Multilevel Projection Methods for First-Order System Least Squares[1]

T. A. Manteuffel and Steve McCormick
Program in Applied Mathematics
University of Colorado
Boulder, Colorado 80309-0526

ABSTRACT. This paper focuses on the multilevel projection method (PML) applied to numerical solution of the basic equations of fluid dynamics formulated as least-squares problems for first-order systems. The fundamental approach taken here is to pose the fluid flow equations in their first-order form, incorporate additional (usually redundant) equations where they improve the character of the formulation, and define the objective as that of minimizing the L_2 norm of the defect (or residual) of these equations. This is the basic idea behind first-order system least squares (FOSLS). The Rayleigh-Ritz variational form of PML is applied to discretize and solve the FOSLS minimization problem. This paper summarizes the basic principles and concepts of PML and FOSLS, and briefly describes some of the special forms that this methodology takes for convection-diffusion, Helmholtz, Cauchy-Riemann, Stokes, and hyperbolic equations.

Key words: least-squares, second-order elliptic problems, finite elements, Rayleigh-Ritz, computational fluid dynamics.

AMS(MOS) subject classifications: 65F10, 65F30.

1. Introduction

The solution of certain physical problems, including many self-adjoint elliptic partial differential equations, can be obtained as the minimizer of a known "energy" functional over an infinite-dimensional space of admissible functions. When properly posed, such optimization problems often have the advantage that they can be discretized by a Rayleigh-Ritz process, based on minimizing the functional over a finite-dimensional subspace of the admissible-function space. The potential benefits of such an approach include the following.

- **an inherent sense of optimality.** Instead of the often misleading and restrictive truncation-error approach used with finite differences (cf. [3]), the basic design issue is to choose the subspace so that it properly reflects the local behavior (e.g., know smoothness) of the target solution. The discrete problem then inherits the property that it determines the best solution from that subspace in the sense of minimization of the functional. In fact, this sense of optimality can and usually should be used to guide the entire solution process. For example, it can be used to ensure that the discrete solution method never diverges and that each of its stages makes optimal changes to the emerging approximate solution.
- **monotonicity of the approximation.** If the subspaces are nested, then the discrete problems exhibit a monotonicity in the sense that "finer" levels

[1]This research was supported by the Air Force Office of Scientific Research under grant number AFOSR-91-0156, by the Department of Energy under grant number DE-FG02-90ER25086, and by the National Science Foundation under grant number DMS 8704169.

provide at least as good an approximation as any "coarser" one. For example, local refinement cannot be deleterious (in the sense of the functional) provided it maintains the Rayleigh-Ritz structure. An important consequence of this monotonicity is that error estimates are relatively easily obtained because issues like interpolation error in the multigrid coarsening process are nonexistent (e.g., there is no interpolation error in coarse-to-fine interval transfers).

- **self-adjointness.** The Hessian of the functional is self-adjoint, so the discretization of linear (or linearized) well-posed physical problems yields self-adjoint positive definite matrix equations. This generally allows for more efficient and reliable algebraic solvers (e.g., conventional conjugate gradient methods). Most importantly here, it provides a natural framework in which to apply the simpler forms of multigrid solvers. In particular, such matrix equations admit the usual variational multigrid schemes (cf. [4]) with several inherent advantages: they require fewer design choices because fine-to-coarse interlevel transfers are the adjoints of the coarse-to-fine transfers, for example; they are better supported by theory, e.g., estimates of optimal V-cycle type exist only for essentially self-adjoint problems; and they tend to be more easily controlled and more robust.

These advantages make it compelling to study ways to formulate general physical problems in an optimization framework. One possibility is to apply least squares to the original equations. Consider, for example, the linear operator equation $Lp = f$, defined on some normed vector space. A least-squares approach would reformulate this problem as that of minimizing the functional $\|Lp - f\|^2$. The disadvantages of this general approach is that it tends to produce a more restricted class of problems that is less well posed. For example, for second-order differential boundary value problems on the usual Sobolev spaces of functions, this approach effectively leads to a fourth-order differential equation that tends to require much more smoothness of the solution, to eliminate the usual finite element spaces from admissibility, and to produce discrete problems with squared condition numbers. It would therefore seem, at first glance, that least squares is not a viable approach to converting general equations to optimization problems. On the contrary, it is viable for a large class of, say, second-order problems, but only if the original equations are expanded to first-order systems before the least-squares principles are applied. Such an approach, which we refer to as first-order system least squares (FOSLS; cf. [1], [2]), is the main focus of this paper.

