Least-Squares Finite Element Method for Fluid Dynamics

Bo-nan Jiang  
*Institute for Computational Mechanics in Propulsion*  
*Lewis Research Center*  
*Cleveland, Ohio*

and

Louis A. Povinelli  
*Lewis Research Center*  
*Cleveland, Ohio*

August 1989
LEAST-SQUARES FINITE ELEMENT METHOD
FOR FLUID DYNAMICS

Bo-nan Jiang*
Institute for Computational Mechanics in Propulsion
Lewis Research Center
Cleveland, Ohio 44135

and

Louis A. Povinelli
National Aeronautics and Space Administration
Lewis Research Center
Cleveland, Ohio 44135

This paper gives an overview of new developments of the least-squares finite element method (LSFEM) in fluid dynamics. Special emphasis is placed on the universality of LSFEM; the symmetry and positiveness of the algebraic systems obtained from LSFEM; the accommodation of LSFEM to equal-order interpolations for incompressible viscous flows; and the natural numerical dissipation of LSFEM for convective transport problems and high-speed compressible flows. The performance of LSFEM is illustrated by numerical examples.

1. Introduction

It has been more than two decades since people started applying finite element methods to fluid dynamics problems [1]. Subsequently, a large number of papers and many books on finite elements in fluids have been published, and numerous fluid dynamics problems have been successfully solved by finite element methods. Currently, however, only two commercial general-purpose finite element packages for incompressible flows, i.e. FIDAP and FLOTRAN [2], are available. This situation differs from that in solid mechanics. Twenty years after the first finite element papers were published, several dozen of finite element commercial packages for solid mechanics problems were on the market [3]. Perhaps, one reason for the lack of fluid dynamics finite element codes was that there was no unified method which could cover a wide range of fluid problems. For example, the classic Galerkin method is used for potential flows, the mixed Galerkin method and the penalty method are dominant for incompressible viscous flows [4,5], the Taylor-Galerkin method and the Petrov-Galerkin method are developed for convective transport problems and compressible flow problems [6-8]. Because the principles and structures of these methods are different, it is extremely difficult to implement these methods in a general-purpose code.

*Work funded under Space Act Agreement C99066G
The purpose of this paper is to advocate an unified method – LSFEM for fluid dynamics problems, to provide a conceptual framework for the method, and to review recent developments. In addition, comparisons are made with other methods and some representative numerical results are presented.

An outline of this paper follows: In Section 2, a brief review of the literature is given. The construction of LSFEM for first-order partial differential equations is presented in Section 3. Application of LSFEM for incompressible viscous flows is discussed in Section 4. Attention is fixed on the accommodation of equal-order interpolations. In Section 5, LSFEM for a convection problem is described. It is shown that the method naturally contains an upwinding mechanism to stabilize the solution. In section 6, LSFEM for the compressible Euler equations is presented. The capability of LSFEM to capture shocks is shown by numerical results. Conclusions are drawn in Section 7.

2. Review of Literature

The least-squares method of interest here is based on minimizing the residual in differential equations and boundary conditions in a least-squares sense. The use of least-squares method for the approximate solution of partial differential equations has a long history[9]. The modern theory of least-squares method for the numerical solution of elliptic boundary-value problems was started with the papers of Bramble and Schatz[10,11], and discussed by Varga[12]. This work uses a finite-dimensional space of approximating functions, similar to the spaces used in finite element methods.

When compared with the classic Galerkin finite element method, LSFEM generally leads to more stringent continuity requirements for the trial functions. Lynn and Arya[13] and Zienkiewicz et al [14] demonstrated that it is possible to reduce the order of continuity requirements at the expense of introducing more unknowns by first transforming the original differential equations into an equivalent system of first-order differential equations. Hence, the smoothness requirement for the space of approximate functions would be reduced, and the $C^0$ elements would be applicable. Subsequently, LSFEM were applied to certain simple flow problems (see Lynn[15], Lynn and Alani[16], and Polk and Lynn[17]). Fix and Guzburger[18] also studied LSFEM for systems of first-order equations and results were reported for the Tricomi equation.

A theoretical analysis of LSFEM for the elliptic system of the Petrovsky type was developed by Wendland[19], and optimal error estimates were obtained. Petrovsky systems are an important subclass of the class of elliptic systems, in which the different equations and unknowns appearing in the system have the same “differentiability order”.

A LSFEM for the Euler equations governing inviscid compressible flows was proposed by Fletcher[20]. The feature of his formulation is the designation of groups of variables

Nguyen and Reynen [22] presented a space-time LSFEM for the advection-diffusion problems. Their numerical results show that the use of upwinding techniques or the Taylor-Galerkin approach in finite elements turns out to be unnecessary when the least-squares weak formulation is extended into the time domain using standard shape functions.

An error estimate for the least-squares finite element solution of Cauchy-Riemann type equations was given by Fix and Rose [23].

A more general mathematical theory of LSFEM for the approximate solution of an elliptic system of Agmon-Douglis-Nirenberg (ADN) type was presented by Aziz, Kellogg and Stephens [24]. The method minimizes a least squares functional that consists of a weighted sum of the residuals occurring in the equations and the boundary conditions of the system. The weights are determined by the indices that enter into the definition of the ADN boundary-value problem.

It is shown that for the Cauchy-Riemann equations the optimal weighting factor is unity [25,26]. The LSFEM for incompressible and compressible flow analysis is implemented in conjunction with an element-by-element preconditioned conjugate gradient method and an adaptive refinement strategy [25,27-29]. A new preconditioner based on the $H^1$ gradient was formulated [25,26].

The recent advance of application of LSFEM in fluid dynamics will be discussed in the following sections.

3. Construction of LSFEM for First-Order Systems

Almost all problems arising in fluid dynamics, solid mechanics, heat transfer, electromagnetics and other mathematical physics can be recast in the form of first-order systems. For convenience, here we shall restrict our treatment initially to steady-state linear problems. It is straightforward to extend the treatment to quasi-linear problems.

