# A New Numerical Approach for the Advective-Diffusive Transport Equation

**Michael A. Celia**
*Department of Civil Engineering, Rm. 48-207, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139*

**Ismael Herrera**
*Instituto de Geofisica, UNAM, Apdo. Postal 22-582, 14000 Mexico D.F., Mexico*

**Efthimios Bouloutas**
*Department of Civil Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139*

**J. Scott Kindred**
*Golder Associates, Inc., Redmond, Washington 98052*

A new numerical solution procedure is presented for the one-dimensional, transient advective-diffusive transport equation. The new method applies Herrera's algebraic theory of numerical methods to the spatial derivatives to produce a semi-discrete approximation. The semi-discrete system is then solved by standard time marching algorithms. The algebraic theory, which involves careful choice of test functions in a weak form statement of the problem, leads to a numerical approximation that inherently accommodates different degrees of advection domination. Algorithms are presented that provide either nodal values of the unknown function or nodal values of both the function and its spatial derivative. Numerical solution of several test problems demonstrates the behavior of the method.

## INTRODUCTION

Numerical solution of the advective-diffusive transport equation is of great importance in many fields of engineering and science. When the diffusive process dominates, the transport equation is relatively easy to solve by virtually any standard numerical scheme. However, when the problem is dominated by advection, standard numerical approximations become problematic. Either nonphysical oscillations appear in the vicinity of sharp fronts, or excessive numerical diffusion is introduced and the ability to capture a sharp front is precluded.

Considerable effort has been expended in developing discretization formulas for the advective-diffusive transport equation. Most formulations involve some sort of upstream weighting to accommodate the advective nature of the transport. These include classical upstream weighted finite differences [20], high-order upstream finite differences [21], collocation methods [1, 22], and a variety of finite element, or Petrov-Galerkin methods [3, 8, 10, 16–18]. A fundamental

Numerical Methods for Partial Differential Equations, 5, 203–226 (1989)
 CCC 0749-159X/89/03203-24$04.00

criticism of many of these methods is the presence of an arbitrary parameter, the choice of which must be decided by the analyst. Criteria to assist in this choice are provided by Hughes and co-workers [17, 18].

This paper proposes an alternative approach, based on the original ideas of Herrera [12–14]. In this approach, finite difference and boundary element techniques are incorporated in a new "algebraic theory." The methodology leads to an optimal approximation in the following sense. Several researchers [9] have considered test functions used in a weak formulation to be optimal when they yield exact values at the nodes. This notion of optimality can be generalized by considering test functions to be optimal when they yield exact values at inter-element boundaries. One can then prove [11] that for a system of test functions to possess this optimality property, for arbitrary forcing functions, it is necessary that the test functions be solutions of the homogeneous adjoint differential equation associated with the original governing differential operator. In the approach of Herrera, test functions are constructed based on this criterion. The resulting test functions have the following advantages: (1) there are no arbitrary parameters that appear in their definition; (2) the functions vary continuously with the coefficients of the governing equation (for example, velocity, diffusion coefficient, reaction rate); and (3) definition of the functions results from a systematic and mathematically sound formulation. We refer to this method as an Optimal Test Function (OTF) method.

The presentation in this paper is organized as follows: The underlying numerical procedure is first developed for the one-dimensional, steady-state, constant coefficient case. Optimal test functions are chosen to satisfy the homogeneous adjoint equation, and thus yield exact nodal values of the sought solution (and its derivative, if desired). Treatment of variable coefficients is discussed briefly, and an example calculation is presented. The procedure is then extended to time-dependent problems by using a semidiscretization is space. Details are presented regarding treatment of the temporal derivative. Next, a detailed discussion on choices of test functions is presented. This includes a comparison of test functions based on OTF to test functions used by Hughes and co-workers [16–18] in their Petrov-Galerkin formulation. Numerical results are then presented, including numerical comparisons between OTF and upwind finite differences, standard Galerkin finite elements, orthogonal collocation, and Hughes Petrov-Galerkin method. Finally, a general discussion is presented that includes comments on future directions for this research. This focuses on extension to multiple dimensions.

## DEVELOPMENT OF THE APPROXIMATION

This section details the salient features of the numerical approximation method. The underlying theory has been presented by Herrera and co-workers [5, 11–15] for general boundary value problems. The presentation that follows focuses specifically on the transport equation.

Let us begin by developing the approximation for the one-dimensional steady-state transport equation, given by

$$\mathcal{L}u \equiv D\frac{d^2u}{dx^2} - V\frac{du}{dx} = f(x), \qquad 0 \le x \le \ell \tag{1}$$

$$u(0) = g_0 \tag{2a}$$

$$u(\ell) = g_\ell \tag{2b}$$

First type boundary conditions, and constant coefficients, are chosen for convenience of presentation only. The numerical procedure is applicable to general differential equations and general boundary conditions.

The numerical algorithm developed herein is based on a weak form of the governing differential equation. For eq. (1), this is expressed as

$$\int_0^\ell (\mathcal{L}u)w(x)\,dx = \int_0^\ell f(x)w(x)\,dx \tag{3}$$

where $w(x)$ is a weight, or test, function. Next, the domain $[0, \ell]$ is subdivided into $E$ subintervals, or elements, $\{[x_0, x_1], [x_1, x_2], \ldots, [x_{E-1}, x_E]\}$, where $x_0 = 0$ and $x_E = \ell$. These subintervals are separated by $E + 1$ node points, $\{x_0, x_1, \ldots, x_E\}$. Let us now assume that the solution $u(x)$ possesses continuity of both function and derivative at all points in $[0, \ell]$, $u(x) \in \mathbb{C}^1[0, \ell]$. In addition, let the test function $w(x)$ exhibit discontinuities within $[0, \ell]$, $w(x) \in \mathbb{C}^{-1}[0, \ell]$. Finally, let any discontinuities in $w(x)$ be assumed to occur only at node points. Under these rather general conditions, the integral on the left side of eq. (3) can be written equivalently as

$$\int_0^\ell (\mathcal{L}u)w(x)\,dx = \sum_{j=0}^{E-1} \int_{x_j}^{x_{j+1}} (\mathcal{L}u)w(x)\,dx \tag{4}$$

Equation (4), which introduces elementwise integration, can be written because of the continuity assumptions on $u(x)$, $w(x)$. The case of lower continuity on $u(x)$ has been treated in the general development of [11, 12]. For the current presentation, $u \in \mathbb{C}^1$ suffices since no solutions of interest are precluded.

There are two key steps in the numerical development. The first is application of integration by parts twice to each element integration of eq. (4). The second is a special choice of test function, $w(x)$; this choice is based on the adjoint operator that arises from integration by parts.

Let each element integral in eq. (4) be integrated by parts twice. Given the definition of $\mathcal{L}$ in eq. (1), the first application of integration by parts yields

$$\int_{x_j}^{x_{j+1}} \left(D\frac{d^2u}{dx^2} - V\frac{du}{dx}\right)w(x)\,dx = \left[Dw\frac{du}{dx} - Vwu\right]_{x_j}^{x_{j+1}} \int_{x_j}^{x_{j+1}} \left(D\frac{du}{dx} - Vu\right)\frac{dw}{dx}\,dx \tag{5a}$$

A second application of integration by parts produces

$$\int_{x_j}^{x_{j+1}} \left(D\frac{d^2u}{dx^2} - V\frac{du}{dx}\right) w(x)\,dx = \left[Dw\frac{du}{dx} - D\frac{dw}{dx}u - Vwu\right]_{x_j}^{x_{j+1}} + \int_{x_j}^{x_{j+1}} \left(D\frac{d^2w}{dx^2} + V\frac{dw}{dx}\right) u(x)\,dx$$

$$= \left[Dw\frac{du}{dx} - D\frac{dw}{dx}u - Vwu\right]_{x_j}^{x_j} + \int_{x_j}^{x_{j+1}} (\mathcal{L}^*w)u(x)\,dx$$

where $\mathcal{L}^*$ is the formal adjoint of the operator $\mathcal{L}$. Equation (5b) replaces the original spatial integral by nodal evaluations (these are the boundary terms in the integration by parts, since the boundaries are nodes $x_j$, $x_{j+1}$), and a spatial integration involving the adjoint operator applied to the test function.

