Eulerian-Lagrangian Localized Adjoint Method: The Theoretical Framework

Ismael Herrera
Instituto de Geofísica, UNAM, Apdo. Postal 22-582, 14000 Mexico, Distrito Federal, Mexico

Richard E. Ewing
College of Science, Texas A&M University, College Station, Texas 77843-3257

Michael A. Celia
Water Resources Program, Department of Civil Engineering and Operations Research, Princeton University, New Jersey 08544

Thomas F. Russell
Department of Mathematics, University of Colorado at Denver, Denver, Colorado 80204-5300

Received 1 October 1992; approved without modification 15 October 1992

This is the second of a sequence of papers devoted to applying the localized adjoint method (LAM), in space-time, to problems of advective-diffusive transport. We refer to the resulting methodology as the Eulerian-Lagrangian localized adjoint method (ELLAM). The ELLAM approach yields a general formulation that subsumes many specific methods based on combined Lagrangian and Eulerian approaches, so-called characteristic methods (CM). In the first paper of this series the emphasis was placed in the numerical implementation and a careful treatment of implementation of boundary conditions was presented for one-dimensional problems. The final ELLAM approximation was shown to possess the conservation of mass property, unlike typical characteristic methods. The emphasis of the present paper is on the theoretical aspects of the method. The theory, based on Herrera's algebraic theory of boundary value problems, is presented for advection-diffusion equations in both one-dimensional and multidimensional systems. This provides a generalized ELLAM formulation. The generality of the method is also demonstrated by a treatment of systems of equations as well as a derivation of mixed methods.

I. INTRODUCTION

This is the second of a sequence of papers devoted to the application of the localized adjoint method (LAM) to problems of advection-diffusive transport [1]. The LAM is a new and promising methodology for discretizing partial differential equations. It is based on Herrera's algebraic theory of boundary value problems [2–6]. Applications have been made successively to ordinary differential equations, for which highly accurate algorithms

Numerical Methods for Partial Differential Equations, 9, 431–457 (1993)

CCC 0749-159X/93/040431-27

were developed [5, 7–9], multidimensional steady-state problems [10], and optimal spatial methods for advection-diffusion equations [11–18]. Finally, this article and the companion one [1] (to be referred to as Paper I) provide generalizations of characteristic methods that we refer to as the Eulerian-Lagrangian localized adjoint method (ELLAM). Related work has been published separately by some of the authors [19–22] and some more specific applications have already been made [23–28].

The numerical solution of the advective-diffusive transport equation is a problem of great importance because many problems in science and engineering involve such mathematical models. When the process is advection dominated the problem is especially difficult. In Paper I a brief review of the methods available was presented, from which we draw here. The methods derive from two main approaches: optimal spatial methods (OSM) and characteristic methods (CM).

The first of these procedures employs an Eulerian approach and develops an accurate solution of the spatial problem. For example, in the pioneering work of Allen and Southwell [29], a finite difference approximation was developed for the advection and diffusion terms that gives exact nodal values for the simplified case of one-dimensional, steady-state, constant-coefficient advective-diffusive transport without sources, sinks, or reaction terms. More general and systematic results in this direction have been developed using the LAM approach [5, 7–10]. However, this kind of approximations tend to be ineffective in transient simulations because of the strong influence of the time derivative. The salient features of this class of approximations may be summarized as follows: (i) Time truncation error dominates the solutions; (ii) Solutions are characterized by significant numerical diffusion and some phase errors; and (iii) The Courant number ($\mathrm{Cu} \equiv V\Delta t/\Delta x$) is generally restricted to be less than 1, and sometimes much less than 1. A general comparison of some of these methods was provided by Bouloutas and Celia [30].

Other Eulerian methods can be developed that perform significantly better than OSM approximations. These methods attempt to use the nonzero spatial truncation error (thereby differing from OSM) to cancel temporal errors and thereby reduce the overall truncation error (see, for example, [31–33]). While improved accuracy results from these formulations, they still suffer from strict Courant number limitations.

Finally, characteristic methods include many related approximation techniques that are called by a variety of names, such as Eulerian-Lagrangian methods (ELM) [34–37], transport diffusion method [38], characteristic Galerkin methods [39], method of characteristics (MOC) [40], modified method of characteristics (MMOC) [41–43], and operator splitting methods [44–46]. Each of these methods has in common the feature that the advective component is treated by a characteristic tracking algorithm (a Langrangian frame of reference) and the diffusive step is treated separately using a more standard (Eulerian) spatial approximation. These methods have the significant advantage that Courant number restrictions of purely Eulerian methods are alleviated because of the Langrangian nature of the advection step. Furthermore, because the spatial and temporal dimensions are coupled through the characteristic tracking, the influence of time truncation error present in OSM approximations is greatly reduced.

This paper and its companion [1] provide a generalization of characteristic methods using the approach of the localized adjoint method. In Paper I, a specific space-time LAM formulation that naturally leads to generalized CM approximations (to be referred to as Eulerian-Lagragian LAM:ELLAM) was introduced that is consistent and does not rely on any operator splitting or equation decomposition. In addition, all relevant boundary terms arise naturally in the formulation; a systematic and complete treatment of boundary conditions was presented which was shown to possess the global mass-conservation

property. The numerical treatment and boundary conditions implementation was developed with considerable detail.

The present Paper II dwells more on the general LAM theory, placing ELLAM more thoroughly in the general perspective of LAM. This contributes to a more complete picture of the possibilities that should be explored and the problems that must be tackled in order to make ELLAM a more effective modeling tool. The present paper begins by reviewing Herrera's algebraic theory as well as the LAM procedure. Then a general discussion of the approach is presented, including illustrations of its application to ordinary differential equations. After this, the LAM formulation of the space-time of advection-diffusion equation is derived in one space dimension, and the choice of continuity and boundary conditions to be satisfied by test functions is briefly discussed. The ELLAM formulation is then extended to several space dimensions. To illustrate the generality of LAM, the manner in which systems of equations are incorporated in this framework is explained, and as an example, mixed methods are developed in this setting.

II. GENERAL BACKGROUND

In this section, Herrera's algebraic theory of boundary value problems [2–6] is briefly explained.

Consider a region Ω and the linear spaces D_1 and D_2 of trial and test functions defined in Ω, respectively. Assume further that functions belonging to D_1 and D_2 may have jump discontinuities across some internal boundaries whose union will be denoted by Σ. For example, in applications of the theory to finite element methods, the set Σ could be the union of all the interelement boundaries. In this setting the general boundary value problem to be considered is one with prescribed jumps across Σ. The differential equation is

$$\mathcal{L}u = f_\Omega \quad \text{in } \Omega\,, \tag{2.1}$$

where Ω may be a purely spatial region or, more generally, a space-time region. Certain boundary and jump conditions are specified on the boundary $\partial\Omega$ of Ω and on Σ, respectively. When Ω is a space-time region, such conditions generally include initial conditions. In the literature on mathematical modeling of macroscopic physical systems, there are a variety of examples of initial-boundary value problems with prescribed jumps. To mention just one, problems of elastic wave diffraction can be formulated as such [47, 48]. The jump conditions that the sought solution must satisfy across Σ, in order to define a well posed problem, depend on the specific application and on the differential operator considered. For example, for elliptic problems of second order, continuity of the sought solution and its normal derivative is usually required, but the problem in which the solution and its normal derivative jump across Σ in a prescribed manner is also well posed [48].

The definition of a formal disjoint requires that a differential operator $\mathcal{L}$ and its formal adjoint $\mathcal{L}^*$ satisfy the condition that $w\mathcal{L}u - u\mathcal{L}^*w$ be a divergence, i.e.,

$$w\mathcal{L}u - u\mathcal{L}^*w = \nabla \cdot \{\underline{\mathfrak{D}}(u, w)\} \tag{2.2}$$

for a suitable vector-valued bilinear function $\underline{\mathfrak{D}}(u, w)$. Integration of Eq. (2.2) over Ω and application of the generalized divergence theorem [49] yields

$$\int_\Omega \{w\mathcal{L}u - u\mathcal{L}^*w\}\,dx = \int_{\partial\Omega} \mathcal{R}_\partial(w, u)\,dx + \int_\Sigma \mathcal{R}_\Sigma(u, w)\,dx \tag{2.3}$$

where

$$\mathcal{R}_\partial(u,w) = \underline{\mathfrak{D}}(u,w)\cdot\underline{n} \quad \text{and} \quad \mathcal{R}_\Sigma(u,w) = -[\underline{\mathfrak{D}}(u,w)]\cdot\underline{n}\,. \tag{2.4}$$

Here the square brackets stand for the "jumps" across Σ of the function contained inside, i.e., limit on the positive side minus limit on the negative side. Here, as in what follows, the positive side of Σ is chosen arbitrarily and then the unit normal vector $\underline{n}$ is taken pointing towards the positive side of Σ. Observe that generally $\mathcal{L}u$ will not be defined on Σ, since u and its derivatives may be discontinuous. Thus, in this article, it is understood that integrals over Ω are carried out excluding Σ. Consequently, differential operators will always be understood in an elementary sense and not in a distributional sense.

In the general theory of partial differential equations, Green's formulas are used extensively. For the construction of such formulas it is standard to introduce a decomposition of the bilinear function $\mathcal{R}_\partial$ (see, for example, Lions and Magenes [50]). Indicating transposes of bilinears forms by means of an asterisk, the general form of such decomposition is

$$\mathcal{R}_\partial(u,w) \equiv \underline{\mathfrak{D}}(u,w)\cdot\underline{n} = \mathfrak{B}(u,w) - \mathfrak{C}^*(u,w)\,, \tag{2.5}$$

where $\mathfrak{B}(u,w)$ and $\mathfrak{C}(w,u) = \mathfrak{C}^*(u,w)$ are two bilinear functions. When considering initial-boundary value problems, the definitions of these bilinear forms depend on the type of boundary and initial conditions to be prescribed. A basic property required of $\mathfrak{B}(u,w)$ is that for any u that satisfies the prescribed boundary and initial conditions, $\mathfrak{B}(u,w)$ is a well-defined linear function of w, independent of the particular choice of u. This linear function will be denoted by g_∂ [thus its value for any given function w will be $g_\partial(w)$], and the boundary conditions can be specified by requiring that $\mathfrak{B}(u,w) = g_\partial(w)$ for every $w \in D_2$ (or more briefly: $\mathfrak{B}(u,\cdot) = g_\partial$). For example, for the Dirichlet problem of the Laplace equation, it will be seen later that $\mathfrak{B}(u,w)$ can be taken to be $u\partial w/\partial n$ on $\partial\Omega$. Thus, if u_∂ is the prescribed value of u on $\partial\Omega$, one has $\mathfrak{B}(u,w) = u_\partial \partial w/\partial n$ for any function u that satisfies the boundary conditions. Thus $g_\partial(w) = u_\partial \partial w/\partial n$ in this case.

The linear function $\mathfrak{C}^*(u,\cdot)$, on the other hand, cannot be evaluated in terms of the prescribed boundary values, but it also depends exclusively on certain boundary values of u (the "complementary boundary values"). Generally, such boundary values can only be evaluated after the initial-boundary value problem has been solved. Taking the example of the Dirichlet problem for the Laplace equation, as before, $\mathfrak{C}^*(u,w) = w\partial u/\partial n$ and the complementary boundary values correspond to the normal derivative on $\partial\Omega$.