At first we consider the boundary-value problem:

\[
\begin{align*}
  L_u &= f \quad \text{in } \Omega \\
  B_u &= g \quad \text{on } \Gamma
\end{align*}
\]

(3.1a) (3.1b)

where $L$ is a first-order partial differential operator:

\[
L_u = \sum_{i=1}^{n} A_i \frac{\partial u}{\partial x_i} + A_u
\]

(3.2)
\( \Omega \subset \mathbb{R}^{n_d} \) is a bounded domain with a piecewise smooth boundary \( \Gamma \), \( n_d = 2 \) or \( 3 \) represents the number of space dimensions, \( \mathbf{u}^T = (u_1, u_2, \ldots, u_m) \) is a vector of \( m \) unknown functions of \( \mathbf{x} \), \( \mathbf{A}_i \) and \( \mathbf{A} \) are \( m \times m \) matrices which depend on \( \mathbf{x} \), \( f \) is a given vector-valued function, \( \mathbf{B} \) is a boundary operator, and \( g \) is a given vector-valued function on the boundary. Without loss of generality we assume that \( g \) is a zero vector.

Here, we do not discuss the existence and uniqueness of the solution to (3.1), because these depend on the structure and properties of \( L \) and \( \mathbf{B} \), and the vector \( f \). In the following discussion, it is assumed that the problem (3.1) has a unique solution. We indicate that if there is a solution to (3.1), then the following least-squares method produces an approximate solution (see, e.g., [30,31]).

Throughout, \( L^2(\Omega) \) denotes the space of square-integrable functions defined on \( \Omega \) with inner product
\[
(u,v) = \int_{\Omega} uv d\Omega \quad u,v \in L^2(\Omega)
\]
and norm
\[
\|u\|_0^2 = (u,u) \quad u \in L^2(\Omega)
\]
We define the Sobolev space as:
\[
H^1(\Omega) = \{u \in L^2(\Omega), \partial^\alpha u \in L^2(\Omega), \forall |\alpha| \leq 1\}
\]
where \( \alpha = (\alpha_1, \alpha_2, \ldots, \alpha_{n_d}) \in \mathbb{N}^{n_d} \) and \( |\alpha| = \alpha_1 + \alpha_2 + \ldots + \alpha_{n_d} \), and define their associated norms by
\[
\|u\|_1^2 = \sum_{|\alpha| \leq 1} \|\partial^\alpha u\|_0^2
\]
For the vector-valued function \( \mathbf{u} \) with \( m \) components, we have the product spaces
\[
L^2(\Omega) = (L^2(\Omega))^m
\]
\[
H^1(\Omega) = (H^1(\Omega))^m
\]
and the corresponding norm
\[
\|\mathbf{u}\|_0^2 = \sum_{j=1}^m \|u_j\|_0^2
\]
\[
\|\mathbf{u}\|_1^2 = \sum_{j=1}^m \|u_j\|_1^2
\]
Considering the boundary condition of the boundary-value problem, we also define the function space
\[
\mathcal{S} = \{\mathbf{u} \in (H^1(\Omega))^m; \mathbf{B}\mathbf{u} = 0 \quad \text{on} \; \Gamma\}
\]
Let us suppose that \( f \in L^2 \) and \( L : S \to L^2 \). For an arbitrary trial function \( u \in S \), we define the residual function \( R = Lu - f \). The LSFEM is based on minimizing the residual function in a least-squares sense.

We construct the least-squares functional

\[
I(u) = \|Lu - f\|^2_0 = \langle Lu - f, Lu - f \rangle
\]

Taking variation of \( I \) with respect to \( u \), and setting \( \delta I = 0 \) and \( \delta u = w \), lead to the least-squares weak statement: Find \( u \in S \) such that

\[
(Lw, Lu) = (Lw, f) \quad \forall w \in S
\]

In the approximate analysis, we first discretize the domain as a union of finite elements and then introduce an appropriate finite element basis. Let \( N_e \) denote the number of nodes for one element and \( \psi_j \) denote the element shape functions. If equal-order interpolations are employed, that is, for all unknown variables the same finite element is used, we can write the expansion

\[
u_h(x) = \sum_{j=1}^{N_e} \psi_j(x) \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_m \end{pmatrix}_j
\]

where \((u_1, u_2, ..., u_m)_j \) are the nodal values at the \( j \)th node, and \( h \) denotes the mesh parameter.

Introducing the finite element approximation defined in (3.14) into the weak statement (3.13), we have the linear algebraic equations

\[
KU = F
\]

where the \( U \) is the global vector of nodal values. The global matrix \( K \) is assembled from the element matrices

\[
K_e = \int_{\Omega_e} (L\psi_1, L\psi_2, ..., L\psi_{N_e})^T (L\psi_1, L\psi_2, ..., L\psi_{N_e}) d\Omega
\]

in which \( \Omega_e \subset \Omega \) is the domain of the \( e \)th element, and \( T \) denotes the transpose, and the vector \( F \) is assembled from the element vectors

\[
F_e = \int_{\Omega_e} (L\psi_1, L\psi_2, ..., L\psi_{N_e})^T f d\Omega
\]
We observe that the matrix $K$ is symmetric and positive definite. Therefore, the iterative methods, such as the preconditioned element-by-element conjugate gradient method[27], can be employed. This property means that large-scale problems can be effectively solved on vector and parallel computers by using LSFEM.

Here it is important to emphasize that there is no added weighting parameter in our LSFEM. Each equation in the first-order system is equally treated. Usually, the governing equations are nondimensionalized by using the characteristic quantities of the problem. The nondimensionalization or scaling is equivalent to weighting the residuals in the least-squares functional. Although the choice of the characteristic quantities is arbitrary, the exact (physical) solution at which the least-squares functional takes the minimum value (zero) is independent of the scaling. In other words, the weighting or scaling can not change the minimum point. However, the weighting can change the condition number of the resulting algebraic system obtained from discretization. For the Cauchy-Riemann equations, we did both numerical experiments and a theoretical analysis to investigate the effect of the weighting factor[25,26]. It was observed that, when the weighting factor was wildly changed from $10^{-8}$ to $10^8$, the numerical solutions were almost fixed, except the number of conjugate gradient iterations was different. When the weighting factor was equal to one, the number of iterations to convergence was minimized.