The next step is elimination of the spatial integral on the right side of eq. (5b). Examination of eq. (5b) indicates that elimination of the integral allows the original spatial integral to be written equivalently in terms of only nodal values. As such, all information is effectively concentrated at node points. To achieve this concentration of information, the test function $w(x)$ is chosen so that it satisfies the homogeneous adjoint equation $\mathcal{L}^*w = 0$ within each element. This eliminates the spatial integral on the right side of eq. (5b) by forcing the integrand to zero. Upon elimination of this integral, eqs. (3), (4), and (5b) combine to form the following equality,

$$\int_0^{\ell} (\mathcal{L}u)w(x)\,dx = \sum_{j=0}^{E-1} \left\{\left[Dw\frac{du}{dx} - D\frac{dw}{dx}u - Vwu\right]_{x_j}^{x_{j+1}}\right\} \tag{6a}$$

Since $u(x) \in \mathbb{C}^1[0, \ell]$, nodal evaluations of both $u$ and $du/dx$ are unique, because these functions are continuous. However, since $w(x) \in \mathbb{C}^{-1}[0, \ell]$, nodal values of $w$ and $dw/dx$ are generally nonunique and depend on the element from which the node is approached. Given these observations, eq. (6a) can be written equivalently as

$$\int_0^{\ell} (\mathcal{L}u)w(x)\,dx = \sum_{j=1}^{E-1} \left\{\left[\left[D\frac{dw}{dx} + Vw\right]\right]_{x_j} u_j - [[Dw]]_{x_j}\frac{du}{dx}j + \left[\left(-D\frac{dw}{dx} - Vw\right)u + (Dw)\frac{du}{dx}\right]_0^{\ell}\right. \tag{6b}$$

where the double bracket denotes a jump operator, $[[\cdot]]_{x_j} \equiv (\cdot)_{x_j^+} - (\cdot)_{x_j^-}$. Since $D$, $V$, and $w$ are known (explicit functional forms for $w(x)$ are derived below), the only unknowns in eq. (6b) are the nodal values of function $u$ and its derivative, $\{u_j, (du/dx)j\}_{j=0}^{E}$. Equation (6b) can be written more succinctly as

$$\int_0^{\ell} (\mathcal{L}u)w(x)\,dx = \sum_{j=0}^{E} A_j u_j + B_j\frac{du}{dx}j \tag{7}$$

The only step that remains is explicit definition of a suitable set of test functions. The formal adjoint operator associated with the operator $\mathcal{L}$ of eq. (1) is

$$\mathcal{L}^* \equiv D\frac{d^2}{dx^2} + V\frac{d}{dx} \tag{8}$$

There are two linearly independent solutions to the homogeneous equation, $\mathcal{L}^*\Psi = 0$. These solutions are given by

$$\Psi_1(x) = 1 \tag{9a}$$

$$\Psi_2(x) = \exp\left[-\left(\frac{V}{D}\right)x\right] \tag{9b}$$

Any linear combination of $\Psi_1(x)$, $\Psi_2(x)$ also satisfies the homogeneous adjoint operator equation. The computational procedure uses two linearly independent solutions to $\mathcal{L}^*w = 0$ within each subinterval. When the test functions are fully discontinuous, $w(x) \in \mathbb{C}^{-1}[0, \ell]$, the following choice of test functions is taken, with superscript $e$ referring to the element defined by $[x_j, x_{j+1}]$,

$$w_1^e(x) = \begin{cases} \dfrac{\exp\left[-\left(\dfrac{V}{D}\right)(x - x_j)\right] - \exp\left[-\left(\dfrac{V}{D}\right)(x_{j+1} - x_j)\right]}{1 - \exp\left[-\left(\dfrac{V}{D}\right)(x_{j+1} - x_j)\right]} & x_j < x < x_{j+1} \\ 0 & x < x_j, \quad x > x_{j+1} \end{cases} \tag{10a}$$

$$w_2^e(x) = \begin{cases} \dfrac{1 - \exp\left[-\left(\dfrac{V}{D}\right)(x - x_j)\right]}{1 - \exp\left[-\left(\dfrac{V}{D}\right)(x_{j+1} - x_j)\right]} & x_j < x < x_{j+1} \\ 0 & x < x_j, \quad x > x_{j+1} \end{cases} \tag{10b}$$

This particular linear combination of the fundamental solutions (9a), (9b) are chosen simply because they possess the properties $w_1^e(x_j) = 1$, $w_1^e(x_{j+1}) = 0$, $w_2^e(x_j) = 0$, $w_2^e(x_{j+1}) = 1$. This is a choice based on convenience only; any other linear combination is equally acceptable.

The choice of test functions given by eqs. (10) produces $2E$ equations for the $2E + 2$ nodal unknowns $\{u_j, (du/dx)j\}_{j=0}^{E}$. These equations are linear algebraic equations of the form of eq. (7), with coefficients $A_j$ and $B_j$ evaluated from the definitions (10). Two additional equations are obtained from the boundary conditions. This closes the system, producing a $(2E + 2) \times (2E + 2)$ matrix equation for the nodal unknowns.

Solution of this matrix equation yields exact nodal values for both function and derivative. This occurs because no approximation has been introduced in

the numerical development, and because the test functions form a T-complete (Herrera [11–13]) system. Detailed algorithms for constant coefficient, ordinary differential equations are presented by Herrera et al [14].

Exact nodal values are generally not obtainable when the coefficients are nonconstant. This is because the homogeneous adjoint equation exhibits nonconstant coefficients and cannot, in general, be solved exactly. Celia and Herrera [5] and Herrera [15] have presented high-order methods to approximate solutions to $\mathcal{L}^*w = 0$, thereby providing good estimates for the test functions. Celia and Herrera [5] used piecewise Lagrange polynomials to approximate the spatially variable coefficients. General power series solutions were obtained as approximations for the test functions (except in the case of piecewise constant approximations, when series solution is unnecessary). Herrera [15] applied collocation to solve $\mathcal{L}^*w = 0$ within each element.

To demonstrate the numerical procedure, consider the following example of transport with nonconstant coefficients,

$$(e^{-4x})\frac{d^2u}{dx^2} - \frac{du}{dx} = 0, \qquad 0 \le x \le 1 \qquad u(0) = 1 \qquad u(1) = 0 \qquad (11)$$

Let the variable coefficient be approximated with a piecewise constant function, so that constant values pertain within each element, with different values from element to element. Solution of $\mathcal{L}^*w = 0$ within any element then involves a constant coefficient ordinary differential equation, with solutions given by eqs. (10), with $D$ and $V$ interpreted as values within the element of interest.

Figure 1 illustrates numerical results for nodal values of both function and derivative. Solutions are plotted for both the approach outlined above, using

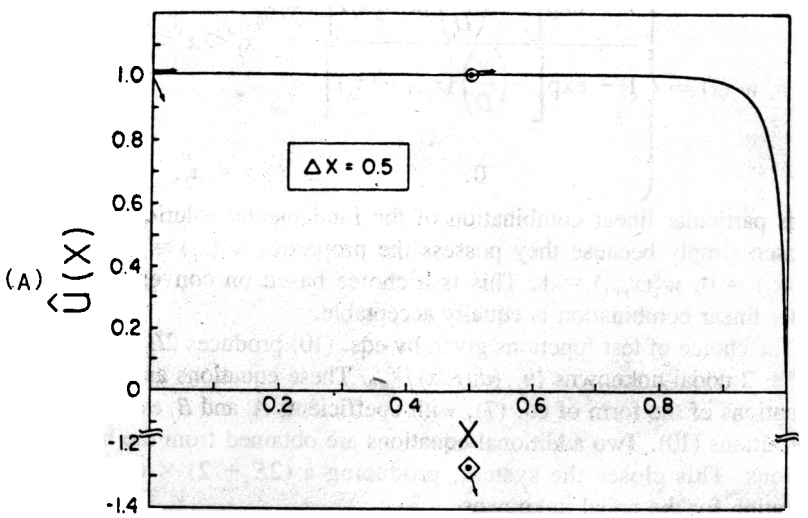


FIG. 1. Solutions to equation (11) using $C^{-1}$ OTF (⊙) and Hermite collocation (◈). Arrows indicate direction of the nodal gradients. Solid line corresponds to the exact solution.

