DEPARTMENT OF MATHEMATICS
UNIVERSITY OF NIJMEGEN The Netherlands

A POSTERIORI ERROR ESTIMATES IN L_2-NORM FOR THE LEAST SQUARES FINITE ELEMENT METHOD APPLIED TO A FIRST ORDER SYSTEM OF DIFFERENTIAL EQUATIONS

O. Axelsson and I.E. Kaporin

Report No. 9945 (December 1999)

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF NIJMEGEN
Toernooiveld
6525 ED Nijmegen
The Netherlands

A Posteriori error estimates in L_2-norm for the Least Squares Finite Element method applied to a First Order System of differential equations

O. Axelsson* I.E. Kaporin†

Abstract

A class of boundary value problems for the first order systems is considered and a least squares finite element method for their solution is analysed. The solution method is based on the adaptation of a nonuniform grid aimed at the reduction of the L_2 norm of the residual.

The fundamental aspect in the solution of linear boundary value problems is the error control, whereby the error estimate must include both discretization and algebraic solver errors. In the present paper it is shown how this can be done when the equations are written as a system of first order partial differential equations, discretized by a least squares finite element (FE) method, and solved approximately, e.g. by some iterative method.

An *a posteriori* estimate of the L_2-norm of the error based on the L_2-norm of the residual is derived under rather mild assumptions on the smoothness of the solution. This error estimate is essential for the understanding of such key points as (a) correct termination of the iterative algebraic solver for the current FE space; and (b) proper adaptive mesh refinement needed to construct the next FE space.

We mainly consider the case of a linear differential problem and then use the obtained result for the analysis of the nonlinear case under certain natural assumptions on the nonlinearity properties.

The obtained estimates and the related solution method are illustrated by numerical results for 2D Navier-Stokes equations with a known analytical solution.

*Faculty of Mathematics and Informatics, University of Nijmegen, Nijmegen, The Netherlands

†Supercomputing and Massively Parallel Applications Research Center, Computing Center of Russian Academy of Science, Moscow, Russia

The research for the first author was supported by the NWO grant 047003017, and for the second author by the grants from the Netherlands Organisation of Scientific Research, NWO B61 - 298 and NWO 047003017, and by the INTAS-RFBR 0095-98 grant

1 Introduction

When solving linear boundary value problems, two basic aspects must be considered:
a) solution of the linear algebraic systems which arise after discretization;
b) error control.

The present paper concerns mainly the second aspect, i.e. the error control.
There are two major error sources in the solution:
(i) the error due to stopping the iterations at some stage of the solution of the discretized problem, for instance, based on the norm of the residual;
(ii) the error due to discretizing the problem by the finite element method.

It will be shown how both types of errors can be estimated and controlled in the solution process based on a certain unified framework.

In order to simplify the analysis, we assume that the differential equation has been rewritten as a first order system of differential equations, so that the residual function obtained using a standard FE space can be readily evaluated. This makes it possible to let the reduction of the (weighted) Euclidean norm of the residual function be the main objective at each stage of the algorithm. However, it is shown that certain more subtle issues are fundamental as well for getting a small solution error. Such issues include the choice of a proper first-order reformulation of a given differential problem and the use of a proper stopping criterion for the discrete solver over the current FE space.

Recent years have shown a renewed strong interest in the least-squares finite element methods for systems of first order partial differential equations, see [5], [6], [7] [11], and [12], for instance. These papers considered mainly linear problems and were concerned with error estimates for various components of the systems. Since any (system of) partial differential equations can be rewritten as a system of first order equations, such least squares methods are generally applicable. Their disadvantage is that many different unknown (physical) variables are introduced, such as all components of the gradient vectors. Moreover, the optimal order of convergence for the finite element error may not be obtained for all variables concerned. Besides, since a first order reformulation is never unique, there exists the nontrivial problem of finding a correct setting of a first order problem which should not involve too many additional unknown functions and equations, but, at the same time, should not suffer from bad regularity and nonlinearity properties.

On the other hand, as we shall see, the least-squares methods have many advantages which may outweigh these disadvantages.
The main advantages are:
(i) The choice of finite elements is much simplified. For instance, there is no need to find stable finite element pairs for (divergence) constrained problems (for the pressure and velocity components, for instance). It is known that finding such stable pairs satisfying the Ladyzhenskaya-Babuska-Brezzi condition is difficult, in particular for problems in three space dimensions. On the other hand, if one accepts that not all variables are always approximated with the best possible order of accuracy, in the least squares methods one can use the same finite element space for all variables involved. Hence, one may also use a single finite element mesh.

