Estimates for first-order homogeneous linear characteristic problems

2005 J. Phys. A: Math. Gen. 38 4209

(http://iopscience.iop.org/0305-4470/38/19/010)

INSTITUTE OF PHYSICS PUBLISHING JOURNAL OF PHYSICS A: MATHEMATICAL AND GENERAL

J. Phys. A: Math. Gen. **38** (2005) 4209–4221 doi:10.1088/0305-4470/38/19/010

Estimates for first-order homogeneous linear characteristic problems

Simonetta Frittelli

Department of Physics, Duquesne University, Pittsburgh, PA 15282, USA
and
Department of Physics and Astronomy, University of Pittsburgh, Pittsburgh, PA 15260, USA

E-mail: simo@mayu.physics.duq.edu

Received 3 August 2004, in final form 24 February 2005
Published 25 April 2005
Online at stacks.iop.org/JPhysA/38/4209

Abstract

An algebraic criterion that is sufficient to establish the existence of certain *a priori* estimates for the solution of first-order homogeneous linear characteristic problems is derived. Estimates of such kind ensure the stability of the solutions under small variations of the data. Characteristic problems that satisfy this criterion are, in a sense, *manifestly well posed*.

PACS number: 02.30.Jr

1. Introduction

Under appropriate (quite general) circumstances, first-order systems of hyperbolic partial differential equations admit characteristic (hyper)surfaces, namely: data surfaces for which the standard Cauchy problem cannot be solved [1]. In the case of the standard three-dimensional wave equation the characteristic surfaces are the null surfaces of flat spacetime and can be interpreted as the hypersurfaces generated by a two-dimensional wavefront propagating at the speed of light, or as surface-forming congruences of null rays. By extension, characteristic surfaces of generic hyperbolic systems are sometimes referred to as *null* or even *lightlike* surfaces, and one can interpret them, in a loose sense, as hypersurfaces that *contain* the evolution of variables that travel at the speed of light (or at the characteristic speed associated with the hypersurface, in any case).

The existence of characteristic surfaces allows for hyperbolic systems to be written in *characteristic form* [2]. In order to do this, a one-parameter family of outgoing characteristic surfaces is chosen as the level surfaces of a coordinate u, referred to as the null coordinate or *retarded time*. The choice of retarded time leads directly to a split of the equations into two sets: equations that involve a derivative with respect to the retarded time, and equations that have no retarded-time derivative. The retarded-time-dependent equations prescribe the evolution of variables that travel at a speed other than the characteristic speed; the values of such variables on one characteristic slice can be obtained from the values on an earlier characteristic slice. These are *normal variables* [2], in the sense that they still behave as variables of a standard Cauchy problem. They require the prescription of data at an initial retarded time.

The retarded-time-independent equations, on the other hand, take the form of constraints internal to the characteristic surface for the variables that travel at the characteristic speed of the surface. These are referred to as *null variables* [2]. Because the action of the differential operator on the null variables is internal to the characteristic slices, not only are the values of the null variables on any given characteristic surface not arbitrary, but they are not related to the values on another characteristic slice. As a result, instead of evolving forward in retarded time, such variables are evolved internally within each characteristic slice. The null variables require the prescription of data on any surface that is transverse to the characteristic slicing.

As an example, consider the first-order equation

$$\phi_{,t} + \phi_{,r} = 0.$$

This equation can be solved uniquely by prescribing an arbitrary function of r as the value of ϕ for $t = 0$. This is the initial-value problem of this equation.

The charaecteristic surfaces of the equation are the lines along which $t - r$ takes a constant value. A change of coordinates $(t, r) \to (u, r)$ with $u = t - r$ results in the characteristic form of the equation:

$$\phi_{,r} = 0.$$

The characteristic form thus involves no derivative with respect to the retarded time u, this meaning that ϕ is a null variable. Not only does this equation preclude arbitrariness in the value of ϕ on any slice $u =$ constant, but precludes as well the evolution from one value of u to the next. A unique solution will be found, however, by prescribing an arbitrary function of u as the value of ϕ at any fixed value of r. This is the characteristic problem corresponding to the original equation.

The characteristic problem makes sense because the general solution of the original equation is $\phi = f(t - r)$ for arbitrary f. Thus ϕ takes a constant value along the characteristic surfaces, which can be prescribed by the values at either $t = 0$ or $r = 0$ (or, in fact, any surface that is transverse to $u = 0$). In a loose sense, the values of ϕ travel along the characteristic surface, thus, they travel at the characteristic speed of the surface.

As a more complete illustrative example at the same level of simplicity, consider now the case of a system

$$\phi_{,t} = \psi_{,r} \qquad \psi_{,t} = \phi_{,r}\,.$$

This system has two different sets of characteristic surfaces: the lines of constant value of $t - r$, as before, and also the lines of constant value of $t + r$. A change of coordinates $(t, r) \to (u, r)$ with $u = t - r$, as before, results in the following two equations,

$$\phi_{,u} = -\psi_{,u} + \psi_{,r} \qquad \psi_{,u} = -\phi_{,u} + \phi_{,r}\,.$$

which can be arranged to read, equivalently,

$$(\psi - \phi)_{,r} = 0 \qquad 2(\psi + \phi)_{,u} - (\psi + \phi)_{,r} = 0.$$

A redefinition of variables $q \equiv \psi + \phi$, $w \equiv \psi - \phi$ finally results in the characteristic form of the original system

$$w_{,r} = 0 \qquad 2q_{,u} - q_{,r} = 0.$$

The characteristic form has one retarded-time-dependent equation for q which is thus a normal variable, and one retarded-time-independent equation for w, the null variable. The system has