In a similar fashion, convenient formulations of boundary value problems with prescribed jumps requires constructing Green's formulas in discontinuous fields. This can be done by introducing a general decomposition of the bilinear function $\mathcal{R}_\Sigma(u,w)$ whose definition is pointwise. The general theory includes the treatment of differential operators with discontinuous coefficients [5]. However, in this article, only continuous coefficients will be considered. In this case, such decomposition is easy to obtain, and it stems from the algebraic identity:

$$[\underline{\mathfrak{D}}(u,w)] = \underline{\mathfrak{D}}([u],\dot{w}) + \underline{\mathfrak{D}}(\dot{u},[w])\,, \tag{2.6}$$

where

$$[u] = u_+ - u_-, \quad \dot{u} = (u_+ u_-)/2\,. \tag{2.7}$$

The desired decomposition is obtained by combining the second of Eqs. (2.4) with (2.6):

$$\mathcal{R}_\Sigma(u,w) = \mathcal{J}(u,w) - \mathcal{X}^*(u,w)\,, \tag{2.8}$$

with

$$\mathcal{J}(u, w) = -\underline{\mathfrak{D}}([u], \dot{w}) \cdot \underline{n} \tag{2.9a}$$

$$\mathcal{K}^*(u, w) = \mathcal{K}(w, u) = \underline{\mathfrak{D}}(\dot{u}, [w]) \cdot \underline{n}. \tag{2.9b}$$

An important property of the bilinear function $\mathcal{J}(u, w)$ is that, when the jump of u is specified, it defines a unique linear function of w, which is independent of the particular choice of u. When considering initial-boundary value problems with prescribed jumps, the linear function defined by the prescribed jumps in this manner will be denoted by j_Σ [thus its value for any given function w will be $j_\Sigma(w)$] and the jump conditions at any point of Σ can be specified by means of the equation $\mathcal{J}(u, \cdot) = j_\Sigma$. In problems with prescribed jumps, the linear functional $\mathcal{K}^*(u, \cdot)$ plays a role similar to that of the complementary boundary values $\mathcal{C}^*(u, \cdot)$. It can only be evaluated after the initial-boundary value problem has been solved and certain information about the average of the solution and its derivatives on Σ is known. Such information will be called the "generalized averages."

Introducing the notation

$$\langle Pu, w\rangle = \int_\Omega w\mathcal{L}u\, dx, \quad \langle Q^*u, w\rangle = \int_\Omega u\mathcal{L}^*w\, dx, \tag{2.10a}$$

$$\langle Bu, w\rangle = \int_{\partial\Omega} \mathcal{B}(u, w)\, dx, \quad \langle C^*u, w\rangle = \int_{\partial\Omega} \mathcal{C}(w, u)\, dx \tag{2.10b}$$

$$\langle Ju, w\rangle = \int_\Sigma \mathcal{J}(u, w)\, dx, \quad \text{and} \quad \langle K^*u, w\rangle = \int_\Sigma \mathcal{K}(w, u)\, dx, \tag{2.10c}$$

Eq. (2.3) can be written as

$$\langle Pu, w\rangle - \langle Q^*u, w\rangle = \langle Bu, w\rangle - \langle C^*u, w\rangle + \langle Ju, w\rangle - \langle K^*u, w\rangle. \tag{2.11}$$

This is an identity between bilinear forms and can be written more briefly, after rearranging, as

$$P - B - J = Q^* - C^* - K^*. \tag{2.12}$$

This is the *Green-Herrera formula for operators in discontinuous fields* [2, 6].

The initial-boundary value problem with prescribed jumps can be formulated pointwise by means of Eq. (2.1) together with

$$\mathcal{B}(u, \cdot) = g_\partial \text{and} \quad \mathcal{J}(u, \cdot) = j_\Sigma. \tag{2.13}$$

In order to associate a variational formulation with this problem, define the linear functionals $f, g, j \in D_2^*$ by means of

$$\langle f, w\rangle = \int_\Omega w f_\Omega\, dx, \quad \langle g, w\rangle = \int_{\partial\Omega} g_\partial(w)\, dx, \quad \langle j, w\rangle = \int_\Sigma j_\Sigma(w)\, dx. \tag{2.14}$$

Then a variational formulation of the initial-boundary value problem with prescribed jumps is

$$Pu = f, \quad Bu = g, \qquad Ju = j. \tag{2.15}$$

The bilinear functional J just constructed, as well as B, are *boundary operators* for P, which are *fully disjoint*. (For the definitions of the concepts that appear in italics here, the reader is referred to Herrera's original papers [2–6]). When this is the case, the system of equations (2.15) is equivalent to the single variational equation

$$\langle (P - B - J)u, w\rangle = \langle f - g - j, w\rangle \quad \forall\, w \in D_2. \tag{2.16}$$

This is said to be "the variational formulation in terms of the data of the problem," because Pu, Bu, and Ju are prescribed. Making use of formula (2.12), the variational formulation (2.16) is transformed into

$$\langle (Q^* - C^* - K^*)u, w\rangle = \langle f - g - j, w\rangle \quad \forall w \in D_2 . \tag{2.17}$$

This is said to be "the variational formulation in terms of the sought information," because Q^*u, C^*u, and K^*u are not prescribed. The variational formulations (2.16) and (2.17) are equivalent by virtue of the identity (2.12). The linear functionals Q^*u, C^*u, and K^*u supply information about the sought solution at points in the interior of the region Ω, the complementary boundary values at $\partial\Omega$, and the generalized averages of the solution at Σ, respectively, as can be verified by inspection of Eqs. (2.10), and as will be illustrated in the examples that follow.

Localized adjoint methods are based on the following observations. When the method of weighted residuals is applied, an approximate solution $\hat{u} \in D_1$ satisfies

$$\langle (P - B - J)\hat{u}, w^\alpha\rangle = \langle f - g - j, w^\alpha\rangle, \qquad \alpha = 1, \ldots, N , \tag{2.18}$$

where $\{w^1, \ldots, w^N\} \subset D_2$ is a given system of weighting functions. However, these equations, when they are expressed in terms of the sought information, become

$$\langle (Q^* - C^* - K^*)\hat{u}, w^\alpha\rangle = \langle f - g - j, w^\alpha\rangle, \quad \alpha = 1, \ldots, N . \tag{2.19}$$

Since the exact solution satisfies (2.17) it must be that

$$\langle (Q^* - C^* - K^*)\hat{u}, w^\alpha\rangle = \langle (Q^* - C^* - K^*)u, w^\alpha\rangle, \quad \alpha = 1, \ldots, N . \tag{2.20}$$

Either in this form or in the form

$$\langle (Q^* - C^* - K^*)(\hat{u} - u), w^\alpha\rangle = 0, \quad \alpha = 1, \ldots, N , \tag{2.20'}$$

Eqs. (2.20) can be used to analyze the information about the exact solution that is contained in an approximate one. In localized adjoint methods, these observations have been used as a framework for selecting more convenient test functions.

III. ILLUSTRATIVE RESULTS

As has already been mentioned, K^*u supplies information about the average of the solution and its derivatives across the surface of discontinuity Σ. Such information can be classified further. In particular, it is useful to decompose the averages K^*u into averages of the function, the first derivative, etc. This is achieved by writing K^* as the sum of operators K^{0*}, $K^{1*}, \ldots$, each one containing the information about the average of the derivative of the corresponding order. Such decomposition is induced when $\mathscr{X}^*(u, w)$ is decomposed pointwise into the sum of bilinear functions $\mathscr{X}^{0*}(u, w)$, $\mathscr{X}^{1*}(u, w), \ldots$, each one containing the corresponding information pointwise. Similarly, J will be written as the sum of operators $J^0, J^1, \ldots$, each containing the jump of the derivative of corresponding order, and $\mathscr{J}(u, w)$ will be the sum of $\mathscr{J}^0(u, w), \mathscr{J}^1(u, w)$, etc. When this is done,

$$K = \sum_i K^i, \quad J = \sum_i J^i, \quad \mathscr{X} = \sum_i \mathscr{X}^i, \quad \mathscr{J} = \sum_i \mathscr{J}^i . \tag{3.1}$$

In view of Eqs. (2.20), it is clear that the information about the exact solution contained in an approximate one depends in an essential manner on the system of weighting functions chosen. The systematic classification of such information introduced by the algebraic

theory can be used to develop weighting functions that concentrate the information in a desired manner. In very general terms, if one wishes to eliminate all the information about the sought solution in the interior of the region $\Omega - \Sigma$ (thus concentrating it in the complementary boundary values and averages across Σ), the weighting functions must be chosen so that $0 = \langle Q^* u, w^\alpha \rangle = \langle Q w^\alpha, u \rangle$. This condition is satisfied if $Q w^\alpha = 0$, i.e., $\mathcal{L}^* w^\alpha = 0$. Similarly, the information about the complementary boundary values can be eliminated if the weighting functions satisfy $C w^\alpha = 0$ globally, i.e., $\mathcal{C}(w^\alpha, \cdot) = 0$ pointwise. Elimination of the average of the function requires $\mathcal{X}^0(w^\alpha, \cdot) = 0$; elimination of the average of the first derivative requires that $\mathcal{X}^1(w^\alpha, \cdot) = 0$, etc. Note that boundary methods are obtained when $\mathcal{L}^* w^\alpha \equiv 0$ and simultaneously $\mathcal{X}^*(w^\alpha, \cdot) \equiv 0$. Also, if all the information is to be concentrated at one point, Green's functions must be constructed. However, development of such weighting functions requires solution of nonlocal problems. As a further general comment, it should be mentioned that, for partial differential equations of second order with continuous coefficients, the condition $\mathcal{X}^1(w^\alpha, \cdot) = 0$ is equivalent to $[w] = 0$ (i.e., w is C^0). This result, which will be shown later in this section for the special case of ordinary differential equations, implies that C^0 methods concentrate the information on the values of the function, both at the interelement boundaries and in the interior of the finite elements.

As an illustration of the procedure, let us review the results for general second-order differential equations that are linear, which were derived by Herrera and co-workers [5, 7–9] for one-dimensional equations. Actually, the methodology is applicable to multidimensional equations of any order which may be time dependent, and to systems of differential equations, as is explained in Sec. VII. However, for partial differential equations (i.e., when more than one independent variable is involved), the implementation of the procedures is considerably more complicated and it is not possible to predict the exact values of the solution in general.

A physical situation that the general ordinary differential equation of second order mimics is transport in the presence of advection, diffusion, and linear sources; a notation related with such processes will be adopted. The general equation to be considered is

$$\mathcal{L}u \equiv -\frac{d}{dx}\left(D\frac{du}{dx} - Vu\right) + Ru = f_\Omega, \quad \text{in } \Omega \equiv [0, l]. \tag{3.2a}$$

Attention will be restricted to the case when D and V are continuous (discontinuous coefficients have been treated previously [5]). When the function u is assumed to be continuous, so that

$$[u] = 0 \quad \text{on } \Sigma, \tag{3.2b}$$

the smoothness condition implied by conservation of mass [49]

$$\left[Vu - D\frac{\partial u}{\partial x}\right] = 0 \quad \text{on } \Sigma \tag{3.2c}$$

reduces to

$$\left[\frac{\partial u}{\partial x}\right] = 0 \quad \text{on } \Sigma \tag{3.2d}$$

by virtue of the assumed continuity of V and D.

A partition $\{0 = x_0, x_1, \ldots, x_{E-1}, x_E = l\}$ is introduced, which for simplicity is assumed to be uniform, i.e., $x_\alpha - x_{\alpha-1} = h$ is independent of α. Trial and test functions will be assumed to be sufficiently differentiable in the interior of each of the subintervals of the partition, so that the differential operator is defined there and the jump discontinuities can

only occur at internal nodes. This corresponds to taking $\Sigma = \{x_1, \ldots, x_{E-1}\}$ in the general framework presented in Sec. II. Observe that the normal vector $\underline{n} = 1$ at l, and $\underline{n} = -1$ at 0. On Σ the choice $\underline{n} = 1$ is convenient, because in this manner the positive side of Σ is the side that is determined by the sense of the x axis. Suitable boundary conditions are assumed to be satisfied at 0 and l in order to have a well defined boundary value problem. The boundary conditions can be Dirichlet, Neumann, or Robin [5], but they are left unspecified, since the following developments accommodate any of them.