(ii) The finite element matrices can be stored in element-by-element form. If one uses the same space for all variables, there is only one such space to store. This greatly simplifies the data structure and makes computation on (massively) parallel computers much easier.
(iii) Even if one uses standard finite element spaces in the function space H^1, the residuals are directly computable when the least squares method is applied to first order operator problems. This simplifies *a posteriori* error analysis and adaptive refinement of the underlying finite element mesh to obtain the optimal convergence order of the errors, and not only of the residuals.

The present paper is mainly concerned with the general error estimates that can be used as an efficient global error indicator, and not with "asymptotically optimum" estimates referring to the maximum mesh stepsize h and some assumptions on the smoothness of individual vector components of the unknown solution. It is a follow-up of a previous paper [2] by the authors. Our main result is that the L_2-norm of the error decreases asymptotically faster that the L_2-norm of the residual as the FE space is refined whenever a rather standard Regularity property holds true for the first order reformulation. In the nonlinear case, a proper Nonlinearity condition should also be assumed. The effect of refinement of the FE space is described by a certain Approximation condition involving no derivatives of the degree higher than one. A similar result was shown in [14] under some special least-squares bilinear form regularity assumptions.

The remainder of the paper is organized as follows. In Section 2 we formulate our main result for the case of the linear problem, and in Section 3 its proof is given. In Section 4 we present a generalization of our results to the nonlinear case. In Section 5 we present a comparative numerical study of the residual and error convergence for different first order reformulations of 2D Navier-Stokes equations with a nontrivial analytical solution. In Section 6 some conclusions are drawn.

Standard notations are used; for instance $\|\cdot\|$ denotes the norm in $L_2(\Omega)$, i.e. in the space of square integrable functions. We will also use the notation

$$\cos[f,g] = \frac{|(f,g)|}{\|f\|\|g\|} \tag{1.1}$$

for the cosine of the acute angle between f and g, and

$$\cos[f,AW] = \max_{w\in W} \frac{(f,Aw)}{\|f\|\|Aw\|}, \tag{1.2}$$

where W is a finite-dimensional linear subspace of functions, and A is a linear operator which is nonsingular on this subspace. Obviously, these cosines lie between 0 and 1, and $\cos[f,g] = 1$ is equivalent to $\|f - \alpha g\| = 0$ for some nonzero scalar α while $\cos[f,AW] = 1$ yields $f \in AW$.

2 Problem setting, basic assumptions, and the main result

Let us consider a linear problem

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

$$\mathcal{M}u = g \quad \text{on } \partial\Omega, \tag{2.2}$$

where $u = (u_1, \ldots, u_m)^T \in (H^1(\Omega))^m$ is a weakly differentiable unknown vector function, $f = (f_1, \ldots, f_n)^T \in (L_2(\Omega))^n$, $n \geq m$, is the right hand side, and $\mathcal{L}$ is a first order linear differential operator, i.e., involving no derivatives of the order higher than one and defined on a domain Ω in $\mathbb{R}^d$, $1 \leq d \leq 3$. The boundary condition operator $\mathcal{M}$ is given by a $l \times n$ matrix, $l \leq n$, with elements typically given by piecewise constant functions defined on $\partial\Omega$. Hence $\mathcal{L} : V \to V'$, where $V = (H^1(\Omega))^m$, $V' = (L_2(\Omega))^n$. Let $(\cdot, \cdot)$ be the scalar product in V'. We shall show that the discretization error in L_2-norm can be bounded by the L_2-norm of the residual multiplied by a certain quantity (typically of the order $O(h)$).

Further, let V_h be a finite element space, where we choose u_h, an approximation to the exact solution u. We will introduce the error function as

$$e = u - u_h \in \mathring{V} \tag{2.3}$$

and define the residual as

$$r = \mathcal{L}u_h - f \equiv -\mathcal{L}e. \tag{2.4}$$

Here the subspace $\mathring{V}$ contains typically functions $v \in V$ satisfying the homogeneous boundary conditions $\mathcal{M}v = 0$ on $\partial\Omega$. The condition $u - u_h \in \mathring{V}$ greatly simplifies the analysis; hence, in the following we will restrict our considerations to polygonal regions Ω (in a simplest case) and the boundary conditions given by the (piecewise) linear or quadratic functions (according to the order of the FE bases used).

