clearly,

$$\frac{dE}{ds} \le C \int_{\Sigma_2(s)} E_{Z,Z}(du)\, dS, \tag{8.14}$$

and hence, by (8.10)–(8.12), we have an estimate of the form

$$\frac{dE}{ds} \le CE(s) + F(s), \tag{8.15}$$

where

$$F(s) = C \int_{\Sigma_1} \big[E_{Z,Z}(\omega) + |g|^2\big]\, dS + C \int_{\mathcal{O}(s)} |f|^2 \, dV. \tag{8.16}$$

Note that (8.15) is equivalent to

$$\frac{d}{ds}\big(e^{-Cs} E(s)\big) \le e^{-Cs} F(s), \tag{8.17}$$

and since $E(s_0) = 0$, we have

$$e^{-Cs} E(s) \le \int_{s_0}^{s} e^{-Cr} F(r)\, dr. \tag{8.18}$$

In view of (8.16), this establishes the following "energy estimate."

Proposition 8.1. *If u solves the hyperbolic equation $Lu = f$ of the form (8.3), with initial data (8.11) on Σ_1, and if $\mathcal{O}(s)$ satisfies the geometrical hypotheses made above and illustrated in Fig. 8.1, then*

$$\int_{\mathcal{O}(s)} E_{Z,Z}(du)\, dV \le C(s - s_0) \int_{\Sigma_1} \big[|g|^2 + |\omega|^2\big]\, dS + C \int_{\mathcal{O}(s)} |f|^2 \, dV, \tag{8.19}$$

for $s \in [s_0, s_1]$.

In particular, if g and ω vanish on Σ_1 and f vanishes on $\mathcal{O}$, then (8.19) implies $du = 0$ on $\mathcal{O}$, so u is constant on $\mathcal{O}$, that constant being $g = 0$. This gives the local uniqueness (finite propagation speed) for solutions to the homogeneous hyperbolic equation $Lu = 0$, extending the result of §7.

We note that, using (8.10) and (8.12), we deduce from (8.19) that

$$\int_{\Sigma_2(s)} E_{Z,Z}(du)\, dS \le C \int_{\Sigma_1} \big[|g|^2 + |\omega|^2\big]\, dS + C \int_{\mathcal{O}(s)} |f|^2 \, dV. \tag{8.20}$$

Exercises

1. Prove the estimate

$$(1-\varepsilon)\int_0^1 |u(s)|^2\,ds \le |u(0)|^2 + C_\varepsilon \int_0^1 |u'(s)|^2\,ds.$$

   What is the best value of C_ε that will work?
2. Give a detailed proof of the estimate (8.12).
3. Sharpen the estimate (8.19) to

$$\int_{\mathcal{O}(s)} E_{Z,Z}(du)\,dV \le C(s-s_0)\int_{\Sigma_1}\big[|g|^2+|\omega|^2\big]\,dS + C(s-s_0)\int_{\mathcal{O}(s)} |f|^2\,dV, \tag{8.21}$$

   under the hypotheses of Proposition 8.1. (*Hint*: Use (8.18) more carefully.)
4. Work out generalizations of the energy estimates (8.10)–(8.19) when u satisfies the *semilinear* PDE

$$\Box u = f(x,u,du). \tag{8.22}$$

   Formulate and prove a finite propagation speed result in this case.
   (*Hint*: Given solutions u_1 and u_2 to (8.22), derive a *linear* PDE for $w = u_1 - u_2$, to get the finite propagation speed result.)

9. The symbol of a differential operator and a general Green–Stokes formula

Let P be a differential operator of order m on a manifold M; P could operate on sections of a vector bundle. In local coordinates, P has the form

$$Pu(x) = \sum_{|\alpha|\le m} p_\alpha(x) D^\alpha u(x), \tag{9.1}$$

where $D^\alpha = D_1^{\alpha_1}\cdots D_n^{\alpha_n},\ D_j = (1/i)\,\partial/\partial x_j$. The coefficients $p_\alpha(x)$ could be matrix valued. The homogeneous polynomial in $\xi \in \mathbb{R}^n$ $(n = \dim M)$,

$$p_m(x,\xi) = \sum_{|\alpha|=m} p_\alpha(x)\xi^\alpha, \tag{9.2}$$

is called the *principal symbol* (or just the symbol) of P. We want to give an intrinsic characterization, which will show that $p_m(x,\xi)$ is well defined on the cotangent bundle of M. For a smooth function ψ, a simple calculation, using the product rule and chain rule of differentiation, gives

$$P\big(u(x)e^{i\lambda\psi}\big) = \big[p_m(x,d\psi)u(x)\lambda^m + r(x,\lambda)\big]e^{i\lambda\psi}, \tag{9.3}$$

where $r(x,\lambda)$ is a polynomial of degree $\le m-1$ in λ. In (9.3), $p_m(x,d\psi)$ is evaluated by substituting $\xi = (\partial\psi/\partial x_1,\dots,\partial\psi/\partial x_n)$ into (9.2). Thus the formula

$$p_m(x,d\psi)u(x) = \lim_{\lambda\to\infty} \lambda^{-m} e^{-i\lambda\psi} P\big(u(x)e^{i\lambda\psi}\big) \tag{9.4}$$

provides an intrinsic characteristization of the symbol of P as a function on T^*M. We also use the notation

$$\sigma_P(x,\xi) = p_m(x,\xi). \tag{9.5}$$

If

$$P : C^\infty(M,E_0) \longrightarrow C^\infty(M,E_1), \tag{9.6}$$

where E_0 and E_1 are smooth vector bundles over M, then, for each $(x,\xi) \in T^*M$,

$$p_m(x,\xi) : E_{0x} \longrightarrow E_{1x} \tag{9.7}$$

is a linear map between fibers. It is easy to verify that if P_2 is another differential operator, mapping $C^\infty(M,E_1)$ to $C^\infty(M,E_2)$, then

$$\sigma_{P_2P}(x,\xi) = \sigma_{P_2}(x,\xi)\sigma_P(x,\xi). \tag{9.8}$$

If M has a Riemannian metric, and the vector bundles E_j have metrics, then the formal adjoint P^t of a differential operator of order m like (9.6) is a differential operator of order m:

$$P^t : C^\infty(M,E_1) \longrightarrow C^\infty(M,E_0),$$

defined by the condition that

$$(Pu,v) = (u,P^tv) \tag{9.9}$$

if u and v are smooth, compactly supported sections of the bundles E_0 and E_1. If u and v are supported on a coordinate patch $\mathcal{O}$ on M, over which E_j are trivialized, so u and v have components u^σ, v^σ, and if the metrics on E_0 and E_1 are denoted $h_{\sigma\delta}$, $\tilde{h}_{\sigma\delta}$, respectively, while the Riemannian metric is g_{jk}, then we have

$$(Pu,v) = \int_{\mathcal{O}} \tilde{h}_{\sigma\delta}(x)\big(Pu\big)^\sigma\, \overline{v}^\delta \sqrt{g(x)}\, dx. \tag{9.10}$$

Substituting (9.1) and integrating by parts produce an expression for P^t, of the form

$$P^tv(x) = \sum_{|\alpha|\le m} p^t_\alpha(x) D^\alpha v(x). \tag{9.11}$$

In particular, one sees that the principal symbol of P^t is given by

$$\sigma_{P^t}(x,\xi) = \sigma_P(x,\xi)^t. \tag{9.12}$$

Compare this with the specific formula (2.22) for the formal adjoint of a real vector field, which has a purely imaginary symbol.