The formal adjoint of the operator $\mathscr{L}$, as defined by (3.2a), is

$$\mathscr{L}^* w \equiv -\frac{d}{dx}\left(D\frac{dw}{dx}\right) - V\frac{dw}{dx} + Rw. \tag{3.3}$$

Therefore

$$w\mathscr{L}u - u\mathscr{L}^* w \equiv \frac{d}{dx}\left\{u\left(D\frac{dw}{dx} + Vw\right) - wD\frac{du}{dx}\right\} \tag{3.4}$$

and

$$\underline{\mathscr{B}}(u, w) \equiv u\left(D\frac{dw}{dx} + Vw\right) - wD\frac{du}{dx}. \tag{3.5}$$

Application of Eqs. (2.9) yields

$$\mathscr{J}^0(u, w) = -[u]\left(D\frac{\overline{dw}}{dx} + V\dot{w}\right), \quad \mathscr{J}^1(u, w) = \dot{w}D\left[\frac{du}{dx}\right], \tag{3.6a}$$

$$\mathscr{X}^0(w, u) = \dot{u}\left[D\frac{dw}{dx} + Vw\right], \quad \mathscr{X}^1(w, u) = -[w]D\frac{\overline{du}}{dx}, \tag{3.6b}$$

from which $\mathscr{J}$ and $\mathscr{X}$ are obtained by means of Eqs. (3.1). In Eqs. (3.6), an overbar is used to indicate that the dot on top refers to the whole expression covered by the bar.

The definitions of the bilinear functions $\mathfrak{B}(u, w)$ and $\mathfrak{C}(w, u)$, depend on the type of boundary conditions to be satisfied. For Dirichlet problems, u is prescribed and one possibility is

$$\mathfrak{B}(u, w) = u\left(D\frac{dw}{dx} + Vw\right) \cdot \underline{n}, \quad \mathfrak{C}(w, u) = wD\frac{du}{dx} \cdot \underline{n}. \tag{3.7}$$

Another possibility is

$$\mathfrak{B}(u, w) = -uD\frac{dw}{dx} \cdot \underline{n}, \quad \mathfrak{C}(w, u) = -w\left(D\frac{du}{dx} - Vu\right) \cdot \underline{n}.$$

To be specific, only Eq. (3.7) will be used here. For Neuman problems, du/dx is the datum, so that

$$\mathfrak{B}(u, w) = -wD\frac{du}{dx} \cdot \underline{n}, \quad \mathfrak{C}(w, u) = -u\left(D\frac{dw}{dx} + Vw\right) \cdot \underline{n}. \tag{3.8}$$

In the most general form of Robin's boundary conditions, a linear combination of the derivative and the value of the solution are prescribed. This general case was developed in [5]. Here, attention is restricted to the case when the total flux $Ddu/dx - Vu$ is prescribed. Then

$$\mathfrak{B}(u, w) = -w\left(D\frac{du}{dx} - Vu\right) \cdot \underline{n}, \quad \mathfrak{C}(w, u) = -uD\frac{dw}{dx} \cdot \underline{n}. \tag{3.9}$$

In the case of ordinary differential equations, Herrera *et al.* [5] developed algorithms for which all the information was concentrated in the values of the solution at internal

nodes; by so doing, they were able to obtain algorithms that yield the exact values of the solution at those nodes. Using the developments already presented, derivation of such algorithms is a simple exercise. The information about the sought solution supplied by the weighted system of equations is exhibited in Eqs. (2.20). In particular, the information that is supplied at internal nodes refers to the averages of the solution and its first derivative there, given by $\mathscr{X}^* u$ [see Eqs. (3.6b)]. Thus, in order to concentrate the information there, it is necessary to remove $\langle Q^* u, w^\alpha \rangle$ and $\langle C^* u, w^\alpha \rangle$ in Eqs. (2.20). This is achieved if $Qw^\alpha = 0$ and $Cw^\alpha = 0$. The first of these conditions is equivalent to $\mathscr{L}^* w \equiv 0$, and the other is the boundary condition $\mathscr{C}(w, \cdot) \equiv 0$, which in the case considered here must be satisfied at $\partial\Omega = \{0, l\}$. In view of Eqs. (3.7)–(3.9), the boundary conditions to be satisfied by the test functions are

$$w = 0, \quad D\frac{dw}{dx} + Vw = 0, \quad D\frac{dw}{dx} = 0, \tag{3.10}$$

wherever Dirichlet, Neuman, and flux conditions, respectively, are prescribed on u. Observe the "symmetry":

> To eliminate the complementary boundary values from the weighted equations, the value of the test functions must vanish wherever the value of the sought solution is prescribed. However, it is the "reserve flux" $(Ddw/dx + Vw)$ which must vanish, wherever the first derivative of the sought solution is prescribed and conversely, it is the derivative (dw/dx) which must vanish wherever the flux $(Ddu/dx - Vu)$ is prescribed.

Since our goal is to concentrate the information on the value of the sought solution at internal nodes, taking into account that $\mathscr{X}(w, u) = \mathscr{X}^0(w, u) + \mathscr{X}^1(w, u)$, it is clear that it is still necessary to remove the information about the first derivative. This will be achieved if the condition $\mathscr{X}^1(w, \cdot) = 0$ is imposed on the test functions. In view of (3.6b) this condition is $[w] = 0$, on Σ.

In summary, the weighting functions that concentrate all the information in the values of the sought solution at internal nodes are solutions of the homogeneous boundary value problem with prescribed jumps:

$$\begin{aligned} &\mathscr{L}^* w = 0 \quad \text{on } \Omega, \quad \mathscr{C}(w, \cdot) = 0 \quad \text{on } \partial\Omega = \{0, l\}, \\ &[w] = 0 \quad \text{on } \Sigma. \end{aligned} \tag{3.11}$$

Clearly, the condition $[w] = 0$ on $\sum$ implies that $w \in C^0$. Such weighting functions can be taken having local support, because $[\partial w/\partial x] \neq 0$ is admissible [5, 8]. Generally, the dimension of the space of solutions of Eqs. (3.11) is $E - 1$, and if that space is used to form the system of weighting functions in Eq. (2.18), an $(E - 1)(E - 1)$ system of equations possessing a unique solution is obtained for the average $\hat{u}$ of the approximate solution at internal nodes. The uniqueness of solution of this system of equations implies that the averages predicted in this manner are equal to the averages of the exact solution at internal nodes, since $\mathscr{X}^0(\hat{u} - u) = 0$ in that case, by virtue of (2.20′). Also, at any internal node, the value of the sought solution is equal to its average there, since it is continuous (Eqs. 3.2b), so that the exact value of the solution at internal nodes is predicted in this manner.

A rigorous discussion of the conditions under which the resulting system of equations possesses a unique solution requires the use of the concept of TH completeness. This concept was introduced by Herrera in [2, 51], where he presented a rigorous discussion of

this question in an abstract setting allowing considerable generality, since the conclusions that he obtained are independent of the order of the differential equations and the number of independent variables involved. However, attention was restricted to the case when the differential operator is symmetric and the corresponding analysis for nonsymmetric operators is wanting. This matter is being studied at present and will be addressed elsewhere.

A similar procedure can be developed for obtaining the exact values of the first derivative at internal nodes, the main difference being that one must require that $\mathscr{R}^0(w,\cdot) = 0$ instead of $\mathscr{R}^1(w,\cdot) = 0$. In view of Eq. (3.6b), Eq. (3.11) is replaced by

$$\mathscr{L}^* w = 0 \quad \text{on } \Omega, \quad \mathscr{C}(w,\cdot) = 0 \quad \text{on } \partial\Omega = \{0, l\},$$
$$[Ddw/\frac{dw}{dx}] = 0 \quad \text{on } \Sigma. \tag{3.12}$$

The corresponding algorithm was developed in [5, 8]. Combinations in which the value of the solution is obtained at some nodes and its derivative at others, or algorithms that simultaneously yield the exact values of the solution and its derivative at internal nodes (with a correspondingly larger system of equations to be solved), are also possible [5].

Until now, no specific representation of the approximate solution has been adopted. Let $\{\Phi^0, \Phi^1, \ldots, \phi^N\}$ be a system of trial functions, generally fully discontinuous (i.e., the function and its derivative have jump discontinuities at internal nodes), and let $\hat{u} = \sum A_j \Phi^j$ be an approximate solution satisfying Eqs. (2.18) or, equivalently, (2.19). When the system of weighting (or test) functions fulfilling (3.11) is TH complete, then

$$\hat{u}(x_j) = u(x_j), \quad j = 1, \ldots, E-1, \tag{3.13a}$$

where $u(x)$ is the exact solution. Correspondingly, if the system of weighting functions satisfies (3.12) and is TH complete, then

$$\frac{d\hat{u}}{dx}(x_j) = \frac{du}{dx}(x_j), \quad j = 1, \ldots, E-1. \tag{3.13b}$$

It is important to observe that either Eq. (3.13a) or (3.13b) holds independently of the trial functions used. In particular, they can be fully discontinuous and they can also violate the prescribed boundary conditions, although this would produce poor approximations, except at the nodes. This is discussed further below.

It is worth exploring some of the implications of these results. To relate the results thus far obtained to standard variational formulations used in finite element methods, let us consider the one-dimension (1D) version of the Poisson equation [i.e., Eq. (3.2a), with $D \equiv 1$, $V \equiv R \equiv 0$], subject to homogeneous Dirichlet boundary conditions [$u(0) = u(l) = 0$]. A standard variational formulation for this problem is as follows: $u \in H_0^1([0, l])$ is a weak solution if and only if

$$-\int_0^l \frac{du}{dx}\frac{dw}{dx}\,dx = \int_0^l f_\Omega w\,dx \quad \forall\, \omega \in H_0^1([0, l]), \tag{3.14}$$

where $H_0^1([0, l])$ is the subspace of the Sobolev space $H^1([0, l])$, whose members have vanishing traces.

Let us look for an approximate solution $\hat{u}$ using trial functions $\{\Phi^1, \ldots, \Phi^{E-1}\}$ which are locally linear and globally continuous [Fig. 1 (a)]. In addition, use the same collection as test functions, as is usually done when applying Galerkin method. Let $\hat{u}(x) = \sum_{j=1}^{E-1} U_j \Phi^j(x)$.

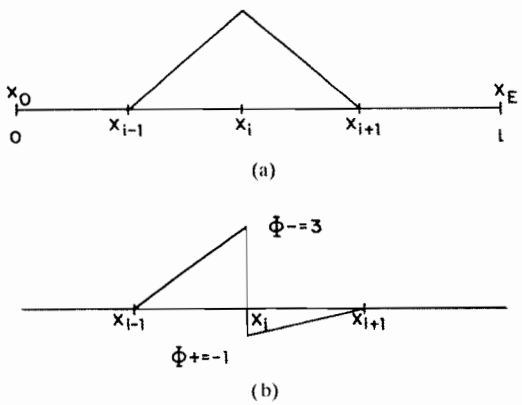

FIG. 1. (a) Base function: Locally linear, globally continuous. (b) Base function: Locally linear, discontinuous.

Then Φ^j, for $j = 1, \ldots, E-1$, as well as $\hat{u}$, belong to $H_0^1([0, l])$. Observe that for this class of functions one has

$$-\int_0^l \frac{d\hat{u}}{dx}\frac{dw}{dx}\,dx = \int_0^l w\frac{d^2\hat{u}}{dx^2}\,dx + \sum_{j=1}^{E-1} w_j\left[\frac{d\hat{u}}{dx}\right]_j = \langle (P - B - J)\hat{u}, w\rangle \quad (3.15)$$

Here, Eqs. (3.6a) and the facts that u and x are continuous and vanish on the boundary were used. Equation (3.15) illustrates the fact that the standard variational formulation for the Laplace operator is a particular case of the general variational principle (2.16) in terms of the data of the problem. However, the standard variational formulation can only be applied when both trial and test functions are continuous, while the algebraic theory supplies a systematic manner of extending it to cases where trial and test functions are fully discontinuous.