The formulation of LSFEM is simple and systematic. Once a LSFEM based on a first-order system is coded, adaptation of the program to other problem is simply carried out by modifying the subroutines associated with the coefficient matrices $A_1, A_2, A_3$ and $A$ and the vector function $f$. In this way, one may develop a general-purpose program.

4. Incompressible Viscous Flow

In the past two decades numerous finite element schemes have arisen for solving the Navier-Stokes equations of incompressible viscous flow(see, e.g.,[5]). To date, most of these scheme require the use of non-equal-order interpolations and the solution of linear system of equations with non-symmetric and non-positive-definite matrix. In this section, we discuss an application of LSFEM which can circumvent these difficulties.

4.1 Stokes Problem

Consider the Stokes problem: Find the velocity $u = (u_1, u_2, u_3)$ and the pressure $p$
such that
\begin{align}
\nabla \cdot u &= 0 \quad \text{in } \Omega \quad (4.1a) \\
-\frac{1}{Re} \Delta u + \nabla p &= f \quad \text{in } \Omega \quad (4.1b) \\
\quad u &= 0 \quad \text{on } \Gamma \quad (4.1c)
\end{align}

where all variables are nondimensionalized, and \( Re \) is the Reynolds number.

The most popular method for the Stokes problem is based on the Galerkin mixed formulation (see, e.g., [4,5]): Find \( \{u, p\} \in (H^1_0(\Omega))^3 \times (L^2(\Omega)/\mathbb{R}) \) such that

\begin{align}
\nabla \cdot u &= 0 \quad \forall q \in L^2(\Omega) \quad (4.2a) \\
\quad a(u, v) - (\nabla \cdot u, q) &= (f, v) \quad \forall v \in (H^1_0(\Omega))^3 \quad (4.2b)
\end{align}

where

\[ H^1_0(\Omega) = \{ u \in H^1(\Omega); u = 0 \text{ on } \Gamma \} \quad (4.2c) \]

\[ a(u, v) = \frac{1}{Re} \sum_{i=1}^3 (\nabla u_i, \nabla v_i) \quad (4.2d) \]

The weak statement (4.2) can be obtained directly from the stationary condition for the Lagrangian

\[ L(u, p) = \frac{1}{2} a(u, u) - (f, u) - (\nabla \cdot u, p) \quad (4.3) \]

The solution \( \{u, p\} \) defines a saddle-point of this functional.

The existence of finite element approximate solution to (4.2) depends on choosing a pair of space \( \mathcal{V}_h \subset (H^1_0(\Omega))^3 \) and \( Q_h \subset L^2(\Omega) \) such that the Ladyzhenskaya-Babuška-Brezzi (LBB) "inf-sup" condition (see [32-34,4])

\[ \inf_{q \in Q_h} \sup_{v \in \mathcal{V}_h} \frac{(\nabla \cdot v, q)}{\| v \|_1 \| q \|_0} \geq \alpha > 0 \quad (4.4) \]

where \( \alpha \) is independent of the mesh size \( h \), holds. This condition precludes the use of equal-order interpolations. Although for two-dimensional problems quite a few convergent pairs of velocity and pressure elements have been developed (see, e.g., [4,5,34]), most of these combinations employ some basis functions that are inconvenient to be implemented. For three-dimensional problems, this difficulty becomes more severe and only rather elaborate constructions can pass the LBB test. The other basic difficulty associated with the mixed method is the non-positiveness of the linear algebraic systems.
features of the LSFEM are more clearly explained by a simple one-dimensional two-point boundary-value problem in [41].

In the following, we describe another LSFEM developed by Jiang and Chang[41]. Introducing an auxiliary variable, the vorticity \( \omega = \nabla \times u \), the Stokes equations can be written as

\[
\nabla \cdot u = 0
\]
\[
\frac{1}{Re} \nabla \times \omega + \nabla p = f \quad \text{in} \quad \Omega
\]
\[
\omega - \nabla \times u = 0
\]

We shall consider the two-dimensional problem only:

\[
\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} = 0
\]
\[
\frac{\partial p}{\partial x} + \frac{1}{Re} \frac{\partial \omega}{\partial y} = f_x
\]
\[
\frac{\partial p}{\partial y} - \frac{1}{Re} \frac{\partial \omega}{\partial x} = f_y
\]
\[
\omega + \frac{\partial u}{\partial y} - \frac{\partial v}{\partial x} = 0
\]

where \((f_x, f_y)\) is the body force components, and \(\Omega\) is a bounded domain in \(\mathbb{R}^2\) with a piecewise smooth boundary \(\Gamma\). Let \((\Gamma_1, \Gamma_2, \Gamma_3, \Gamma_4, \Gamma_5)\) denote the sides of \(\Gamma\). The unit outward normal vector to \(\Gamma\) is denoted by \(n\), and the tangential vector to \(\Gamma\) by \(t\).

We now write (4.8) in a general form of first-order system (3.1a), in which

\[
A_1 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -\frac{1}{Re} \\ 0 & -1 & 0 & 0 \end{pmatrix} \quad A_2 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{Re} \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}
\]
\[
A = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad f = \begin{pmatrix} f_x \\ f_y \end{pmatrix} \quad u = \begin{pmatrix} u \\ v \\ p \\ \omega \end{pmatrix}
\]

The boundary conditions should be supplemented to complete the definition of the boundary-value problem. We may consider the following boundary conditions:

\((a) u, v \text{ given on } \Gamma_1\)
(b) $u_n, \omega \text{ given on } \Gamma_2$

(c) $p, u_t \text{ given on } \Gamma_3$

(d) $p, \omega \text{ given on } \Gamma_4$

(e) $p, u_n \text{ given on } \Gamma_5$ \hfill (4.10)

For example, $\Gamma_1$ could be the inflow, outflow or wall boundary; $\Gamma_2, \Gamma_3, \Gamma_4,$ and $\Gamma_5$ could be the free surface, inflow, or outflow boundary.

The solvability of this elliptic boundary value problem depends on the combination of the boundary conditions. The system of equations is of the Petrovsky type. However, for many practical problems, the boundary operator does not satisfy the Lopatinski condition [19]. Thus, the proof of solvability is not trivial. This problem and a theoretical error analysis are addressed in [42].