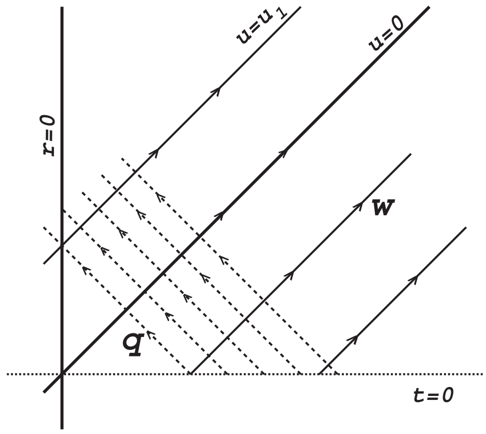

Figure 1. In the (r, t) plane, the characteristics of the wave equation $X,_{tt} - X,_{rr}$ are the two perpendicular sets of lines at 45° and 135° with the horizontal. The characteristic variable $w = X,_t - X,_r$ travels along the outgoing characteristics, whereas the characteristic variable $q = X,_t + X,_r$ travels inwardly along the second set of characteristics. Using $u = t - r$ as the chosen characteristic slicing, the evolution of the variable w becomes internal to the slices, whereas the evolution of the variable q takes place across slices. The characteristic problem can thus be solved by prescribing arbitrary data for q on the slice $u = 0$ and for w on the transverse surface $r = 0$.

a unique solution if one prescribes $q_0 = g(r)$ on a surface $u = 0$ and $w_0 = f(u)$ on a surface $r = 0$, both arbitrarily.

One can make sense of this characteristic problem by introducing a variable X such that $\phi = X,_r$ and $\psi = X,_t$. Then the original system translates into the standard unidimensional wave equation $X,_{tt} - X,_{rr} = 0$, with general solution $X = f(t - r) + g(t + r)$ for arbitrary functions f and g. In relation to X, we have $w = 2f'(t - r)$ and $q = 2g'(t + r)$, with a prime denoting the ordinary derivative. Clearly, the null variable w represents the outgoing wave travelling along the retarded-time slice at the speed of light, whereas the normal variable q represents the incoming wave, travelling across slices at the speed of light. The characteristic and initial-value problems for this example are illustrated in figure 1.

The examples make the point that the essence of a characteristic problem is the presence of equations that have no retarded-time derivative, which must occur because the equations must have, by definition, at least one characteristic variable associated with the chosen characteristic slicing.

The examples also hint at the fact, known for quite some time [2], that the equivalent of the Cauchy problem for hyperbolic systems in characteristic form can be solved: for each complete set of normal data and null data a solution exists and is unique (see [2] for analyticity assumptions). An issue that has attracted less attention is: if the data are perturbed slightly, under what circumstances is the variation of the solution under control? Equivalently, will nearly zero data evolve into a solution that is also close to zero?

We set up a framework in which to address this question by defining certain types of estimates of the solution in terms of the free data, after [3]. Subsequently we derive an algebraic criterion that is sufficient to determine whether the solutions satisfy such an *a priori* estimate, thus establishing their stability with respect to small variations of the free data. This kind of stability is of relevance to numerical applications. A prominent instance of the use of the characteristic problem for numerical applications is that of the simulation of gravitational waves by numerically integrating the characteristic formulation of the Einstein equations [4].

As argued extensively in [3], characteristic problems for which the estimate can be established may be considered to be *well posed* in the sense that for each set of data the solution exists, is unique and depends continuously on the free data. In addition, characteristic problems that satisfy the algebraic criterion developed here can be thought of as *manifestly well posed*. Manifest well-posedness in the sense defined here is to characteristic problems as symmetric hyperbolicity [5] is to initial-value problems.

Section 2 describes first-order linear characteristic problems after Duff [2]. The estimates of interest are defined in section 3 where the algebraic criterion is developed as well. A relationship between manifestly well posed characteristic problems and well posed initial-value problems for the same system of equations up to coordinate transformations is pointed out in section 4. Concluding remarks are offered in section 5.

2. Homogeneous linear characteristic problems in canonical form

Consider a generic homogeneous hyperbolic system of linear partial differential equations for m functions $v = (v^\alpha)$ of n variables y^a, which can be written in matrix form as follows,

$$\boldsymbol{A}^a \frac{\partial v}{\partial y^a} + \boldsymbol{D} v = 0, \tag{1}$$

where summation over repeated indices is understood. A characteristic surface $\mathcal{N}$ is a surface given by $\phi(y^a) = 0$ such that

$$\det\left(\boldsymbol{A}^a \frac{\partial \phi}{\partial y^a}\right) = 0. \tag{2}$$

Denote by m the multiplicity of this characteristic surface (so that the rank of $\boldsymbol{A}^a \partial\phi/\partial y^a$ is $n - m$). Suppose $\mathcal{T}$ given by $\psi(y^a) = 0$ is another surface intersecting $\mathcal{N}$ at a submanifold of dimension $n - 2$, whose further properties are to be determined. We choose a suitable coordinate system (u, x, x^i), $i = 1, \ldots, n-2$, for $\mathbb{R}^n$ adapted to these two surfaces; i.e., such that

$$u \equiv \phi(y^a), \tag{3a}$$

$$x \equiv \psi(y^a). \tag{3b}$$

In these coordinates (1) reads

$$\boldsymbol{B}^u \partial_u v + \boldsymbol{B}^x \partial_x v + \boldsymbol{B}^i \partial_i v + \boldsymbol{D} v = 0, \tag{4}$$

with

$$\boldsymbol{B}^u \equiv \boldsymbol{A}^a \frac{\partial \phi}{\partial y^a}, \tag{5a}$$

$$\boldsymbol{B}^x \equiv \boldsymbol{A}^a \frac{\partial \psi}{\partial y^a}, \tag{5b}$$

$$\boldsymbol{B}^i \equiv \boldsymbol{A}^a \frac{\partial x^i}{\partial y^a}. \tag{5c}$$