The basic idea behind FOSLS can be traced back several decades However, not until recently was the full power of this methodology realized. In fact, earlier least-squares methods were unable either to avoid certain limitations like the restrictive Ladyhenskaya-Babǔska-Brezzi condition, to obtain a strong sense of ellipticity, to achieve relatively well-conditioned discrete problems, or to realize a robust and systematic approach. Not before the multigrid scheme developed in [2] was a general fast solver for the resulting algebraic equations established. (See the historical discussions in [1] and [2].) The multilevel FOSLS methodology developed in [1] and [2] is a unified approach that carefully formulates a well-posed solution process that is faithful to the least-squares principles, and therefore avoids the disadvantages of earlier attempts.

This short note gives an overview of the basic FOSLS approach and its multigrid treatment by the multilevel projection method (PML, [5]) in the next two sections, respectively. The fourth section describes extensions to other applications of the methods developed in [1] and [2] for second-order elliptic-type problems. The final section contains a few brief remarks.

2. FOSLS

In this section, we illustrate the basic concepts behind FOSLS for formulating general equations as functional minimization problems, using a simple convection-diffusion equation as the model and making specific choices for the method. The original paper [1] should be consulted for treatment of more general elliptic-type equations and for variations on the choices we make here. Section 4 below includes brief comments on other types of equations.

Let Ω be a bounded domain in $\mathbb{R}^n$, $n = 2$ or 3, with piecewise smooth boundary $\partial\Omega$ having no re-entrant corners. Consider the convection-diffusion equation

$$\left.\begin{aligned} -\nabla\cdot(A\nabla p + \mathbf{b}p) = f &\quad \text{in } \Omega,\\ p = 0 &\quad \text{on } \Gamma_D,\\ \mathbf{n}\cdot A\nabla p = 0 &\quad \text{on } \Gamma_N \end{aligned}\right\} \tag{2.1}$$

for $f \in L^2(\Omega)$, where $A(x)$ is an $n\times n$ symmetric matrix of functions in $C^{1,1}(\Omega)$, $\mathbf{b} \in (H^1)^n$, $\partial\Omega = \Gamma_D \cup \Gamma_N$ is the boundary of Ω, and $\mathbf{n}$ is the outward unit normal on $\partial\Omega$. Assume that A is uniformly bounded and positive definite on $\overline{\Omega}$. Our assumptions guarantee H^2 regularity of (2.1).

Let $\|\cdot\|$ denote the $L^2(\Omega)$ norm, H^1 the usual Sobolev space of scalar functions with norm $\|\cdot\|_1$, and $(H^1)^n$ the corresponding space of vector functions whose norm we also denote by $\|\cdot\|_1$. Define

$$V = \{p \in H^1 : p = 0 \text{ on } \Gamma_D\}$$

and

$$\mathbf{W} = \left\{\mathbf{u} \in (H^1)^n : \mathbf{n}\cdot\mathbf{u} = 0 \text{ on } \Gamma_N,\ \mathbf{n}\times A^{-1}\mathbf{u} = \mathbf{0} \text{ on } \Gamma_D\right\}.$$

Introducing the flux vector $\mathbf{u} = A\nabla p$, note that it satisfies the curl-free constraint $\nabla\times A^{-1}\mathbf{u} = \mathbf{0}$ and, when p solves (2.1), the tangential boundary condition $\mathbf{n}\times A^{-1}\mathbf{u} = \mathbf{0}$ on Γ_D. We therefore arrive at the following system that is equivalent to (2.1):

$$\left.\begin{aligned} A\nabla p - \mathbf{u} = 0 &\quad \text{in } \Omega,\\ -\nabla\cdot(\mathbf{u} + \mathbf{b}p) = f &\quad \text{in } \Omega,\\ -\nabla\times A^{-1}\mathbf{u} = \mathbf{0} &\quad \text{in } \Omega,\\ (\mathbf{u}, p) \in \mathbf{W}\times V. & \end{aligned}\right\} \tag{2.2}$$