Let us now make the following two basic assumptions. They are similar to the standard ones made in the theory of finite element methods (see [10], for instance) and involve regularity-like and approximation-like assumptions.

Regularity assumption of the problem

In our case, the regularity assumption simply reads as follows:

There exists a constant $\sigma > 0$ such that for any $v \in \mathring{V}$ the inequality

$$\|v\| \leq \sigma \|\mathcal{L}v\| \tag{2.5}$$

holds.

This is a condition widely used in the LSFE analysis. In practice one should use only first order reformulations that possess this property. This explains why some additional (redundant) equations are often added to obtain (2.1)-(2.2) with $n > m$,

see, e.g., [12], [5]. Note that condition (2.5) essentially coincides (up to the treatment of boundary conditions) with the "Assumption H1" used in [16] to characterize regularity properties of so called Friedrichs type first order systems [9].

Approximation assumption on the F.E. space

The approximation assumption is related to the finite element space V_h used and the type of boundary conditions (2.2) imposed.

We relate this assumption to the solution $w \in \mathring{V}$ of the following auxiliary problem:

$$(\mathcal{L}w, \mathcal{L}v) = (e, v) \quad \forall v \in \mathring{V}. \tag{2.6}$$

and formulate it as follows:

$$\sin[\mathcal{L}w, \mathcal{L}\mathring{V}_h] \equiv \sqrt{1 - \cos^2[\mathcal{L}w, \mathcal{L}\mathring{V}_h]} \to 0 \tag{2.7}$$

where

$$\mathring{V}_h = V_h \cap \mathring{V}$$

as the subspace V_h is refined. Further we will give some upper bounds for this quantity which can be useful for the understanding of its behavior.

Remark 2.1 Typically, this "sin" is of the order $O(h^\alpha)$, where h is the maximum mesh stepsize and $\alpha > 0$ depends on the regularity of the auxiliary problem (cf. Remark 3.3 below) and on the space V_h used.

Inexactness assumption on the F.E. solution

Finally, we will make an additional assumption on the closeness of the finite element approximation $u_h \in V_h$ to the exact solution of the discrete LSFE problem:

For the residual r defined by (2.4), the following condition holds:

$$\cos[r, \mathcal{L}\mathring{V}_h] \ll 1. \tag{2.8}$$

Recalling (2.4), the condition $\cos[r, \mathcal{L}\mathring{V}_h] = 0$ obviously implies that standard discrete LSFE problem is solved exactly:

$$(\mathcal{L}u_h - f, \mathcal{L}v_h) = 0, \qquad \forall v_h \in \mathring{V}_h. \tag{2.9}$$

In the general case, this "cos" takes into account both numerical integration and/or discrete numerical solution errors.

The main result

Under the assumptions given above, the following error estimate holds:

$$\|e\| \leq \left(\sin[\mathcal{L}w, \mathcal{L}\mathring{V}_h] + \cos[r, \mathcal{L}\mathring{V}_h]\right) \sigma\|r\| \tag{2.10}$$

provided that the solution $w \in V$ of the auxiliary problem (2.6) given above, exists.

Remark 2.2 It should be stressed that the above assumptions impose no additional requirements on the regularity of the functions involved. Essentially, we are working with continuous functions $w, e, v_h, \ldots$ which are weakly differentiable and satisfy conditions of the type $w \in L_2(\Omega)$ and $\mathcal{L}w \in L_2(\Omega)$. However, in order to prove the validity of the first two assumptions, a somewhat more restrictive condition $w, e, v_h, \ldots \in W_2^1(\Omega)$ is usually imposed.

Remark 2.3 The following trivial estimate can readily be obtained from (2.4) and (2.5) with $v = e$:

$$\|e\| \leq \sigma\|r\|.$$

Adaptive FE space refinement startegies which are actually based on this estimate were presented, e.g., in [4]. The drawback of its use is that it does not show that under certain conditions the error may converge to zero faster than the residual as the FE space is refined. The additional assumptions underlying our error estimate may comprise a basis of a solution technique for obtaining an error norm even smaller than that of the residual.

In sufficiently regular cases, the new estimate demonstrates the order

$$\|e\| = O(h\|r\|),$$

where h is the maximum stepsize of the FE grid (cf. Remark 3.3. below).