For the case where the prescribed boundary conditions are nonhomogeneous, a suitable representation of the approximate solution is

$$\hat{u}(x) = U_0\Phi^0 + U_E\Phi^E + \sum_{j=1}^{E-1} U_j\Phi^j(x). \quad (3.16)$$

Let us apply the variational formulation (2.18) using the weighting functions $\{\Phi^1, \ldots, \Phi^{E-1}\}$ associated with internal nodes. This system concentrates all the information at internal nodes because it satisfies Eqs. (3.11); moreover, it is TH complete. The resulting system of equations is

$$\frac{U_{\alpha+1} + U_{\alpha-1} - 2U_\alpha}{h} = \int_{x_{\alpha-1}}^{x_{\alpha+1}} f_\Omega w^\alpha\,dx, \quad \alpha = 2, \ldots, E-2, \quad (3.17a)$$

$$\frac{U_2 - 2U_1}{h} = \int_0^{x_2} f_\Omega w^1\,dx - \frac{u_0}{h}, \quad \frac{U_{E-2} - 2U_{E-1}}{h} = \int_{x_{E-2}}^{l} f_\Omega w^{E-1}\,dx - \frac{u_E}{h}, \quad (3.17b)$$

where u_0 and u_E are the prescribed values for u at 0 and l, respectively, and bear no relation with U_0 and U_E, which are unknown. Since the same system of equations is

fulfilled by the exact solution, one has

$$\frac{U_{\alpha+1} + U_{\alpha-1} - 2U_\alpha}{h} = \frac{u_{\alpha+1} + u_{\alpha-1} - 2u_\alpha}{h}, \quad \alpha = 2,\ldots,E-2 \tag{3.18a}$$

together with

$$\frac{U_{E-2} - 2U_{E-1}}{h} = \frac{u_{E-2} - 2u_{E-1}}{h}, \quad \frac{U_2 - 2U_1}{h} = \frac{u_2 - 2u_1}{h}, \tag{3.18b}$$

which implies $U_\alpha = u(x_\alpha)$ for $\alpha = 1,\ldots,E-1$ (i.e., at all the interior nodes), as predicted by the theory.

The above properties depend solely on the weighting functions and are independent of the trial functions used. In particular, when the boundary values are not satisfied by the trial functions [i.e., $U_0 \neq u_0$, $U_E \neq u_l$ in Eq. (3.16)], one would expect to obtain an $(E+1)(E-1)$ system of equations that would be undetermined, since the number of unknowns is greater than the number of equations. However, U_0 and U_E do not occur in the system that is obtained, and the resulting system can be interpreted as an $(E-1)(E-1)$ system for $\{U_1,\ldots,U_{E-1}\}$, whose only solution, as already mentioned, is the values of the exact solution at the internal nodes, leaving U_0 and U_E undetermined. This latter fact is natural, since the system of equations (3.18) was derived using the variational principle (2.17), in terms of the sought information, and the values of the sought solution at the boundary are not included in the sought information. Indeed, in the boundary only the derivative is included in the sought information.

Moreover, the trial functions themselves can be changed arbitrarily and still the nodal values will be predicted correctly. In particular, let us illustrate the use of fully discontinuous trial functions by exhibiting these results when such trial functions are used. To this end, keeping the same weighting functions as before, change the trial functions in the representation (3.16) of the approximate solution, to [see Fig. 1(b)]

$$\Phi^j(x) = \begin{cases} \dfrac{3(x - x_{j-1})}{h}, & x_{j-1} \leq x \leq x_j \\ \dfrac{x - x_{j+1}}{h}, & x_j \leq x \leq x_{j+1} \\ 0, \quad \text{elsewhere.} \end{cases} \tag{3.19}$$

Then the same system of equations (3.17) is obtained. Observe that the averages of the trial functions given by (3.19) are zero at internal nodes, except $\dot{\Phi}^j(x_j)$ which is equal to one. Thus (3.17) is equivalent to

$$\frac{\dot{\hat{u}}_{\alpha+1} + \dot{\hat{u}}_{\alpha-1} - 2\dot{\hat{u}}_\alpha}{h} = \int_{x_{\alpha-1}}^{x_{\alpha+1}} f_\Omega w^\alpha \, dx, \quad \alpha = 2,\ldots,E-2, \tag{3.20a}$$

$$\frac{\dot{\hat{u}}_2 - 2\dot{\hat{u}}_1}{h} = \int_0^{x_2} f_\Omega w^1 \, dx - \frac{u_0}{h}, \quad \frac{\dot{\hat{u}}_{E-2} - 2\dot{\hat{u}}_{E-1}}{h} = \int_{x_{E-2}}^{x_E} f_\Omega w^{E-1} \, dx - \frac{u_E}{h}, \tag{3.20b}$$

whose only solution is $\dot{\hat{u}}(x_\alpha) = u(x_\alpha)$ for $\alpha = 1,\ldots,E-1$ [i.e., Eq. (3.13a)]. Thus, even if discontinuous trial functions are used, the values of the exact solution are predicted correctly by the averages at internal nodes of the discontinuous approximate solutions.

Although very simple examples were chosen to illustrate the results presented in this section, the conclusions are valid for the general Eq. (3.2a) and also for the different kinds of algorithms that were introduced. They exhibit the general fact that the information about

the sought solution contained in an approximate one is independent of the trial functions used. However, the reader must not be misled to conclude that the choice of trial functions is irrelevant. On the contrary, as is well known, a judicious choice of the trial functions is essential to develop satisfactory approximate solutions. For example, if the trial functions (3.19) are used as interpolators in present examples, very poor estimations of the solutions would be obtained in the interior of the subintervals $[x_{i-1}, x_i]$, in spite of the fact that its values at internal nodes are predicted with unlimited precision.

However, an important conclusion that can be drawn from the preceding examples is that in the construction of approximate solutions, there are two processes, equally important but different, that must be clearly distinguished. They are (i) gathering information about the sought solution, and (ii) interpolating the information about the sought solution which is available.

These two processes are distinct, although in many numerical methods they are not differentiated clearly. The information that is gathered is determined by the weighting functions, while the manner in which it is interpolated depends of the trial functions chosen. A peculiarity of the examples that have been given in this section is that here, those processes are not only independent, but they need not be carried out simultaneously. Indeed, since the exact values at the nodes are obtained independently of the trial functions used, given that the requirements in Eq. (3.11) are satisfied, one can obtain them first and choose the interpolator afterwards.

Also, the two processes mentioned above are to a large extent independent, exhibiting some of the severe limitations associated with methods such as the Galerkin method, in which trial and test functions are required to be the same. The conditions that test functions must satisfy in order to be effective for gathering information, in general, will be quite different from those that must be satisfied by trial functions in order to be effective interpolators.

A point that deserves further attention is the criteria that must be used to judiciously select effective trial functions. Taking into account their role as interpolators, it is clear that approximation theory must be applied. However, in many cases the matters involved may go beyond approximation theory. For example, in the illustrations presented thus far in this section, which in some sense are extreme cases since the exact values of the solution are obtained at the nodes, a very efficient way of interpolating the available information would be to solve the boundary value problem that is defined by that information on each of the subintervals $[x_{i-1}, x_i]$. These questions, although important, are complex, and it would not be appropriate to explore them in all their generality at this point. Thus we leave the matter here, but intend to resume it elsewhere.

IV. ADVECTION-DIFFUSION EQUATION

When applying the methods of Sec. II to time-dependent problems, it will be necessary to consider a region Ω in space-time. Also, the surface Σ on which discontinuities can occur will be a surface in space-time, and a suitable notation will be required. Space-time vectors $\underline{M}$ will be written as pairs:

$$\underline{M} = (\underline{m}, m_t), \tag{4.1}$$

where $\underline{m}$ is the vector made by its spatial components and m_t corresponds to its temporal component. Let $\underline{V}_\Sigma$ be the vectorial velocity of the discontinuity $\Sigma(t)$, where $\Sigma(t)$ is the set of points of Σ whose time coordinate is t. This is a vector in space, which can be

written as

$$\underline{V}_\Sigma = V_\Sigma \underline{n}, \tag{4.2}$$

where $\underline{n}$ is the unit normal vector to $\Sigma(t)$. Generally, V_Σ can be positive or negative, depending on the sense of motion of $\Sigma(t)$ and the choice of $\underline{n}$. In particular, in the one-dimensional case, $\underline{n}$ will be taken to be equal to 1 on Σ. Observe that the space-time vector $(\underline{V}_\Sigma, 1) = (V_\Sigma \underline{n}, 1)$ is tangent to Σ. Using this fact, it is easy to see that a space-time unit normal vector $\underline{N}$ to Σ is given by

$$\underline{N} = (1 + V_\Sigma^2)^{-1/2}(\underline{n}, -V_\Sigma), \tag{4.3}$$

In this section, we consider the one-dimensional transient advection-diffusion equation in conservative form:

$$\mathcal{L}u \equiv \frac{\partial u}{\partial t} - \frac{\partial}{\partial x}\left(D\frac{\partial u}{\partial x} - Vu\right) + Ru = f_\Omega(x,t) \quad \text{in } \mathfrak{U}, \tag{4.4}$$

$$x \in \Omega_x = [0, l],$$
$$t \in \Omega_t = [t^n, t^{n+1}],$$
$$(x,t) \in \Omega = \Omega_x X \Omega_t,$$

subject to initial conditions

$$u(x, t^n) = u^n(x) \tag{4.5}$$

and suitable boundary conditions at $x = 0$ and l. The following development accommodates any combination of boundary conditions. The manner in which the region Ω and the initial conditions were chosen in Eqs. (4.4) and (4.5) is convenient when applying a numerical integration procedure step by step in time.

The adjoint operator is

$$\mathcal{L}^* w \equiv -\frac{\partial w}{\partial t} - \frac{\partial}{\partial x}\left(D\frac{\partial w}{\partial x}\right) - V\frac{\partial w}{\partial x} + Rw, \tag{4.6}$$

and $\underline{\mathfrak{D}}(u,w)$ as defined by Eq. (2.2) is

$$\underline{\mathfrak{D}}(u,w) = \left\{u\left(D\frac{\partial w}{\partial x} + Vw\right) - wD\frac{\partial u}{\partial x}, uw\right\}. \tag{4.7}$$

Therefore

$$\mathfrak{R}_\Sigma(u,w) = -[\underline{\mathfrak{D}}(u,w)] \cdot \underline{N} = (1 + V_\Sigma^2)^{-1/2}\left[u\left(D\frac{\partial w}{\partial x} + (V - V_\Sigma)w\right) - wD\frac{\partial u}{\partial x}\right]. \tag{4.8}$$

Assuming that the physical process which Eq. (4.4) mimics is that of transport with Fickian diffusion of a solute whose concentration is u in a free fluid moving with velocity V, the smoothness conditions implied by mass balance are [48]

$$\left[u(V - V_\Sigma) - D\frac{\partial u}{\partial x}\right] = 0 \quad \text{on } \Sigma. \tag{4.9}$$

In addition, Fickian diffusion implies

$$[u] = 0 \quad \text{on } \Sigma. \tag{4.10a}$$

When the coefficients V and D are continuous, Eq. (4.9) in the presence of (4.10a) can be replaced by

$$\left[\frac{\partial u}{\partial x}\right] = 0 \quad \text{on } \Sigma\,. \tag{4.10b}$$