Now we are ready to use LSFEM as described in Section 3.1. As usual we use the Gaussian quadrature technique to evaluate the coefficients of $K_e$ and $F_e$. The least-squares method with Gaussian quadrature is equivalent to the collocation least-squares method [43]. Therefore, the total number of collocation points ('Gauss' points) should be compatible with that of unknowns to get good results. For this reason, in our numerical experiments we use reduced integration [14].

We chose a model problem studied by Oden and Jacquotte [44]. This model problem corresponds to the polynomial divergence-free velocity-pressure-vorticity field:

\begin{align*}
  u(x, y) &= x^2(1 - x)^2(2y - 6y^2 + 4y^3) \\
  v(x, y) &= y^2(1 - y)^2(-2x + 6x^2 - 4x^3) \\
  p(x, y) &= x^2 - y^2 \\
  \omega(x, y) &= -x^2(1 - x)^2(2 - 12y + 12y^2) + y^2(1 - y)^2(-2 + 12x - 12x^2)
\end{align*} \hfill (4.11)

with the boundary conditions shown in Figure 4.1.

We have tested the bilinear element and the eight-node quadratic element using uniform meshes. The one-point Gaussian quadrature for the bilinear element and the $2 \times 2$ quadrature for the eight-node quadratic element are employed. The numerical results of the rate of convergence are shown in Figure 4.2. It is found that in all tested cases

\begin{align*}
  \|u - u_h\|_0^2 + \|v - v_h\|_0^2 \leq ch^{k+1}, \quad \|p - p_h\|_0 \leq ch^{k+1}, \quad \|\omega - \omega_h\|_0 \leq ch^{k+1}
\end{align*} \hfill (4.12)
where \( h \) is the mesh size; as a subscript, \( h \) denotes the approximate solution; \( k \) is the order of the polynomial. That is, all variables \( u, v, p, \omega \) converge in \( L^2 \) norm at the optimal rate, as predicted by a theoretical error estimate in [42].

4.2 Incompressible Navier-Stokes Problem

The method presented in Section 4.1 can be generalized to solve the two-dimensional incompressible Navier-Stokes equations written in the following first-order quasi-linear velocity-pressure-vorticity formulation[45]:

\[
\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} = 0
\]

\[
\begin{align*}
\frac{u}{\partial x} + \frac{v}{\partial y} + \frac{\partial p}{\partial x} + \frac{1}{Re} \frac{\partial \omega}{\partial y} &= f_x \\
u \frac{\partial v}{\partial x} + v \frac{\partial v}{\partial y} + \frac{\partial p}{\partial y} - \frac{1}{Re} \frac{\partial \omega}{\partial x} &= f_y \\
\omega + \frac{\partial u}{\partial y} - \frac{\partial v}{\partial x} &= 0
\end{align*}
\]

in \( \Omega \) (4.13)

The boundary conditions should be supplemented to complete the boundary-value problem.

We now write (4.13) in a general form of first-order system (3.1a), in which

\[
A_1 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ u & 0 & 1 & 0 \\ 0 & u & 0 & -\frac{1}{Re} \\ 0 & -1 & 0 & 0 \end{pmatrix} \quad A_2 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ v & 0 & 0 & \frac{1}{Re} \\ 0 & v & 1 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}
\]

\[
A = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad f = \begin{pmatrix} 0 \\ f_x \\ f_y \\ 0 \end{pmatrix} \quad u = \begin{pmatrix} u \\ v \\ p \\ \omega \end{pmatrix}
\]

A successive substitution method is used to solve this quasi-linear problem. The corresponding solution of the Stokes problem is taken as the initial guess. The velocity field at the previous step is used to calculate the coefficient matrices.

The method is tested for the driven cavity problem by using \( 40 \times 40 \) nonuniform bilinear elements. The difference between the results of two successive steps is defined as

\[
e = \max_{i=1, N_n} |u_i^n - u_i^{n+1}|
\]
where \( i \) denotes the node, \( N \) is the total number of nodes, and \( n \) denotes the substitution level. The substitution continues until the difference \( e \) becomes less than the tolerance \( 10^{-4} \). The required number of iterations are 8, 12, 15, 57, and 160 for the Reynolds number \( Re \) of 100, 400, 1000, 3200, and 5000, respectively. The numerical results compare favorably with a finite difference solution using fine grid. Figure 4.3 illustrates the mesh used and Figure 4.4 compares the LSFEM results with those obtained by Ghia[46]. Figure 4.5 through 4.8 show the flow pattern, the streamlines, the pressure contours and the vorticity contours obtained with the LSFEM for \( Re = 5000 \).

We conclude this section by comparing LSFEM with other methods. In contrast to the Galerkin mixed method based on the velocity- pressure formulation[5,34], LSFEM based on velocity-pressure-vorticity formulation does not depend on the LBB condition, can accommodate equal-order interpolations, and always produces a symmetric, positive and definite matrix. In contrast to the penalty method[5], LSFEM does not have any added parameter in the scheme. In other words, LSFEM is robust. In contrast to the stream function-vorticity formulation[5,34], LSFEM produces the velocity and pressure directly and can be extended to three-dimensional problems.

5. Transport

In this section, we consider the two- or three-dimensional convection(hyperbolic) problems:

\[
\begin{align*}
\frac{\partial \phi}{\partial t} + u \cdot \nabla \phi &= 0 \quad \text{in } \Omega \times (0, T) \\
\phi &= g \quad \text{on } \Gamma_+ \times (0, T) \\
\phi &= \phi_0 \quad \text{in } \Omega \text{ for } t = 0
\end{align*}
\]

where \( \phi \) is, for example, the concentration, the convective velocity vector \( u \) is assumed known and divergence-free in \( \Omega \), and the boundary data \( g \) is prescribed only on the inflow boundary \( \Gamma_- = \{ \mathbf{x} \in \Gamma; \mathbf{n}(\mathbf{x}) \cdot \mathbf{u}(\mathbf{x}) < 0 \} \) in which \( \mathbf{n}(\mathbf{x}) \) is the outward unit normal to \( \Gamma \) at \( \mathbf{x} \in \Gamma \).