By (2), there are m linearly independent left null vectors $\tilde{z}_{(\nu)}$ and also m linearly independent right null vectors $z_{(\nu)}$ (with $\nu = 1, \ldots, m$) of the matrix $\boldsymbol{B}^u$, namely

$$\tilde{z}_{(\nu)} \boldsymbol{B}^u = 0, \tag{6a}$$

$$\boldsymbol{B}^u z_{(\nu)} = 0. \tag{6b}$$

We choose the right null vectors to be orthonormal in the sense that

$$z^{\alpha}_{(\nu)} z^{\alpha}_{(\mu)} = \delta_{\mu\nu}. \tag{7}$$

Multiplying (4) on the left with $\tilde{z}_{(\nu)}$ we find that m of the equations in the system do not involve derivatives with respect to u:

$$\tilde{z}_{(\nu)} \boldsymbol{B}^{x} \partial_x v + \tilde{z}_{(\nu)} \boldsymbol{B}^{i} \partial_i v + \tilde{z}_{(\nu)} \boldsymbol{D} v = 0. \tag{8}$$

Our aim is now to find a convenient transformation of variables that takes advantage of this split of the equations. We start by noting that, using the m right null vectors as the first m legs of an orthonormal basis e'_{α} of $\mathbb{R}^m$, we have a unitary transformation

$$e'_{\alpha} = S_{\alpha\beta} e_{\beta} \tag{9}$$

from the trivial basis $e_{\beta} = \{(1, 0, \ldots, 0), \ldots, (0, \ldots, 0, 1)\}$ to the new orthonormal basis, with $S_{\alpha\gamma} S_{\beta\gamma} = \delta_{\alpha\beta}$ and such that $S_{\nu\alpha} = z_{(\nu)\alpha}$ for $\nu = 1, \ldots, m$. The components of $v = v_{\alpha} e_{\alpha}$ in the new orthonormal basis are

$$v'_{\alpha} = S_{\alpha\beta} v_{\beta}. \tag{10}$$

In particular, the first m components are the scalar products of v with the right null vectors $z_{(\mu)}$, which we denote by w_{μ}

$$w_{\mu} \equiv z_{(\mu)\alpha} v_{\alpha} = v'_{\mu} \qquad \text{for} \quad \mu = 1, \ldots, m. \tag{11}$$

Multiplying (4) on the left with $\boldsymbol{S}$, the system transforms into

$$\boldsymbol{B}'^{a} \partial_a v' + \boldsymbol{D}' v' = 0. \tag{12}$$

From now on the index a refers to the characteristic coordinates, namely: $a = u, x, i$. The matrices have transformed according to $\boldsymbol{B}'^{a} = \boldsymbol{S}\boldsymbol{B}^{a}\boldsymbol{S}^{t}$ and $\boldsymbol{D}' = \boldsymbol{S}\boldsymbol{D}\boldsymbol{S}^{t}$, and $\boldsymbol{S}^{t}$ is the transpose of $\boldsymbol{S}$. Because $(\boldsymbol{B}^{u}\boldsymbol{S}^{t})_{\alpha\nu} = B^{u}_{\alpha\beta} S_{\nu\beta} = B^{u}_{\alpha\beta} z_{(\nu)\beta} = 0$ for $\nu \leqslant m$, the matrix $\boldsymbol{B}'^{u}$ has a Jordan form with all vanishing coefficients in the first m columns. This means that the u-derivatives of the m new variables w_{ν} are not involved, and consequently, the remaining variables v'_{ν} with $\nu > m$ are the normal variables of the problem, which we denote by $q = (q_{\mu})$, $\mu = m + 1, \ldots, n$. We have thus split the new fundamental variables into

$$v' = \boldsymbol{S} v \equiv (w_1, \ldots, w_m, q_{m+1}, \ldots, q_n). \tag{13}$$

Inverting (10) we have $v_{\alpha} = S_{\beta\alpha} v'_{\beta}$, which can be used into (8) to obtain a set of equations in the transformed variables:

$$\tilde{z}_{(\nu)} \boldsymbol{B}^{x} \boldsymbol{S}^{t} \partial_x v' + \tilde{z}_{(\nu)} \boldsymbol{B}^{i} \boldsymbol{S}^{t} \partial_i v' + \tilde{z}_{(\nu)} \boldsymbol{D} \boldsymbol{S}^{t} v' = 0. \tag{14}$$

We would like these equations to be solvable for the x-derivatives of all the variables w_{μ}, that is: the ones that do not evolve out of the initial characteristic surface. The first m terms in each equation for fixed ν are

$$\tilde{z}_{(\nu)\alpha} (\boldsymbol{B}^{x} \boldsymbol{S}^{t})_{\alpha\mu} \partial_x w_{\mu} = \tilde{z}_{(\nu)\alpha} B^{x}_{\alpha\beta} z_{(\mu)\beta} \partial_x w_{\mu}. \tag{15}$$

Thus the set of m equations (14) can be solved for the m variables $\partial_x w_{\mu}$ if and only if

$$\det \left(\tilde{z}_{(\nu)\alpha} B^{x}_{\alpha\beta} z_{(\mu)\beta} \right) \neq 0. \tag{16}$$

This is a restriction on the choice of $\psi(y^a)$. For this restriction to hold it is sufficient, but not necessary, that the level surfaces of $\psi(y^a)$ be non-characteristic. In many applications, the level surfaces of ψ are chosen to be timelike. For now on we assume that (16) holds. This allows us to interpret the m variables w_{ν} as the null variables of the problem.