Now suppose M is a compact, smooth manifold with smooth boundary. We want to obtain a generalization of formula (2.24), that is,

$$(Xu,v) - (u, X^t v) = \int_{\partial M} \langle \nu, X\rangle u\overline{v}\, dS, \tag{9.13}$$

to the case where P is a general *first-order* differential operator, acting on sections of a vector bundle as in (9.6). Using a partition of unity, we can suppose that u and v are supported in a coordinate patch $\mathcal{O}$ in M. If the patch is disjoint from ∂M, then of course (9.9) holds. Otherwise, suppose $\mathcal{O}$ is a patch in $\mathbb{R}^n_+$. If the first-order operator P has the form

$$Pu = \sum_{j=1}^n a_j(x)\frac{\partial u}{\partial x_j} + b(x)u, \tag{9.14}$$

then

$$\int_{\mathcal{O}} \langle Pu, v\rangle\, \sqrt{g}\, dx = \int_{\mathcal{O}} \Big[\sum_{j=1}^n \langle a_j(x)\frac{\partial u}{\partial x_j}, v\rangle + \langle b(x)u, v\rangle\Big]\sqrt{g}\, dx. \tag{9.15}$$

If we apply the fundamental theorem of calculus, the only boundary integral comes from the term involving $\partial u/\partial x_n$. Thus we have

$$\int_{\mathcal{O}} \langle Pu, v\rangle \sqrt{g}\, dx = \int_{\mathcal{O}} \langle u, P^t v\rangle \sqrt{g}\, dx - \int_{\mathbb{R}^{n-1}} \langle a_n(x',0)u, v\rangle \sqrt{g(x',0)}\, dx', \tag{9.16}$$

where $dx' = dx_1 \cdots dx_{n-1}$. If we pick the coordinate patch so that $\partial/\partial x_n$ is the unit inward normal at ∂M, then $\sqrt{g(x',0)}\, dx'$ is the volume element on ∂M, and we are ready to establish the following Green–Stokes formula:

Proposition 9.1. *If M is a smooth, compact manifold with boundary and P is a first-order differential operator (acting on sections of a vector bundle), then*

$$(Pu, v) - (u, P^t v) = \frac{1}{i}\int_{\partial M} \langle \sigma_P(x,\nu)u, v\rangle\, dS. \tag{9.17}$$

Proof. The formula (9.17), which arose via a choice of local coordinate chart, is invariant and hence valid independent of choices.

As in (9.13), ν denotes the *outward*-pointing unit normal to ∂M; we use the Riemannian metric on M to identify tangent vectors and cotangent vectors.

We will see an important application of (9.17) in the next section, where we consider the Laplace operator on k-forms.

Exercises

1. Consider the divergence operator acting on (complex-valued) vector fields:

$$\mathrm{div} : C^\infty(\Omega, \mathbb{C}^n) \longrightarrow C^\infty(\Omega), \quad \Omega \subset \mathbb{R}^n.$$

Show that its symbol is given by

$$\sigma_{\mathrm{div}}(x, \xi)v = i\langle v, \xi\rangle.$$

2. Consider the gradient operator acting on (complex valued) functions:

$$\mathrm{grad} : C^\infty(\Omega) \longrightarrow C^\infty(\Omega, \mathbb{C}^n), \quad \Omega \subset \mathbb{R}^n.$$

Show that its symbol is

$$\sigma_{\mathrm{grad}}(x, \xi) = i\xi.$$

3. Consider the operator

$$L = \mathrm{grad}\,\mathrm{div} : C^\infty(\Omega, \mathbb{C}^n) \longrightarrow C^\infty(\Omega, \mathbb{C}^n).$$

Show that its symbol is

$$\sigma_L(x, \xi) = -|\xi|^2 P_\xi,$$

where $P_\xi \in \mathrm{End}(\mathbb{C}^n)$ is the orthogonal projection onto the (complex) linear span of ξ.

4. What is the symbol of the operator

$$P = \mu\Delta + (\lambda + \mu)\,\mathrm{grad}\,\mathrm{div},$$

which appears in the equation (1.59) of linear elasticity? What are the eigenvalues of the symbol, for given $\xi \in \mathbb{R}^n$?

5. Generalize Exercises 1–3 to the case of a Riemannian manifold.
6. Let L be a constant-coefficient, second-order, homogeneous, linear differential operator acting on functions on $\mathbb{R}^n$ with values in $\mathbb{C}^k$, of the form

$$Lu = \sum_{|\alpha|=2} A_\alpha\, D^\alpha u, \quad A_\alpha \in \mathrm{End}(\mathbb{C}^k).$$

Let $\xi \in \mathbb{R}^n \setminus 0$. A "plane wave" solution to $u_{tt} - Lu = 0$ is a $\mathbb{C}^k$-valued function $u(t, x)$ of the form

$$u(t, x) = v(t, x \cdot \xi),$$

with $v(t, y)$ a $\mathbb{C}^k$-valued function on $\mathbb{R} \times \mathbb{R}$. Show that the PDE for v becomes

$$v_{tt} - M v_{yy} = 0,$$

with

$$M = -\sigma_L(x, \xi).$$

In case $\sigma_L(x, \xi)$ is negative-definite with eigenvalues $-c_j^2 = -c_j(\xi)^2$, show that the initial-value problem for v can be solved in terms of the formula for the one-dimensional wave equation derived in §1.

7. Consider the equation of linear elasticity from (1.59):

$$m w_{tt} - \mu \Delta w - (\lambda + \mu) \operatorname{grad} \operatorname{div} w = 0.$$

Suppose $\mu > 0, 2\mu + \lambda > 0$. Fix $\xi \in \mathbb{R}^n \setminus 0$. Using the results of Exercises 4 and 6, analyze plane wave solutions $w(t, x) = v(t, x \cdot \xi)$. Show that if $n \geq 2$, there are *two* propagation speeds. The faster and slower waves are called "p-waves" (pressure waves) and "s-waves" (shear waves), respectively. If $n = 1$, only p-waves arise.