Using Eqs. (2.9), it is seen that

$$\mathcal{J}^0(u,w) = -(1+V_\Sigma^2)^{-1/2}[u]\left\{\overline{D\frac{\partial w}{\partial x}} + \dot{w}(V - V_\Sigma)\right\}, \tag{4.11a}$$

$$\mathcal{J}^1(u,w) = (1+V_\Sigma^2)^{-1/2}\dot{w}D\left[\frac{\partial u}{\partial x}\right], \tag{4.11b}$$

$$\mathcal{K}^0(w,u) = (1+V_\Sigma^2)^{-1/2}\dot{u}\left\{\left[D\frac{\partial w}{\partial x}\right] + (V - V_\Sigma)[w]\right\}, \tag{4.12a}$$

$$\mathcal{K}^1(w,u) = -(1+V_\Sigma^2)^{-1/2}[w]D\overline{\frac{\partial u}{\partial x}}\,, \tag{4.12b}$$

It will be useful to decompose the boundary $\partial\Omega$ into $\partial_0\Omega$, $\partial_l\Omega$, $\partial_n\Omega$, and $\partial_{n+1}\Omega$, which are defined as the subsets of Ω for which (x,t) satisfies $x = 0$, $x = l$, $t = t^n$, and $t = t^{n+1}$, respectively. The initial conditions given by Eq. (4.5) are to be satisfied at $\partial_n\Omega$, and the boundary conditions pertain to $\partial_0\Omega \cup \partial_l\Omega$. These latter conditions can be of Dirichlet ($u = u_a$), Neuman [$D(\partial u/\partial x)\underline{n} = q$], or Robin type, or a combination of them. Because of the special role that the total flux $D\partial u/\partial x - Vu$ plays in mass conservation, the only boundary conditions of Robin type that will be considered will be those for which $(D\partial u/\partial x - Vu)\underline{n}$ $(= F)$ is prescribed. In what follows, the notations $\partial_D\Omega$, $\partial_N\Omega$, and $\partial_F\Omega$ refer to that part of the boundary where Dirichlet, Neuman, and total-flux boundary conditions are prescribed, respectively.

The bilinear functions $\mathcal{B}(u,w)$ and $\mathcal{C}(w,u)$ implied by the initial and boundary conditions are

$$\mathcal{B}(u,w) = -uw \quad \text{on } \partial_n\Omega\,, \tag{4.13a}$$

$$\mathcal{C}(w,u) = -uw \quad \text{on } \partial_{n+1}\Omega\,, \tag{4.13b}$$

$$\mathcal{B}(u,w) = uD\frac{\partial w}{\partial x}\cdot\underline{n}, \quad \mathcal{C}(w,u) = w\left(D\frac{\partial u}{\partial x} - Vu\right)\cdot\underline{n} \quad \text{on } \partial_D\Omega\,, \tag{4.13c}$$

$$\mathcal{B}(u,w) = -wD\frac{\partial u}{\partial x}\cdot\underline{n}, \quad \mathcal{C}(w,u) = -u\left(D\frac{\partial w}{\partial x} + Vw\right)\cdot\underline{n} \quad \text{on } \partial_N\Omega\,, \tag{4.13d}$$

$$\mathcal{B}(u,w) = -w\left(D\frac{\partial u}{\partial x} - Vu\right)\cdot\underline{n}, \quad \mathcal{C}(w,u) = -uD\frac{\partial w}{\partial x}\cdot\underline{n} \quad \text{on } \partial_F\Omega\,, \tag{4.13e}$$

where the unit normal vector $\underline{n}$ can take the values $+1$ or -1. Observe that $\mathcal{C}^*(u,\cdot) \equiv 0$ on $\partial_n\Omega$, while $\mathcal{B}(u,\cdot) \equiv 0$ on $\partial_{n+1}\Omega$, i.e., no information is sought at $t = t^n$, which is

natural for an initial value problem. Also, in the case of Dirichlet boundary conditions, alternative expressions to (4.13c) are

$$\mathfrak{B}(u,w) = u\left(D\frac{\partial w}{\partial x} + Vw\right)\underline{n}, \quad \mathfrak{C}(w,u) = wD\frac{\partial u}{\partial x}\underline{n} \quad \text{on } \partial_D\Omega\,. \tag{4.13f}$$

However, in this paper, we use (4.13c) only. In view of (4.11) and (4.13), it is clear that

$$g_\partial(w) = -u^n w \quad \text{on } \partial_n\Omega, \quad g_\partial(w) = u_\partial\left(D\frac{\partial w}{\partial x} + Vw\right)\cdot\underline{n} \quad \text{on } \partial_D\Omega\,, \tag{4.14a}$$

$$g_\partial(w) = -wq\cdot\underline{n} \quad \text{on } \partial_N\Omega, \quad g_\partial(w) = -wF\cdot\underline{n} \quad \text{on } \partial_F\Omega\,, \tag{4.14b}$$

while $j_\Sigma(w) \equiv 0$ on Σ. The expressions for the bilinear functionals B, C, J, and K, are obtained by integration of $\mathfrak{B}, \mathfrak{C}, \mathfrak{J}$, and $\mathfrak{K}$, respectively, the first two on $\partial\Omega$ and the latter two on Σ. Similarly, according to (2.17), the expressions for f, g, and j are obtained by integrating f_Ω, g_∂, and j_Σ, in Ω, $\partial\Omega$, and Σ, respectively. In the present case $j_\Sigma \equiv 0$, so that $j = 0$ also.

As in Sec. III, a partition of $[0,l]$ is introduced and the region Ω is decomposed into a collection of subregions $\Omega_1,\ldots,\Omega_E$ (Fig. 2), limited by space-time curves Σ_i, $i = 1,\ldots,E$, whose parametric representations are given by the functions $\sigma_i(t)$; it will be assumed that discontinuities occur exclusively on these lines i.e.,

$$\Sigma = \bigcup_{i=1}^{E}\Sigma_i\,.$$

In addition, it is assumed that each such curve passes through its corresponding node at time t^{n+1} [i.e., $\sigma_i(t^{n+1}) = x_i$] and the notation $x_i^* = \sigma_i(t^n)$ is adopted. The velocity of propagation V_Σ of each of these lines of discontinuity is $d\sigma_i/dt$.

Using Eqs. (4.11) and (4.12), the expressions for J and K^* can be obtained by integration; they are the sum of the contributions of each of the curves Σ_i. Thus, one can write

$$J = \sum_{\alpha=1}^{E} J_\alpha \quad \text{and} \quad K^* = \sum_{\alpha=1}^{E} K_\alpha^*, \tag{4.15}$$

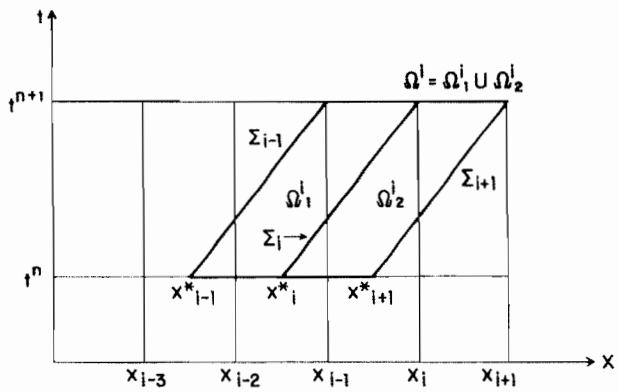

FIG. 2. Space-time supports of weighting functions.

where

$$\langle J_\alpha u, w\rangle = \int_{\Sigma_\alpha} \left\{ [u]\overline{D}\frac{\overline{\partial w}}{\partial x} - \dot{w}\left(D\left[\frac{\partial u}{\partial x}\right] - (V - V_\Sigma)[u]\right)\right\}_\alpha dt\,, \tag{4.16a}$$

$$\langle K^*_\alpha u, w\rangle = \int_{\Sigma_\alpha} \left\{ \dot{u}\left[D\frac{\partial w}{\partial x}\right] - [w]\left(\overline{D}\frac{\overline{\partial u}}{\partial x} - (V - V_\Sigma)\dot{u}\right)\right\}_\alpha dt\,. \tag{4.16b}$$

Here the subindex Σ_α means that the line integral is to be carried out on Σ_α. To obtain Eqs. (4.16), use has been made of the fact that on each line Σ_α, the element of time dt is $(1 + V_\Sigma^2)^{-1/2}$ times the length of the element in space-time.

In a similar fashion, it is convenient to decompose the bilinear functions B and C^* into the contributions which stem from $\partial_n\Omega, \partial_{n+1}\Omega, \partial_D\Omega, \partial_N\Omega$, and $\partial_F\Omega$. In this manner one can write

$$B = B_n + B_D + B_N + B_F \quad \text{and} \quad C^* = C^*_{n+1} + C^*_D + C^*_N + C^*_F\,, \tag{4.17}$$

where

$$\langle B_n u, w\rangle = -\int_0^l (uw)_{t=t^n}\,dx, \quad \langle C^*_{n+1}u, w\rangle = -\int_0^l (uw)_{t=t^{n+1}}\,dx\,, \tag{4.18a}$$

$$\langle B_D u, w\rangle = \int_{\partial_D\Omega} uD\frac{\partial w}{\partial x}\cdot \underline{n}\,dt, \quad \langle C^*_D u, w\rangle = \int_{\partial_D\Omega} w\left(D\frac{\partial u}{\partial x} - Vu\right)\cdot \underline{n}\,dt\,, \tag{4.18b}$$

$$\langle B_N u, w\rangle = -\int_{\partial_N\Omega} wD\frac{\partial u}{\partial x}\cdot \underline{n}\,dt, \quad \langle C^*_N u, w\rangle = -\int_{\partial_N\Omega} u\left(D\frac{\partial w}{\partial x} + Vw\right)dt\,, \tag{4.18c}$$

$$\langle B_F u, w\rangle = -\int_{\partial_F\Omega} w\left(D\frac{\partial u}{\partial x} - Vu\right)\cdot \underline{n}\,dt, \quad \langle C^*_F u, w\rangle = -\int_{\partial_F\Omega} uD\frac{\partial w}{\partial x}\cdot \underline{n}\,dt\,. \tag{4.18d}$$

To complete the formulation of the problem, it remains to define the linear functionals f, g, and j. The last one is zero, while $g = g_n + g_D + g_N + g_F$, with

$$\langle g_n, w\rangle = -\int_0^l u^n w(t^n)\,dx, \quad \langle g_D, w\rangle = \int_{\partial_D\Omega} u_\partial\left(D\frac{\partial w}{\partial x} + Vw\right)\cdot \underline{n}\,dt\,, \tag{4.19a}$$

$$\langle g_N, w\rangle = -\int_{\partial_N\Omega} wq\cdot \underline{n}\,dt, \quad \langle g_F, w\rangle = -\int_{\partial_F\Omega} wF\cdot \underline{n}\,dt\,. \tag{4.19b}$$

V. EULERIAN-LAGRANGIAN LAM

In what follows, the variational formulation in terms of the sought information, Eq. (2.17),

$$\langle (Q^* - C^* - K^*)u, w\rangle = \langle f - g - j, w\rangle\,, \tag{5.1}$$

will be applied. In addition, weighting functions w will be chosen satisfying $Qw = 0$, i.e.,

$$\mathcal{L}^* w \equiv -\frac{\partial w}{\partial t} - \frac{\partial}{\partial x}\left(D\frac{\partial w}{\partial x}\right) - V\frac{\partial w}{\partial x} + Rw = 0 \quad \text{in } \Omega\,. \tag{5.2}$$

In this case (5.1) becomes $\langle (C^* + K^*)u, w\rangle = \langle g + j - f, w\rangle$.