3 The LSFE error estimate in L_2-norm

We now show how the above assumptions can be used in order to obtain the required estimate (2.10). Let u_h be an approximate least squares FE solution and $e \in \mathring{V}$ be the corresponding error as defined in (2.3).

The auxiliary problem

Let $w \in \mathring{V}$ be the solution of the problem (2.6).

Using the regularity assumption (2.5), setting $v = w$ in (2.6), and applying the Cauchy-Schwarz inequality, one finds that the estimates

$$\|w\| \leq \sigma^2\|e\| \tag{3.1}$$

and

$$\|\mathcal{L}w\| \leq \sigma\|e\| \tag{3.2}$$

hold. Indeed, one has by (2.5)

$$\sigma^{-2}\|w\|^2 \leq \|\mathcal{L}w\|^2 = (e,w) \leq \|e\|\|w\| \leq \sigma\|e\|\|\mathcal{L}w\|,$$

and both (3.1) and (3.2) readily follow.

We now prove the following two auxiliary propositions. To simplify the notations, we will further denote

$$\delta = \cos[r, \mathcal{L}\mathring{V}_h].$$

Lemma 3.1 *For any $v_h \in \mathring{V}_h$ the following error estimate holds,*

$$\|e\|^2 \leq (\|\mathcal{L}(w - v_h)\| + \delta\|\mathcal{L}v_h\|)\,\|r\|, \tag{3.3}$$

where w is the solution of the auxiliary problem (2.6).

Proof. Taking any $v_h \in \mathring{V}_h$, setting $v = e$ in (2.6) and using inexactness condition (2.8) one obtains

$$\begin{aligned}
\|e\|^2 &= (\mathcal{L}w, \mathcal{L}e) \\
&\equiv (\mathcal{L}(w - v_h), \mathcal{L}e) + (\mathcal{L}v_h, \mathcal{L}e) \\
&= -(\mathcal{L}(w - v_h), r) - (\mathcal{L}v_h, r) \\
&\leq \|\mathcal{L}(w - v_h)\|\|r\| + \delta\|\mathcal{L}v_h\|\|r\|,
\end{aligned}$$

which is (3.3).
Q.E.D.

Lemma 3.2 *Let w be the solution of the auxiliary problem (2.6). Then the following error estimate holds:*

$$\|e\|^2 \leq \left(\sin[\mathcal{L}w, \mathcal{L}\mathring{V}_h] + \delta\cos[\mathcal{L}w, \mathcal{L}\mathring{V}_h]\right)\|\mathcal{L}w\|\|r\|. \tag{3.4}$$

Proof. Let the maximizer of the cosine be the function

$$y_h = \arg\max_{v \in \mathring{V}_h} \cos[\mathcal{L}w, \mathcal{L}v]$$

normalized by the condition

$$\|\mathcal{L}y_h\| = \|\mathcal{L}w\|.$$

To simplify notations, let us also denote the value of this maximum cosine by

$$\theta = \cos[\mathcal{L}w, \mathcal{L}y_h].$$

We have then, setting $v_h = \theta y_h$ and inserting it into the estimate of Lemma 3.1,

$$\begin{aligned}
\|e\|^2/\|r\| &\leq \|\mathcal{L}(w - v_h)\| + \delta\|\mathcal{L}v_h\| \\
&= \sqrt{\|\mathcal{L}w\|^2 - 2\theta(\mathcal{L}w, \mathcal{L}y_h) + \theta^2\|\mathcal{L}y_h\|^2} + \delta\theta\|\mathcal{L}y_h\| \\
&= \sqrt{\|\mathcal{L}w\|^2 - 2\theta^2\|\mathcal{L}w\|\|\mathcal{L}y_h\| + \theta^2\|\mathcal{L}y_h\|^2} + \delta\theta\|\mathcal{L}y_h\| \\
&= \left(\sqrt{1 - \theta^2} + \delta\theta\right)\|\mathcal{L}w\|
\end{aligned}$$

Q.E.D.

Using now (3.2) and recalling that $\theta \leq 1$ one obtains the required error estimate (2.10)

Remark 3.1 Since typically $\sin[\mathcal{L}w, \mathcal{L}\mathring{V}_h]$ tends to zero as $h \to 0$, the error estimate shows that the discretization error in L_2-norm converges faster than the L_2-norm of the residual $f - \mathcal{L}u_h$. The latter can be related to an H^1-norm of the error and (2.10) is an example of what is sometimes called "L_2 lifting". The corresponding estimate for second order problems is also referred to as the Aubin-Nitsche trick, see [10].