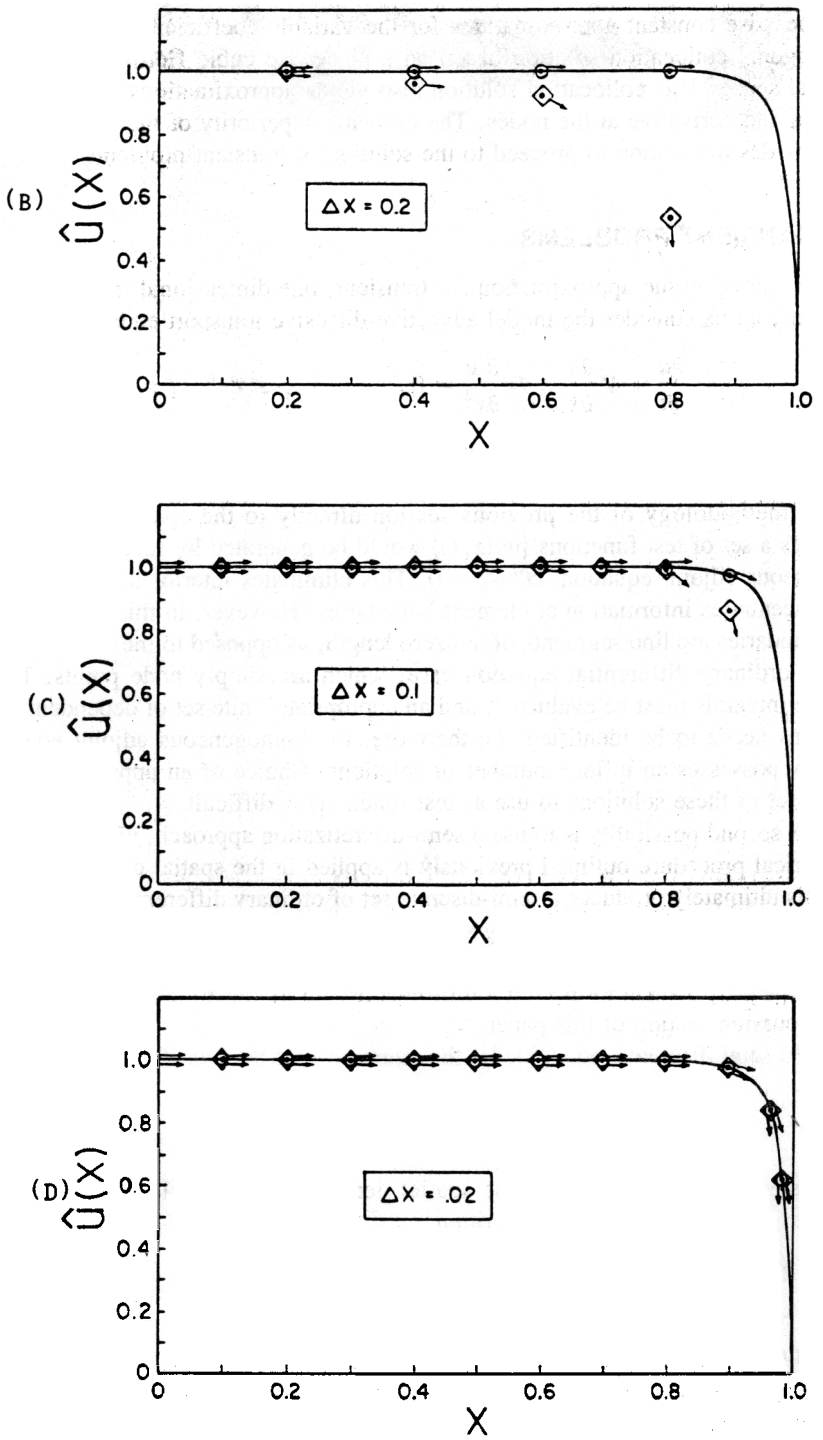


FIG. 1. (Continued)

piecewise constant approximations for the variable coefficient, and for an orthogonal collocation solution based on a piecewise cubic Hermite polynomial trial space. This collocation solution also yields approximations for both function and derivative at the nodes. The dramatic superiority of the new approach provides motivation to proceed to the solution of transient problems.

## TRANSIENT PROBLEMS

To develop the approximation for transient, one-dimensional transport equations, let us consider the model advective-diffusive transport equation,

$$\frac{\partial u}{\partial t} + V\frac{\partial u}{\partial x} - D\frac{\partial^2 u}{\partial x^2} = 0, \qquad 0 \le x \le \ell \qquad t > 0 \tag{12}$$

subject to suitable initial and boundary conditions. Several approaches to this multi-dimensional (space-time) problem can be proposed. The first is to extend the methodology of the previous section directly to the space-time domain. Thus a set of test functions $\{w_k(x,t)\}$ would be generated by solving the homogeneous adjoint equation, $\mathcal{L}^* w_k = 0$. This eliminates interior integration, and concentrates information at element boundaries. However, in this case element boundaries are line segments of nonzero length, as opposed to the boundaries in the ordinary differential equation case, which are simply node points. These line integrals must be evaluated, and an appropriate finite set of dependent variables needs to be identified. Furthermore, the homogeneous adjoint equation now possesses an infinite number of solutions. Choice of an appropriate finite subset of these solutions to use as test functions is difficult.

A second possibility is to use a semi-discretization approach, wherein the numerical procedure outlined previously is applied in the spatial dimension only. This ultimately produces a semi-discrete set of ordinary differential equations in time, which can be solved using standard time marching algorithms. This latter approach is taken herein, and an outline of its development follows. Several comments pertinent to the full multi-dimensional treatment are presented in the Discussion section of this paper.

To semi-discretize eq. (12), let the equation first be rewritten as

$$\mathcal{L}_x u \equiv D\frac{\partial^2 u}{\partial x^2} - V\frac{\partial u}{\partial x} = \frac{\partial u}{\partial t} - f(x,t) \tag{13}$$

where operator $\mathcal{L}_x$ refers to the spatial derivatives of eq. (12). Let the weak form analogous to eq. (3) be written,

$$\int_0^\ell \left(D\frac{\partial^2 u}{\partial x^2} - V\frac{\partial u}{\partial x}\right) w(x)\,dx = \int_0^\ell \left(\frac{\partial u}{\partial t} - f(x,t)\right) w(x)\,dx \tag{14}$$

Under the assumption that $u \in \mathbb{C}^1[0,\ell]$, $w \in \mathbb{C}^{-1}[0,\ell]$, the left side of eq. (14) can be written as a sum of elementwise integrations, exactly as eq. (4). Application of integration by parts twice to each element integration then leads to an equation analogous to eq. (5), viz.,

$$\int_0^\ell \left(\frac{D\,\partial^2 u}{\partial x^2} - V\frac{\partial u}{\partial x}\right) w(x)\,dx = \sum_{j=0}^{E-1} \left\{ \left[ Dw\frac{\partial u}{\partial x} - D\frac{dw}{dx}u - Vwu \right]_{x_j}^{x_{j+1}} + \int_{x_j}^{x_{j+1}} (\mathcal{L}_x^* w) u dx \right\} \tag{15}$$

Choice of $w(x)$ such that $\mathcal{L}_x^* w = 0$, and appropriate accounting of discontinuities in $w(x)$, makes eq. (15) equivalent to the following.

$$\int_0^\ell \left(D\frac{\partial^2 u}{\partial x^2} - V\frac{\partial u}{\partial x}\right) w(x)\,dx = \sum_{j=0}^{E-1} \left\{ \left[ \left[ D\frac{dw}{dx} + Vw \right] \right]_{x_j} u_j - [[Dw]]_{x_j} \frac{\partial u}{\partial x} j \right\} + \left[ \left( -D\frac{dw}{dx} - Vw \right) u + (Dw)\frac{\partial u}{\partial x} \right]_0^\ell \tag{16}$$

$$= \sum_{j=0}^{E} A_j u_j + B_j \frac{\partial u}{\partial x} j$$

Equation (16) uses the fact that $\mathcal{L}_x^* w = 0$ to eliminate all interior integrals. Test functions that conform to this condition are identical to those of eqs. (9) and (10), since the spatial operator is unchanged by the addition of transience. Equation (16) also reflects the fact that all variables in the equation are known except the nodal values $\{u_j, (\partial u/\partial x) j\}_{j=0}^E$. Coefficients $A_j, B_j$ are known scalars.

The key to numerical solution of transient problems is effective treatment of the temporal derivative that appears on the right side of eq. (13). Examination of eq. (13) indicates that a spatial integration of the product of $\partial u/\partial t(x,t)$ and $w(x)$ must be evaluated. To evaluate this integral, the temporal derivative $\partial u/\partial t$ is approximated using a polynomial expansion that involves the nodal values appearing in eq. (16). Given that both function and derivative appear as nodal values, the natural choice for polynomial interpolator is a piecewise cubic Hermite polynomial. Such an interpolator has the form

$$\frac{\partial u}{\partial t}(x,t) \simeq \frac{\partial \hat{u}}{\partial t}(x,t) \equiv \sum_{j=0}^{E} U_j(t)\phi_{0j}(x) + \frac{\partial U}{\partial x} j(t)\phi_{1j}(x) \tag{17}$$

In eq. (17), $\{U_j, (\partial U/\partial x) j\}_{j=0}^E$ are time-dependent nodal values of function and spatial derivative, while $\phi_{0j}, \phi_{1j}$ are standard piecewise cubic Hermite polynomials (see, for example, Carey and Oden [4], pp. 63–65). Substitution of expansion (17) into the integral of interest yields

$$\int_0^\ell \frac{\partial u}{\partial x}(x,t) w(x)\,dx \simeq \sum_{j=0}^{E} \frac{d}{dt}[U_j(t)] \int_0^\ell \phi_{0j}(x) w(x)\,dx + \frac{d}{dt}\left[\frac{\partial U}{\partial x} j(t)\right] \int_0^\ell \phi_{1j}(x) w(x)\,dx \tag{18a}$$