The classic Galerkin variational problem is to find \( \phi \in S = \{ H^1(\Omega); v = g \text{ on } \Gamma_- \} \) for each \( t \), such that

\[
(w, \frac{\partial \phi}{\partial t} + u \cdot \nabla \phi) = 0 \quad \forall w \in V = \{ H^1(\Omega); v = 0 \text{ on } \Gamma_- \}
\]

together with the initial condition

\( \phi(0) = \phi_0 \)
For convective transport problems, the classic Galerkin finite element method suffers from spurious oscillations in the same way as the central finite differencing. Motivated by backward differencing to reflect the "upstream" dependence of the solution, the Petrov-Galerkin method and the streamline upwind Petrov-Galerkin (SUPG) method were developed (see, e.g., [47,8]). The weak statement of the SUPG method is to find $\phi \in S$ for each $t$, such that

$$ (w + ru \cdot \nabla w, \frac{\partial \phi}{\partial t} + u \cdot \nabla \phi) = 0 \quad \forall w \in V $$ (5.3)

where $r$ is a parameter. The term "Petrov-Galerkin" is used to indicate that the test functions are no longer the same as the trial functions for the approximation. The term "streamline upwind" comes from the term $1 - \nu$ in the test functions.

In the steady case, the SUPG formulation is: Find $\phi \in S$ such that

$$ (w + ru \cdot \nabla w, u \cdot \nabla \phi) = 0 \quad \forall w \in V $$ (5.4)

Note that this formulation combines the usual Galerkin term and the least-squares term together. For this reason, Hughes has renamed the SUPG method as the Galerkin/least-squares method [48]. By fine-tuning the upstream weighting parameter $r$ in the scheme, excellent results can be obtained.

Recently, the ideas of the Lax-Wendroff scheme in finite difference method have motivated development of the related Taylor-Galerkin finite element method [49,6]. Let us briefly describe the Taylor-Galerkin method. Since equation (5.1a) defines an evolution statement, there must exist the Taylor expansion

$$ \phi^{n+1} = \phi^n + \Delta t \phi_t^n + \frac{1}{2} \Delta t^2 \phi_{tt}^n + ... $$ (5.5)

where the subscript 't' denotes the order of temporal derivative at $t_n$, and $t_{n+1} = t_n + \Delta t$. Equation (5.1a) permits restatement of the first derivative term and the second derivative term in (5.5) as

$$ \phi_t = -u \cdot \nabla \phi $$ (5.6)

$$ \phi_{tt} = -u \cdot \nabla \phi_t = u \cdot \nabla (u \cdot \nabla \phi) $$ (5.7)

Substituting (5.6) and (5.7) into (5.5) yields

$$ \phi^{n+1} = \phi^n - \Delta tu \cdot \nabla \phi^n + \frac{1}{2} \Delta t^2 u \cdot \nabla (u \cdot \nabla \phi^n) + ... $$ (5.8)

Testing (5.8) against $w \in V$ and integrating by parts we obtain the Taylor-Galerkin weak statement: Find $\phi \in S$ such that

$$ (w, \phi^{n+1}) = (w, \phi^n - \Delta tu \cdot \nabla \phi^n) - \frac{1}{2} \Delta t^2 (u \cdot \nabla w, u \cdot \nabla \phi^n) $$
where $\Gamma_+$ denotes the outflow boundary.

Since the term $(w, \phi^{n+1})$ introduces the mass matrix, the Taylor-Galerkin method (5.9) requires inversion of the mass matrix. To maintain the explicit character of the method, one may use the lumped (diagonal) mass matrix. The method thus suffers from more severe stability limitations (Courant number $C = \frac{\Delta t}{h} \leq 1/\sqrt{3}$). An implicit Taylor-Galerkin method [50] can be constructed, but it requires solution of nonsymmetric matrix. Furthermore, the Taylor-Galerkin method is essentially a high-order scheme, and hence may promote oscillations at discontinuities. Artificial dissipation can be introduced to suppress oscillations, but the form of this added term is not unique and the associated parameters must be specified.

For first-order hyperbolic systems, another choice is LSFEM [51-54]. This method uses the backward scheme or the Crank-Nicolson scheme of finite differences in the temporal domain and finite elements in the spatial domain. For one-dimensional problem, it was demonstrated that LSFEM is unconditionally stable for all Courant numbers, and naturally acts in a manner similar to upwinding.

Here let us apply LSFEM to the problem (5.1). Backward differencing (5.1) through a time step and writing $\phi^n$ for $\phi(x, t_n)$, the resulting semidiscrete problem is

$$\phi^{n+1} - \phi^n + \Delta t \; u \cdot \nabla \phi^{n+1} = 0 \tag{5.10}$$

Now we are ready to use LSFEM as described in Section 3. In order to explain how the method works, we deduce the least-squares weak formulation.

For any admissible $\phi^{n+1}(x) \in S$, we obtain from (5.10) the residual $R$, and the least-squares functional $I = \int_{\Omega} R^2 d\Omega$ becomes

$$I(\phi^{n+1}) = ||\phi^{n+1} - \phi^n + \Delta t u \cdot \nabla \phi^{n+1}||_0^2 \tag{5.11}$$

Applying the stationary condition $\delta I = 0$, and setting $\delta \phi^{n+1} = w$, we obtain the weak statement for (5.10): Find $\phi^{n+1} \in S$ such that

$$(w + \Delta t u \cdot \nabla w, \phi^{n+1} + \Delta t u \cdot \nabla \phi^{n+1}) = (w + \Delta t u \cdot \nabla w, \phi^n) \quad \forall w \in V \tag{5.12}$$

Clearly (5.12) can be interpreted as a Petrov-Galerkin weighted-residual statement for (5.10) with test function $\tilde{w} = (1 + \Delta t u \cdot \nabla)w$. More importantly, we note that the convective velocity $u$ enters explicitly in this test function so that this form is analogous to SUPG methods. Hence, the effect of the least-squares procedure in two or three-dimensions is to add numerical dissipation preferentially in the local flow direction in a manner similar to SUPG method.
Introducing the increment $\Delta \phi = \phi^{n+1} - \phi^n$, the formulation (5.12) may be rewritten in the $\Delta$-form: Find $\Delta \phi \in V$ such that

$$(w + \Delta t u \cdot \nabla w, \Delta \phi + \Delta t u \cdot \nabla (\Delta \phi)) = -\Delta t (w + \Delta t u \cdot \nabla w, u \cdot \nabla \phi^n) \quad \forall w \in V \quad (5.13)$$

The formulation (5.13) clearly shows that the resulting matrix of LSFEM is symmetric.