Two important differences between the present case and the simple developments for ordinary differential equations that were presented in Sec. III deserve attention. In the case of ordinary differential equations, obtaining information about the sought solution (or possibly its derivative) at internal nodes was our main goal. Thus algorithms for which all the information was concentrated at internal nodes were developed, and by doing so, it was possible to predict the exact values of the sought solution at such nodes. This was feasible because TH-complete systems for 1D problems are finite dimensional. As opposed to such simple developments, TH-complete systems for a partial differential equation such as (4.4) are infinite dimensional. However, only a finite number of test functions can be used. Different choices of test functions that satisfy Eq. (5.2) lead to different classes of approximations, including optimal spatial methods and general characteristic methods.

On the other hand, when a numerical integration procedure is applied to Eq. (4.4), step by step in time, the objective is to predict the values of u at time t^{n+1}, when the values at time t^n and the boundary conditions are given. Ideally, all the information about the sought solution should be concentrated in the value of the solution u at each one of the subintervals $[x_{i-1}, x_i]$, $i = 1, \ldots, E$, at time t^{n+1}. For example, our goal could be obtaining the $\mathcal{L}^2([0, l])$ projection of the exact solution $u(x, t^{n+1})$ on the subspace of piecewise linear interpolators that are globally continuous. This subspace of $\mathcal{L}^2([0, l])$ is generated by the system of functions

$$w^i(x_1, t^{n+1}) = \begin{cases} \dfrac{x - x_{i-1}}{\Delta x}, & x_{i-1} \leq x \leq x_i \\ \dfrac{x_{i+1} - x}{\Delta x}, & x \leq x \leq x_{i+1}. \end{cases} \tag{5.3}$$

This requires elimination of all information about the sought solution except at such subinterval and time. The weighting functions that do such job, in addition to satisfying $\mathcal{L}^* w^i = 0$ in Ω, must be smooth (i.e., $[w^i] = [\partial w^i / \partial x] = 0$) in the interior $(0, \ell) \times (t^n, t^{n+1})$ of Ω and must fulfill the boundary conditions $\mathcal{C}(w, \cdot) = 0$ on the lateral boundary of Ω, where $\mathcal{C}(w, \cdot)$ is given by Eqs. (4.13c)–(4.13e). Also, $w^i(x, t^{n+1}) = 0$, except when $x \in [x_{i-1}, x_{i+1}]$. In the interval $[x_{i-1}, x_{i+1}]$, one requires that $w^i(x, t^{n+1})$ be given by Eq. (5.3), by virtue of (4.13a). Then the resulting initial boundary value problem generally will be well posed [49], but such a weighting function would be nonlocal.

Generally, in numerical applications, localized weighting functions are sought. At a general interior node x_i, as the one illustrated in Fig. 2, such localization can be achieved by introducing nonsmooth weighting functions. Thus, if the condition $\mathcal{L}^* w^i \equiv 0$ is sustained, then either $[\partial w^i / \partial x] \neq 0$, or $[w] \neq 0$, or both, and some information about the solution u, or its normal derivative, or both, on the curves Σ_j, where discontinuities occur, will be incorporated in the final system of equations. This is so in spite of the fact that the actual objective is to obtain information about the sought solution at time t^{n+1}. The classification of numerical methods into OSM and CM can be related to the speed of propagation of discontinuity lines. If time-independent solutions of Eq. (5.2) are chosen as weighting functions, then $V_\Sigma = 0$ necessarily, and one is led to optimal spatial methods, to which several papers have been devoted using the LAM approach [12–18]. On the other hand, if the lines Σ_j satisfy $V_\Sigma = V$, characteristic methods are obtained.

As mentioned above, there are also several possibilities for the degree of smoothness of the weighting functions. In Paper I, weighting functions satisfying the condition $[w] = 0$ were chosen. In view of Eq. (4.12b), it is clear that for this special choice, $\mathcal{X}^1(w, u)$ vanishes identically, and that, in the lines of discontinuity Σ_i $(i = 1, \ldots, E)$,

all the information is concentrated in the sought solution u. In this case Eq. (4.12a) becomes

$$\mathscr{X}^0(w,u) = (1 + V_\Sigma^2)^{-1/2}\dot{u}\left[D\frac{\partial w}{\partial x}\right]. \tag{5.4}$$

Assume that our goal is, as before, to obtain the $\mathscr{L}^2([0,l])$ projection of the exact solution $u(x,t^{n+1})$, on the subspace of piecewise linear interpolators that are globally continuous. In the spirit of the previous developments and taking into account the limitation $[\partial w/\partial w] \neq 0$, which is unavoidable, a suitable set of properties for the test functions $w^i(x,t)$, is the following:

(a) Support of w^i is $\Omega^i = \Omega_1^i \cup \Omega_2^i$, where $\Omega_1^i = \Omega_i$ and $\Omega_2^i = \Omega_{i+1}$ $(i = 1,\ldots, E-1)$. See Fig. 2.
(b) w^i satisfies Eq. (5.2), i.e., $\mathscr{L}^* w^i = 0$ in Ω.
(c) At $t = t^{n+1}$, w^i reduces to the piecewise linear interpolator given by Eq. (5.3).
(d) w^i is continuous.
(e) The jump $[\partial w^i/\partial x]$ is constant on Σ_i.
(f) At the lateral boundary of Ω, boundary conditions which eliminate all the boundary information [i.e., $\mathscr{C}(w^i,\cdot) = 0$] are imposed.

By inspection of Eqs. (4.13c)–(4.13e), it is seen that the last condition is

$$w^i = 0 \quad \text{on } \partial_D\Omega\,, \tag{5.5a}$$

$$D\frac{\partial w^i}{\partial x_i} + Vw^i = 0 \quad \text{on } \partial_N\Omega\,, \tag{5.5b}$$

$$\frac{\partial w^i}{\partial x} = 0 \quad \text{on } \partial_F\Omega\,. \tag{5.5c}$$

Observe that in the case of Dirichlet boundary conditions, Eq. (5.5a) is to be applied even if the option (4.13f) for defining $\mathscr{B}$ and $\mathscr{C}$ is used.

The development of test functions with these properties is not easy, in general, when the coefficients are nonconstant, even if the domain Ω^i does not intersect the lateral boundaries, but may become particularly involved when the domain intersects one of the lateral boundaries. For the case when the coefficients of Eq. (4.4) are constant, the source term vanishes ($R \equiv 0$), and the partition is uniform, the test functions used in Paper I, were

$$w^i(x,t) = \begin{cases} \dfrac{x - x_{i-1}}{\Delta x} + V\dfrac{t^{n+1} - t}{\Delta x}, & (x,t) \in \Omega_1^i \\ \dfrac{x_{i+1} - x}{\Delta x} + V\dfrac{t^{n+1} - t}{\Delta x}, & (x,t) \in \Omega_2^i \\ 0, \quad \text{all other } (x,t)\,. \end{cases} \tag{5.6}$$

If the domain Ω^i does not intersect the lateral boundaries, these weighting functions satisfy all the required properties, (a)–(f) above; however, if the lateral boundary is intersected by the corresponding domain, then (f) is violated.

An important advantage of the ELLAM approach is precisely its ability to deal with boundary conditions effectively. As was demonstrated through numerical examples in Paper I, the ELLAM approach provides a systematic and consistent methodology for the proper incorporation of boundary conditions. This allows construction of an overall

approximation that possesses the conservative property, thereby assuring conservation of mass in the numerical solution.

Observe that for such weighting functions, $\mathcal{J}$ and $\mathcal{X}$ do not vanish on three lines of discontinuity, at most Σ_{i-1}, Σ_i, and Σ_{i+1}. Thus

$$\langle K^* u, w^i \rangle = \sum_{j=i-1}^{i+1} \left\langle K_j^* u, w^i \right\rangle. \tag{5.7}$$

The jumps are

$$\left[\frac{\partial w}{\partial x} \right]_{i-1} = \frac{1}{\Delta x}, \quad \left[\frac{\partial w}{\partial x} \right]_i = \frac{-2}{\Delta x}, \quad \left[\frac{\partial w}{\partial x} \right]_{i+1} = \frac{1}{\Delta x}. \tag{5.8}$$

When the region Ω^i, which includes the support of w^i, does not intersect the lateral boundaries, the boundary terms vanish and the variational principle in terms of the sought information (5.1) reduces to Eq. (9) of Paper I:

$$\int_{x_{i-1}}^{x_{i+1}} u(x, t^{n+1}) w^i(x, t^{n+1})\, dx - \frac{D}{\Delta x}\left(\int_{\Sigma_{i-1}} u\, dt - 2 \int_{\Sigma_i} u\, dt + \int_{\Sigma_{i+1}} u\, dt \right)$$
$$= \int_{x_{i-1}^*}^{x_{i+1}^*} u(x, t^n) w^i(x, t^n)\, dx + \int_\Omega f_\Omega w^i\, dx\, dt. \tag{5.9}$$

When the region Ω^i does intersect the lateral boundary, as illustrated in Fig. 3, for an inflow boundary, the equation associated with node x_i in the figure, is

$$\int_{x_{i-1}}^{x_{i+1}} u(x, t^{n+1}) w^i(x, t^{n+1})\, dx - \frac{D}{\Delta x}\left(\int_{\Sigma_{i-1}} u\, dt - 2 \int_{\Sigma_i} u\, dt + \int_{\Sigma_{i+1}} u\, dt \right)$$
$$- \int_{t^n}^{t^*} \left(Vu - D \frac{\partial u}{\partial x} \right)\Big|_{x=0} w^i(0, t)\, dt = \int_{x_0}^{x_{i+1}^*} u(x, t^n) w^i(x, t^n)\, dx$$
$$+ \frac{D}{\Delta x} \int_{t^n}^{t^*} u(0, t)\, dt + \int_\Omega f_\Omega w^i\, dx\, dt. \tag{5.10}$$

This equation follows from Eq. (15b) of Paper I. The additional terms relative to Eq. (5.9) are due to nonzero w^i at $x = 0$ [i.e., they are due to violation of property (f) above]. However, this leads naturally to the presence of the total flux term $(\partial Vu - D\partial u/x)$ at the inflow boundary. This is physically appropriate, leading to global mass conservation,

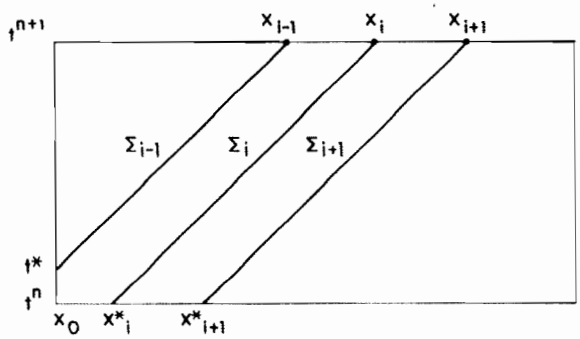

FIG. 3. Case when the support of w^i intersects the inflow boundary.

and the resulting set of equations yields accurate numerical results [1]. An alternative formulation with very similar properties, based on integrating by parts once instead of twice, is developed in [19], also with accurate numerical results; error estimates of optimal order based on this formulation are proved in [52]. As mentioned in Paper I, the integrals that appear in these equations may be approximated in many different ways. Different approximations of these integrals lead to different CM algorithms reported in the literature.