Remark 3.2 It can be readily shown that

$$\sin[\mathcal{L}w, \mathcal{L}\mathring{V}_h] = \min_{v \in \mathring{V}_h} \frac{\|\mathcal{L}(w - v)\|}{\|\mathcal{L}w\|}$$

and the minimizer

$$w_h = \arg\min_{v \in \mathring{V}_h} \|\mathcal{L}(w - v)\|$$

coincides with the solution of the *discretized* auxiliary problem: find $w_h \in \mathring{V}_h$ such that

$$(\mathcal{L}w_h, \mathcal{L}v) = (e, v) \quad \forall v \in \mathring{V}_h. \tag{3.5}$$

This easily follows from the main property of w_h presented by

$$(\mathcal{L}(w - w_h), \mathcal{L}v) = 0 \quad \forall v \in \mathring{V}_h.$$

Thus, the quantity $\sin[\mathcal{L}w, \mathcal{L}\mathring{V}_h] = \|\mathcal{L}(w - w_h)\|/\|\mathcal{L}w\|$ is nothing but the relative solution error in $\mathcal{L}$-norm obtained when applying the FE method to the auxiliary problem (2.6).

Remark 3.3 The techniques of [14] can easily be applied to estimate the error/residual norm ratio, and gives

$$\frac{\|e\|}{\|r\|} = O(h^{\min(\alpha-1,k)} + \cos[r, \mathcal{L}\mathring{V}_h]),$$

where h is the maximum stepsize of the FE mesh, k is the polynomial degree of the FE space used, and $\alpha \geq 1$ determines the regularity of the least-squares bilinear form $(\mathcal{L}w, \mathcal{L}v)$ in the sense that for any $e \in L_2$ the solution of the auxiliary problem (2.6) lies in H^α and satisfies $\|w\|_\alpha \leq Const\|e\|$.

4 A generalization of the LSFE error L_2-norm estimate to the nonlinear case

Suppose now that we seek a solution $u \in V$ of the nonlinear operator equation

$$Fu = 0. \tag{4.1}$$

Since F is supposed to be a nonlinear first order differential operator, and the inclusion $Fu \in L_2$ will further be used, we will assume V to be an appropriate subset of H^1. Let $u_h \in V_h \subset V$ be a FE approximation to this solution that satisfies a certain *Nonlinear Inexactness condition* to be specified later.

In order to estimate the solution error $u - u_h$ let us assume that the nonlinear mapping F is Frechet differentiable at u_h, so that

$$Fv = Fu_h + F'(u_h)(v - u_h) + Q(u_h, v) \tag{4.2}$$

and the nonlinear term Q satisfies the following *Nonlinearity condition*: for a certain set of functions v containing the solution u (e.g., the level set $\{v : \|Fv\| \le \|Fu_h\|\}$) the inequality

$$\|Q(u_h, v)\| \le c_0 \|Fv - Fu_h\| \|v - u_h\| \tag{4.3}$$

holds. Note that we assume here that the space V is chosen in such a way that both F and $F'(u_h)$ map V into certain subsets of L_2.

In order to relate this setting to the result of Section 2, let us denote

$$e = u - u_h, \quad \mathcal{L} = F'(u_h), \quad r = -\mathcal{L}e, \tag{4.4}$$

which correspond to (2.3) and (2.4). Therefore, taking $v = u$ in (4.2) and using (4.1), one has

$$0 = Fu_h - r + Q(u_h, u). \tag{4.5}$$

A natural characterization of the error resulting from the finite-dimensional minimization of $\|Fu_h\|$ over the FE space V_h is given by the following *Nonlinear Inexactness condition*:

$$\delta_0 \equiv \cos[Fu_h, \mathcal{L}\mathring{V}_h] \ll 1.$$

Let us now suppose that for the operator $\mathcal{L}$ so defined, the Regularity assumption (2.5) and FE Approximation assumption (2.6)-(2.7) hold true. Note that by (1.2) our main estimate (2.10) can be rewritten as follows:

$$\|e\| \le \sin[\mathcal{L}w, \mathcal{L}\mathring{V}_h]\sigma\|r\| + \sigma \max_{v_h \in \mathring{V}_h} \frac{|(r, \mathcal{L}v_h)|}{\|\mathcal{L}v_h\|}. \tag{4.6}$$