Consider the associated least-squares functional defined on $\mathbf{W}\times V$ for a given $f \in L^2(\mathbb{R})$ by

$$G(\mathbf{u}, p; f) = \|A\nabla p - \mathbf{u}\|^2 + \|\nabla\cdot(\mathbf{u} + \mathbf{b}p) + f\|^2 + \|\nabla\times A^{-1}\mathbf{u}\|^2. \tag{2.3}$$

For fixed $f \in L^2(\Omega)$, assuming that (2.1) has a unique solution p in H^2, it is easy to see that the unique minimum of $G(\mathbf{u}, p; f)$ over $(\mathbf{u}, p) \in \mathbf{W}\times V$ corresponds to the solution of (2.1) in the sense that p solves these equations and $\mathbf{u} = A\nabla p$. The main result of [1] and [2] was to show that $G(\mathbf{u}, p; f)$ is elliptic with respect to the $\mathbf{W} \times V$ norm:

$$0 \ll \frac{G(\mathbf{u}, p; 0)}{\|\mathbf{u}\|_1^2 + \|p\|_1^2} \ll \infty \tag{2.4}$$

for every $(\mathbf{u}, p) \in \mathbf{W} \times V$. This in turn implies (see [1]) that the usual Rayleigh-Ritz-type discretization of (2.3) by standard finite element subspaces of $\mathbf{W} \times V$ (e.g., piecewise bilinear elements on a nonstaggered grid) leads to stable and optimally accurate (e.g., $O(h^2)$) discrete problems with $O(h^{-2})$ condition numbers, for any bounded scale of convection. This sense of ellipticity also allows for an optimal multigrid solver of the discrete system, that is, the solver converges with a factor that is bounded uniformly in h (see [2]). This is the subject of the next section.

3. PML

The FOSLS functional minimization problem is to find $(\mathbf{u}, p) \in \mathbf{W} \times V$ such that

$$G(\mathbf{u}, p; f) = \min_{(\mathbf{v}, q) \in \mathbf{W} \times V} G(\mathbf{v}, q; f). \tag{3.1}$$

The Rayleigh-Ritz version of PML is ideally suited to this numerical task because it designs the relaxation and coarsening processes in concert with the Rayleigh-Ritz discretization of (3.1). To explain this formally and in brief (see [5] for more detail, and for other incarnations of PML, together with several applications), consider a *discrete* subspace $\mathbf{W}^h \times V^h \subset \mathbf{W} \times V$ indexed by the mesh parameter h on which (3.1) is to be solved. To implement the PML methodology, we need *relaxation* subspaces $\mathbf{W}_k^h \times V_k^h$, for $1 \leq k \leq n^h$, so that

$$\sum_{k=1}^{n} \mathbf{W}_k^h \times V_k^h \subset \mathbf{W} \times V,$$

and a *coarse* subspace $\mathbf{W}^{2h} \times V^{2h} \subset \mathbf{W}^h \times V^h$.

A two-level PML scheme applied to (3.1) is then defined by the following basic components.

- **discretization.** Find $(\mathbf{u}^h, p^h) \in \mathbf{W}^h \times V^h$ such that

$$G\left(\mathbf{u}^h, p^h; f\right) = \min_{(\mathbf{v}^h, q^h) \in \mathbf{W}^h \times V^h} G\left(\mathbf{v}^h, q^h; f\right).$$

- **relaxation.** Given the current approximation $(\mathbf{u}^h, p^h) \in \mathbf{W}^h \times V^h$ to the solution of the discretization, for each $k \in \{1, 2, \ldots, n^h\}$ in turn, find $(\mathbf{u}_k^h, p_k^h) \in \mathbf{W}_k^h \times V_k^h$ such that

$$G\left(\mathbf{u}^h + \mathbf{u}_k^h, p^h + p_k^h; f\right) = \min_{(\mathbf{v}_k^h, q_k^h) \in \mathbf{W}_k^h \times V_k^h} G\left(\mathbf{u}^h + \mathbf{v}_k^h, p^h + q_k^h; f\right),$$

  and make the replacement

$$\left(\mathbf{u}^h, p^h\right) \leftarrow \left(\mathbf{u}^h + \mathbf{u}_k^h, p^h + p_k^h\right).$$