We have shown that under very weak conditions for the surface $\mathcal{T}$, the most general characteristic problem takes the following form:

$$\boldsymbol{N}^{u} \partial_u q + \boldsymbol{N}^{x} \partial_x q + \boldsymbol{N}^{i} \partial_i v' + \boldsymbol{N}^{0} v' = 0 \tag{17a}$$

$$\partial_x w + \boldsymbol{L}^{x} \partial_x q + \boldsymbol{L}^{i} \partial_i v' + \boldsymbol{L}^{0} v' = 0. \tag{17b}$$

Clearly the null variables w can be redefined by $\widehat{w} \equiv w + \boldsymbol{L}^x q$ so that none of equations (17b) contains x-derivatives of the normal variables. Additionally, since $\boldsymbol{N}^u$ is non-singular, we can choose normal variables $\widehat{q} \equiv \boldsymbol{N}^u q$. In terms of these special choices of null and normal variables, equations (17a) and (17b) assume what is referred to as the *canonical form*:

$$\partial_u \widehat{q} + \widehat{\boldsymbol{N}}^x \partial_x \widehat{q} + \widehat{\boldsymbol{N}}^i \partial_i \widehat{v} + \widehat{\boldsymbol{N}}^0 \widehat{v} = 0 \tag{18a}$$

$$\partial_x \widehat{w} + \widehat{\boldsymbol{L}}^i \partial_i \widehat{v} + \widehat{\boldsymbol{L}}^0 \widehat{v} = 0 \tag{18b}$$

where $\widehat{v} \equiv (\widehat{w}, \widehat{q})$. We refer to (18$a$) as the *evolution equations*, and to (18b) as the *hypersurface equations*. For a unique solution to exist, one must prescribe the values of $\widehat{w}$ on the surface $x = 0$ and the values of $\widehat{q}$ on the surface $u = 0$. The solution can then be constructed in a hierarchical manner. Since q is a known source for (18b) at $u = 0$, then $\widehat{w}$ can be found on the entire surface $u = 0$. Once $\widehat{w}$ is known at $u = 0$, it can be used as a given source for (18a) in order to find the values of the normal variables $\widehat{q}$ on the next surface at $u = \mathrm{d}u$. These are then used into (18b) to obtain $\widehat{w}$ on the surface $u = \mathrm{d}u$. And so forth. In fact, Duff proves a theorem of existence and uniqueness of the solution given the canonical form of the characteristic problem [2].

As an example, consider the following equations for four unknowns v as functions of four variables $x^a = (t, x, y, z)$:

$$\partial_t v_1 = \partial_x v_2 + \partial_y v_3 + \partial_z v_4, \tag{19a}$$

$$\partial_t v_2 = \partial_x v_1, \tag{19b}$$

$$\partial_t v_3 = \partial_y v_1, \tag{19c}$$

$$\partial_t v_4 = \partial_z v_1. \tag{19d}$$

These equations constitute a first-order version of the wave equation in three spatial dimensions (if we interpret the variables v_α as the derivatives of a single function f). However, this first-order version of the wave equation has characteristic speeds of 0 (rest) in addition to 1 (light). The level surfaces of $\phi \equiv t - x$ are null planes, so they are characteristic and intersect the surfaces of fixed value of x. We change coordinates $(t, x, y, z) \to (u, x, y, z)$ with

$$u = t - x \tag{20}$$

which implies that $\partial_t \to \partial_u$ and $\partial_x \to \partial_x - \partial_u$. System (19) turns into

$$\partial_u v^1 + \partial_u v^2 = \partial_x v^2 + \partial_y v^3 + \partial_z v^4, \tag{21a}$$

$$\partial_u v^2 + \partial_u v^1 = \partial_x v^1, \tag{21b}$$

$$\partial_u v^3 = \partial_y v^1, \tag{21c}$$

$$\partial_u v^4 = \partial_z v^1. \tag{21d}$$

We can read off the matrix $\boldsymbol{B}^u$:

$$\boldsymbol{B}^u = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{22}$$

which is obviously singular of rank 3, so we have $m = 1$ in this example. We expect only one null variable, and three normal variables for this problem in canonical form. $\boldsymbol{B}^u$ has only one right null vector $z = 2^{-1/2}(1, -1, 0, 0)$, and only one left null vector $\tilde{z} = 2^{-1/2}(1, -1, 0, 0)$, which coincides with z because $\boldsymbol{B}^{\boldsymbol{u}}$ is symmetric. An orthonormal basis for $\mathbb{R}^4$ can be chosen as $\{2^{-1/2}(1, -1, 0, 0), 2^{-1/2}(1, 1, 0, 0), (0, 0, 1, 0), (0, 0, 0, 1)\}$. So the unitary transformation is

$$\boldsymbol{S} = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 & 0 \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{23}$$

The null variable of the problem is $w = (v_1 - v_2)/\sqrt{2}$ and the normal variables are $q_v = ((v_1 + v_2)/\sqrt{2}, v_3, v_4)$, in terms of which system (21) takes the almost canonical form:

$$2\partial_u q_2 - \partial_x q_2 - \frac{1}{\sqrt{2}}\partial_y q_3 - \frac{1}{\sqrt{2}}\partial_z q_4 = 0, \tag{24a}$$

$$\partial_u q_3 - \frac{1}{\sqrt{2}}\partial_y q_2 - \frac{1}{\sqrt{2}}\partial_y w = 0, \tag{24b}$$

$$\partial_u q_4 - \frac{1}{\sqrt{2}}\partial_z q_2 - \frac{1}{\sqrt{2}}\partial_z w = 0, \tag{24c}$$

$$\partial_x w - \frac{1}{\sqrt{2}}\partial_y q_3 - \frac{1}{\sqrt{2}}\partial_z q_4 = 0. \tag{24d}$$

For a unique solution, we need to prescribe the value of w on the surface $x = 0$, and the values of q_2, q_3 and q_4 on the surface $u = 0$. Note that, in this example, the surface $x = 0$ is not timelike with respect to the hyperbolic operator $\boldsymbol{A}^{\boldsymbol{a}}$, but is also characteristic, as can be seen by inspection of the matrix $\boldsymbol{B}^{\boldsymbol{x}}$:

$$\boldsymbol{B}^{\boldsymbol{x}} = \begin{pmatrix} 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \tag{25}$$

However, we have $\tilde{z}\boldsymbol{B}^{\boldsymbol{x}}z = 2 \neq 0$. Therefore condition (16) is satisfied in spite of the fact that the surfaces of fixed value of x are not timelike.