Given that $\phi_{0j}(x)$, $\phi_{1j}(x)$, and $w(x)$ are each well defined and known functions, the spatial integrals of eq. (18a) can be evaluated directly. Thus eq. (18a) represents a linear combination of known scalar coefficients and time derivatives of unknown, time-dependent coefficients, viz.,

$$\int_0^\ell \frac{\partial u}{\partial t}(x,t)w(x)\,dx \simeq -\sum_{j=0}^{E}\left\{(\alpha_j)\frac{d}{dt}[U_j(t)] + (\beta_j)\frac{d}{dt}\left[\frac{\partial U}{\partial x}j(t)\right]\right\} \quad (18b)$$

Equations (18), (16), and (14) can be combined to produce the semi-discretized equation of interest.

$$\sum_{j=0}^{E}\alpha_j\frac{d}{dt}(U_j) + \beta_j\frac{d}{dt}\left(\frac{\partial U}{\partial x}j\right) + A_jU_j + B_j\frac{\partial U}{\partial x}j = -\int_0^\ell f(x,t)w(x)\,dx \quad (19)$$

Equation (19) is written for each of the $2E$ choices of $w(x)$ given in eqs. (10). The two boundary conditions required for the second order equation (13) provide two additional equations, so that $2E + 2$ equations result for the $2E + 2$ nodal unknowns. The semi-discrete system is thus

$$\underset{\approx}{\mathbf{G}} \cdot \frac{d\underset{\sim}{\mathbf{U}}}{dt} + \underset{\approx}{\mathbf{H}} \cdot \underset{\sim}{\mathbf{U}} = \underset{\sim}{\mathbf{F}} \quad (20)$$

where matrix $\underset{\approx}{\mathbf{G}}$ contains the coefficients $\alpha_j, \beta_j$, $\underset{\approx}{\mathbf{H}}$ is composed of coefficients $A_j, B_j$, and $\underset{\sim}{\mathbf{F}}$ contains the forcing function terms and boundary condition information. Equation (20) can be solved by any standard time integrator. For example, a variably weighted Euler approximation produces,

$$\underset{\approx}{\mathbf{G}} \cdot \frac{\underset{\sim}{\mathbf{U}}^{n+1} - \underset{\sim}{\mathbf{U}}^{n}}{\Delta t} + \underset{\approx}{\mathbf{H}} \cdot [\theta\underset{\sim}{\mathbf{U}}^{n+1} + (1-\theta)\underset{\sim}{\mathbf{U}}^{n}] = \underset{\sim}{\mathbf{F}}^{n+\theta} \quad (21a)$$

or,

$$\left[\left(\frac{1}{\Delta t}\right)\underset{\approx}{\mathbf{G}} + \theta\underset{\approx}{\mathbf{H}}\right] \cdot \underset{\sim}{\mathbf{U}}^{n+1} = \left[\left(\frac{1}{\Delta t}\right)\underset{\approx}{\mathbf{G}} - (1-\theta)\underset{\approx}{\mathbf{H}}\right] \underset{\sim}{\mathbf{U}}^{n} + \underset{\sim}{\mathbf{F}}^{n+\theta}$$

where $\theta$ is a time weighting parameter, usually taken as $0 \le \theta \le 1$. Imposition of the initial condition, and subsequent marching through time, yields the discrete approximation of interest.

## CHOICE OF TEST FUNCTIONS

There are several important and distinguishing features of the numerical procedure and, in particular, the choice of test functions. The method itself is distinct in its nontraditional choice of test functions as well as the absence of a trial function. This latter fact clearly distinguishes this method from traditional finite element or Petrov-Galerkin techniques.

Figure 2 shows typical $\mathbb{C}^{-1}$ test functions defined by eqs. (10). Examination of the function definitions shows the following behavior. When $(V/D)$ is small, diffusion dominates and the test functions approach piecewise linear. As such, they become analogous to standard "hat" functions of finite elements; however, in the present case the functions are fully discontinuous and are thus not equivalent to hat functions. As advection becomes progressively dominant, the test functions become increasingly skewed in the upstream direction. A sort of upstream weighting is inherent in the functions. The degree of upstreaming is automatically set by the requirement that test functions satisfy the homogeneous

FIGURE 2A: $C^{-1}$ TEST FUNCTIONS

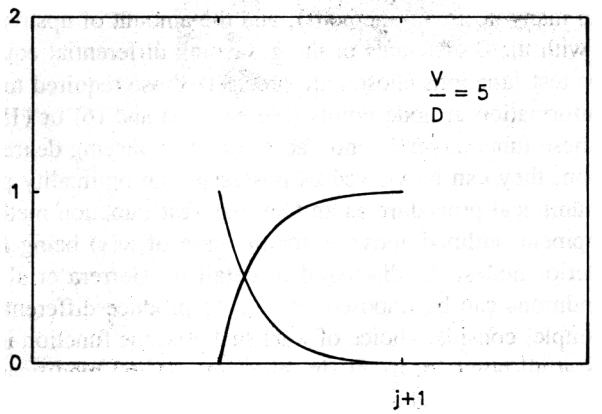


FIGURE 2B: $C^0$ TEST FUNCTIONS

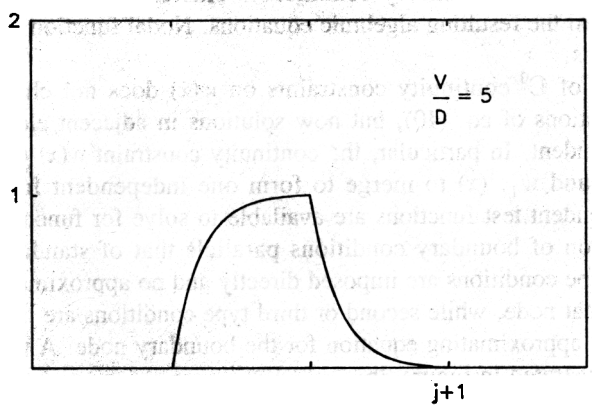


FIGURE 2C: DISCONTINUOUS TEST FUNCTIONS

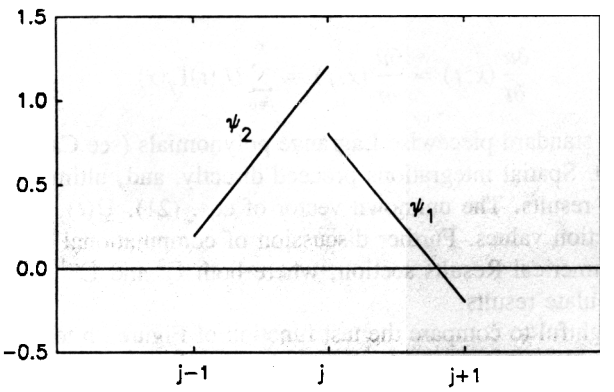


FIG. 2. Test functions used for (A) $C^{-1}$ Optimal Test Function method, (B) $C^0$ Optimal Test Function method, and (C) Petrov-Galerkin formulation of Hughes and coworkers.

adjoint equation. No arbitrary parameters appear in the test function definition (as opposed to many upstream methods), and the amount of upstreaming varies continuously with the coefficients of the governing differential equation.

Because the test functions chosen are precisely those required to analytically concentrate information at node points [see eqs. (5) and (6) or (15) and (16)], and because these functions inherently accommodate varying degrees of advection domination, they can be viewed as possessing an optimality property. We refer to this numerical procedure as an Optimal Test Function method.

The development outlined above is for the case of $w(x)$ being fully discontinuous at interior nodes. As discussed in detail by Herrera et al. [14], other continuity conditions can be imposed on $w(x)$ to produce different approximations. For example, consider choice of $w(x)$ such that the function is required to be continuous at all points in $[0, \ell]$. In this case, several modifications pertain to the development presented above, although the overall procedure is analogous. First, since $w(x) \in \mathbb{C}^0[0, \ell]$, $[[w]]_{x_j} = 0$ for all $j$. Therefore coefficients $B_j$ in eq. (7) are zero for $j = 1, 2, \ldots, E - 1$ ($B_0, B_E$ remain to accommodate second and third type boundary conditions). This eliminates nodal derivative unknowns from the resulting algebraic equations. Nodal function values are the only unknowns.

Imposition of $\mathbb{C}^0$ continuity constraints on $w(x)$ does not change the fundamental solutions of eq. (10), but now solutions in adjacent elements are no longer independent. In particular, the continuity constraint $w(x) \in \mathbb{C}^0[0, \ell]$ requires $w_2^e(x)$ and $w_1^{e+1}(x)$ to merge to form one independent function. Thus $E + 1$ independent test functions are available to solve for function nodal values. Imposition of boundary conditions parallels that of standard finite elements: first type conditions are imposed directly and no approximating equation is written at that node, while second or third type conditions are imposed in the context of the approximating equation for the boundary node. A typical $\mathbb{C}^0$ test function is illustrated in Figure 2b.