- **coarsening.** Given the current approximation $(\mathbf{u}^h, p^h) \in \mathbf{W}^h \times V^h$, find $(\mathbf{u}^{2h}, p^{2h}) \in \mathbf{W}^{2h} \times V^{2h}$ such that

$$G\left(\mathbf{u}^h + \mathbf{u}^{2h}, p^h + p^{2h}; f\right) = \min_{(\mathbf{v}^{2h}, q^{2h}) \in \mathbf{W}^{2h} \times V^{2h}} G\left(\mathbf{u}^h + \mathbf{v}^{2h}, p^h + q^{2h}; f\right)$$

and make the replacement

$$\left(\mathbf{u}^h, p^h\right) \leftarrow \left(\mathbf{u}^h + \mathbf{u}^{2h}, p^h + p^{2h}\right).$$

This is just a formal definition of the primary ingredients of a PML scheme. Actual algorithms require determination of secondary components, like the representation (e.g., nodal values), the basic cycling scheme (e.g., a $V(1,1)$ cycle), and various implementation details. The primary choices are the various subspaces that are used in discretization, relaxation, and coarsening.

Typically, the relaxation subspaces are defined by the coordinates. For example, $\mathbf{W}_k^h \times V_k^h$ might be the space of functions $(\mathbf{v}, q)$ in $\mathbf{W}^h \times V^h$ that are zero at all grid points but the k^{th}, $1 \le k \le n^h$. The sum of such spaces spans $\mathbf{W}^h \times V^h$, and this choice means that the PML relaxation scheme is just nonlinear block Gauss-Seidel, with the k^{th} block corresponding to the $n+1$ variables at the k^{th} grid point. However, many other possibilities exist that lead, for example, to point, line, or plane Gauss-Seidel, or local grid refinement. (See [5] for some detail.) Similarly, the coarse subspace is typically that which is analogous to the scale h discrete subspace but defined on a $2h$ grid. Other possibilities exist here as well, including semi-coarsening and choices that lead to local grid refinement. Obvious modifications of the PML framework defined here also allow for parallel relaxation schemes like Jacobi's method and for nonconforming multigrid methods that violate the nested subspace property. The point here is that the PML methodology provides a framework within which to develop multigrid algorithms, and it focuses on simplifying the major design choices, namely, the discrete, relaxation, and coarse subspaces.

The multigrid scheme treated in [2] is a Rayleigh-Ritz PML method, simplified because the target problem, (2.1), is linear and essentially elliptic. Allowable discrete and coarse subspaces for that scheme include the usual bilinear finite-element spaces on a (quasi-)uniform grid of mesh spacing h, with the same space (except for boundary conditions) used for each of the $n+1$ variables. Allowable relaxation subspaces include the individual coordinate subspaces that lead to the usual point Gauss-Seidel relaxation. For such a scheme, the theory in [2] established optimal multigrid convergence under the usual regularity and smooth coefficient assumptions, again for any bounded scale of convection.

4. Other applications

Here we illustrate the potential for the least-squares approach by introducing possible forms that FOSLS might take for a few typical physical applications. These examples are considered in simplified forms (e.g., simple regions and two dimensions). In all but the first example, we ignore the (critical) issue of boundary conditions. The aim here is to suggest approaches, not establish a rigorous and detailed technique. The illustrations are therefore brief and very formal. Moreover, we will not comment on other work in this direction, although several others have pursued least-squares method for a wide variety of fluid flow equations. (See the discussions in [1] and [2].)

We have had successful early results for another version of the first application we describe, but each are nonetheless at the conjecture stage and must be studied carefully to determine viability. No doubt major alterations will be suggested as their characteristics are exposed. The point is that these are highly speculative, if not compelling, ideas.