VI. MULTIDIMENSIONAL ADVECTION-DIFFUSION EQUATION

Application of the algebraic theory to multidimensional problems is straightforward. Only slight modifications have to be made in Eqs. (4.4)–(4.12). The region Ω_x is multidimensional in this case, and the equations are

$$\mathscr{L}u \equiv \frac{\partial u}{\partial t} + \nabla \cdot (\underline{V}u) - Ru - \nabla \cdot (D \cdot \nabla u) = f_\Omega(x,t) \quad \text{in } \Omega\,. \tag{6.1}$$

The initial conditions (4.5) are sustained. The adjoint operator is

$$\mathscr{L}^* w \equiv -\frac{\partial w}{\partial t} - \underline{V} \cdot \nabla u - Rw - \nabla \cdot (D \cdot \nabla w)\,. \tag{6.2}$$

$\underline{\mathscr{D}}(u,w)$, as defined by Eq. (2.1), is

$$\underline{\mathscr{D}}(u,w) = \{uD\nabla w + Vw) - wD\nabla u,\ uw\}\,. \tag{6.3}$$

Therefore

$$\mathscr{R}_\Sigma(u,w) = -[\mathscr{D}(u,w)] \cdot \underline{N} = \left[u\left(D\frac{\partial w}{\partial n} + (V - V_\Sigma)w\right) - wD\frac{\partial u}{\partial n}\right], \tag{6.4}$$

with the same assumptions as before. The smoothness conditions implied by mass balance and Fickian diffusion are

$$[u] = \left[\frac{\partial u}{\partial n}\right] = 0 \quad \text{on } \Sigma\,. \tag{6.5}$$

Using Eqs. (2.8), it is seen that

$$\mathscr{J}^0(u,w) = -(1 + V_\Sigma^2)^{-1/2}[u]\left\{\overline{D\frac{\partial w}{\partial n}} + \dot{w}(V - V_\Sigma)\right\}, \tag{6.6a}$$

$$\mathscr{J}^1(u,w) = (1 + V_\Sigma^2)^{-1/2}\dot{w}D\left[\frac{\partial u}{\partial n}\right], \tag{6.6b}$$

$$\mathscr{X}^0(w,u) = (1 + V_\Sigma^2)^{-1/2}\left\{\dot{u}\left[D\frac{\partial w}{\partial n}\right] + (V - V_\Sigma)[w]\right\}, \tag{6.7a}$$

$$\mathscr{X}^1(w,u) = -(1 + V_\Sigma^2)^{-1/2}[w]D\overline{\frac{\partial u}{\partial n}}\,. \tag{6.7b}$$

Finally, the bilinear functionals $\mathscr{B}(u,w)$, and $\mathscr{C}(w,u)$, associated with initial and boundary conditions, are

$$\mathscr{B}(u,w) = -uw \quad \text{on } \partial_n\Omega, \quad \mathscr{C}(w,u) = -uw \quad \text{on } \partial_{n+1}\Omega\,, \tag{6.8a}$$

$$\mathscr{B}(u,w) = uD\frac{\partial w}{\partial n}, \quad \mathscr{C}(w,u) = w\left(D\frac{\partial u}{\partial n} - Vu\right) \quad \text{on } \partial_D\Omega\,, \tag{6.8b}$$

$$\mathfrak{B}(u,w) = -wD\frac{\partial u}{\partial n}, \quad \mathfrak{C}(w,u) = -u\left(D\frac{\partial w}{\partial n} + Vw\right) \quad \text{on } \partial_N\Omega\,, \tag{6.8c}$$

$$\mathfrak{B}(u,w) = -w\left(D\frac{\partial u}{\partial n} - Vu\right), \quad \mathfrak{C}(w,u) = -uD\frac{\partial w}{\partial n} \quad \text{on } \partial_F\Omega\,. \tag{6.8d}$$

Although the theoretical extension to multiple dimensions has been straightforward, the numerical implementation for this case is considerably more complicated.

VII. SYSTEMS OF EQUATIONS: MIXED METHODS

As mentioned above, localized adjoint methods are very general. In particular, they are applicable to systems of equations. This section is devoted to explaining such applications and to developing mixed methods as an illustration.

A. Systems of Equations

Let $\mathbf{u}$ be a vector valued function defined in the region Ω. A linear system of equations can be written as

$$\underline{\mathscr{L}}\mathbf{u} = \mathbf{f}_\Omega \quad \text{in } \Omega\,, \tag{7.1}$$

where $\underline{\mathscr{L}}$ is a linear differential transformation. The adjoint $\underline{\mathscr{L}}^*$ of $\underline{\mathscr{L}}$ is defined by the condition

$$\mathbf{w}\cdot\underline{\mathscr{L}}\mathbf{u} - \mathbf{u}\cdot\underline{\mathscr{L}}^*\mathbf{w} = \nabla\cdot\{\underline{\mathfrak{D}}(\mathbf{u},\mathbf{w})\}, \tag{7.2}$$

where $\underline{\mathfrak{D}}(\mathbf{u},\mathbf{w})$ is a vector-valued bilinear function.

Application of generalized divergence theorem yields

$$\int_\Omega \{\mathbf{w}\cdot\underline{\mathscr{L}}u - \mathbf{u}\cdot\underline{\mathscr{L}}^*\mathbf{w}\}\,dx = \int_{\partial\Omega}\mathfrak{R}_\partial(\mathbf{u},\mathbf{w})\,dx + \int_\Sigma \mathfrak{R}_\Sigma(\mathbf{u},\mathbf{w})\,dx\,, \tag{7.3}$$

where

$$\mathfrak{R}_\partial(\mathbf{u},\mathbf{w}) = \underline{\mathfrak{D}}(\mathbf{u},\mathbf{w})\cdot\underline{n} \quad \text{and} \quad \mathfrak{R}_\Sigma(\mathbf{u},\mathbf{w}) = -[\underline{\mathfrak{D}}(\mathbf{u},\mathbf{w})]\cdot\underline{n}\,. \tag{7.4}$$

These equations are very similar to Eqs. (2.2) and (2.3), and the other equations of Sec. II [(2.4)–(2.8)] are essentially the same. Thus, associated with every kind of boundary conditions, one has a decomposition

$$\mathfrak{R}_\partial(\mathbf{u},\mathbf{w}) \equiv \underline{\mathfrak{D}}(\mathbf{u},\mathbf{w})\cdot\underline{n} = \mathfrak{B}(\mathbf{u},\mathbf{w}) - \mathfrak{C}(\mathbf{w},\mathbf{u})\,, \tag{7.5}$$

where $\mathfrak{B}(\mathbf{u},\mathbf{w})$ and $\mathfrak{C}(\mathbf{w},\mathbf{u})$ are two bilinear functions. Using the identity

$$[\underline{\mathfrak{D}}(\mathbf{u},\mathbf{w})] = \underline{\mathfrak{D}}([\mathbf{u}],\dot{\mathbf{w}}) + \underline{\mathfrak{D}}(\dot{\mathbf{u}},[\mathbf{w}])\,, \tag{7.6}$$

it is seen that

$$\mathfrak{R}_\Sigma(\mathbf{u},\mathbf{w}) = \mathscr{J}(\mathbf{u},\mathbf{w}) - \mathscr{K}(\mathbf{w},\mathbf{u})\,, \tag{7.7}$$

with

$$\mathscr{J}(\mathbf{u},\mathbf{w}) = -\underline{\mathfrak{D}}([\mathbf{u}],\dot{\mathbf{w}})\cdot\underline{n}\,, \tag{7.8a}$$

$$\mathscr{K}(\mathbf{w},\mathbf{u}) = \underline{\mathfrak{D}}(\dot{\mathbf{u}},[\mathbf{w}])\cdot\underline{n}\,. \tag{7.8b}$$

B. Mixed Methods

As an illustration of how LAM can be used to provide insight into mixed methods, consider the steady state of the general advection-diffusion equation of Sec. IV,

$$\nabla \cdot (D\nabla u) - \nabla \cdot (\underline{V}u) + Ru = f_\Omega(x) \quad \text{in } \Omega\,. \tag{7.9}$$

This can be replaced by the system of equations

$$\underline{p} - D^{1/2}\nabla u + D^{-1/2}\underline{V}u = 0\,, \tag{7.10a}$$

$$\nabla \cdot \left(D^{1/2}\underline{p}\right) + Ru = f_\Omega(\underline{x})\,. \tag{7.10b}$$

In three spatial dimensions, Eqs. (7.10) constitute a system of four equations (underlined quantities are 3 D vectors). Thus, a four-dimensional vector $\mathbf{u} = \left\{\underline{p}, u\right\}$ will be considered, and the linear differential operator $\mathcal{L}$

$$\underline{\mathcal{L}}\mathbf{u} \equiv \left\{\underline{p} - D^{1/2}\nabla u + D^{-1/2}\underline{V}u,\ \nabla \cdot \left(D^{1/2}\underline{p}\right) + Ru\right\}. \tag{7.11}$$

Let $\mathbf{w}$ be a function whose values are the four-dimensional vectors $\left\{\underline{q}, w\right\}$. Then the adjoint differential operator is defined by

$$\underline{\mathcal{L}}^*\mathbf{w} \equiv \left\{\underline{q} - D^{1/2}\nabla w,\ \nabla \cdot \left(D^{1/2}\underline{q}\right) + D^{-1/2}\underline{V} \cdot \underline{q} + Rw\right\}. \tag{7.12}$$

The identity

$$\mathbf{w} \cdot \underline{\mathcal{L}}\mathbf{u} - \mathbf{u} \cdot \underline{\mathcal{L}}^*\mathbf{w} \equiv \nabla \cdot \left\{wD^{1/2}\underline{p} - uD^{1/2}\underline{q}\right\} \tag{7.13}$$

implies that

$$\underline{\mathcal{D}}(\mathbf{u}, \mathbf{w}) = wD^{1/2}\underline{p} - uD^{1/2}\underline{q}\,. \tag{7.14}$$

The smoothness conditions implied by conservation of mass are

$$[u] = 0 \quad \text{and} \quad \left[D^{1/2}\underline{p}\right] \cdot \underline{n} = 0\,. \tag{7.15}$$

When the coefficients are continuous, the latter of these equations is equivalent to $[\underline{p}] \cdot \underline{n} = 0$. Use of Eqs. (7.8) and (7.14) leads to

$$\mathcal{J}(\mathbf{u}, \mathbf{w}) = [u]D^{1/2}\dot{\underline{q}} \cdot \underline{n} - \dot{w}D^{1/2}\left[\underline{p}\right] \cdot \underline{n}\,, \tag{7.16a}$$

$$\mathcal{K}(\mathbf{w}, \mathbf{u}) = [w]D^{1/2}\dot{\underline{p}} \cdot \underline{n} - \dot{u}D^{1/2}\left[\underline{q}\right] \cdot \underline{n}\,. \tag{7.16b}$$

Equation (7.16b) has interesting implications. Because the exact solution $\mathbf{u} = \left\{\underline{p}, u\right\}$ satisfies (7.15), one has $\dot{\underline{p}} \cdot \underline{n} = \underline{p} \cdot \underline{n}$ and $\dot{u} = u$. If it is desired to concentrate all the information in the flux at interelement boundaries, then the weighting functions $\mathbf{w} = \left\{\underline{q}, w\right\}$, in addition to satisfy the complementary boundary conditions $\mathcal{C}(\mathbf{w}, \cdot) = 0$, must satisfy the adjoint system of differential equations $\underline{\mathcal{L}}^*\mathbf{w} = 0$, which is

$$\underline{q} - D^{1/2}\nabla w = 0\,, \tag{7.17a}$$

$$\nabla \cdot \left(D^{1/2}\underline{q}\right) + D^{1/2}\underline{V} \cdot \underline{q} + Rw = 0\,. \tag{7.17b}$$

By inspection of Eq. (7.16b), it is seen that elimination of the information about the function u requires that $[\underline{q}] \cdot \underline{n} = 0$, i.e., that the normal component of $\underline{q}$ be continuous. However, the weighting function w must be discontinuous. This is essentially the Raviart-

Thomas [52] result, which constitutes the basis of mixed methods. This result has been quite useful for approximating the total flux directly as an independent variable. In particular when $D = 1, \underline{V} = 0$, and $R = f_\Omega = 0$, Eq. (7.9) becomes the Laplace equation while (7.17) can be written as

$$\underline{q} = \nabla w, \qquad \nabla \cdot \underline{q} = 0. \tag{7.18}$$

Thus $\underline{q}$ must be incompressible with continuous normal component across interelement boundaries, while w is discontinuous.