A final point regarding $\mathbb{C}^0$ test functions is treatment of temporal derivatives. Since nodal values of spatial derivatives no longer appear in eq. (16), the Hermite expansion of eq. (17) is replaced by a piecewise Lagrange polynomial expansion. This takes the form

$$\frac{\partial u}{\partial t}(x,t) \simeq \frac{\partial \hat{u}}{\partial t}(x,t) = \sum_{j=0}^{E} U_j(t)\Gamma_j(x) \tag{22}$$

where $\Gamma_j(x)$ are standard piecewise Lagrange polynomials (see Carey and Oden [4], pp. 38–41). Spatial integrations proceed directly, and, ultimately, a semidiscrete system results. The unknown vector of eqs. (21), $\underline{U}(t)$, now involves only nodal function values. Further discussion of computational details is provided in the Numerical Results section, where both $\mathbb{C}^0$ and $\mathbb{C}^{-1}$ test functions are used to calculate results.

It is also insightful to compare the test function of Figure 2b to the test function used by Hughes and co-workers (Brooks and Hughes [3]). The Hughes method is a Petrov-Galerkin formulation for which test functions are chosen to be perturbations of the standard Galerkin functions, viz.,

$$\tilde{w}(x) = w(x) + \pi V \frac{dw}{dx} \tag{23}$$

where $w(x)$ is the standard Galerkin test function (for example, piecewise linear), $V$ is velocity, and $\pi$ is a free parameter. Hughes gives a criterion for "optimal" choice of $\pi$, based on results of Christie et al. [6]. This latter reference derived a procedure that yields exact nodal values for the one-dimensional, steady-state, constant coefficient, nonreactive, homogeneous advective-diffusive transport equation. (A similar result was originally presented by Allen and Southwell [1]). A typical test function, $\tilde{w}(x)$ of eq. (23), is illustrated in Figure 2c. Based on the definition of optimality discussed in the Introduction, these test functions must be considered optimal for the one-dimensional, steady-state, constant coefficient, nonreactive, homogeneous transport equation.

However, the functions of eq. (23) do not yield exact nodal values when the equation is nonhomogeneous, when any other terms are present in the governing equation (for example, a first-order reaction term), or when the coefficients of the equation are spatially variable. This is to be compared to our Optimal Test Function method, which yields exact nodal values when the equation is nonhomogeneous and when reaction terms are present. We have the option to obtain only nodal function values or both function and derivative values at the nodes. In addition, for the case of variable coefficients, procedures have been developed (Celia and Herrera [5], Herrera [15]) to obtain nodal values to any desired order of accuracy. As such, OTF maintains optimality properties over a broad class of equations. The reason is that the test functions are based on the adjoint equation, which reflects all terms that are present in the governing differential operator. The test functions thus incorporate all relevant information for any given operator, not just homogeneous advection-diffusion.

## NUMERICAL RESULTS

Two test problem results are presented to demonstrate the performance of the proposed method. The $\mathbb{C}^{-1}$ formulation using Hermite polynomials and the $\mathbb{C}^0$ formulation using quadratic Lagrange polynomials are implemented. Results are compared to other weighted residual methods. The governing equation that is solved is eq. (12).

In computing the integrals associated with the temporal derivative [see eqs. (18) and (22)], use is made of the fact that each test function $w(x)$ is nonzero over only one (in the $\mathbb{C}^{-1}$ case) or two (the $\mathbb{C}^0$ case) elements. As such, the only integrals calculated are those for which both $w(x)$ and the interpolating functions ($\phi_{0j}, \phi_{1j}$ for the Hermite case, $\Gamma_j$ for the Lagrange case) have intersecting nonzero regions. When calculating solutions for the $\mathbb{C}^0$ quadratic case, a unique quadratic Lagrange polynomial is defined based on the three nodes that define the span of the given test function $w(x)$. This means that the quadratic polynomial changes with different test functions. The procedure essentially uses a "sliding template" to define the relevant quadratic polynomial. This differs from standard quadratic finite element interpolation, which defines

fixed quadratic (three node) elements, with interpolation independent of test function.

## Propogation of an initial step discontinuity

This example problem uses the following initial and boundary conditions: $u(x,0) = 0$, $0 \leq x \leq \ell$, $u(0,t) = 1$, $t > 0$, $u(\ell,t) = 0$, $t > 0$, and the right boundary is maintained sufficiently far downstream to not influence the solution. Parameters of interest are chosen as $V = 1$, $D = 0.0002$, $\Delta x = 0.02$, $\Delta t = 0.01$, leading to Courant and element Peclet numbers of $Cu = 0.5$ and $Pe = V\Delta x/D = 100$, respectively. For time integration, a Crank Nicolson method [$\theta = 0.5$ in eq. (21)] is chosen. Results for this problem are shown in Figure 3*a–e*. In each plot, the analytical solution is given as solid line and the numerical solution by dashed line. Figure 3a and c show simulations using the

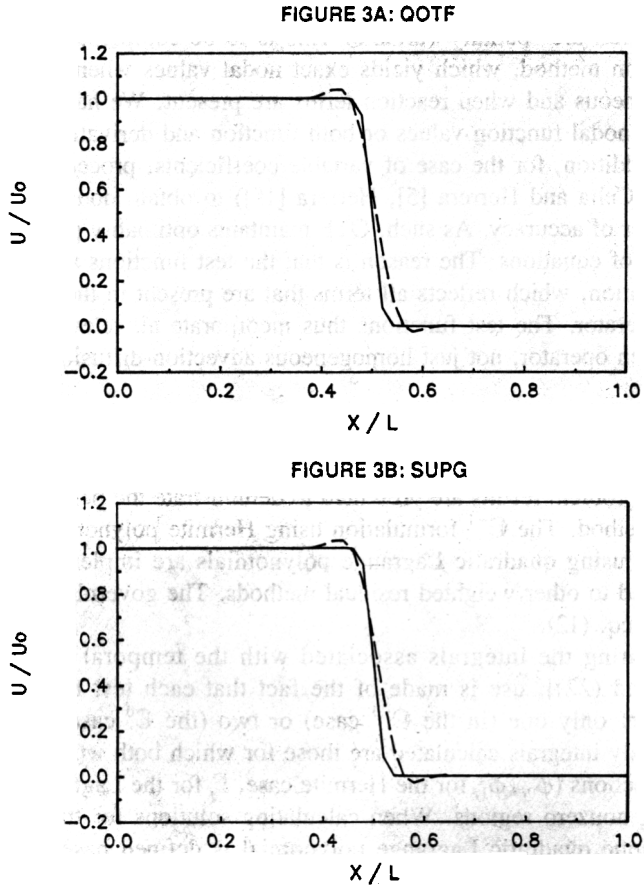


FIG. 3. Numerical results for propagation of an initial step discontinuity using (A) Quadratic ($C^0$) Optimal Test Function method (QOTF), (B) Streamline Upwind Petrov Galerkin (SUPG), (C) Hermite ($C^{-1}$) Optimal Test Function Method (HOTF), (D) Hermite collocation (HCOL), and (E) Linear Finite Elements (LFEM).

FIGURE 3C: HOTF

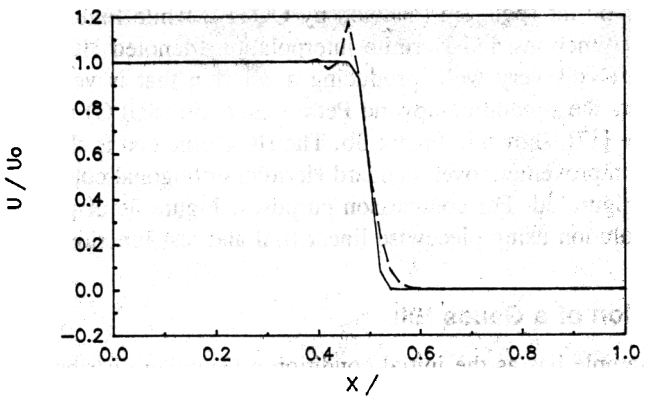


FIGURE 3D: HCOL

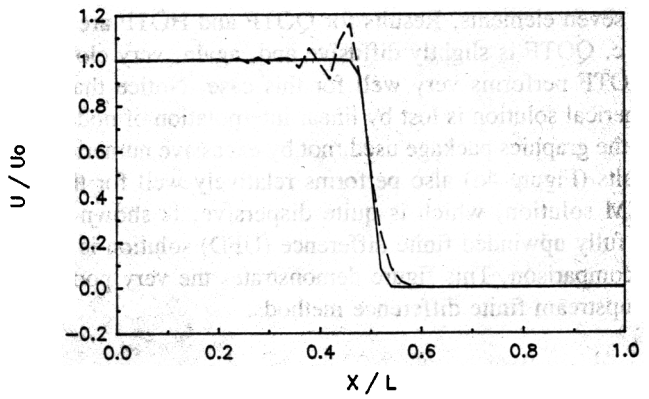


FIGURE 3E: LFEM

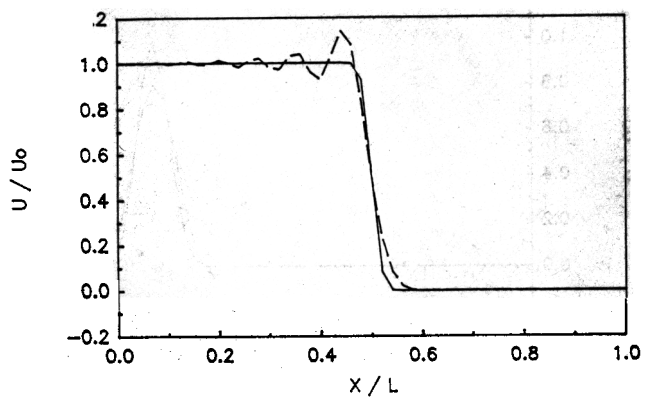


FIG. 3. (Continued)