4.1. Helmholtz. Consider an annulus Ω in $\mathbb{R}^2$ centered at the origin, with Γ_D as its inner boundary (unit circle) and Γ_N as its outer boundary (circle of radius $R > 1$). A simplified form of the Helmholtz equation is given in terms of a Helmholtz term $k \in \mathbb{R}$ and in polar coordinates by

$$\left.\begin{aligned} \Delta p + k^2 p = f \quad & \text{in } \Omega \\ p = 0 \quad & \text{on } \Gamma_D \\ p_r - ikp = 0 \quad & \text{on } \Gamma_N \end{aligned}\right\} \tag{4.1}$$

Note that this problem involves complex functions introduced by way of the purely imaginary boundary equation coefficient ik. (More common is to pose the Helmholtz problem with the Dirichlet boundary equation being inhomogeneous and the interior equation being homogeneous, but the form in (4.1) is more in line with our development and, in any case, the principles are the same.) One of the many possible forms for reducing (4.1) to a system is given by

$$\left.\begin{aligned} u - p_r + ikp = 0 \quad & \text{in } \Omega, \\ v - \tfrac{1}{r} p_\theta = 0 \quad & \text{in } \Omega, \\ u_r + \left(\tfrac{1}{r} + ik\right) u + \tfrac{1}{r} v_\theta + \tfrac{ik}{r} p + f = 0 \quad & \text{in } \Omega, \\ \tfrac{1}{r} u_\theta - v_r - \left(\tfrac{1}{r} - ik\right) v = 0 \quad & \text{in } \Omega, \\ p = v = 0 \quad & \text{on } \Gamma_D, \\ u = 0 \quad & \text{on } \Gamma_N. \end{aligned}\right\} \tag{4.2}$$

This is an interesting system for the Helmholtz problem. While fast solvers already exist for this simplified problem, we have in mind more general boundaries and variable Helmholtz terms, and our approach should extend fairly easily to such problems. Nevertheless, this simplified case should guide the development of our multigrid scheme, just as the simplified Poisson problem guides elliptic solver development. What seems compelling here is that each of the three diagonal terms in the least-squares functional for (4.2) (that is, the terms formed by setting two of the variables u, v, p to zero) exhibit H^1 ellipticity. Moreover, the introduced fluxes u and v are physically relevant, and the polar coordinate system naturally allows for concentration of grid points at the internal boundary Γ_D and expansion to the far field Γ_N. The central difficulty in developing an effective PML scheme will be to determine the correct coarse subspace: for $kh \ll R$, components that are slow to converge in relaxation must be smooth, so standard PML coarsening should work; when kh is comparable to R, however, certain oscillatory components cannot be smoothed in relaxation, so they must be well represented in the coarse subspace. The basic idea is to coarsen by choosing these subspaces to match closely the oscillatory components locally. While these can be subtle, early results with a Cartesian coordinate version of (4.2) were encouraging and suggest that such a strategy will lead to an optimal PML solver for what seems to be a more amenable polar coordinate formulation.

4.2. Cauchy-Riemann. These equations are already in an appropriate first-order form:

$$\left.\begin{array}{ll} u_x - v_y & = f \\ u_y + v_x & = g \end{array}\right\} \tag{4.3}$$

The least-squares functional

$$G(u, v; f, g) = \|u_x - v_y - f\|^2 + \|u_y + v_x - g\|^2$$

has a minimum at the solution of the gradient or Euler equations

$$\begin{pmatrix} \frac{\partial}{\partial x} & -\frac{\partial}{\partial y} \\ \frac{\partial}{\partial y} & \frac{\partial}{\partial x} \end{pmatrix}^* \left[\begin{pmatrix} \frac{\partial}{\partial x} & -\frac{\partial}{\partial y} \\ \frac{\partial}{\partial y} & \frac{\partial}{\partial x} \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} - \begin{pmatrix} f \\ g \end{pmatrix} \right] = 0,$$

where * denotes operator adjoint. Formally (that is, with sufficient smoothness and under the boundary conditions that make the formal and actual adjoints the same; see [1]), we are led to the elliptic system

$$\begin{pmatrix} Delta & 0 \\ 0 & -\Delta \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} -f_x + g_y \\ -f_y - g_x \end{pmatrix}.$$

This suggests that the least-squares approach generally leads to a well-posed elliptic problem, which our numerical experiments confirm.