3. Well-posedness of homogeneous linear characteristic problems

The canonical system (18) can be written in the compact form

$$\boldsymbol{C}^{\boldsymbol{a}}\partial_a v + \boldsymbol{D}v = 0 \tag{26}$$

where $\boldsymbol{C}^{\boldsymbol{u}}$ and $\boldsymbol{C}^{\boldsymbol{x}}$ have block-diagonal forms of a special type:

$$\boldsymbol{C}^{\boldsymbol{u}} = \begin{pmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}, \qquad \boldsymbol{C}^{\boldsymbol{x}} = \begin{pmatrix} \boldsymbol{N}^{\boldsymbol{x}} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} \end{pmatrix}, \tag{27}$$

where $\mathbf{1}$ is the identity of dimension $n - m$ in the case of $\boldsymbol{C}^{\boldsymbol{u}}$, and of dimension m in the case of $\boldsymbol{C}^{\boldsymbol{x}}$. The matrix $\boldsymbol{N}^{\boldsymbol{x}}$ is square, of dimension $n - m$, and the various rectangular blocks $\mathbf{0}$ are vanishing matrices whose dimensions are clear from the context. Dropping the hats ($\widehat{\ }$) for ease of notation, the variable v represents the set of normal variables q and null variables w of the characteristic problem in canonical form. Additionally, m functions $w_0 \equiv \left(w_0^\nu(u, x^i)\right)$ are given as data on the surface $\mathcal{T}$ and $n - m$ functions $q_0 \equiv \left(q_0^\nu(x, x^i)\right)$ are given as data on the surface $\mathcal{N}$.

For the remainder of this section, we make the strong assumption that $\boldsymbol{N}^x$ and $\boldsymbol{C}^i$ are symmetric. Multiplication of (26) by v on the left then leads to a 'conservation law' of the form

$$\partial_a(v\boldsymbol{C}^a v) + v\boldsymbol{R}v = 0 \tag{28}$$

where $\boldsymbol{R} \equiv 2\boldsymbol{D} - \partial_a \boldsymbol{C}^a$. We now integrate this conservation law in an appropriate volume $\mathcal{V}$ of $\mathbb{R}^n$. Our volume is a 'hyperprism' limited by the surface $u = 0$ from 'below', the surface $x = 0$ on the 'left', and the surface $u + x = T$, for an arbitrary constant T, on the 'top'. We assume there are no boundaries in the remaining coordinate directions, in the sense that the solutions v will either be periodic functions of x^i or will decay sufficiently fast at large values of x^i in order for the integrals of their squares to exist. The integration yields

$$\int_{\Sigma_T} v(\boldsymbol{C}^u + \boldsymbol{C}^x)v\,\mathrm{d}\Sigma_T - \int_{\mathcal{N}}(v\boldsymbol{C}^u v)\,\mathrm{d}\mathcal{N} - \int_{\mathcal{T}}(v\boldsymbol{C}^x v)\,\mathrm{d}\mathcal{T} + \int_{\mathcal{V}} v\boldsymbol{R}v\,\mathrm{d}\mathcal{V} = 0. \tag{29a}$$

Clearly

$$\int_{\mathcal{N}} v\boldsymbol{C}^u v\,\mathrm{d}\mathcal{N} = \int_{\mathcal{N}} \sum_{\nu=m+1}^{n} \left(q_0^\nu\right)^2 \mathrm{d}\mathcal{N} \equiv \|q_0\|^2, \tag{30}$$

and

$$\int_{\mathcal{T}} v\boldsymbol{C}^x v\,\mathrm{d}\mathcal{T} = \int_{\mathcal{T}} q\boldsymbol{N}^x q\,\mathrm{d}\mathcal{T} + \int_{\mathcal{T}} \sum_{\nu=1}^{m} \left(w_0^\nu\right)^2 \mathrm{d}\mathcal{T}$$

$$\equiv \int_{\mathcal{T}} q\boldsymbol{N}^x q\,\mathrm{d}\mathcal{T} + \|w_0\|^2 \tag{31a}$$

Thus equation (29) is rearranged to read

$$\int_{\Sigma_T} v(\boldsymbol{C}^u + \boldsymbol{C}^x)v\,\mathrm{d}\Sigma_T = \|q_0\|^2 + \|w_0\|^2 + \int_{\mathcal{T}} q\boldsymbol{N}^x q\,\mathrm{d}\mathcal{T} - \int_{\mathcal{V}} v\boldsymbol{R}v\,\mathrm{d}\mathcal{V}. \tag{32a}$$

If $\boldsymbol{N}^x$ is non-positive definite but also such that $\boldsymbol{1} + \boldsymbol{N}^x$ is positive definite, we can define the norm of the solution v on the surface Σ_T by

$$\|v\|_T^2 \equiv \int_{\Sigma_T} v(\boldsymbol{C}^u + \boldsymbol{C}^x)v\,\mathrm{d}\Sigma_T, \tag{33}$$

and equation (32) implies

$$\|v\|_T^2 \leqslant \|q_0\|^2 + \|w_0\|^2 - \int_{\mathcal{V}} v\boldsymbol{R}v\,\mathrm{d}\mathcal{V}. \tag{34}$$

In special case of constant coefficients with no undifferentiated terms, namely $\boldsymbol{R} = \boldsymbol{0}$, equation (34) takes the form

$$\|v\|_T^2 \leqslant \|q_0\|^2 + \|w_0\|^2, \tag{35}$$

which represents an *a priori* estimate of the solution in terms of the free data. It implies that the 'size' of the solution is controlled by the 'size' of the data. We may interpret it as a statement of well-posedness of the characteristic problem. Clearly the estimate holds in the presence of non-constant coefficients and undifferentiated terms as long as $\boldsymbol{R}$ is non-negative definite.

