Linear and quadratic least squares finite element method for the solution of hyperbolic problems

M. H. Afshar\textsuperscript{a} and K. Morgan\textsuperscript{b}  
\textsuperscript{a}Civil Eng. Dept., Imam Higher Education Center, Tehran, Iran  
\textsuperscript{b}Civil Eng. Dept. University College of Swansea, Swansea, UK

Abstract

A least-squares finite element method has been developed for the solution of nonlinear hyperbolic problems. The formulation of the problem is achieved via a backward differencing in time leading to the semi-discretized form of the governing differential equations. Temporal discretization of the governing equation produces a residual which is then minimized in a least-square sense resulting in the variational form of the problem. The discretization of the variational statement by finite elements leads to a set of algebraic equations to be solved for nodal unknowns. The possibilities of using linear and quadratic elements is investigated and examined numerically. Application of the method to solve Burgers equation, nozzle and dam-break problems are also presented.

Introduction

Finite element methods are now well established as a basic technique for solving boundary-value problems. For self-adjoint operators, the equivalence of the Galerkin method with a variational principle \cite{1,2,3} implies that the method has the best approximation property according to a prescribed norm. Many problems in solid mechanics can be described by such operators, which justifies the extensive use of the Galerkin approach in this field.

Many of the interesting problems encountered in fluid dynamics, however, are governed by partial differential equations which exhibit hyperbolic features. It has been long observed that, in contrast to elliptic and parabolic cases, standard application of the Galerkin finite element method to hyperbolic problems, if the exact solution has a jump discontinuity, produces results that, in general, exhibit large spurious oscillations, even far from the jump, and will then not be close to the exact solution anywhere. Many interesting cases of fluid flow are characterized by the hyperbolic features of the governing partial differential equation. This exhibits itself through the formation of shock waves in compressible flows and the formation of hydraulic jump and bores \cite{4} in incompressible flows. The success of upwinding in finite difference methods has encouraged the introduction of this concept into finite element methods. This has led to the development of various Petrov-Galerkin \cite{4,5,6} finite element methods in which the weighting function is different from the trial function and is constructed in such a way that the upwinding effect is introduced.
method of streamline, or balancing, diffusion and Streamline-Upwind Petrov-Galerkin (SUPG) \[7\] and the commonly used Taylor-Galerkin method are typical examples of these methods.

In this paper we present a least squares finite element method for the solution of hyperbolic problems. For clarity we first describe the application of the method to scalar model problem and then develop the formulation to consider the system of differential equations. It will be shown that the method naturally provides the necessary upwinding mechanism and requires no 'free' parameter. The method is unconditionally stable and hence most suitable for steady state computations. The method has a first order accuracy in time but, it can also be used for the solution of transient problems and this is demonstrated by the dam-break problem.

Scalar Equation

Hyperbolic partial differential equations (HPDE) are an important class of partial differential equations. Many physical laws can be expressed by this class of equations. Conservation laws are a particular and important example in the fluid dynamics field. In this section, we will outline the proposed finite element method for the prototype one dimensional scalar first-order hyperbolic equation in its conservative form

\[ u_t + f_x = 0 \]  

or the equivalent Quasi-linear form

\[ u_t + au_x = 0 \]  

subject to appropriate initial and boundary conditions. In general \( a \) is considered to be a function of \( x,u \) and \( t \). Equation (2) is first discretized in time by backward differencing to give

\[ u^{n+1} - u^n + \Delta t a^{n+1} u_x^{n+1} = 0 \]  

A residual can then be defined as

\[ r^{n+1}_\Omega = u^{n+1} - u^n + \Delta t a^n u_x^{n+1} \]  

where a linearization has been carried out by replacing \( a^{n+1} \) with \( a^n \). Now we can construct a least squares minimization problem. We seek to find \( u^{n+1} \) which minimizes the least squares functional

\[ I(u^{n+1}) = \int_\Omega (r^{n+1}_\Omega)^2 \, d\Omega \]  

with respect to variations in \( u^{n+1} \). Setting \( \delta I = 0 \), implies that

\[ \int_\Omega (\delta u^{n+1} + \Delta t a^n \delta u_x^{n+1}) (u^{n+1} - u^n + \Delta t a^n u_x^{n+1}) \, d\Omega = 0 \]