10. The Hodge Laplacian on k-forms

If M is an n-dimensional Riemannian manifold, recall the exterior derivative

$$d : \Lambda^k(M) \longrightarrow \Lambda^{k+1}(M), \tag{10.1}$$

satisfying

$$d^2 = 0. \tag{10.2}$$

The Riemannian metric on M gives rise to an inner product on T_x^* for each $x \in M$, and then to an inner product on $\Lambda^k T_x^*$, via

$$\langle v_1 \wedge \cdots \wedge v_k, w_1 \wedge \cdots \wedge w_k \rangle = \sum_\pi (\operatorname{sgn} \pi) \langle v_1, w_{\pi(1)} \rangle \cdots \langle v_k, w_{\pi(k)} \rangle, \tag{10.3}$$

where π ranges over the set of permutations of $\{1, \ldots, k\}$. Equivalently, if $\{e_1, \ldots, e_n\}$ is an orthonormal basis of $T_x^* M$, then $\{e_{j_1} \wedge \cdots \wedge e_{j_k} : j_1 < j_2 < \cdots < j_k\}$ is an orthonormal basis of $\Lambda^k T_x^* M$. Consequently, there is an inner product on k-forms (that is, sections of Λ^k) given by

$$(u, v) = \int_M \langle u, v \rangle \, dV(x). \tag{10.4}$$

Thus there is a first-order differential operator

$$\delta : \Lambda^{k+1}(M) \longrightarrow \Lambda^k(M), \tag{10.5}$$

which is the formal adjoint of d, that is, δ is characterized by

$$(du, v) = (u, \delta v), \quad u \in \Lambda^k(M),\ v \in \Lambda^{k+1}(M),\ \text{compactly supported.} \tag{10.6}$$

We set $\delta = 0$ on 0-forms. Of course, (10.2) implies

$$\delta^2 = 0. \tag{10.7}$$

There is a useful formula for δ, involving d and the "Hodge star operator," which will be derived in Chap. 5, §8.

The Hodge Laplacian on k-forms,

$$\Delta : \Lambda^k(M) \longrightarrow \Lambda^k(M), \tag{10.8}$$

is defined by

$$-\Delta = (d + \delta)^2 = d\delta + \delta d. \tag{10.9}$$

Consequently,

$$(-\Delta u, v) = (du, dv) + (\delta u, \delta v), \quad \text{for } u, v \in C_0^\infty(M, \Lambda^k). \tag{10.10}$$

Since $\delta = 0$ on $\Lambda^0(M)$, we have $-\Delta = \delta d$ on $\Lambda^0(M)$.

We will obtain an analogue of (10.10) for the case where M is a compact manifold with boundary, so a boundary integral appears. To obtain such a formula, we specialize the general Green–Stokes formula (9.17) to the cases $P = d$ and $P = \delta$. First, we compute the symbols of d and δ. Since, for a k-form u,

$$d(u\, e^{i\lambda\psi}) = i\lambda e^{i\lambda\psi}(d\psi) \wedge u + e^{i\lambda\psi}\, du, \tag{10.11}$$

we see that

$$\frac{1}{i}\, \sigma_d(x, \xi)u = \xi \wedge u. \tag{10.12}$$

As a special case of (9.12), we have

$$\sigma_\delta(x, \xi) = \sigma_d(x, \xi)^t. \tag{10.13}$$

The adjoint of the map (10.12) from $\Lambda^k T_x^*$ to $\Lambda^{k+1} T_x^*$ is given by the interior product

$$\iota_\xi u = u \rfloor X, \tag{10.14}$$

where $X \in T_x$ is the vector corresponding to $\xi \in T_x^*$ under the isomorphism $T_x \approx T_x^*$ given by the Riemannian metric. Consequently,

$$\frac{1}{i}\, \sigma_\delta(x, \xi)u = -\iota_\xi u. \tag{10.15}$$

Now, the Green–Stokes formula (9.17) implies, for M a compact Riemannian manifold with boundary,

$$
\begin{aligned}
(du, v) &= (u, \delta v) + \frac{1}{i} \int\limits_{\partial M} \langle \sigma_d(x,\nu)u, v\rangle \, dS \\
&= (u, \delta v) + \int\limits_{\partial M} \langle \nu \wedge u, v\rangle \, dS,
\end{aligned}
\tag{10.16}
$$

and

$$
\begin{aligned}
(\delta u, v) &= (u, dv) + \frac{1}{i} \int\limits_{\partial M} \langle \sigma_\delta(x,\nu)u, v\rangle \, dS \\
&= (u, dv) - \int\limits_{\partial M} \langle \iota_\nu u, v\rangle \, dS.
\end{aligned}
\tag{10.17}
$$

Recall that ν is the outward-pointing unit normal to ∂M.

Consequently, our generalization of (10.10), and also of (4.8), is

$$
\begin{aligned}
-(\Delta u, v) = (du, dv) &+ (\delta u, \delta v) \\
&+ \frac{1}{i} \int\limits_{\partial M} \big[\langle \sigma_d(x,\nu)\, \delta u, v\rangle + \sigma_\delta(x,\nu)\, du, v\rangle\big]\, dS
\end{aligned}
\tag{10.18}
$$

or, equivalently,

$$
\begin{aligned}
-(\Delta u, v) = (du, dv) &+ (\delta u, \delta v) \\
&+ \int\limits_{\partial M} \big[\langle \nu \wedge (\delta u), v\rangle - \langle \iota_\nu (du), v\rangle\big]\, dS.
\end{aligned}
\tag{10.19}
$$

Taking adjoints of the symbol maps, we can also write

$$
\begin{aligned}
-(\Delta u, v) = (du, dv) &+ (\delta u, \delta v) \\
&+ \int\limits_{\partial M} \big[\langle \delta u, \iota_\nu v\rangle - \langle du, \nu \wedge v\rangle\big]\, dS.
\end{aligned}
\tag{10.20}
$$

Let us note what the symbol of Δ is. By (10.12) and (10.15),

$$
-\sigma_\Delta(x,\xi)u = \iota_\xi \xi \wedge u + \xi \wedge \iota_\xi u.
\tag{10.21}
$$

If we perform the calculation by picking an orthonormal basis for T_x^* of the form $\{e_1, \dots, e_n\}$ with $\xi = |\xi| e_1$, we see that

$$
\sigma_\Delta(x,\xi)u = -|\xi|^2 u.
\tag{10.22}
$$

In other words, in a local coordinate system, we have, for a k-form u,

$$\Delta u = g^{j\ell}(x)\,\partial_j\partial_\ell u + Y_k u, \tag{10.23}$$

where Y_k is a first-order differential operator.

A differential operator $P : C^\infty(M, E_0) \to C^\infty(M, E_1)$ is said to be *elliptic* provided $\sigma_P(x,\xi) : E_{0x} \to E_{1x}$ is invertible for each $x \in M$ and each $\xi \neq 0$. By (10.22), the Laplace operator on k-forms is elliptic.

Of course, the definition $-\Delta = \delta d$ for the Laplace operator on 0-forms coincides with the definition given in §4. In this regard, it is useful to note explicitly the following result about δ on 1-forms. Let X be a vector field and ξ the 1-form corresponding to X under a given metric:

$$g(Y, X) = \langle Y, \xi\rangle. \tag{10.24}$$

Then

$$\delta\xi = -\operatorname{div} X. \tag{10.25}$$

This identity is equivalent to (2.18) and the definition of δ as the formal adjoint of d.

We end this section with some algebraic implications of the symbol formula (10.21)–(10.22) for the Laplace operator. If we define $\wedge_\xi : \Lambda^* T_x^* \to \Lambda^* T_x^*$ by $\wedge_\xi(\omega) = \xi \wedge \omega$, and define ι_ξ as above, by (10.14), then the content of this calculation is

$$\wedge_\xi \iota_\xi + \iota_\xi \wedge_\xi = |\xi|^2. \tag{10.26}$$

As we have mentioned, this can be established by picking $\xi/|\xi|$ to be the first member of an orthonormal basis of T_x^*. Extending the identity (10.26), we have

$$\wedge_\xi \iota_\eta + \iota_\eta \wedge_\xi = \langle \xi, \eta\rangle, \tag{10.27}$$

a result that follows from the formula (13.37) of Chap. 1. Note also the connection with (2.26).