An estimate can still be drawn in the presence of a negative definite bounded $\boldsymbol{R}$, but it is weaker and holds only for small values of T, as we show next.

Since $\boldsymbol{R}$ is negative definite, then

$$-v\boldsymbol{R}v \leqslant r\sum_{\nu=1}^{n}(v^\nu)^2 \tag{36}$$

where $r = \max(|R_{ij}|)$ in the volume $\mathcal{V}$, assuming that such a number r exists. On the other hand, since $\boldsymbol{C}^{\boldsymbol{u}} + \boldsymbol{C}^{\boldsymbol{x}}$ is positive definite and symmetric then all its eigenvalues are positive and we have

$$v(\boldsymbol{C}^{\boldsymbol{u}} + \boldsymbol{C}^{\boldsymbol{x}})v \geqslant c \sum_{\nu=1}^{n} (v^{\nu})^2 \tag{37}$$

with c being the smallest eigenvalue of $\boldsymbol{C}^{\boldsymbol{u}} + \boldsymbol{C}^{\boldsymbol{x}}$. This implies

$$-v\boldsymbol{R}v \leqslant \frac{r}{c} v(\boldsymbol{C}^{\boldsymbol{u}} + \boldsymbol{C}^{\boldsymbol{x}})v. \tag{38}$$

Thus

$$-\int_{\mathcal{V}} v\boldsymbol{R}v \, \mathrm{d}\mathcal{V} \leqslant \frac{r}{c} \int_0^T \|v\|_t^2 \, \mathrm{d}t \tag{39}$$

where $\|v\|_t^2$ is the norm of the solution on the surface Σ_t given by $u + x = t$ for fixed value of $t < T$. Thus inequality (34) implies

$$\|v\|_T^2 \leqslant \|q_0\|^2 + \|w_0\|^2 + \frac{r}{c} \int_0^T \|v\|_t^2 \, \mathrm{d}t. \tag{40}$$

Here $\|q_0\|^2$ and $\|w_0\|^2$ are the norms of the normal and null variables with respect to the surfaces $\mathcal{N}$ and $\mathcal{T}$ both bounded by the spatial surface at $u + x = T$. For any value of $t \leqslant T$ we can write the same inequality

$$\|v\|_t^2 \leqslant \int_{\mathcal{N}_t} \sum (q^{\nu})^2 \, \mathrm{d}\mathcal{N}_t + \int_{\mathcal{T}_t} \sum (w^{\nu})^2 \, \mathrm{d}\mathcal{T}_t + \frac{r}{c} \int_0^t \|v\|_{t'}^2 \, \mathrm{d}t' \tag{41}$$

where $\mathcal{N}_t$ and $\mathcal{T}_t$ are the subsets of $\mathcal{N}$ and $\mathcal{T}$ bounded by Σ_t, respectively. Since both integrals indicated are less than the norms $\|q_0\|^2$ and $\|w_0\|^2$ respectively, this implies

$$\|v\|_t^2 \leqslant \|q_0\|^2 + \|w_0\|^2 + \frac{r}{c} \int_0^t \|v\|_{t'}^2 \, \mathrm{d}t'. \tag{42}$$

Using this inequality recursively into the right-hand side of (40) we have

$$\begin{aligned} \|v\|_T^2 \leqslant & \left(1 + \frac{rT}{c} + \frac{(rT)^2}{2c^2} + \cdots + \frac{(rT)^j}{j!c^j}\right) \\ & \times \left(\|q_0\|^2 + \|w_0\|^2 + (r/c)^{j+1} \int_0^T \mathrm{d}t_1 \int_0^{t_1} \mathrm{d}t_2 \cdots \int_0^{t_j} \|v\|_{t_{j+1}}^2 \, \mathrm{d}t_{j+1}\right) \end{aligned} \tag{43}$$

for any given non-negative integer j. In the limit for $j \to \infty$ the sequence on the right-hand side converges if $(rT/c) < 1$, in which case we have

$$\|v\|_T^2 \leqslant \mathrm{e}^{(r/c)T} (\|q_0\|^2 + \|w_0\|^2). \tag{44}$$

This is our final estimate for the solution in terms of the free data on the surfaces $\mathcal{N}$ and $\mathcal{T}$. The estimate involves an exponential factor essentially due to the presence of undifferentiated terms. The exponential factor depends on the properties of the system of equations (the principal matrices and the undifferentiated terms), but not on the choice of data. This is analogous to the *a priori* estimates for Cauchy problems with undifferentiated terms. As usual in such cases, the estimate is useless for large T, and, in particular, our proof only guarantees the estimate for $T < c/r$. Perhaps with greater care the estimate can be extended to longer values of T.

Because the *a priori* estimates (44) are independent of the choice of data, we can say that our characteristic problem is well-posed. The conditions under which we are able to derive *a priori* estimates thus become our *criteria for well-posedness of linear homogeneous characteristic problems in canonical form:*

(i) The principal matrices $\boldsymbol{C}^{a}$ are symmetric.
(ii) The normal block of the principal x-matrix, denoted $\boldsymbol{N}^{x}$, is non-positive definite but such that $\mathbf{1} + \boldsymbol{N}^{x}$ is positive definite (namely, $-\mathbf{1} < \boldsymbol{N}^{x} \leqslant \mathbf{0}$).