OTF procedure developed herein. Figure 3a corresponds to $\mathbb{C}^0$ test functions with quadratic interpolation (denoted by QOTF), while Figure 3c corresponds to $\mathbb{C}^{-1}$ test functions and Hermite interpolation (denoted HOTF). The QOTF method performs very well, producing a solution that is very similar to that produced by the streamline-upwind Petrov-Galerkin (SUPG) method of Brooks and Hughes [17], shown in Figure 3b. The HOTF method is slightly oscillatory, but shows improvement over standard Hermite orthogonal collocation (HCOL), shown in Figure 3d. For comparison purposes, Figure 3e displays a traditional Galerkin solution using piecewise linear trial and test functions (LFEM).

## Propagation of a Gauss Hill

This example has as the initial condition a Gaussian distribution centered at $x_0$ with standard deviation $\sigma$, viz., $u(x, 0) = \exp\{-(x - x_0)^2/\sigma^2\}$. Homogeneous Dirichlet boundary conditions are used. Relevant parameters are $x_0 = 0.25$, $\sigma = 0.0333$, $V = 1.0$, $D = 0.00002$, $\Delta x = 0.02$, $\Delta t = 0.01$, so that $Pe = 1000$, $Cu = 0.5$. The chosen values produce a narrow initial distribution that spans only seven elements. Results for QOTF and HOTF are displayed in Figure 4a and c. QOTF is slightly diffusive and, again, very close to SUPG (Figure 4b). HOTF performs very well for this case. Notice that the peak in the HOTF numerical solution is lost by linear interpolation of nodal function values inherent in the graphics package used, not by excessive numerical diffusion. The HCOL results (Figure 4d) also performs relatively well for this case. A traditional LFEM solution, which is quite dispersive, is shown in Figure 4e. In addition, a fully upwinded finite difference (UFD) solution is displayed in Figure 4f for comparison. This figure demonstrates the very poor performance of traditional upstream finite difference methods.

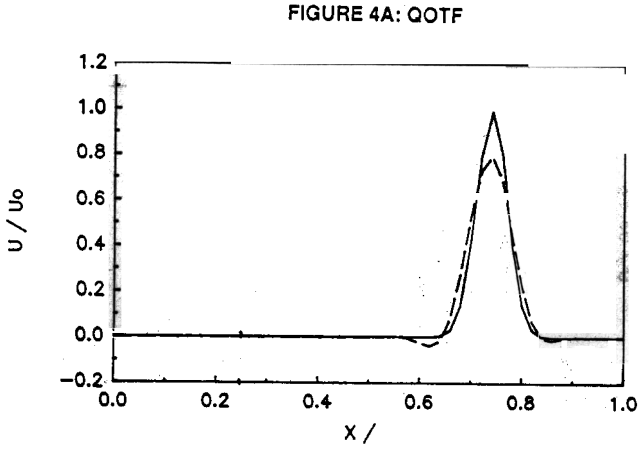


FIG. 4. Numerical results for propagation of a Gauss hill. Plots (A) to (E) correspond to methods of Figure 3; plot (F) is upstream finite difference method.

FIGURE 4B: SUPG

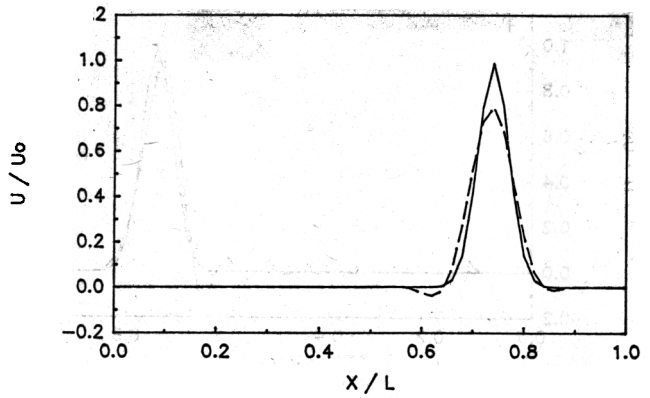


FIGURE 4C: HOTF

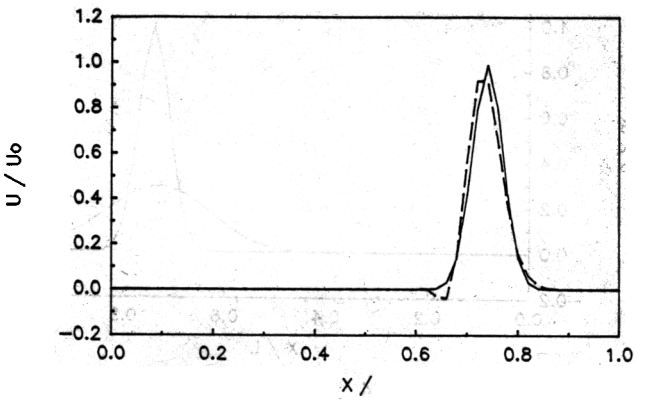


FIGURE 4D: HCOL

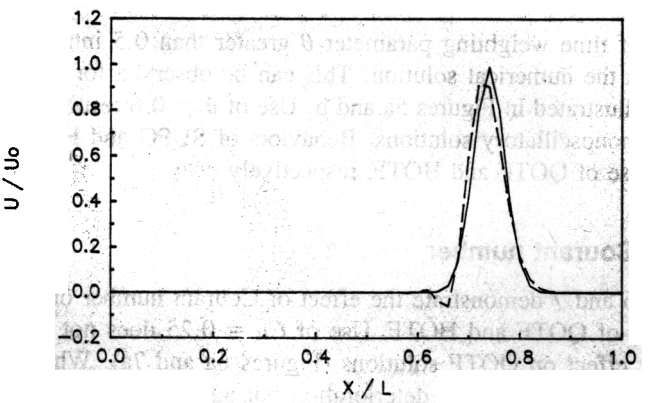


FIG. 4. (Continued)

FIGURE 4E: LFEM

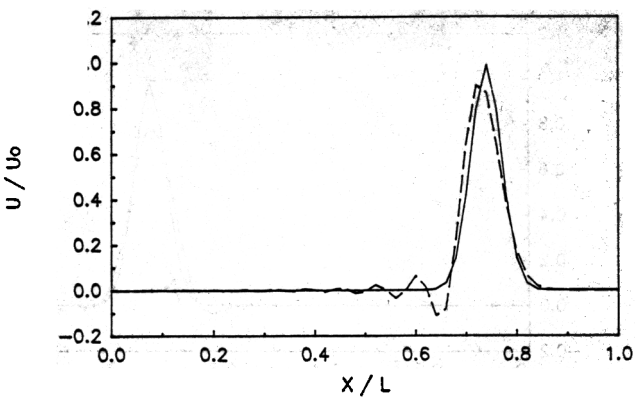


FIGURE 4F: UFD

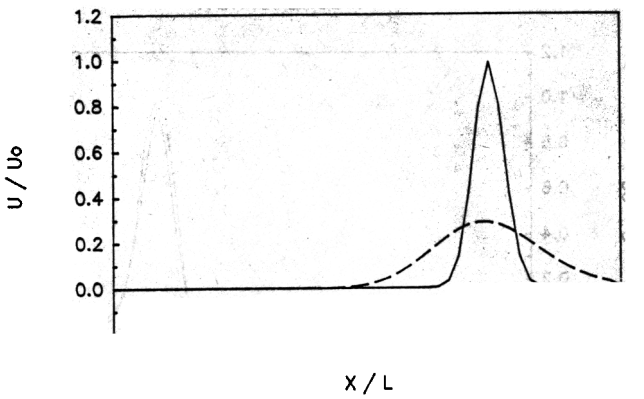


FIG. 4. (Continued)

## Effect of time weighting parameter $\theta$

Choice of time weighting parameter $\theta$ greater than 0.5 introduces artificial diffusion in the numerical solution. This can be observed for both QOTF and HOTF, as illustrated in Figures 5a and b. Use of $\theta = 0.6$ results in slightly diffusive but nonoscillatory solutions. Behaviors of SUPG and HCOL are analogous to those of QOTF and HOTF, respectively.