If the boundary condition (5.1b) is independent of the time $t$. The steady state solution exists. In this case, when the steady state is reached by the time marching, $\Delta \phi$ becomes zero, and the corresponding weak statement is: Find $\phi \in S$ such that

$$(w + \Delta t u \cdot \nabla w, u \cdot \nabla \phi) = 0 \quad \forall w \in V \quad (5.14)$$

Note that the formulation (5.14) is the same as the SUPG or Galerkin/least-squares formulation (5.4). Although the Galerkin/least-squares formulation can be deduced by the least-squares procedure, there is a significant difference between these two methods. Even for steady state cases, LSFEM uses (5.12) or (5.13) as a time marching procedure. For each time-step, the algebraic system obtained by LSFEM is symmetric, while the Galerkin/least-squares method leads to a nonsymmetric matrix.

We should also point out that the success of LSFEM for the hyperbolic systems depends on the choice of time step $\Delta t$. If the time step is too large, the numerical results at discontinuities will be diffused too much; if the time step is too small, the solution will have significant oscillation at discontinuities. Our numerical experiments tells us that one may choose the time step such that the Courant number $C \approx 10 \sim 50$. Of course, a theoretical investigation about this problem is needed. Although LSFEM has this disadvantage, the symmetry of the matrix makes the method attractive.

As a numerical example, we consider the water-oil displacement problem in 5-spot pattern. A displacement front propagates from the water injection well at the lower-left corner towards the oil production well at the top-right corner. The flow domain is assumed homogeneous, and the velocity field is then given by the solution of Laplace’s equation for this source-sink system with symmetry (no flow) conditions on the sides to enforce periodicity as

$$u = \frac{x}{x^2 + y^2} - \frac{x - 1}{(x - 1)^2 + (y - 1)^2}$$

$$v = \frac{y}{x^2 + y^2} - \frac{y - 1}{(x - 1)^2 + (y - 1)^2} \quad (5.15)$$

Using this specified velocity field in the least-squares finite element analysis, the finite-element system follows directly from the introduction of the finite element expansion for $\phi$ into (5.12). The unit square domain is discretized to a uniform 40\times40 grid of bilinear elements with the initial data $\phi(x, y, 0) = 0$ and the injection value $\phi(0, 0, t) = 1$ for $t \geq 0$. 15
The equispaced solution contours for concentration $\phi(x,y,t)$ are shown in Figure 5.1 at $t = 0.05$ and $0.3$ and the fixed time step $\Delta t = 0.0025$. We note that there is numerical dissipation of the propagating front and no oscillation.

6. High-Speed Compressible Flow

Recently, much attention has been focused on the development of finite element algorithms for the analysis of high speed compressible flows especially for treating discontinuities. Although considerable success has been achieved by using the methods such as the streamline-upwind Petrov-Galerkin method[8], the Taylor-Galerkin method[55,7], the Taylor-Galerkin method with flux-corrected transpose(FCT) and mesh refinement[56-59], the block relaxation via Godunov's method[60], the characteristic Galerkin method[61], and the non-oscillatory discontinuous Galerkin method[62], more investigation is still needed to compare the efficiency and accuracy of these methods and alternative schemes.

For the numerical solution of the one- and two-dimensional compressible Euler equations, we proposed a class of least-squares methods[63,64]. We begin by considering the first-order implicit time-differenced non-conservative formulation. The least-squares method is then employed to minimize the residual in the $L^2$ or $H^1$ norm. It was demonstrated that the associated numerical dissipation arises naturally as a consequence of the method.

6.1 $L^2$ Method

Now let us describe LSFEM for the two-dimensional compressible Euler equations in the form of a first-order system:

$$\frac{\partial u}{\partial t} + A_1 \frac{\partial u}{\partial x} + A_2 \frac{\partial u}{\partial y} = 0$$

(6.1)

where $u = (\rho, u, v, p)$, and

$$A_1 = \begin{pmatrix} u & \rho & 0 & 0 \\ 0 & u & 0 & \rho^{-1} \\ 0 & 0 & u & 0 \\ 0 & \gamma p & 0 & u \end{pmatrix} \quad A_2 = \begin{pmatrix} v & 0 & \rho & 0 \\ 0 & v & 0 & 0 \\ 0 & 0 & v & \rho^{-1} \\ 0 & 0 & \gamma p & v \end{pmatrix}$$

in which $\rho$ is the density, $(u, v)$ are the fluid velocity components, $p$ is the pressure, and $\gamma$ is the specific heat ratio. The boundary conditions and the initial condition are needed to complete the description of the problem.
For a given time step $\Delta t = t^{n+1} - t^n$, we linearize the problem by setting $A_1^n = A_1(u^n)$, $A_2^n = A_2(u^n)$. Backward differencing leads to the implicit time-differenced problem

$$u^{n+1} - u^n + \Delta t A_1^n \frac{\partial u^{n+1}}{\partial x} + \Delta t A_2^n \frac{\partial u^{n+1}}{\partial y} = 0 \quad (6.2)$$

Now we are ready to use the formulation of LSFEM in Section 3 to get the solution $u^{n+1}$ for each time-step. Here, the first-order partial differential operator $L$ has the following form:

$$L = \Delta t A_1^n \frac{\partial}{\partial x} + \Delta t A_2^n \frac{\partial}{\partial y} + E \quad (6.3)$$

where $E$ is an identity matrix.

We proceed further to discern why this method can capture shocks. For this reason we deduce the weak statement.