There is, clearly, no obstacle in generalizing the construction slightly to the case where the characteristic problem is cast into 'almost' canonical form, namely, the case when the principal matrices are

$$\boldsymbol{C}^{u} = \begin{pmatrix} \boldsymbol{N}^{u} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}, \qquad \boldsymbol{C}^{x} = \begin{pmatrix} \boldsymbol{N}^{x} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} \end{pmatrix}, \tag{45}$$

which corresponds to a strictly canonical form up to a transformation of normal variables among themselves. In this case, the criterion is

(i) The principal matrices $\boldsymbol{C}^{a}$ are symmetric and
(ii) The normal block of the principal u-matrix, denoted by $\boldsymbol{N}^{u}$, is positive definite. The normal block of the principal x-matrix, denoted by $\boldsymbol{N}^{x}$, is non-positive definite but such that $\boldsymbol{N}^{u} + \boldsymbol{N}^{x}$ is positive definite (namely, $-\boldsymbol{N}^{u} < \boldsymbol{N}^{x} \leqslant \mathbf{0}$).

If (i) and (ii) hold for a linear homogeneous characteristic problem in 'almost' canonical form, then the problem is well posed in the sense that there exist *a priori* estimates of the kind $\|v\|_T^2 \leqslant \mathrm{e}^{KT}(\|q_0\|^2 + \|w_0\|^2)$, where K is a constant independent of the data. This inequality is sufficient to establish the stability of the solutions under small variations of the data. Note that $-\boldsymbol{N}^{u} < \boldsymbol{N}^{x} \leqslant \mathbf{0}$ is equivalent to the requirement that the surfaces given by $\phi(y^a) + \psi(y^a) = T$ with fixed value of T are spatial with respect to the hyperbolic operator $\boldsymbol{A}^{a}$, which in turn means that they can be interpreted as the level surfaces of a time function $t(y^a) \equiv \phi(y^a) + \psi(y^a)$.

As an illustration, we can see that the first-order form of the wave equation, equations (24), is well posed. For equations (24) we have

$$\boldsymbol{N}^{u} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad \boldsymbol{N}^{x} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \tag{46}$$

and also

$$\boldsymbol{C}^{y} = -\frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}, \qquad \boldsymbol{C}^{z} = -\frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}. \tag{47}$$

Therefore all the conditions are satisfied, and the estimates follow. In this case, the estimates are of the form (35) and hold for any chosen T because all the principal matrices have constant coefficients and there are no undifferentiated terms ($\boldsymbol{R} = \mathbf{0}$).

4. Relationship between the properties of well-posedness of characteristic and associated initial-value problems

The following question is of general interest, as well as of particular interest to numerical applications of general relativity. If a given characteristic problem is manifestly well posed in the sense of section 3, does it follow that the corresponding initial-value problem is well posed?

The answer is yes. We have the following theorem:

Theorem 1. *Given a manifestly well-posed characteristic problem, there is a choice of a time coordinate for which the associated initial value problem can be cast into symmetric hyperbolic form.*

The proof is by construction. Suppose we have a system of equations in almost canonical characteristic form

$$\boldsymbol{C}^{\boldsymbol{a}} v_{,a} + \boldsymbol{D} v = 0$$

with $\boldsymbol{C}^{\boldsymbol{a}}$ satisfying the criteria (i) and (ii) of the previous section for manifest well posedness. Consider the coordinate transformation $x^a = (u, x, x^i) \to (t, x, x^i) \equiv y^a$ where the time coordinate is $t = u + x$. With this change of coordinates the system of equations becomes

$$\boldsymbol{A}^{\boldsymbol{t}} v_{,t} + \boldsymbol{A}^{\boldsymbol{x}} v_{,x} + \boldsymbol{A}^{\boldsymbol{i}} v_{,i} + \boldsymbol{D} v = 0$$

where

$$\boldsymbol{A}^{\boldsymbol{t}} = \boldsymbol{C}^{\boldsymbol{u}} + \boldsymbol{C}^{\boldsymbol{x}} \qquad \boldsymbol{A}^{\boldsymbol{x}} = \boldsymbol{C}^{\boldsymbol{x}} \qquad \boldsymbol{A}^{\boldsymbol{i}} = \boldsymbol{C}^{\boldsymbol{i}}.$$

Since the matrices $\boldsymbol{C}^{\boldsymbol{a}}$ are symmetric, so are the matrices $\boldsymbol{A}^{\boldsymbol{a}}$. Furthermore, since $\boldsymbol{C}^{\boldsymbol{u}}+\boldsymbol{C}^{\boldsymbol{x}}$ is positive definite, so is $\boldsymbol{A}^{\boldsymbol{t}}$. Thus the system of equations is now symmetric hyperbolic and the initial-value problem is well posed. This completes the proof.

The converse is also true:

Theorem 2. *Given a symmetric hyperbolic system of partial differential equations with a surface-forming characteristic covector field, there is a choice of a coordinate x on the characteristic slices for which the characteristic problem is manifestly well posed.*

A symmetric hyperbolic system of PDEs in coordinates y^a is given by

$$\boldsymbol{A}^{\boldsymbol{a}} v_{,a} + \boldsymbol{D} v = 0$$

where the matrices $\boldsymbol{A}^{\boldsymbol{a}}$ are symmetric and one of them which we denoted here by $\boldsymbol{A}^{\boldsymbol{t}}$ is positive definite. A symmetric hyperbolic system is hyperbolic in the standard sense, that is: at any given point it admits covectors $\xi_a = (\xi_t, \vec{\xi})$ that solve the characteristic equation

$$\det(\boldsymbol{A}^{\boldsymbol{t}} \xi_a) = 0$$

with real ξ_t and arbitrary $\vec{\xi}$. By a surface-forming characteristic covector we mean a field of characteristic covectors $\xi_a = \xi_a(y^b)$ such that

$$\xi_a = \frac{\partial \phi}{\partial y^a}$$

for some scalar function $\phi(y^a)$. The characteristic covector would thus be normal to the level surfaces of the function ϕ, which would thus be interpreted as the retarded time. Surface-forming characteristic covectors exist for any symmetric hyperbolic system with constant coefficients, in which case they can be chosen to be constant, which results in retarded times $\phi(y^b) = \xi_a y^a$ that are linear functions of the coordinates and can be viewed as null planes.