4.3. Stokes flow. The two-dimensional Stokes equations can be written in the form

$$\begin{pmatrix} \Delta & 0 & \frac{\partial}{\partial x} \\ 0 & \Delta & \frac{\partial}{\partial y} \\ -\frac{\partial}{\partial x} & -\frac{\partial}{\partial y} & 0 \end{pmatrix} \begin{pmatrix} u \\ v \\ p \end{pmatrix} = \begin{pmatrix} f \\ g \\ h \end{pmatrix}. \tag{4.4}$$

Introducing the flux vectors $\mathbf{u} = \nabla u$ and $\mathbf{v} = \nabla v$ and defining M=(I 0) and N=(0 I), one of the many possible associated first-order systems (here we include the curl-free condition, which we write as $\nabla \times \mathbf{w} = \mathbf{0}$) is written as

$$\begin{pmatrix} I & -\nabla & 0 & 0 & 0 \\ \nabla\times & 0 & 0 & 0 & 0 \\ \nabla\cdot & 0 & 0 & 0 & \frac{\partial}{\partial x} \\ 0 & 0 & I & -\nabla & 0 \\ 0 & 0 & \nabla\times & 0 & 0 \\ 0 & 0 & \nabla\cdot & 0 & \frac{\partial}{\partial y} \\ M & 0 & N & 0 & 0 \end{pmatrix} \begin{pmatrix} \mathbf{u} \\ u \\ \mathbf{v} \\ v \\ p \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ f \\ 0 \\ 0 \\ g \\ h \end{pmatrix}.$$

The formal gradient equation for the least-squares functional based on this system is

$$\begin{pmatrix} D_1 & -\nabla & K & 0 & -\frac{\partial}{\partial x}\nabla \\ \nabla\cdot & -\Delta & 0 & 0 & 0 \\ K^* & 0 & D_2 & -\nabla & -\frac{\partial}{\partial y}\nabla \\ 0 & 0 & \nabla\cdot & -\Delta & 0 \\ -\frac{\partial}{\partial x}\nabla\cdot & 0 & -\frac{\partial}{\partial y}\nabla\cdot & 0 & -\Delta \end{pmatrix} \begin{pmatrix} \mathbf{u} \\ u \\ \mathbf{v} \\ v \\ p \end{pmatrix} = \begin{pmatrix} -\nabla f - h \\ 0 \\ -\nabla g - h \\ 0 \\ -f_x - g_y \end{pmatrix},$$

where

$$D_1 \equiv \begin{pmatrix} 2I-\Delta & 0 \\ 0 & I-\Delta \end{pmatrix},$$

$$D_2 \equiv \begin{pmatrix} I-\Delta & 0 \\ 0 & 2I-\Delta \end{pmatrix},$$

and

$$K \equiv \begin{pmatrix} 0 & I \\ 0 & 0 \end{pmatrix}.$$

Within the individual variable groups $\binom{\mathbf{u}}{u}$, $\binom{\mathbf{v}}{v}$, and p, this is a well behaved system, analogous to that for (4.1). However, there is a strong connection between these groups as is evident in the above off-diagonal terms $-\frac{\partial}{\partial x}\nabla\cdot$ and $-\frac{\partial}{\partial y}\nabla\cdot$ (and their adjoints). This is just a reflection of the original strong connection between pressure and the velocities in (4.4). Thus, we cannot expect the usual H^1-type norm equivalence that would allow for standard multigrid methods, but a relatively simple modification to relaxation should be able to account for this inter-variable coupling.

4.4. Hyperbolic equations. It is speculative but interesting to consider least-squares approaches to purely hyperbolic equations. To this end, consider the simple system

$$p_{xx} - p_{yy} = f. \tag{4.5}$$

Letting $\mathbf{u} = \binom{u}{v} = \nabla p$ and incorporating a curl-free constraint, we obtain the first-order system

$$\begin{pmatrix} I & 0 & -\frac{\partial}{\partial x} \\ 0 & I & -\frac{\partial}{\partial y} \\ \frac{\partial}{\partial x} & -\frac{\partial}{\partial y} & 0 \\ \frac{\partial}{\partial y} & -\frac{\partial}{\partial x} & 0 \end{pmatrix} \begin{pmatrix} u \\ v \\ p \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ f \\ 0 \end{pmatrix}.$$