The basic least-squares method for the system (6.2) amounts to minimizing the $L^2$ norm of the residual $R$ for admissible $u^{n+1}$ in (6.2); i.e., minimizing the objective functional

$$I_0 = \|R\|^2_0 \quad (6.4)$$

where

$$R = u^{n+1} - u^n + \Delta t A_1^n \frac{\partial u^{n+1}}{\partial x} + \Delta t A_2^n \frac{\partial u^{n+1}}{\partial y}$$

Taking variations of $I_0$ with respect to $u^{n+1}$ and setting $w = \delta u^{n+1}$, $\delta I_0 = 0$ leads to the least-squares weak statement

$$(w + \Delta t A_1^n \frac{\partial w}{\partial x} + \Delta t A_2^n \frac{\partial w}{\partial y}, u^{n+1} + \Delta t A_1^n \frac{\partial u^{n+1}}{\partial x} + \Delta t A_2^n \frac{\partial u^{n+1}}{\partial y})$$

$$= (w + \Delta t A_1^n \frac{\partial w}{\partial x} + \Delta t A_2^n \frac{\partial w}{\partial y}, u^n) \quad (6.5)$$

When the steady state is reached by time marching, the corresponding weak statement becomes

$$(w, A_1 \frac{\partial w}{\partial x} + A_2 \frac{\partial w}{\partial y}) + \Delta t (A_1 \frac{\partial w}{\partial x} + A_2 \frac{\partial w}{\partial y}, A_1 \frac{\partial u}{\partial x} + A_2 \frac{\partial u}{\partial y}) = 0 \quad (6.6)$$

**Remark** For steady state problems we still use the formulation (6.5) as the time marching procedure, since the resulting algebraic system is symmetric. The formulation
(6.6) is not appropriate for practical calculation, because it is a nonlinear problem and the resulting matrix is nonsymmetric.

Let us examine the structure of (6.6). The first term in (6.6) corresponds to the standard Galerkin weak statement. The second term leads to a symmetric, positive and definite matrix, and always acts as a numerical dissipation to stabilize the solution by smoothing out any discontinuities.

We now demonstrate some features of the $L^2$ LSFEM using three numerical examples.

A standard test problem corresponding to the reflection of a shock from a wall is depicted in Figure 6.1. On the upper boundary of the flow domain $\rho = 1.7, u = 2.6185, v = -0.5082, p = 1.5282$ and on the upstream boundary $\rho = 1.0, u = 2.9, v = 0.0, p = 0.7143$ are prescribed. The lower boundary, where $v = 0$, is the wall from which the generated shock reflects, and the downstream boundary conditions remain free for outflow. The initial data were prescribed as constant at values given on the upper boundary and the specific heat ratio is $\gamma = 1.40$. In the calculation, a uniform $20\times60$ mesh of bilinear elements was used. The solution was integrated with time-step $\Delta t = 0.33333$ until an essentially steady state is obtained in 12 time-steps. The pressure contours for this steady solution are given in Figure 6.1. Qualitatively, it is seen that the flow physics are correctly modelled. Although the shock is somewhat smeared, the oscillations are absent and the calculation is efficient.

The second problem is a Mach 3 flow (with $\gamma = 1.40$) over a $20^\circ$ ramp. The gas enters with uniform flow conditions through the left boundary of the domain and an oblique shock develops at the root of the ramp. A mesh with 824 bilinear elements and the computed pressure contours are illustrated in Figure 6.2. In the calculation, the initial data were prescribed as constant at the value given on the left boundary, and the time-step was $\Delta t = 0.33333$. The steady state was obtained in 16 time-steps.

We also considered a cylinder in a supersonic flow with Mach number $M_\infty = 2, \gamma = 1.40$. Figure 6.3 presents the mesh with 800 bilinear elements and the computed pressure contours.

We note that in all three numerical examples, LSFEM with bilinear elements produces non-oscillatory shock profiles and the results compare favorably with those in the literature based on other finite element methods.

6.2 $H^1$ Method

Numerical experiments show that as long as the time-step $\Delta t$ is large enough (the corresponding Courant number $\approx 10 - 50$), the $L^2$ method with linear elements gives
non-oscillatory shock profiles in the computed steady state. However, for high-order elements, computed solutions based on this method may have oscillations. The behaviour of this method is similar to that of high-order finite difference schemes. To circumvent this problem, one may use low-order (linear) elements around discontinuities and high-order elements elsewhere. However, the implementation is not so easy. Thus, we proposed an alternative technique in which the objective function is modified to control the residual derivatives also\[63-65\]. Accordingly, the following multi-objective optimization problem is constructed\[66\]:

$$
\min \left( \| R \|_0^2, \| \frac{\partial R}{\partial x} \|_0^2, \| \frac{\partial R}{\partial y} \|_0^2 \right)
$$

(6.7)

in which the minimization of $I_0$ is the main objective. As a technique to solve (6.7), we minimize the modified functional

$$
I = \| R \|_0^2 + \beta h^2 \left[ \| \frac{\partial R}{\partial x} \|_0^2 + \| \frac{\partial R}{\partial y} \|_0^2 \right]
$$

(6.8)

where $\beta$ is a small parameter, $0 < \beta \ll 1$. Note that minimizing the functional (6.8) can be interpreted as minimizing the weighted $H^1$ norm of $R$. However, the form in (6.8) contains second derivatives of $u$, which implies that, for conformity, $C^1$ elements are appropriate. Generally, for two- and three-dimensional problems, simpler elements are desirable. An approximate formulation based on (6.8) can be constructed using high order $C^0$ elements. we may approximate (6.8) as

$$
I_1 = (R, R) + \beta h^2 \left[ (\frac{\partial R}{\partial x}, \frac{\partial R}{\partial x}) \tilde{\Omega} + (\frac{\partial R}{\partial y}, \frac{\partial R}{\partial y}) \tilde{\Omega} \right]
$$

(6.9)

where $\tilde{\Omega}$ denotes element interiors. For example, with $C^0$ quadratic or cubic elements, the contributions of the second derivatives in element interiors can be calculated for (6.9).

We now demonstrate the shock capturing ability of the $H^1$ LSFEM using the same shock wave reflection problem as in Section 6.1. The computed pressure contours with 8-node quadratic elements are shown in Figure 6.4. Comparing Figure 6.4 with Figure 6.1, we observe that for the same grids this new technique gives the better resolution of the shocks.