If the system has variable coefficients then the characteristic equation prescribes ξ_t as a function of the coordinates y^a as well as of $\vec{\xi}$. In other words, we have an explicit function $\xi_t = \xi_t(y^a, \vec{\xi})$ determined by a root of the characteristic equation. Substituting $\xi_a = \phi_{,a}$ into this expression we obtain a partial differential equation for the retarded time ϕ of the form

$$\phi_{,t} = \xi_t(y^a, \nabla\phi)$$

where ∇ is the gradient with respect to all the spatial coordinates. The characteristic covector field is surface forming if this equation for the retarded time ϕ can be solved. In general this will be the case, at least within a small neighbourhood.

To set up a characteristic problem we take $u = \phi(y^a)$ as the retarded time coordinate, and define $x = t - \phi(y^a)$ as the coordinate on the surfaces of fixed value of u. The coordinate transformation is completed in a trivial manner by assigning the names of x^i to the $n-2$ coordinates y^i that are linearly independent of u and x. We have thus

$$u = \phi(y^a) \qquad x = t - \phi(y^a) \qquad x^i = y^i.$$

By the procedure delineated in section 2 the system of equations is cast into almost characteristic form

$$\boldsymbol{B}'^a v'_{,a} + \boldsymbol{D}' v' = 0$$

with symmetric matrices $\boldsymbol{B}'^a$ related to symmetric matrices $\boldsymbol{B}^a$ by a unitary transformation. Given our choice of coordinates we have

$$\boldsymbol{B}^u = \boldsymbol{A}^a \phi_{,a}$$

and also

$$\boldsymbol{B}^x = \boldsymbol{A}^a \frac{\partial x}{\partial y^a} = \boldsymbol{A}^a \left(\frac{\partial t}{\partial y^a} - \phi_{,a} \right) = \boldsymbol{A}^t - \boldsymbol{B}^u$$

Thus the matrices $\boldsymbol{B}^u$ and $\boldsymbol{B}^x$ satisfy $\boldsymbol{B}^u + \boldsymbol{B}^x = \boldsymbol{A}^t$, which is positive definite, and since a unitary transformation respects the positive-definiteness of a matrix, then $\boldsymbol{B}'^u + \boldsymbol{B}'^x$ is positive definite as well and the characteristic problem is manifestly well posed. This completes the proof.

5. Concluding remarks

Characteristic problems for hyperbolic equations are rarely discussed in the literature. In fact, prior to Balean's work [6–8], practically nothing was known about the characteristic problem of the simplest hyperbolic equation, that is, the wave equation. Balean discussed how to derive estimates for the solutions of the wave equation in its standard second-order form. Balean's estimates differ markedly from ours. The estimates for general linear characteristic problems of the first order that we present here constitute a direct generalization of the estimates that we recently derived for the particular case of solutions of the characteristic problem of the wave equation as a first-order system of PDEs [3].

The value of the generalization that we present here resides in the formulation of algebraic criteria sufficient for the existence of the *a priori* estimates. We demonstrate elsewhere [9] that these criteria allow us to formulate the characteristic problem of the linearized Einstein equations in a form that is guaranteed to be well posed.

Several issues of interest remain wide open. First, given a general characteristic problem that is well posed in the sense that we introduce here, it is not at all clear as yet whether estimates of the derivatives of the solution in terms of the derivatives of the data would exist as well. We have succeeded in deriving estimates for the derivatives in the particular case of the characteristic problem of the wave equation [3]. However, the derivation depends strongly on the particular form of the hyperbolic operator of the wave equation, and its generalization to arbitrary characteristic problems is far from straightforward, quite unfortunately.

Secondly, a sufficient criterion to establish well posedness of a characteristic problem is useful, but a necessary criterion would, perhaps, be invaluable as a means to rule out unstable problems with an eye towards numerical applications.

Thirdly, we have shown that under very general conditions a first-order system of PDEs admits a manifestly well-posed characteristic problem if and only if it admits a symmetric hyperbolic initial-value problem. But whether or not all well-posed hyperbolic problems admit well-posed characteristic problems in our sense might well be the most intriguing open question at this time.

Acknowledgments

I am greatly indebted to Roberto Gómez for numerous conversations. This work was supported by the NSF under grants nos PHY-9803301, PHY-0070624 and PHY-0244752 to Duquesne University.

References

[1] Courant R and Hilbert D 1962 *Methods of Mathematical Physics* vol II (New York: Interscience)
[2] Duff G F D 1958 *Can. J. Math.* **10** 127
[3] Frittelli S 2004 *J. Phys. A: Math. Gen.* **37** 8639 (*Preprint* math-ph/0408007)
[4] Winicour J 2001 *Living Rev. Relativ.* **4** 3 (http://www.livingreviews.org)
[5] Gustaffson B, Kreiss H-O and Oliger J 1995 *Time-Dependent Problems and Difference Methods* (New York: Wiley)
[6] Balean R M 1996 *PhD Thesis* University of New England, Armidale, NSW
[7] Balean R 1997 *Commun. PDE* **22** 1325
[8] Balean R and Bartnik R 1998 *Proc. R. Soc. Lond.* A **454** 2041
[9] Frittelli S 2005 *Phys. Rev.* D **71** 024021