Exercises

1. Show that the adjoint of the exterior product operator $\xi\wedge$ is ι_ξ, as asserted in (10.14).
2. If $\alpha = \sum a_{jk}(x)\,dx_j \wedge dx_k$ and $a_j{}^k = g^{k\ell} a_{j\ell}$, relate $\delta\alpha$ to the divergence $a_j{}^k{}_{;k}$, as defined in (3.29).
3. Using (10.20), write down an expression for

$$(\Delta u, v) - (u, \Delta v)$$

as a boundary integral, when u and v are k-forms.

4. Relate the characterization (10.3) of the inner product on $\Lambda^* T_x^*$ arising from an inner product on T_x^*, to that given in the following section, before (11.24).
5. Let $\omega \in \Lambda^n(M)$, $n = \dim M$, be the volume form of an oriented Riemannian manifold M. Show that $\delta\omega = 0$. (*Hint*: Compare (10.6)with the special case of Stokes' formula $\int_M du = 0$ for $u \in \Lambda^{n-1}(M)$, compactly supported.)
6. Given the result of Exercise 5, show that Stokes' formula $\int_M du = \int_{\partial M} u$, for $u \in \Lambda^{n-1}(M)$, follows from (10.16).
7. If $f \in C^\infty(M)$ and $u \in \Lambda^k(M)$, show that
$$\delta(fu) = f\delta u - \iota_{(df)}u.$$
8. For a vector field u on the Riemannian manifold M, let $\tilde{u}$ denote the associated 1-form. Show that
$$\delta(\tilde{u} \wedge \tilde{v}) = (\operatorname{div} v)\tilde{u} - (\operatorname{div} u)\tilde{v} - \widetilde{[u, v]},$$
for $\widetilde{u}, \widetilde{v} \in \Lambda^1(M)$. Reconsider this problem after reading Chap. 5, §8.

11. Maxwell's equations

The equations governing the electromagnetic field are one of the major triumphs of theoretical physics. We list them here, for the electric field E and the magnetic field B, in a vacuum:

$$\operatorname{div} B = 0, \tag{11.1}$$

$$\frac{\partial B}{\partial t} + \operatorname{curl} E = 0, \tag{11.2}$$

$$\operatorname{div} E = 4\pi\rho, \tag{11.3}$$

$$\frac{\partial E}{\partial t} - \operatorname{curl} B = -4\pi J. \tag{11.4}$$

Here, ρ is the charge density and J the electric current. Units are chosen so that the speed of light is 1. Here we are glossing over the distinction between two types of electric field, typically denoted E and D, and two types of magnetic field, typically denoted B and H, and their relation via "dielectric constants." Material on this may be found in texts on electromagnetism, such as [Ja].

Of the four equations above, (11.1) and (11.3) have a relatively elementary character. Equation, known as Gauss' law, follows in the case of stationary charges from the statement that a charge e at a point $p \in \mathbb{R}^3$ produces an electric field
$$E(x) = e\,\frac{x - p}{|x - p|^3},$$
which is Coulomb's law. Equation (11.1) is the statement that there are no magnetic charges. Both of these laws are well supported by experiments. We note parenthetically that there is reason to believe that at high energies magnetic

charges might exist. A theoretical framework for this is provided by a modification of the theory of the electromagnetic field, called the "electroweak theory." But that is a story that we will not try to relate in this book. As one reference, we mention [IZ].

The equations (11.2) and (11.4) are more subtle. Equation (11.2), which implies that a changing magnetic field produces an electric field, is called Faraday's law. One implication of (11.4) is that an electric current produces a magnetic field; this is exploited in electric motors. The first quantitative expression of this effect written down was

$$\operatorname{curl} B = 4\pi J,$$

which is valid when all quantities involved are independent of time. It breaks down when variation with time is allowed. Indeed, the left side must have vanishing divergence, but in the time-varying case one has, not $\operatorname{div} J = 0$, but rather the following law of conservation of charge:

$$\frac{\partial \rho}{\partial t} + \operatorname{div} J = 0. \tag{11.5}$$

Maxwell produced the modification (11.4), which completed the set of equations for the electromagnetic field.

Careful thought about the implications of Maxwell's equations, together with the experimental fact that two observers moving with respect to each other would measure the speed of light to be the same, led to the development of Einstein's theory of relativity. We will not discuss how this was done. Rather, following J. Wheeler, we will reverse the historical order. We will rewrite (11.1)–(11.4) in an invariant fashion, depending only on the Lorentz metric $-dx_0^2 + dx_1^2 + dx_2^2 + dx_3^2$ on Minkowski spacetime $\mathbb{R}^4$ rather than a particular Cartesian product decomposition of $\mathbb{R}^4$ into time $\mathbb{R}$ and space $\mathbb{R}^3$. We can then show that, within the relativistic framework, the subtle (11.2) and (11.4) actually follow from the "simple" (11.1) and (11.3).

We bring in the 2-form (with $t = x_0$)

$$\mathcal{F} = \sum_1^3 E_j \, dx_j \wedge dt + B_1 \, dx_2 \wedge dx_3 + B_2 \, dx_3 \wedge dx_1 + B_3 \, dx_1 \wedge dx_2. \tag{11.6}$$

In §18 of Chap. 1 it was shown how this form arises naturally in the relativistic expression of how the electromagnetic field acts on a charged particle to make it move. A calculation of the exterior derivative gives

$$d\mathcal{F} = \sum_1^3 \Big(\frac{\partial B}{\partial t} + \operatorname{curl} E \Big)_j (*dx_j) \wedge dt + (\operatorname{div} B) \, dx_1 \wedge dx_2 \wedge dx_3, \tag{11.7}$$

where, for $1 \le j \le 3$, we set

$$*dx_j = dx_k \wedge dx_\ell, \quad (j,k,\ell) \text{ a cyclic permutation of } (1,2,3).$$

Consequently, (11.1) and (11.2) together are equivalent to the equation

$$d\mathcal{F} = 0. \tag{11.8}$$

On the other hand, (11.1) alone is equivalent to the following. For fixed T, define $\kappa_T : \mathbb{R}^3 \to \mathbb{R}^4$ by $\kappa_T(x') = (T, x')$. Then (11.1) holds at $t = T$ if and only if

$$\kappa_T^* d\mathcal{F} = 0. \tag{11.9}$$

Now, in the relativistic set-up, any physical law that is valid on all surfaces $t =$ const. in $\mathbb{R}^4$ should be valid on *all* spacelike hyperplanes in $\mathbb{R}^4$. But the following result is easy to establish.

Lemma 11.1. *Let $0 \le k \le 3$, and suppose $\alpha \in \Lambda^k(\mathbb{R}^4)$ has the property that*

$$\kappa^*\alpha = 0, \tag{11.10}$$

for every inclusion $\kappa : S \to \mathbb{R}^4$ of spacelike hyperplanes in $\mathbb{R}^4$. Then $\alpha = 0$.

Applying this to $\alpha = d\mathcal{F}$, we see how (11.1) yields (11.2).