## Effect of Courant number

Figures 6 and 7 demonstrate the effect of Courant number on the predictive capabilities of QOTF and HOTF. Use of $Cu = 0.25$ does not appear to have significant effect on QOTF solutions (Figures 6a and 7a). When $Cu = 0.83$, QOTF solutions show some deterioration but again changes are not dramatic (Figures 6c and 7c). Sensitivity of SUPG to Courant number is analogous to

FIGURE 5A: QOTF θ = 0.6

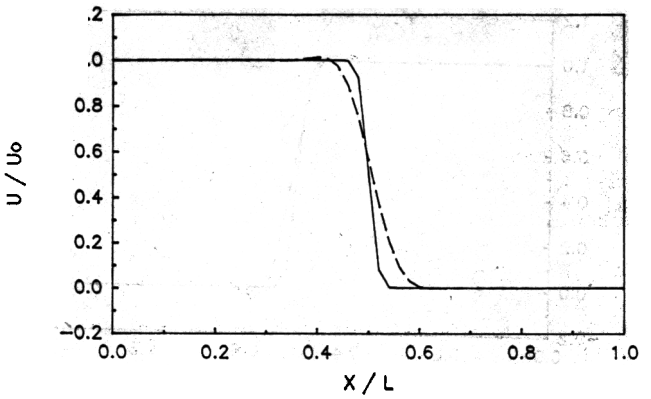


FIGURE 5B: HOTF θ = 0.6

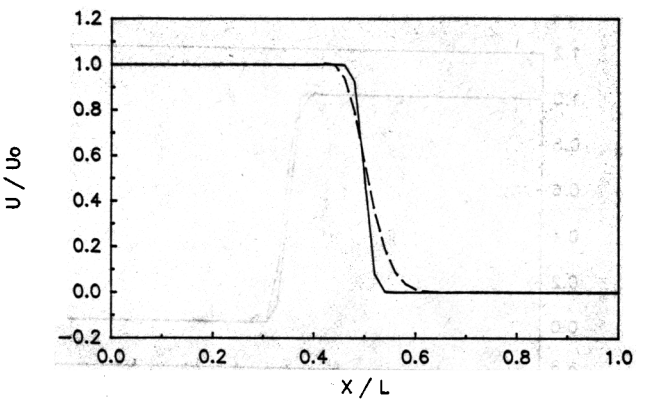


FIG. 5. Effect of time weighting parameter ($\theta$) on OTF solutions. Results are for $\theta = 0.6$.

that of QOTF. HOTF solutions are much more sensitive to Courant number. Figures 6b and 7b show dramatic improvements in HOTF solutions when $Cu = 0.25$ is used. The Gaussian hill is practically exact, while the steep front simulation is excellent (a shift of approximately one-half grid step occurs in this latter case due to Hermite interpolation of the step initial condition). On the other hand, use of $Cu = 0.83$ results in significant deterioration in HOTF solutions (Figures 6d and 7d). HCOL behaves in a similar way.

## DISCUSSION AND FUTURE DIRECTIONS

Results presented herein for the case of one-dimensional, transient, nonreactive advective-diffusive transport show that Optimal Test Function solutions are at least as good as those from the best available interior methods. These results represent our initial efforts to apply OTF to transient problems. While our re-

FIGURE 6A: QOTF, Cu = 0.25

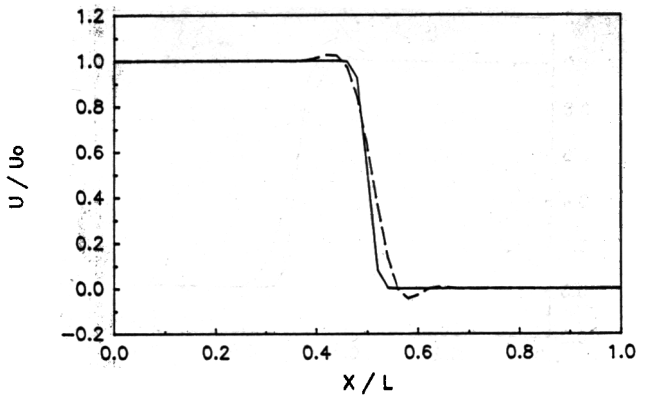


FIGURE 6B: HOTF, Cu = 0.25

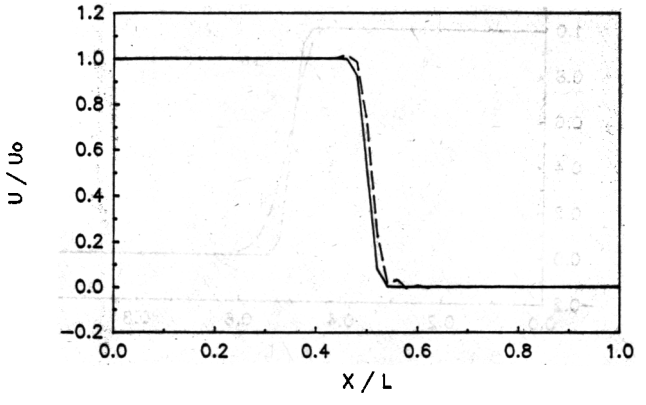


FIGURE 6C: QOTF, Cu = 0.83

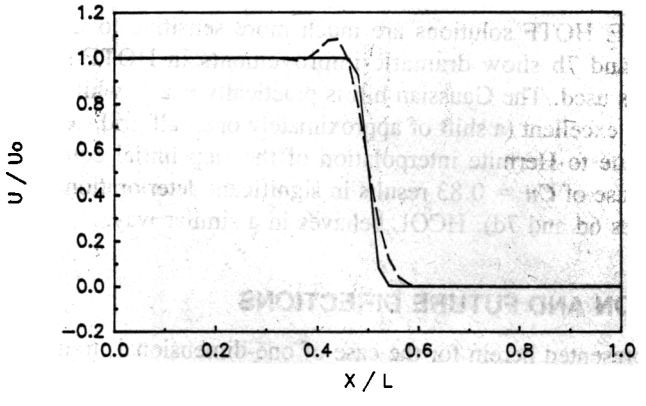


FIG. 6. Effect of Courant number on OTF solutions for the propagation of an initial step discontinuity.

FIGURE 6D: HOTF, Cu = 0.83

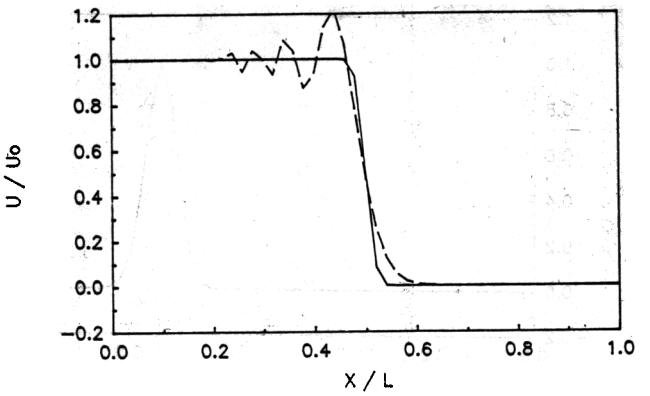


FIG. 6. (Continued)

sults are similar to those of, for example, optimal Petrov-Galerkin formulations (equivalent to Hughes' method), we believe that our method offers advantages in its generality and its systematic and mathematically rigorous underpinnings. This provides a firm foundation on which extensions to higher dimensions can be built, and from which improvements can be formulated.

There are currently two directions that we are pursuing in seeking advancements and improvements in the methodology. The first is improved treatment of temporal discretization, and the second is extension to multiple spatial dimensions. Improvements in temporal treatment is focusing on choosing space-time test functions that satisfy the homogeneous space-time adjoint operator while effectively incorporating the first order hyperbolic (characteristic) behavior of

FIGURE 7A: QOTF, Cu = 0.25

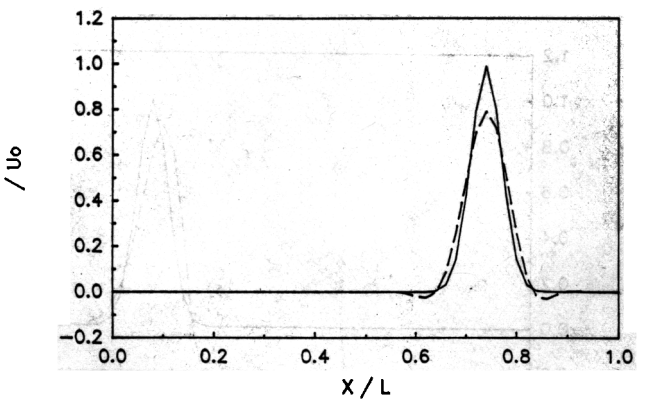


FIG. 7. Effect of Courant number on OTF solutions for the propagation of a Gauss hill.