The last term of (4.7) can be readily estimated using (4.5) and (4.6):

$$|(r, \mathcal{L}v_h)| = |(Fu_h + Q(u_h, u), \mathcal{L}v_h)| \le \delta_0 \|Fu_h\| \|\mathcal{L}v_h\| + \|Q(u_h, u)\| \|\mathcal{L}v_h\|.$$

Taking into account the nonlinearity condition (4.3) written for $v = u$,

$$\|Q(u_h, u)\| \le c_0 \|Fu_h\| \|e\|, \tag{4.7}$$

one obtains

$$|(r, \mathcal{L}v_h)| \le \delta_0 \|Fu_h\| \|\mathcal{L}v_h\| + c_0 \|Fu_h\| \|e\| \|\mathcal{L}v_h\|.$$

Hence we get

$$\max_{v_h \in \mathring{V}_h} \frac{|(r, \mathcal{L}v_h)|}{\|\mathcal{L}v_h\|} \le (\delta_0 + c_0\|e\|)\|Fu_h\|. \tag{4.8}$$

It remains to note that (4.5) and (4.7) yield

$$\|r\| = \|Fu_h + Q(u_h, u)\| \le \|Fu_h\| + c_0\|Fu_h\|\|e\| = (1 + c_0\|e\|)\|Fu_h\|. \tag{4.9}$$

Substituting (4.8) and (4.9) into (4.6), one has

$$\|e\| \le \sin[\mathcal{L}w, \mathcal{L}\mathring{V}_h]\sigma\Big(1 + c_0\|e\|\Big)\|Fu_h\| + \sigma\Big(\delta_0 + c_0\|e\|\Big)\|Fu_h\|.$$

Using the latter inequality, one can easily obtain the following result: if the Euclidean norm of the nonlinear residual Fu_h is sufficiently small,

$$\|Fu_h\| < \frac{1}{2c_0\sigma}$$

then the error estimate

$$\|e\| \le \frac{(\sin[\mathcal{L}w, \mathcal{L}\mathring{V}_h] + \cos[Fu_h, \mathcal{L}\mathring{V}_h])\sigma\|Fu_h\|}{1 - 2c_0\sigma\|Fu_h\|}$$

follows.

Hence, we obtain almost the same error estimate as in the linear case, with the only additional requirement for the norm of the nonlinear residual to be sufficiently small. The larger the regularity constant σ and nonlinearity constant c_0 are, the more restrictive is the condition imposed on the residual error norm. However, using the multilevel Gauss-Newton type procedure [3] for the minimization of $\|Fu_h\|$ one can eventually reduce the residual to a level sufficient for this error estimate to be applicable.

Remark 4.1 In the nonlinear case, taking properly account of the nonzero value of $\cos[Fu_h, \mathcal{L}\mathring{V}_h]$ seems to be essential since this term may never be negligible in practice. Indeed, for each FE subspace we actually have a large-scale nonlinear least squares problem with nonzero minimum to solve, and the numerical solution of it with high precision would often require an excessively large computational cost.

5 Numerical Experiments

In this section we compare two first order reformulations of the 2-D Navier Stokes equations with Dirichlet boundary conditions for velocities. The first one, the velocity-vorticity-Bernoulli pressure formulation, is in general incompatible with the original boundary conditions and thus does not even satisfy the Regularity condition. The second one is a certain velocity-pressure-flux type formulation (which is a simplified and modified version of the one used e.g., in [5]) possessing considerably better regularity properties.

5.1 A nontrivial analytical solution of Navier-Stokes equations

.

In our tests we used a family of solutions of the 2-D Navier-Stokes equations for incompressible viscous flow with zero driving forces

$$\begin{aligned} &p_x + \nu(-u_{xx} - u_{yy}) + uu_x + vu_y = 0, \\ &p_y + \nu(-v_{xx} - v_{yy}) + uv_x + vv_y = 0, \\ &u_x + v_y = 0 \end{aligned}$$

in $S = (0,1)^2$, with u, v given on ∂S, p fixed at the origin, $p(0,0) = 0$. A special case of analytical solution (with no regard to the boundary conditions) is given by

$$\begin{aligned} u = U(x,y) &= \tfrac{-y}{2}\left(1 + \log\tfrac{x^2+y^2}{\nu}\right), \\ v = V(x,y) &= \tfrac{x}{2}\left(1 + \log\tfrac{x^2+y^2}{\nu}\right), \\ p = P(x,y) &= -2\nu \arctan\tfrac{x}{y} + \tfrac{x^2+y^2}{8}\left(1 + \log^2\tfrac{x^2+y^2}{\nu}\right). \end{aligned}$$