We will be able to rewrite(11.3)–(11.4) using the "adjoint" to d:

$$d^\star : \Lambda^k(\mathbb{R}^4) \longrightarrow \Lambda^{k-1}(\mathbb{R}^4), \tag{11.11}$$

defined like $\delta = d^*$ in §10, but using an inner product coming from the Lorentz metric. Thus, for compactly supported u,

$$L(du, v) = L(u, d^\star v), \tag{11.12}$$

for a $(k-1)$-form u and a k-form v, where the inner product of two k-forms v_j is

$$L(v_1, v_2) = \int \langle v_1, v_2\rangle \, dx_0 \cdots dx_3, \tag{11.13}$$

the integral of the pointwise inner product, characterized as follows.

A form $dx_{j_1} \wedge \cdots \wedge dx_{j_k}$ with distinct j_ν's has square norm $\varepsilon_{j_1} \cdots \varepsilon_{j_k}$, where $\varepsilon_0 = -1,\ \varepsilon_1 = \varepsilon_2 = \varepsilon_3 = 1$. Two such forms are orthogonal unless their sets of indices $\{j_1, \dots, j_k\}$ coincide. A straightforward calculation yields

$$d^\star g_{k\ell}(x)\, dx_k \wedge dx_\ell = -\sum_{i,j} \varepsilon_j \varepsilon(i,j;k,\ell) \frac{\partial g_{k\ell}}{\partial x_i}\, dx_j, \tag{11.14}$$

where

$$\varepsilon(i,j;k,\ell) = \langle dx_i \wedge dx_j, dx_k \wedge dx_\ell \rangle \tag{11.15}$$

is characterized above. This is 0 unless $\{i,j\} = \{k,\ell\}$, and we can rewrite (11.14) as

$$d^\star g_{k\ell}(x)\, dx_k \wedge dx_\ell = \varepsilon(k,\ell;k,\ell)\Big[\varepsilon_k \frac{\partial g_{k\ell}}{\partial x_\ell}\, dx_k - \varepsilon_\ell \frac{\partial g_{k\ell}}{\partial x_k}\, dx_\ell\Big]. \tag{11.16}$$

This implies

$$d^\star \sum_1^3 E_j\, dx_j \wedge dx_0 = -(\operatorname{div} E)\, dx_0 - \sum_1^3 \frac{\partial E_j}{\partial t}\, dx_j, \tag{11.17}$$

and

$$d^\star\big(B_1\, dx_2 \wedge dx_3 + B_2\, dx_3 \wedge dx_1 + B_3\, dx_1 \wedge dx_2\big) = \sum_1^3 (\operatorname{curl} B)_j\, dx_j. \tag{11.18}$$

Thus (11.3) and (11.4) together are equivalent to the equation

$$d^\star \mathcal{F} = 4\pi \mathcal{J}^b, \tag{11.19}$$

where

$$\mathcal{J}^b = -\rho\, dt + \sum_1^3 J_k\, dx_k. \tag{11.20}$$

Thus $\mathcal{J}^b$ is the 1-form associated via the Lorentz metric to the vector

$$\mathcal{J} = (\rho, J), \tag{11.21}$$

called the charge-current 4-vector.

In this case, (11.3) alone is equivalent to the identity

$$\big(d^\star \mathcal{F} - 4\pi \mathcal{J}^b\big) \rfloor \frac{\partial}{\partial t} = 0. \tag{11.22}$$

Again, in the relativistic set-up, such a physical law ought to be independent of the choice of timelike vector field with which to take the interior product. Thus, if we assume that $\mathcal{F}$ has an invariant significance as a 2-form and also that $\mathcal{J}^b$ has an invariant significance as a 1-form, we are in a position to apply the following.

Lemma 11.2. *If* $1 \le k \le 4$ *and* $\alpha \in \Lambda^k(\mathbb{R}^4)$ *has the property that*

$$\alpha \rfloor V = 0 \tag{11.23}$$

for all timelike vectors V, *then* $\alpha = 0$.

Applying this to $\alpha = d^\star \mathcal{F} - 4\pi \mathcal{J}^b$, we see how (11.3) yields (11.4).

The pair of Maxwell equations (11.8), (11.19) make sense on any Lorentz manifold of dimension 4 and provide the appropriate equations for an electromagnetic field in curved spacetime. To define $d^\star$, one uses the formula (11.12), replacing $dx_0 \cdots dx_3$ by the natural volume element on a general Lorentz manifold M in (11.13).

This construction defines $d^\star$ for Lorentz manifolds of any dimension. The inner product in the integrand in (11.13) can be characterized as follows. To the Lorentz inner product on $V = T_xM$ corresponds an isomorphism $Q : V \to V'$ satisfying $Q' = Q$ (with $V'' = V$). This induces isomorphisms

$$Q_k : \Lambda^k V \to \Lambda^k V' \approx \big(\Lambda^k V\big)',$$

with the same symmetry property, yielding inner products on $\Lambda^k V$, $0 \le k \le m = \dim M$. Equivalently, if you pick an "orthonormal" basis $\{v_0, \dots, v_{m-1}\}$ of V, satisfying $\langle v_0, v_0 \rangle = -1$, $\langle v_j, v_j \rangle = 1$ for $1 \le j \le m-1$, then the characterization given after (11.13) is easily extended.

In analogy with (10.9), it is of interest to form the second-order operator

$$-\Box = (d + d^\star)^2 = dd^\star + d^\star d. \tag{11.24}$$

A calculation similar to (10.23) gives

$$\Box u = h^{j\ell}(x)\, \partial_j \partial_\ell u + Y_k u, \tag{11.25}$$

for a k-form u, where $(h^{j\ell})$ is formed from the Lorentz metric tensor, as in (7.7), and Y_k is a first-order differential operator. On 0-forms, this operator is exactly (7.7). For Minkowski spacetime $\mathbb{R}^4$, $\Box$ is just $-\partial^2/\partial x_0^2 + \sum_1^3 \partial^2/\partial x_j^2$, acting on each component of a k-form.

The equations $d\mathcal{F} = 0$, $d^\star \mathcal{F} = 4\pi \mathcal{J}^b$ imply that $\mathcal{F}$ satisfies the "wave equation"

$$\Box \mathcal{F} = -4\pi\, d\mathcal{J}^b. \tag{11.26}$$

The results developed in §8 for scalar hyperbolic operators of the type (8.2) are easily extended to cover the operator $\Box$ constructed here, which by (11.25) has scalar principal part.

In particular, finite propagation speed arguments apply to solutions to Maxwell's equations. Existence of solutions, including propagation of electromagnetic waves in regions bounded by perfect conductors, is studied in Chap. 6.

The energy in an electromagnetic field in $\mathbb{R}^4 = \mathbb{R} \times \mathbb{R}^3$ is

$$V(t) = \frac{1}{8\pi} \int\limits_{\mathbb{R}^3} \big[|E(t,x)|^2 + |B(t,x)|^2\big]\,dx. \tag{11.27}$$

If (11.1)– (11.4) hold, then

$$\begin{aligned} 4\pi\,\frac{dV}{dt} &= \Big(\frac{\partial E}{\partial t}, E\Big) + \Big(\frac{\partial B}{\partial t}, B\Big) \\ &= (\operatorname{curl} B, E) - (\operatorname{curl} E, B) - 4\pi(J, E). \end{aligned} \tag{11.28}$$

If $E(t,x)$ and $B(t,x)$ decrease sufficiently rapidly as $|x| \to \infty$, we have

$$(\operatorname{curl} B, E) = (B, \operatorname{curl} E), \tag{11.29}$$

as can be established by integration by parts. Hence

$$\frac{dV}{dt} = -\int\limits_{\mathbb{R}^3} J(t,x)\cdot E(t,x)\,dx. \tag{11.30}$$

In particular, for $J = 0$ we have conservation of $V(t)$.