VIII. DISCUSSION AND CONCLUSIONS

In a sequence of two papers, localized adjoint method has been applied in space-time to problems of advective-diffusive transport. The approach is based on space-time discretizations in which specialized test functions are applied. These functions satisfy the homogenous adjoint equation locally within each element. The resulting method is referred to as the Eulerian-Lagragian localized adjoint method. The ELLAM approach, in addition to providing a unification of characteristic methods (CM), supplies a systematic framework for incorporation of boundary conditions in CM approximations. Any type of boundary conditions can be accommodated in a mass conservative manner. This seems to be the first complete treatment of boundary conditions in Eulerian-Lagragian methods that leads to a conservative scheme for the general transport equations. Additionally, the ELLAM approach provides a framework within which LAM concepts can be applied to advection-dominated problems, handling time-dependent situations more accurately than OSM. Thus ELLAM combines Eulerian-Lagrangian ideas and the LAM framework to their mutual benefit.

In Paper I [1], the numerical implementation was developed and discussed thoroughly. In this second paper of the series, the theoretical aspects are covered in a more complete form, and the ELLAM procedures are more clearly related with the general LAM framework. This provides a more systematic development of the ELLAM methodology, making it possible to establish a more complete picture of the possibilities that should be explored and the problems that must be tackled in order to make ELLAM a more effective modeling tool. In particular, the LAM framework has been demonstrated to be very suitable for motivating specialized test functions. The effect that different boundary and continuity (or smoothness) conditions, satisfied by test functions, have on approximate solutions is clearly exhibited. Also, the LAM framework leads in a natural manner to a definition of suitable unknowns for a given problem. For example, when developing the numerical implementation of ELLAM in Paper I, it became apparent that in some cases it was necessary to introduce the total flux as an additional unknown at the boundaries, in spite of the fact that the main goal was to predict the value of the function at time t^{n+1}. The generality of the theory was corroborated, once more, by applying it to systems of equations and deriving mixed methods.

However, there are many points that should be studied in more depth. We need a more extensive study of both the theory and implementation of ELLAM techniques for variable coefficients, particularly in multidimensional applications. Implementation of boundary conditions for variable-coefficient problems in multiple dimensions is also an important problem. In addition, the treatment of nonlinear problems deserves further study. Since the unknown variables appear in nonlinear coefficients that are usually evaluated in the interior of mesh blocks via numerical quadrature, greater attention must be placed on

the full approximation-theoric properties of the trial functions in these applications. The potential of local refinement in both space and time holds enormous potential for ELLAM and is the object of ongoing research.

Finally, we want to emphasize that ELLAM forms a general and powerful framework for investigating and comparing a wide variety of numerical methods for problems that have important advective properties. The framework motivates different choices of test functions to approximate different properties of the unknowns or even different unknowns, such as fluxes. The general theory is expanding to provide more insight. These techniques appear to have enormous flexibility and potential for treating general advection-diffusion-reaction problems.

This work was partially supported by the International Atomic Energy Agency under Contract No. 6088/RB, by the International Development Research Centre, Canada, under Contract Nos. 89-1029-02, and by the National Science Foundation under Contracts Nos. DMS8712021, RII-8610680, DMS8821330, and CES8657419.

References

1. M. A. Celia, T. F. Russell, I. Herrera, and R. E. Ewing, "An Eulerian-Lagrangian localized adjoint method for the advection-diffusion equation," *Adv. Water Res.* **13**, 187 (1990).
2. I. Herrera, *Boundary Methods: An Algebraic Theory,* Pitman, London, 1984.
3. I. Herrera, "Unified formulation of numerical methods. I Green's formulas for operators in discontinuous fields," *Numer. Methods Partial Different. Equ.* **1**, 25 (1985).
4. I. Herrera, "Unified approach to numerical methods, Part 2. Finite elements, boundary methods, and its coupling," *Numer. Methods Partial Different. Equ.* **3**, 159 (1985).
5. I. Herrera, L. Chargoy, and G. Alduncin, "Unified approach to numerical methods. III. Finite differences and ordinary differential equations," *Numer. Methods Partial Different. Equ.* **1**, 241 (1985).
6. I. Herrera, "Some unifying concepts in applied mathematics," in *The Merging of Disciplines: New Directions in Pure, Applied, and Computational Mathematics,* R. E. Ewing, K. I. Gross, and C. F. Martin, Eds., Springer-Verlag, New York, 1986, pp. 79–88.
7. M. A. Celia and I. Herrera, "Solution of general ordinary differential equations using the algebraic theory approach," *Numer. Methods Partial Different. Equ.*, **3**, 117 (1987).
8. I. Herrera, "The algebraic theory approach for ordinary differential equations: Highly accurate finite differences," *Numer. Methods Partial Different. Equ.* **3**, 199 (1987).
9. I. Herrera and L. Chargoy, *An Overview of the Treatment of Ordinary Differential Equations by Finite Differences,* Pergamon, Oxford, 1987, Vol. 8, pp. 17–19.
10. M. A. Celia, I. Herrera, and E. T. Bouloutas, "Adjoint Petrov-Galerkin methods for multi-dimensional flow problems," in *Finite Element Analysis in Fluids,* T. J. Chung and Karr R., Eds., UAH Press, Huntsville, AL, 1989, pp. 953–958.
11. I. Herrera, "New method for diffusive transport," in *Groundwater Flow and Quality Modelling,* Reidel, Dordrecht, 1988, pp. 165–172.
12. I. Herrera, "New approach to advection-dominated flows and comparison with other methods," in *Computational Mechanics '88,* Springer-Verlag, Heidelberg, 1988, Vol. 2.
13. M. A. Celia, I. Herrera, E. T. Bouloutas, and J. S. Kindred, "A new numerical approach for the advective-diffusive transport equation," *Numer. Methods Partial Different. Equ.* **5**, 203 (1989).
14. M. A. Celia, J. S. Kindred, and I. Herrera, "Contaminant transport and biodegradation: 1. A numerical model for reactive transport in porous media," *Water Resourc. Res.*, **25**, 1141 (1989).
15. I. Herrera and G. Hernández, "Advances on the numerical simulation of steep fronts," in *Numerical Methods for Transport and Hydrologic Processes,* Vol. 2, M. A. Celia, L. A. Ferrand, and G. Pinder, Eds., Developments in Water Science Computational Mechanics Publications, Elsevier, Amsterdam, 1988, Vol. 36, pp. 139–145.
16. I. Herrera, M. A. Celia, and J. D. Martinez, "Localized adjoint method as a new approach to advection dominated flows." in *Recent Advances in Ground-Water Hydrology,* J. E. Moore, A. A. Zaporozec, S. C. Csallany, and T. C. Varney, Eds., American Institute of Hydrology, 1989, pp. 321–327.

17. I. Herrera, "Localized adjoint methods: Application to advection-dominated flows," in *Groundwater Management: Quantity and Quality,* IAHS Publ. No 188, 1989, pp. 349–357.
18. I. Herrera, "Advances in the numerical simulation of steep fronts," in *Finite Element Analysis in Fluids,* T.J. Chung and R. Karr, Eds. University of Alabama Press, Huntsville, 1989, pp. 965–970.
19. T.F. Russell, "Eulerian-Lagrangian localized adjoint methods for advection-dominated problems," in *Numerical Analysis 1989,* G.A. Watson and D.F. Griffiths, Eds., Pitmman Research Notes in Mathematics Vol. 228, Longman Scientific and Technical, Harlow, U.K., 1990, pp. 206–228.
20. I. Herrera, "Localized adjoint methods in water resources problems," in *Computational Methods in Surface Hydrology,* G. Gambolati, A. Rinaldo, and C.A. Brebbia, Eds., Springer-Verlag, Berlin, 1990, pp. 433–440.
21. I. Herrera, "Localized adjoint methods: A new discretization methodology," in *Computational Methods in Geosciences,* W.E. Fitzgibon and M.F. Wheeler, Eds., SIAM, Bellingham, WA, 1992, pp. 66–77.
22. T.F. Russell and R.V. Trujillo, "Eulerian-Lagrangian localized adjoint methods with variable coefficients in multiple dimensions," in *Computational Methods in Surface Hydrology,* G. Gambolati *et al.,* Eds., Springer Verlag, Berlin, 1990, pp. 357–363.
23. R.E. Ewing, "Operator splitting and Eulerian-Lagrangian localized adjoint methods for multiphase flow" J. Whiteman (ed.), MAFELAP 1990, Academic Press, San Diego, 1991, pp. 215–232.
24. I. Herrera and R.E. Ewing, "Localized adjoint methods: Applications to multiphase flow problems," Proceedings, Fifth Wyoming Enhanced Oil Recovery Symposium, 10–11 May 1989, Enhanced Oil Recovery Institute, University of Wyoming, 1990, pp. 155–173.
25. R.E. Ewing and M.A. Celia, "Multiphase flow simulation in groundwater hydrology and petroleum engineering," in *Computational Methods in Subsurface Hydrology,* G. Gambolati *et al.,* Eds., Springer Verlag, Berlin, 1990, pp. 195–202.
26. S. Zisman, "Simulation of contaminant transport in groundwater systems using Eulerian-Lagrangian localized adjoint methods," M.S. thesis, MIT, 1989.
27. M.A. Celia and S. Zisman, "Eulerian-Lagrangian localized adjoint method for reactive transport in groundwater," in *Computational Methods in Subsurface Hydrology,* Ref. 25, pp. 383–390.
28. S.P. Neuman, "Adjoint Petrov-Galerkin method with optimum weight and interpolation functions defined on multidimensional nested grids," in *Computational Methods in Surface Hydrology* (Ref. 22), pp. 347–356.
29. D. Allen and R. Southwell, "Relaxation methods applied to determining the motion in two dimensions of a fluid past a fixed cylinder," Q. *J. Mech. Appl. Math.* **8**, 129 (1955).
30. E.T. Bouloutas and M.A. Celia, "An analysis of a class of Petrov-Galerkin and optimal test functions methods," Proc. Seventh Int. Conf. Computational Methods in Water Resources, Celia *et al.,* Eds., 1988, pp. 15–20.
31. E.T. Bouloutas and M.A. Celia, "An improved cubic Petrov-Galerkin method for advection-dominated flows in rectangularly decomposable domains," *Comput. Methods* Appl. Mech. Eng. (to appear).
32. M.E. Cantekin and J.J. Westerink, "Nondiffusive $N + 2$ degree Petrov-Galerkin methods for two-dimensional transient transport computations," *Int. J. Numer. Methods Eng.* **30**, 397 (1990).
33. J.J. Westerink and D. Shea, "Consistent higher-degree Petrov-Galerkin methods for the solution of the transient convection-diffusion equation," *Int. J. Numer. Methods Eng.* **28**, 1077 (1989).
34. A.M. Baptista, "Solution of advection-dominated transport by Eulerian-Langrangian methods using the backward methods of characteristics," Ph.D. thesis, MIT, 1987.
35. S.P. Neuman, "An Eulerian-Langrangian numerical scheme for the dispersion-convection equation using conjugate space-time grids," *J. Comput. Phys.* **41**, 270 (1981).
36. S.P. Neuman, "Adaptive Eulerian-Lagrangian finite element method for advection-dispersion," *Int. J. Numer. Eng.* **20**, 321 (1984).
37. S.P. Neuman and S. Sorek, "Eulerian-Lagrangian methods for advection-dispersion," *Methods Proceeding in Fourth International Conference on Finite Elements in Water Resources,* Holz *et al.,* Eds., Springer-Verlag, Berlin, 1982, pp. 14, 41-14-68.
38. O. Pironneau, "On the transport-diffusion algorithm and its application to the Navier-Stokes equations," *Numer. Math.* **38**, 309 (1982).