### 6.3 Conservative Form

The LSFEM described in Section 6.1 is based on the nonconservative formulation. Here we would like to construct a conservative LSFEM (in the sense of the steady state). The Euler equations governing two-dimensional compressible inviscid flows can be written in conservative form as
\[
\frac{\partial q}{\partial t} + \frac{\partial F}{\partial x} + \frac{\partial G}{\partial y} = 0 \tag{6.11}
\]

where

\[
q = \begin{pmatrix} \rho \\ \rho u \\ \rho v \\ \rho e \end{pmatrix}, \quad F = \begin{pmatrix} \rho u \\ \rho u^2 + p \\ \rho u v \\ (\rho e + p)u \end{pmatrix}, \quad G = \begin{pmatrix} \rho v \\ \rho u v \\ \rho v^2 + p \\ (\rho e + p)v \end{pmatrix} \tag{6.12}
\]

in which \( e \) is the total energy, and for the case of perfect gas the equation of state is

\[
p = (\gamma - 1)(e - \frac{1}{2} \rho (u^2 + v^2)) \tag{6.13}
\]

Retaining the conservative variables as an intermediate step, we first convert (6.11) (temporarily) into the following nonconservative form:

\[
\frac{\partial q}{\partial t} + \mathbf{A}_1 \frac{\partial q}{\partial x} + \mathbf{A}_2 \frac{\partial q}{\partial y} = 0 \tag{6.14}
\]

\[
\mathbf{A}_1 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ \frac{1}{2} \gamma (u^2 + v^2) - u^2 & (3 - \gamma) u & -\gamma v & \gamma \\ -uv & v & u & 0 \\ (\gamma (u^2 + v^2) - \gamma e) u & -\gamma u^2 & -\gamma uv & \gamma u \end{pmatrix} \tag{6.15}
\]

\[
\mathbf{A}_2 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ -uv & v & u & 0 \\ \frac{1}{2} \gamma (u^2 + v^2) - v^2 & -\gamma u & (3 - \gamma) v & \gamma \\ (\gamma (u^2 + v^2) - \gamma e) v & -\gamma uv & (\gamma - \gamma v^2 & \gamma v \end{pmatrix} \tag{6.16}
\]

in which

\[
\bar{\gamma} = \gamma - 1, \quad \bar{e} = \gamma e - \frac{\gamma (u^2 + v^2)}{2}.
\]

Using the same procedure as before we obtain the corresponding least-squares weak statement similar to (6.5). Introducing \( \Delta q = q^{n+1} - q^n \) and regrouping the terms, we have then a conservative least-squares weak statement:

\[
(w + \Delta t \mathbf{A}_1^n \frac{\partial w}{\partial x} + \Delta t \mathbf{A}_2^n \frac{\partial w}{\partial y}, \Delta q + \Delta t \mathbf{A}_1^n \frac{\partial \Delta q}{\partial x} + \Delta t \mathbf{A}_2^n \frac{\partial \Delta q}{\partial y}) =
\]

\[
-\Delta t(w + \Delta t \mathbf{A}_1^n \frac{\partial w}{\partial x} + \Delta t \mathbf{A}_2^n \frac{\partial w}{\partial y}, \frac{\partial F^n}{\partial x} + \frac{\partial G^n}{\partial y}) \tag{6.17}
\]

20
Here $\Delta q$ is the unknown increment in conservative variables for the time step $\Delta t$. We may use the conservative variables at the previous time step to calculate the nodal values of components of flux $F^n$ and $G^n$, then use a finite element approximation to calculate $\frac{\partial F^n}{\partial x}$ and $\frac{\partial G^n}{\partial y}$. Once the steady conservation law

\[
\frac{\partial F}{\partial x} + \frac{\partial G}{\partial y} = 0
\]

(6.18)

is satisfied, the increment $\Delta q$ becomes zero, and the calculation is terminated. Our numerical experiments on the shock wave reflection problem show that there is no essential difference between the results of nonconservative and conservative LSFEM. More numerical experiments on different problems are needed to compare these two methods.

7. Conclusions

A variety of complex flow problems have been solved using the least-squares finite element method. Perhaps no other single method can solve the same range of flow problems. This paper has demonstrated that the method is simple, general, and reliable. It is hoped that this review will help stimulate the development of a general-purpose flow solver based on the least-squares finite element method, which in turn will speed the development of new applications.

References


Figure 4.1  Boundary conditions for Stokes problem

(a) 8-node quadratic element

(b) Bilinear element

Figure 4.2  Computed convergence rate for Stokes problem.
Figure 4.3  Finite element mesh (40×40 bilinear elements) for the cavity flow.
Figure 4.4 Horizontal velocity profile for cavity flow at Reynolds number 5000

— present (40x40 grids), △ Ghia[47] (257x257 grids).
Figure 4.5  Flow pattern of cavity flow at Re=5000
Figure 4.6 Streamlines for cavity flow at Re=5000
Figure 4.7 Pressure contours for cavity flow at Re=5000
Figure 4.8  Vorticity contours for cavity flow at Re=5000
(a) at $t = 0.05$

(b) at $t = 0.30$

Figure 5.1 Concentration contours for five-spot problem

$(40 \times 40$ bilinear elements, $\Delta t = 0.0025)$
Figure 6.1  Pressure contours for wave reflection problem ( $L^2$ method, 20×60 bilinear elements, $\Delta t = 0.33333$, number of steps = 12 )
Figure 6.2 Finite element mesh and pressure contours for supersonic flow over a 20° ramp (L^2 method, 864 bilinear elements, Δt = 0.33333, number of steps=16)
Figure 6.3  Finite element mesh and pressure contours for the Mach 2 flow past a circular cylinder ($L^2$ method, 800 bilinear elements, $\Delta t = 0.1$, number of steps = 36)

Figure 6.4  Pressure contours for wave reflection problem ($H^1$ method, $\beta = 0.0001$, $10 \times 30$ 8-node quadratic elements, $\Delta t = 0.33333$, number of steps = 14)
This paper gives an overview of new developments of the least-squares finite element method (LSFEM) in fluid dynamics. Special emphasis is placed on the universality of LSFEM; the symmetry and positiveness of the algebraic systems obtained from LSFEM; the accommodation of LSFEM to equal-order interpolations for incompressible viscous flows; and the natural numerical dissipation of LSFEM for convective transport problems and high-speed compressible flows. The performance of LSFEM is illustrated by numerical examples.