One can construct a stress-energy tensor $\mathcal{T}$ due to the electromagnetic field, by an argument similar to that of §7. First note that, with $\mathcal{F}$ given by (11.6), we have

$$\langle \mathcal{F}, \mathcal{F}\rangle = |B|^2 - |E|^2. \tag{11.31}$$

Equivalently,

$$\operatorname{Tr} \widetilde{\mathcal{F}}^2 = 2(|E|^2 - |B|^2), \tag{11.32}$$

where $\widetilde{\mathcal{F}}$ is the tensor field of type $(1,1)$ associated to $\mathcal{F}$. Note also that $\big(\widetilde{\mathcal{F}}^2\big)^0{}_0 = |E|^2$. Thus a natural construction of $\mathcal{T}$ giving rise to $\mathcal{T}_{00} = (1/8\pi)(|E|^2+|B|^2)$ is

$$\widetilde{\mathcal{T}} = -\frac{1}{4\pi}\Big(\widetilde{\mathcal{F}}^2 - \frac{1}{4}(\operatorname{Tr}\widetilde{\mathcal{F}}^2)I\Big) = -\frac{1}{4\pi}\Big(\widetilde{\mathcal{F}}^2 + \frac{1}{2}\langle \mathcal{F}, \mathcal{F}\rangle I\Big), \tag{11.33}$$

where $\widetilde{\mathcal{T}}$ is the tensor field of type $(1,1)$ associated with $\mathcal{T}$. In index notation,

$$\mathcal{T}_{ij} = \frac{1}{4\pi}\big(\mathcal{F}_{im}\mathcal{F}_j{}^m - \frac{1}{4}h_{ij}\mathcal{F}^{mn}\mathcal{F}_{mn}\big), \tag{11.34}$$

where (h_{ij}) is the Lorentz metric tensor. In this case, in analogy with (7.10), one obtains

$$\mathcal{T}^{jk}{}_{;k} = -\mathcal{F}^{j}{}_{k}\mathcal{J}^{k}, \tag{11.35}$$

provided the Maxwell equations (11.8) and (11.19) hold. Equivalently, with $\widehat{\mathcal{T}}$ denoting the tensor field of type $(2, 0)$ associated with $\mathcal{T}$,

$$\operatorname{div} \widehat{\mathcal{T}} = -\widetilde{\mathcal{F}}\mathcal{J}. \tag{11.36}$$

If the electromagnetic field $\mathcal{F}$ is defined on a Lorentz 4-manifold which is simply connected, the equation $d\mathcal{F} = 0$ implies the existence of a 1-form $\mathcal{A}$ such that $\mathcal{F} = d\mathcal{A}$. We can define the Lagrangian

$$L = -\frac{1}{8\pi}\langle \mathcal{F}, \mathcal{F}\rangle = -\frac{1}{8\pi}\langle d\mathcal{A}, d\mathcal{A}\rangle, \tag{11.37}$$

with inner product as in (11.31). The action integral $I(\mathcal{A}) = \int L\, dV$ satisfies, for a compactly supported 1-form β,

$$\frac{d}{d\tau} I(\mathcal{A} + \tau\beta)\big|_{\tau=0} = -\frac{1}{4\pi}\int \langle d\beta, d\mathcal{A}\rangle\, dV = -\frac{1}{4\pi}\int \langle \beta, d^{\star} d\mathcal{A}\rangle\, dV, \tag{11.38}$$

so the stationary condition $\delta \int L\, dV = 0$ is equivalent to $d^{\star} d\mathcal{A} = 0$, that is, to the rest of Maxwell's equations (11.19), in case $\mathcal{J} = 0$. Thus (11.37) is the appropriate Lagrangian for the electromagnetic field, in order to produce Maxwell's equations in empty space. If the current $\mathcal{J}$ is *given* (subject to the condition $d^{\star}\mathcal{J} = 0$), and $\mathcal{F} = d\mathcal{A}$, then the (11.19) is the stationary condition $\delta \int L\, dV = 0$ for the Lagrangian

$$L = -\frac{1}{8\pi}\langle \mathcal{F}, \mathcal{F}\rangle + \langle \mathcal{A}, \mathcal{J}\rangle. \tag{11.39}$$

In typical problems the current is not given in advance, but is itself influenced by the electromagnetic force. The nature of the influence involves the masses of the substances that carry charges, whose motion produces the current. Then the Maxwell equations are coupled to other equations, which are often nonlinear. We describe a model for one example.

Suppose we have a diffuse cloud of electrons, in otherwise empty space. We model this as a continuous charged substance, whose motion is described by a 4-velocity vector field u, satisfying $\langle u, u\rangle = -1$, yielding a current $\mathcal{J} = \sigma u$, where $\sigma\, dV$ is the charge density, measured by an observer whose velocity is u.

Taking a cue from the Lagrangian (18.20) of Chap. 1, derived to reflect the relativistic Lorenz force law, we use the Lagrangian

$$L = -\frac{1}{8\pi}\langle \mathcal{F}, \mathcal{F}\rangle + \langle \mathcal{A}, \mathcal{J}\rangle + \frac{1}{2}\mu\langle u, u\rangle = L_1 + L_2 + L_3, \tag{11.40}$$

where $\mu\, dV$ is the mass density, measured by an observer whose velocity is u. We are assuming that only one type of matter is present, so σ is a constant multiple of μ. In more general cases there would be additional terms in the Lagrangian.

We look at $I(\mathcal{A}, u) = I_1 + I_2 + I_3$. The term I_3 is independent of $\mathcal{A}$, and as above we have

$$\frac{\partial}{\partial\tau} I(\mathcal{A} + \tau\beta, u)\big|_{\tau=0} = \int \Big[-\frac{1}{4\pi}\langle \beta, d^{\star} d\mathcal{A}\rangle + \langle \beta, \mathcal{J}\rangle\Big]\, dV. \tag{11.41}$$

The stationary condition this yields is again the Maxwell equation (11.19). Next we compute $(\partial/\partial\tau) I\big(\mathcal{A}, u(\tau)\big)\big|_{\tau=0}$, where $u(\tau)$ is a one-parameter family of velocity fields on M, obtained by varying the electron trajectories. There is no variation in I_1, so we need to consider I_2 and I_3.