On setting \( \delta u^{n+1} = W \), we obtain the following variational (weak) problem: Find \( u^{n+1} \) satisfying the prescribed boundary condition such that
for all admissible test function \( W \). Note that \( W = 0 \) on that part of the boundary where the Dirichlet boundary condition is prescribed, since \( \partial u^{n+1} = 0 \) there. Discretizing the domain and introducing the finite element shape functions \( N_i \), associated with node \( i \), the solution to equation (2) is approximated in the form

\[
\bar{u}^{n+1} = u_i^{n+1} N_i
\]  

where \( N_i \) is a \( C^0 \) - finite element shape function. Substituting equation (7) into equation (8) and using the interpolating function \( N_i \) as the test function \( W \), equation (7) reduces to a linear system of algebraic equations of the form

\[
K U = F
\]  

where

\[
K_{ij} = \int_\Omega (N_i + \Delta t a^n N_{i,x}) (N_j + \Delta t a^n N_{j,x}) d\Omega
\]

\[
F_i = \int_\Omega (N_i + \Delta t a^n N_{i,x}) U^n d\Omega
\]

and \( U \) is the vector of unknowns \([u_1^{n+1}, u_2^{n+1}, \ldots, u_i^{n+1}]^T\). A useful form of the equation (9) can be written which is particularly useful for conservation problems. This form would ensure the conservation property in a steady state sense. In this form we have

\[
K \Delta U = F
\]  

where \( K \) is defined as before and

\[
F_i = \int_\Omega (N_i + \Delta t a^n N_{i,x}) f_x^n d\Omega
\]

and \( \Delta U = U^{n+1} - U^n \). Flux term in equation (13) can be approximated using the same functions which are used to interpolate the unknown. To explore the properties of the algorithm, one can first integrate Eq. (7) in part and then expand proper terms by a Taylor Series to find the corresponding Euler equation in the form of

\[
u_t + f_x = \frac{\Delta t}{2} a^2 u_{xx} - \frac{\Delta t^2}{2} a^3 u_{xxx}
\]  

where \( a \) is assumed to be a constant. The first term on the right hand side is clearly a diffusion term. It is this term which stabilizes the method when problems with discontinuous solutions are considered. The amount of the diffusion introduced is dependent upon the size of the time step such that increasing the time step results in a more diffusive method. This is of particular importance in the solution of flows with different shock intensities since more diffusion is required to stabilize high intensity shocks. The second term is a
third order term representing the effect of dispersion. It should be noted that the method has a first order accuracy in time, and therefore more suitable as a time marching scheme for solving the steady state problems. The proposed algorithm, however, can be used as a transient scheme provided the time step size is kept small enough. At steady state conditions, that is \( u^{n+1} = u^n \), this time marching least squares method is equivalent to a weighted residual method defined by

\[
\int_{\Omega} (N_i + \Delta t a N_{i,x}) f_x^n d\Omega = 0
\]

which is clearly a Petrov-Galerkin method for the solution of the steady state form of the original problem.

**System Of Equations**

In this section we will extend the application of the proposed algorithm for solving system of hyperbolic partial differential equations. Consider the conservative form of a system of HPDE in 1-D

\[
U_t + F_x = 0
\]

subject to appropriate boundary and initial conditions. Here \( U \) is the vector of unknowns and \( F \) is the vector of fluxes, which generally depend on \( U \). The quasi-linear form of equation (16) reads

\[
U_t + A U_x = 0
\]

where \( A = \frac{\partial F}{\partial U} \) is the jacobian matrix. Using backward differencing to discretize equation (17) in time followed by a Least-squares minimization of the linearized residual and finite element discretization as above leads to

\[
K \Delta U = F
\]

where

\[
K_y = \int_{\Omega} (N_i + \Delta t A^n N_{i,x}) (N_j + \Delta t A^n N_{j,x}) d\Omega
\]

\[
F_i = \int_{\Omega} (N_i + \Delta t A^n N_{i,x}) f_x^n d\Omega
\]

The calculation of the jacobian matrix \( A \) in equation (19) needs attention. The natural choice is to calculate the matrices at the Gauss points. It is possible, however, to assume that the matrices are constant over the elements. This assumption eliminates the need for numerical integration, which is of significant importance when higher order elements are used. In addition, it provides a means of improving the stability of the method in the presence of the shocks. This matter will be addressed later when numerical tests are carried out. The constant Jacobians can be constructed in different ways.
a) Calculate the jacobian as a function of the independent variables at the center of the element

\[ A^e = A(U^e) \]  
\[ U^e = \frac{1}{n} \sum_{i} U_i \]  

b) Calculate the element jacobian by averaging the nodal values of the jacobian over the element, where \( n \) is the number of nodes per element

\[ A^e = \frac{1}{n} \sum_{i} A_i \]  
\[ A_i = A(U_i) \]  

The merits of these methods will be demonstrated when numerical tests are carried out.