FIGURE 7B: HOTF, Cu = 0.25

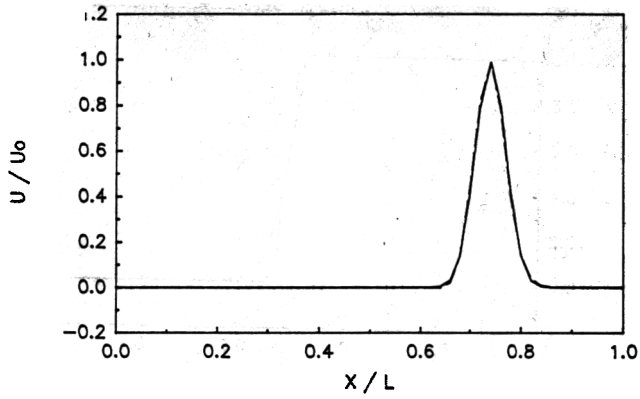


FIGURE 7C: QOTF, Cu = 0.83

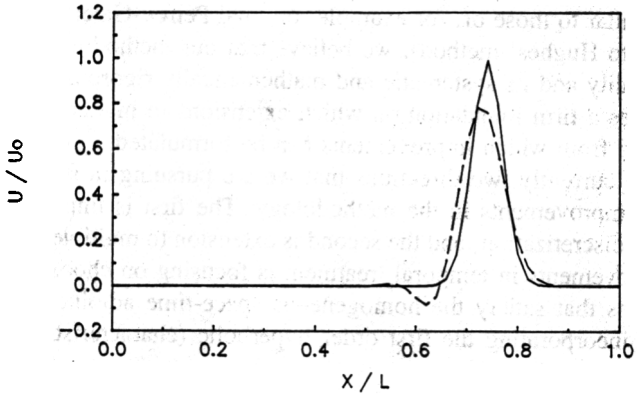


FIGURE 7D: HOTF, Cu = 0.83

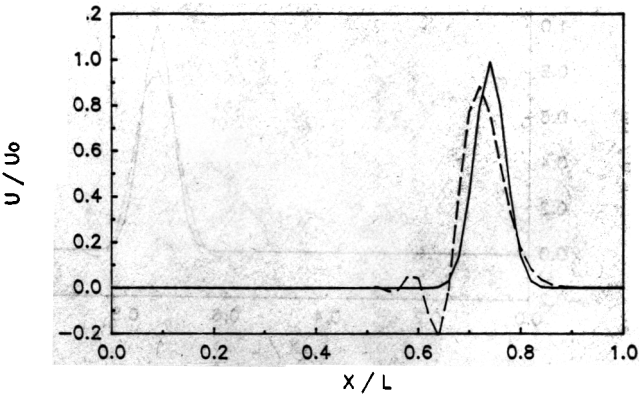


FIG. 7. (Continued)

the governing equation. This approach is not based on a semi-discretization and differs from the approach presented herein in fundamental ways, although the overall philosophy is the same.

Several methodologies for multiple spatial dimensions are currently under investigation. One such method uses tensor products of one-dimensional test functions [eqs. (10)] to form appropriate two-dimensional test functions. For two dimensions, elementwise integration, and subsequent integration by parts, leads to line integrals along element boundaries. If the Laplace equation is considered as an example, and $\mathbb{C}^0$ test functions are used, it is a straightforward calculation to show that different nodal interpolation schemes along element boundaries lead to different approximations. Piecewise constant approximations, in the context of element boundary integrations, produces a standard five-point finite difference approximation; piecewise linear interpolation leads to a standard nine-point (piecewise bilinear) finite element approximation; piecewise quadratic interpolation leads to a nine-point scheme of significantly enhanced accuracy. In fact, this latter nine point scheme is sixth order accurate. Furthermore, Collatz [7] has shown that higher order nine point schemes are impossible. Therefore, this OTF approximation leads to optimal accuracy in terms of a best possible approximation. When advection is added to the governing two-dimensional equation, the procedures remain invariant with the only changes being in definition of test functions. We are currently investigating the case of two-dimensional, advective-diffusive transport.

Overall, we feel that the methodology presented herein provides a new and potentially powerful approach for developing numerical approximations to partial differential equations in general and to transport equations in particular.

## SUMMARY

The presentation introduces a new numerical method for solution of the one-dimensional, transient advection diffusion transport equation. The procedure is based on a weak form of the governing equation. Definition of test functions is provided as part of the numerical algorithm. These test functions vary continuously as the parameters of the governing equation change, thereby automatically accommodating varying degrees of advection domination. The test functions possess optimality properties, and we thus call the method an Optimal Test Function (OTF) method. Initial numerical results show excellent behavior of the method over a wide range of grid Peclet numbers.

This work was supported in part by the Sloan-Cabot research funds of MIT under grant 26950 as well as Project Athena of MIT under grant 24728. The work was also supported by CONOCYT of Mexico.

## References

[1] D. Allen and R. Southwell, "Relaxation methods applied to determining the motion, in two dimensions, of a viscous fluid past a fixed cylinder," *Quart. J. Mech. Appl. Math.*, **8**, 129–145 (1955).

[2] M. B. Allen and G. F. Pinder, "Collocation simulation of multiphase porous me-

dium flow," *Soc. Petr. Eng. J.*, **23**(1), 135–142 (1983).

[3] A. Brooks and T. J. R. Hughes, "Streamline upwind Petrov-Galerkin formulations of convection dominated flows with particular emphasis on the incompressible Navier-Stokes equations," *Comp. Meth. Appl. Mech. Engrg.*, **32**, 199–259 (1982).

[4] G. F. Carey and J. T. Oden, *Finite Elements: A Second Course*, Volume II, The Texas Finite Element Series, Prentice-Hall, 1983.

[5] M. A. Celia and I. Herrera, "Solution of general ordinary differential equations by a unified theory approach," *Numerical Methods for Partial Differential Equations*, **3**, 117–129 (1987).

[6] I. Christie, D. F. Griffiths, A. R. Mitchell, and O. C. Zienkiewicz, "Finite element methods for second order partial differential equations with significant first order derivatives," *Int. J. Num. Meth. Engrg.*, **10**, 1389–1396 (1976).

[7] L. Collatz, *Numerical Treatment of Differential Equations*, Third Edition, Springer, 1960.

[8] C. A. J. Fletcher, *Computational Galerkin Methods*, Springer Series in Computational Physics, Springer-Verlag, 1984.

[9] D. F. Griffiths and A. R. Mitchell, "On generating upwind finite element methods," in *Finite Element Methods for Convection Dominated Flows*, T. J. R. Hughes, ed., AMD Volume **34**, 91–104 (1979).

[10] J. C. Heinrich, P. S. Huyakorn, O. C. Zienkiewicz, and A. R. Mitchell, "An 'upwind' finite element scheme for two-dimensional convective-transport equation," *Int. J. Num. Meth. Engrg.*, **11**, 131–143 (1977).

[11] I. Herrera, *Boundary Methods: An Algebraic Theory*, Pitman Publishing Co., London, 1984.

[12] I. Herrera, "Unified approach to numerical methods, Part I: Green's formula for operators in discontinuous fields," *Numerical Methods for Partial Differential Equations*, **1**, 25–44, 1985.

[13] I. Herrera, "Unified approach to numerical methods, Part II: Finite elements, boundary elements, and their coupling," *Numerical Methods for Partial Differential Equations*, **1**, 159–186, 1985.

[14] I. Herrera, L. Chargoy, and G. Alducin, "Unified approach to numerical methods, Part III: Finite differences and ordinary differential equations," *Numerical Methods for Partial Differential Equations*, **1**, 241–258, 1985.

[15] I. Herrera, "The algebraic theory approach for ordinary differential equations: highly accurate finite differences," *Numerical Methods for Partial Differential Equations*, **3**, 199–218, 1987.

[16] T. J. R. Hughes, "A simple scheme for developing 'upwind' finite elements," *Int. J. Num. Meth. Engrg.*, **15**, 1359–1365, 1978.

[17] T. J. R. Hughes and A. Brooks, "A theoretical framework for Petrov-Galerkin methods with discontinuous weighting functions: Applications to the streamline-upwind procedure," in *Finite Elements in Fluids: Vol. 4*, Gallagher et al., eds., 47–65, 1982.

[18] T. J. R. Hughes and T. E. Tezduyar, "Finite element methods for first-order hyperbolic systems with particular emphasis on the compressible Euler equations," *Comp. Meth. Appl. Mech. Engrg.*, **45**, 217–284, 1984.

[19] P. S. Huyakorn, "Upwind finite element scheme for improved solution of the convection-diffusion equation," Water Resources Program Report 76-WR-7, Princeton University, 1976.

[20] L. Lapidus and G. F. Pinder, *Numerical Solution of Partial Differential Equations in Science and Engineering*, John Wiley, 1982.

[21] B. P. Leonard, "A stable and accurate convective modelling procedure based on quadratic upstream interpolation," *Comp. Meth. Appl. Mech. Eng.*, **19**, 59–98, 1979.

[22] G. F. Pinder and A. Shapiro, "A new collocation method for the solution of the convection-dominated transport equation," *Water Resources Research*, **15**, 1177–1182, 1979.