Using these functions, one can easily obtain a problem with analytical solution depending on an arbitrary point (x_0, y_0) lying outside S as follows:

$$\begin{aligned} u(x,y) &= U(x - x_0, y - y_0), \\ v(x,y) &= V(x - x_0, y - y_0), \\ p(x,y) &= P(x - x_0, y - y_0) - P(-x_0, -y_0), \end{aligned}$$

with the boundary conditions for u and v defined via the traces of these functions on ∂S.

This solution demonstrates a kind of singularity at the point (x_0, y_0), so this point was chosen outside S. We used the values $\nu = 1$ and $x_0 = -y_0 = 0.25$ for our numerical experiments (this was done in order to illustrate the essentially different behavior of the two reformulations even for a rather smooth analytical solution). Unlike the test solution used in [3] this one clearly demonstrates the troubles arising from the incompatibility of the velocity-vorticity-Bernoulli pressure formulation with the velocity Dirichlet boundary conditions. Namely, when using this incompatible first order reformulation, the poor performance of the Least Squares-FE solvers for this test problem closely resembles the erroneous behavior of approximate solutions observed in attempts to solve the Driven cavity problem using the techniques presented in [3].

5.2 First order reformulations of the Navier-Stokes equations

Using the above analytical solution, let us compare the performance of the multilevel adaptive Least Squares FE procedure described in [3] for the following two first order reformulations of the Navier-Stokes equations. Here we use four equations with four unknowns and nine equations with seven unknowns, respectively.

"4 by 4" reformulation

Let us consider the Navier-Stokes problem in the first order velocity-vorticity-Bernoulli pressure formulation for incompressible viscous flow:

$$\begin{aligned} &b_x + \nu\omega_y - v\omega = 0, \\ &b_y - \nu\omega_x + u\omega = 0, \\ &u_y - v_x + \omega = 0, \\ &u_x + v_y = 0 \end{aligned}$$

in $S = (0,1)^2$, with u, v given on ∂S, b fixed at the origin, $p(0,0) = 0$.

Here, the Bernoulli pressure is defined as

$$b = p + \frac{u^2 + v^2}{2},$$

so the analytical expressions for b and ω can readily be obtained from those given for u, v, p.

Let us choose $V \subset (H^1(S))^4$ and $V' = (L_2(S))^4$ and define the standard scalar product

$$(p, q) = \int_0^1 \int_0^1 \sum_{i=1}^{4} p_i(x,y) q_i(x,y) dxdy.$$

Unfortunately, as it turns out in this case, the Regularity condition (2.5) does not hold under any natural choice of V, see e.g. [5].

"9 by 7" reformulation

Another first order formulation involves four additional functions (components of the flux tensor), $a = u_x$, $b = u_y$, $c = v_x$, and $d = v_y$. The equations which we used were

$$\begin{aligned} &u_x - a = 0, \\ &u_y - b = 0, \\ &v_x - c = 0, \\ &v_y - d = 0, \\ &a_y - b_x = 0, \\ &c_y - d_x = 0, \\ &p_x + \nu(-a_x - b_y) + ua + vb = 0, \\ &p_y + \nu(-c_x - d_y) + uc + vd = 0, \\ &a + d = 0 \end{aligned}$$

in $S = (0,1)^2$, with u, v given on ∂S, p fixed at the origin, $p(0,0) = 0$, and with the following additional boundary conditions written in terms of tangential derivatives of the boundary data functions (without which this setting is as bad as the above one):

$$b(x,0) = \frac{\partial u}{\partial y}(x,0), \quad b(x,1) = \frac{\partial u}{\partial y}(x,1),$$

$$d(x,0) = \frac{\partial v}{\partial y}(x,0), \quad d(x,1) = \frac{\partial v}{\partial y}(x,1),$$
$$a(0,y) = \frac{\partial u}{\partial x}(0,y), \quad a(1,y) = \frac{\partial u}{\partial x}(1,y),$$
$$c(0,y) = \frac{\partial v}{\partial x}(0,y), \quad c(1,y) = \frac{\partial v}{\partial x}(1,y).$$