We first treat the variation of I_3, in a manner parallel to the calculations (11.17)–(11.26) in Chap. 1, leading to the geodesic equations. To do this, we parameterize the electron trajectories by $X : \Omega \times I \to M$, $X(y, s) = x$, $u = \partial_s X$. We suppose the mass density is constant in (y, s)-coordinates, say m, so $m\; dy\; ds = \mu\; dV$. Since $u = \partial/\partial s$ in (y, s)-coordinates, this implies $\mathcal{L}_u(\mu\, dV) = 0$, or

$$\operatorname{div}(\mu u) = 0, \tag{11.42}$$

where div is computed using the Lorentz metric on M. Our hypothesis amounts to the law of conservation of matter. If we vary this map, using $X(y, s, \tau)$, then

$$\frac{d}{d\tau}\int \frac{1}{2}\mu\langle u, u\rangle\, dV = \int \frac{1}{2} m\mathcal{L}_w\langle u, u\rangle\, dy\, ds = \int m\langle \nabla_w u, u\rangle\, dy\, ds, \tag{11.43}$$

where $\partial_\tau X = w$. Using $[\partial_s, \partial_\tau] = 0$, convert this last integral to

$$-\int m\langle w, \nabla_u u\rangle\, dy\, ds + m\int \mathcal{L}_u\langle w, u\rangle\, dy\, ds. \tag{11.44}$$

The last integral here vanishes for a compactly supported perturbation, by the fundamental theorem of calculus, so

$$\frac{d}{d\tau} I_3\big(\mathcal{A}, u(\tau)\big)\big|_{\tau=0} = -\int \langle w, \nabla_u u\rangle m\, dy\, ds = -\int \langle w, \mu\nabla_u u\rangle\, dV. \tag{11.45}$$

We now treat the variation of I_2, also using (y, s)-coordinates. Since σ is a constant multiple of μ, we have $\sigma\; dV = e\; dy\; ds$ for some constant e, and, parallel to (11.42), we have conservation of electric charge, $\operatorname{div}(\sigma u) = 0$ (i.e., div $\mathcal{J} = 0$), which is equivalent to (11.5) when M is Minkowski space. We have

$$\frac{d}{d\tau} \int \langle \mathcal{A}, \mathcal{J} \rangle \, dV = \int e\mathcal{L}_w \langle \mathcal{A}, u \rangle \, dy \, ds. \tag{11.46}$$

We use the identity $\mathcal{L}_u \mathcal{A} = d\mathcal{A} \rfloor u + d(\mathcal{A} \rfloor u)$ to write

$$\begin{aligned} \mathcal{L}_w \langle \mathcal{A}, u \rangle &= -(d\mathcal{A})(u, w) + \langle \mathcal{L}_u \mathcal{A}, w \rangle \\ &= -(d\mathcal{A})(u, w) + \mathcal{L}_u \langle \mathcal{A}, w \rangle - \langle \mathcal{A}, \mathcal{L}_u w \rangle. \end{aligned} \tag{11.47}$$

Since $d\mathcal{A} = \mathcal{F}$, $[\partial_s, \partial_\tau] = 0$, and $\mathcal{L}_u \langle \mathcal{A}, w \rangle$ integrates to zero, we have

$$\frac{d}{d\tau} \int \langle \mathcal{A}, \mathcal{J} \rangle \, dV = \int e \langle \widetilde{\mathcal{F}} u, w \rangle \, dy \, ds = \int \langle \widetilde{\mathcal{F}} \mathcal{J}, w \rangle \, dV. \tag{11.48}$$

Together with (11.45), this gives

$$\frac{\partial}{\partial \tau} I\big(\mathcal{A}, u(\tau)\big)\big|_{\tau=0} = -\int \langle \mu \nabla_u u - \widetilde{\mathcal{F}} \mathcal{J}, w \rangle \, dV. \tag{11.49}$$

Thus the stationary condition for variation of u is

$$\mu \nabla_u u - \widetilde{\mathcal{F}} \mathcal{J} = 0 \ \text{ or, equivalently, } \ \nabla_u u - \frac{e}{m} \widetilde{\mathcal{F}} u = 0, \tag{11.50}$$

which is the Lorentz force law in this context.

It is useful to consider what the stress-energy tensor should be when we have the Lagrangian (11.40). It is reasonable to take it to be the sum of the stress-energy tensor $\mathcal{T}_e$ for the electromagnetic field, given by (11.34), and a stress-energy tensor $\mathcal{T}_m$ associated with the "matter field." If we want $\mathcal{T}_m(Z, Z) dV$ to be the mass-energy density of the electrons observed by one moving with velocity Z, then it is natural to set

$$\widehat{\mathcal{T}}_m = \mu u \otimes u, \tag{11.51}$$

(i.e., $\mathcal{T}_m^{jk} = \mu u^j u^k$). Then the total stress-energy tensor is given by

$$\mathcal{T}^{jk} = \frac{1}{4\pi} \Big(\mathcal{F}^j{}_\ell \mathcal{F}^{k\ell} - \frac{1}{4} h^{jk} \mathcal{F}^{i\ell} \mathcal{F}_{i\ell} \Big) + \mu u^j u^k. \tag{11.52}$$

The divergence of $\widehat{\mathcal{T}}_e$ is given by (11.36), provided the Maxwell equation (11.19) holds. Furthermore, $(\mu u^j u^k)_{;k} = (\mu u^k)_{;k} u^j + \mu u^k u^j{}_{;k}$, so

$$\operatorname{div} \widehat{\mathcal{T}}_m = \operatorname{div}(\mu u) u + \mu \nabla_u u. \tag{11.53}$$

Thus, for $\widehat{\mathcal{T}} = \widehat{\mathcal{T}}_e + \widehat{\mathcal{T}}_m$, we have (granted (11.19))

$$\text{div}\,\widehat{\mathcal{T}} = \text{div}(\mu u) + \mu \nabla_u u - \widetilde{\mathcal{F}}\mathcal{J}. \tag{11.54}$$

We have the conservation law div $\widehat{\mathcal{T}} = 0$ for a solution to the coupled Maxwell–Lorentz equations. Indeed, the vanishing of the first term on the right side of (11.54) is equivalent to the matter conservation law (11.42), and the vanishing of the sum of the other terms on the right side of (11.54) is equivalent to the Lorentz force law (11.50).