The formal gradient equation for the associated least-squares functional is

$$\begin{pmatrix} I-\Delta & 2\frac{\partial^2}{\partial x \partial y} & -\frac{\partial}{\partial x} \\ 2\frac{\partial^2}{\partial x \partial y} & I-\Delta & -\frac{\partial}{\partial y} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & -\Delta \end{pmatrix} \begin{pmatrix} u \\ v \\ p \end{pmatrix} = \begin{pmatrix} -f_x \\ -f_y \\ 0 \end{pmatrix}.$$

Note that the matrix for this system can be written as

$$\begin{pmatrix} -\Delta & 2\frac{\partial^2}{\partial x \partial y} & 0 \\ 2\frac{\partial^2}{\partial x \partial y} & -\Delta & 0 \\ 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} I & 0 \\ 0 & I \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} \end{pmatrix} \begin{pmatrix} I & 0 - \frac{\partial}{\partial x} \\ 0 & I - \frac{\partial}{\partial y} \end{pmatrix}.$$

This suggests that any lack of H^1-type norm equivalence would come from the coupling between the velocities, which is expressed by writing the leading 2×2 block in the first term above as follows:

$$\begin{pmatrix} -\Delta & 2\frac{\partial^2}{\partial x \partial y} \\ 2\frac{\partial^2}{\partial x \partial y} & -\Delta \end{pmatrix} = \frac{1}{2} \begin{pmatrix} I & I \\ I & -I \end{pmatrix} \begin{pmatrix} -\Delta + \frac{2\partial^2}{\partial x \partial y} & 0 \\ 0 & -\Delta - \frac{2\partial^2}{\partial x \partial y} \end{pmatrix} \begin{pmatrix} I & I \\ I & -I \end{pmatrix}. \tag{4.6}$$

The u and v are clearly coupled according to this transformation. This coupling can be treated simply by redefining u and v via the transformation in (4.6), or more generally by accounting for the coupling in relaxation by simultaneous updates of u and v at each grid point. This does not eliminate the hyperbolic character of the equations, however, as the nonnegative definite diagonal terms in (4.6) illustrate:

$$-\left(\Delta \pm 2\frac{\partial^2}{\partial x \partial y}\right) = \left(\frac{\partial}{\partial x} \pm \frac{\partial}{\partial y}\right)^* \left(\frac{\partial}{\partial x} \pm \frac{\partial}{\partial y}\right).$$

Thus, the fluxes have inherited the hyperbolic character of the original scalar variable, as we should expect, but the equations are now in a relatively well-posed self-adjoint form. The point that this example illustrates is that the basic nature of the equations are more or less preserved, but the result is a system imbued with a minimization principle, as intended.

5. Concluding remarks

The basic purpose of this brief note is to explore, in loose terms, a relatively new methodology for systematic treatment of general equations of fluid dynamics in a functional minimization framework. Current work is proceeding from the encouraging results for simple elliptic equations into more significant applications, including Helmholtz, Stokes, and convection-diffusion equations. If early numerical results and theoretical observations hold out, this may ultimately provide a unified treatment of a large class of fluid flow and other equations.

References

1. Cai,Z., Lazarov,R., Manteuffel,T.A., and McCormick,S.F. *First-order system least squares for partial differential equations: discretization*, manuscript, 1993.
2. Cai,Z., Manteuffel,T.A., and McCormick,S.F. *First-order system least squares for partial differential equations: multigrid solvers*, manuscript, 1993.
3. Kriess,H.O., Manteuffel,T.A., Schwartz,B., Wendroff,B., and White,A.B. *Supraconvergence schemes on irregular grids.* **Math.Comp.**, vol.47, no.176, 1984, pp.537-554.
4. McCormick,S. (ed.), *Multigrid Methods*, SIAM Frontiers Series in Applied Mathematics, vol.3, SIAM, Philadelphia, 282 pages.
5. McCormick,S. *Multilevel Projection Methods for Partial Differential Equations*, CBMS-NSF Regional Conference Series in Applied Mathematics, SIAM, Philadelphia, 112 pages.