**Numerical Examples**

In this section we present some examples to test the behavior of the method. We will first consider the solution to the well known Burgers equation. A transient example, namely
Formation and Propagation of a discontinuity: The first example to consider is the nonlinear Burgers equation

\[ u_t + uu_x = 0 \quad \text{on } 0 \leq x \leq 0 \]

subject to the initial condition \( u = 2 - 0.5x \) for \( 0 \leq x \leq 0.2 \) and \( u = 1 \) for \( 0.2 \leq x \leq 1 \) and boundary condition \( u = 2 \) at \( x=0 \). The formation and propagation of a shock is depicted in Fig. 1 for different time step sizes. It is seen that the shock is generally captured in fewer elements with smaller time step as suggested before. The accuracy of the transient solution is dominated by the time step size. Nevertheless, quadratic elements produce results which are more accurate than those of linear one.

Isothermal Nozzle Flow: Isothermal flow in a convergent-divergent nozzle is the steady state example to be considered. The governing equations can be written in a divergence form as
Advances in Fluid Mechanics II 99

Fig. 3. Isothermal nozzle flow. a) Linear and b) Quadratic solution with $c=10$. Jacobians are calculated according to Eq. 22.

$$
\begin{bmatrix}
\rho u \\
\rho au
\end{bmatrix}
+
\begin{bmatrix}
\rho u \\
\rho au^2 + \rho ac
\end{bmatrix}
=
\begin{bmatrix}
0 \\
\rho x^2 \frac{da}{dx}
\end{bmatrix}
$$

Here $\rho$ is the density, $u$ is the velocity, $a$ is the cross sectional area of the nozzle and $c$ is the velocity of sound ($c = 1$). The cross sectional area of the nozzle is defined to vary according to $a = 1 + (x-2.5)^2 / 12.5$ for $0 \leq x \leq 1$.

Three possible flow regimes can be recognized subject to different boundary conditions. We will only consider the shocked flow case which is more interesting and difficult to solve. The numerical solutions are obtained on uniform meshes of 40 linear and 20 quadratic elements. Two and three point gauss quadrature is used for the linear and quadratic elements respectively. The numerical solution to this problem has been obtained by Lohner et al [9], using the Taylor-Galerkin method and by Hughes and Tezdiyar [10], using the SUPG method. Artificial diffusion has been used in both methods to stabilize the solution around the shock. Jiang and Carey [11] modified their least-square method to capture the shock. The original, $L^2$ - method, was unstable due to the oscillation around the shock. In the modified $H^1$ - method, a weighted $H^1$ - norm of the residuals was minimized instead of
Breaking of a dam. Linear (a) and quadratic (b) solutions obtained with time step 0.5.

Breaking of Dam: This problem has been previously considered by Lohner et al. [9] using a Taylor-Galerkin scheme and by Carey and Jiang [15] using a Least-Squares approach. The equation governing the problem is defined as

\[
\begin{bmatrix}
H + \eta \\
(H + \eta)u
\end{bmatrix}
+ \begin{bmatrix}
(H + \eta)u \\
(H + \eta)u^2 + g(H + \eta)^2 / 2
\end{bmatrix}_t
= \begin{bmatrix}
0 \\
g(H + \eta)\frac{dH}{dx}
\end{bmatrix}
\]

Where \( H \) is the depth, \( \eta \) the surface elevation, \( u \) the velocity and \( g \) the acceleration due to gravity. Here a uniform mesh of 80 linear and 40 quadratic elements are used. The depth \( H \) and \( g \) are constant and equal to unity. The results obtained with the linear and quadratic element are shown in Fig. 4 at different times where the initial conditions are represented by \( t=0 \). The propagating jump discontinuity is well represented and the solutions are well behaved. The correct steep front in shallow water and the decreasing gradient in the deep water region is noteworthy.

Concluding Remarks

We have developed a least-squares finite element method for the solution of first order hyperbolic differential equation. The method is unconditionally stable and therefore more conveniently used as a time marching scheme for obtaining the steady state solutions. The
method is, however, shown to be capable of producing transient results comparable to existing finite elements methods. The incorporation of the higher order elements into the method is easily achieved, leading to improved results. The proposed jacobian evaluation improves the stability characteristics of the method and, more importantly, eliminates the need for numerical integration when solving problems in more than one dimension. The proposed method has the advantage of producing symmetric positive-definite matrices which makes the way for using efficient iterative methods such as Conjugate Gradient method. This property of the method will be touched upon at future date.

References