Exercises

1. Demonstrate Lemmas 11.1 and 11.2.
2. Verify the calculations (11.14)–(11.18).
3. Show that the inner product of forms defined after (11.13) depends only on the Lorentz metric on $\mathbb{R}^4$, not on the coordinate representation.
4. Show that div curl = 0 is a special case of $dd = 0$.
5. Show that (11.3)–(11.4) imply the "conservation law" (11.5).
   (*Hint*: Apply $\partial/\partial t$ to (11.3) and div to (11.4); use div curl = 0.)
   Show that (11.5) is equivalent to $d^\star \mathcal{J}^b = 0$.
6. Verify the identity (11.29), for any compactly supported vector fields $E(x)$ and $B(x)$ on $\mathbb{R}^3$.
7. Prove the conservation law (11.36), as a consequence of Maxwell's equations.
8. Show that the identity $d\mathcal{F} = 0$ is equivalent to
$$\mathcal{F}_{jk;\ell} + \mathcal{F}_{k\ell;j} + \mathcal{F}_{\ell j;k} = 0.$$
9. Show that the identity $d^\star \mathcal{F} = 4\pi \mathcal{J}^b$ is equivalent to
$$\mathcal{F}^{jk}{}_{;k} = 4\pi \mathcal{J}^j.$$
10. The equation $d\mathcal{F} = 0$ on $\mathbb{R}^4$ implies $\mathcal{F} = \mathcal{A}$ for some 1-form $\mathcal{A}$ on $\mathbb{R}^4$. $\mathcal{A}$ is not unique, as any 1-form du can be added. Show that $\mathcal{A}$ can be picked to satisfy $d^\star \mathcal{A} = 0$ and that, for such $\mathcal{A}$,
$$\Box \mathcal{A} = -4\pi \mathcal{J}^b.$$
    (*Hint*: Set up a PDE for u. Look for the relevant existence theorem in Chap. 3.)
11. The calculation (11.31) of $\langle \mathcal{F}, \mathcal{F} \rangle$ shows that $|B|^2 - |E|^2$ is Lorentz invariant. Calculate $\mathcal{F} \wedge \mathcal{F}$ and show that $E \cdot B$ is also Lorentz invariant.
12. Think about the fact that the tensor $\widetilde{\mathcal{T}}$ given by (11.33) is trace-free, i.e., Tr $\widetilde{\mathcal{T}} = 0$. What is the trace of the stress-energy tensor defined by (7.5) or, equivalently (7.11)?
13. As mentioned in Exercise 5 in §18, Chap. 1, a sign change in the Lorentz metric, from one of signature $(-,+,+,+)$ to one of signature $(+,-,-,-)$ (which some people prefer), leads to a sign change in the formula for the 2-form $\mathcal{F}$ (though no change in the tensor field $\widetilde{\mathcal{F}}$ of type $(1,1)$). Show that it leads to a sign change in the formula (11.34) for the stress-energy tensor of the electromagnetic field.
    What sign changes arise in the formula (11.40) for the Lagrangian of an electromagnetic field coupled to charged matter?

References

[Ar] V. Arnold, *Mathematical Methods of Classical Mechanics*, Springer, New York, 1978.

[Au] T. Aubin, *Nonlinear Analysis on Manifolds. Monge-Ampere Equations*, Springer, New York, 1982.

[Ba] H. Bateman, *Partial Differential Equations of Mathematical Physics*, Dover, New York, 1944.

[BGM] M. Berger, P. Gauduchon, and E. Mazet, *Le Spectre d'une Variété Riemannienne*, LNM no.194, Springer, New York, 1971.

[BJS] L. Bers, F. John, and M. Schechter, *Partial Differential Equations*, Wiley, New York, 1964.

[Car] C. Caratheodory, *Calculus of Variations and Partial Differential Equations of the First Order*, Holden-Day, San Francisco, 1965.

[ChM] A. Chorin and J. Marsden, *A Mathematical Introduction to Fluid Mechanics*, Springer, New York, 1979.

[CK] D. Christodoulu and S. Klainerman, *The Global Nonlinear Stability of the Minkowski Space*, Princeton University Press, Princeton, N. J., 1993.

[CH] R. Courant and D. Hilbert, *Methods of Mathematical Physics II*, Wiley, New York, 1966.

[ES] J. Eells and J. Sampson, Harmonic mappings of Riemannian manifolds, *Am. J. Math.* 86(1964), 109–160.

[FM] A. Fischer and J. Marsden, General relativity, partial differential equations, and dynamical systems, *AMS Proc. Symp. Pure Math.* 23(1973), 309–327.

[Fo1] G. Folland, *Introduction to Partial Differential Equations*, Princeton University Press, Princeton, N. J., 1976.

[Frl] F. G. Friedlander, *Sound Pulses*, Cambridge University Press, Cambridge, 1958.

[Frd] A. Friedman, *Generalized Functions and Partial Differential Equations*, Prentice-Hall, Englewood Cliffs, N. J., 1963.

[Ga] P. Garabedian, *Partial Differential Equations*, Wiley, New York, 1964.

[Go] H. Goldstein, *Classical Mechanics*, Addison-Wesley, New York, 1950.

[Had] J. Hadamard, *Le Probleme de Cauchy et les Equations aux Derivées Partielles Linéaires Hyperboliques*, Hermann, Paris, 1932.

[HE] S. Hawking and G. Ellis, *The Large Scale Structure of Space-Time*, Cambridge University Press, Cambridge, 1973.

[Hel] L. Helms, *Introduction to Potential Theory*, Wiley, New York, 1969.

[HM] T. Hughes and J. Marsden, *A Short Course in Fluid Mechanics*, Publish or Perish, Boston, 1976.

[IZ] C. Itzykson and J. Zuber, *Quantum Field Theory*, McGraw-Hill, New York, 1980.

[Ja] J. Jackson, *Classical Electrodynamics*, Wiley, New York, 1962.

[Jo] F. John, *Partial Differential Equations*, Springer, New York, 1975.

[Jos] J. Jost, *Nonlinear Methods in Riemannian and Kahlerian Geometry*, Birkhauser, Boston, 1988.

[Keg] O. Kellogg, *Foundations of Potential Theory*, Dover, New York, 1954.

[Lax] P. Lax, *The Theory of Hyperbolic Equations*. Stanford Lecture Notes, 1963.

[Lry] J. Leray, *Hyperbolic Differential Equations*, Princeton University Press, Princeton, N. J., 1952.

[MS] H. McKean and I. Singer, Curvature and the eigenvalues of the Laplacian, *J. Diff. Geom*. 1(1967), 43–69.

[Mj] A. Majda, *Compressible Fluid Flow and Systems of Conservation Laws in Several Space Variables*, Appl. Math. Sci. #53, Springer, New York, 1984.

[MH] J. Marsden and T. Hughes, *Mathematical Foundations of Elasticity*, Prentice-Hall, Englewood Cliffs, N. J., 1983.

[Miz] S. Mizohata, *The Theory of Partial Differential Equations*, Cambridge University Press, Cambridge, 1973.

[Mor] C. B. Morrey, *Multiple Integrals in the Calculus of Variations*, Springer, New York, 1966.

[Os] R. Osserman, *A Survey of Minimal Surfaces*, Van Nostrand, New York, 1969.

[Poo] W. Poor, *Differential Geometric Structures*, McGraw-Hill, New York, 1981.

[Rau] J. Rauch, *Partial Differential Equations*, Springer, New York, 1992.

[Ray] L. Rayleigh, On waves propagated along the plane surface of an elastic solid, *Proc. Lond. Math. Soc*. 17(1885), 4–11.

[RS] M. Reed and B. Simon, *Methods of Mathematical Physics*, Academic, New York, Vols. 1,2, 1975; Vols. 3,4, 1978.

[Sm] J. Smoller, *Shock Waves and Reaction-Diffusion Equations*, Springer, New York, 1983.

[So] S. Sobolev, *Partial Differential Equations of Mathematical Physics*, Dover, New York, 1964.

[Stk] I. Stakgold, *Boundary Value Problems of Mathematical Physics*, Macmillan, New York, 1968.

[Stb] S. Sternberg, *Lectures on Differential Geometry*, Prentice Hall, Englewood Cliffs, N. J., 1964.

[Str] M. Struwe, *Variational Problems*, Springer, New York, 1990.

[VMF] R. von Mises and K. Friedrichs, *Fluid Dynamics*, Springer, New York, 1971.

Springer

http://www.springer.com/978-1-4419-7054-1

Partial Differential Equations I
Basic Theory
Taylor, M.
2011, XXII, 654 p., Hardcover
ISBN: 978-1-4419-7054-1