Even better results are observed when adding the boundary conditions obtained from the continuity equation $u_x + v_y = 0$ written for the boundary,

$$a(x,0) = -\frac{\partial v}{\partial y}(x,0), \quad a(x,1) = -\frac{\partial v}{\partial y}(x,1),$$

$$d(0,y) = -\frac{\partial u}{\partial x}(0,y), \quad d(1,y) = -\frac{\partial u}{\partial x}(1,y).$$

In recent publications, e.g., [5], the following three equations are used instead of the last equation $a + d = 0$: the original one $u_x + v_y = 0$ and $a_x + d_x = 0,\ a_y + d_y = 0$. At the same time, the last two additional boundary conditions for a and d are not used there.

5.3 The solution method

Given the (nonlinear) equation

$$F\mathbf{u} = 0,$$

with (typically linear) boundary conditions

$$G\mathbf{u} = g,$$

where $\mathbf{u}$ is a vector function from a proper functional space V, we start from some initial FE space V_h and an initial guess $u_h \in V_h$ such that $Gu_h = g_h$.

Then we perform several updating steps of the form

$$u_h := u_h + \tau d_h$$

with proper stepsizes $\tau > 0$ (however, $\tau = 1$ often works well even when u_h lies far from the exact solution), and the directions defined as an approximate solution of the following linear least squares problem:

$$d_h \approx \arg\min_{d_h \in V_h} \Big(\|Fu_h + F'(u_h)d_h\|^2 + \mu \|Gd_h\|^2_{\partial\Omega} \Big),$$

where μ is a very large positive number, say $\mu = 10^{15}$ (hence d_h appears to be very close to the minimizer of $\|Fu_h + F'(u_h)d_h\|$ and nearly satisfies the homogeneous boundary conditions). The integrals over triangles were approximated using a fourth order quadrature rule presented in [8], while the integral along the boundary was replaced by a simple sum of squares of appropriate components of d_h taken at the boundary nodes. The discretized problem was then solved by the (preconditioned)

Conjugate Gradient method. In our experiments, we used the Robust second order Incomplete Cholesky preconditioning [13] with a bandwidth reduction preordering of the stiffness matrix.

The stopping criterion for these iterations over the same space V_h is

$$\cos[Fu_h, F'(u_h)V_h] \leq \delta,$$

where δ is a small positive number chosen consistently with the maximum stepsize h. (The relative precision parameter for stopping the Conjugate Gradient iterations was chosen as $O(\delta^2)$.)

If this criterion is satisfied but $\|Fu_h\|$ is still large, then a refinement of the space V_h is performed (and a slight update of u_h to conform the refined boundary conditions is done). Some details of the adaptive h-refinement procedure used are given in the next subsection.

As follows from the previous discussion, when the residual norm is sufficiently small and the problem is sufficiently regular, one should have the error estimate of the the same type as in [1], $\|u - u_h\|/\|Fu_h\| = O(\sigma h + \sigma\delta)$, that is, the Euclidean norm of the error should decrease faster than the Euclidean norm of the residual.

5.4 The adaptive refinement procedure

Let the initial finite-dimensional subspaces V_0 be constructed using some coarse initial triangulation of Ω. Since Ω is a unit square in our case, we started the calculations from the standard uniform grid composed of uniformly shaped right isosceles triangles. In our experiments we used for the construction of the sequence of nested finite-dimensional subspaces $V_0, \ldots, V_h, \ldots$ the standard piecewise linear nodal basis functions and the corresponding hierarchical quadratic basis (as is well-known, for each element, the three quadratic elemental basis functions are obtained as pairwise products of the corresponding linear ones). In order to construct the subsequent triangulations, we choose the Longest Edge Bisection procedure [15], so that any subsequent grid is composed only of right isosceles triangles having various sizes and different orientations. Hence, any degeneration of triangles is prevented during the refinement process; moreover, considerably more triangles than was prescribed by the error estimator may be refined on each step, which prevents possible undermeshing. After each refinement, new nodal values were obtained by linear interpolation along the bisected edges (in the same order as they were bisected), and then the velocities and fluxes at the boundary nodes were set equal to their exact values using the boundary conditions.

For each triangle, we use the local error estimator taken as the integral of the squared residual function $(Fu_h)^2$ over the element. Thus, on each refinement step the bisection procedure was applied to those triangles over which the estimator exceeds ϱ^2 times its maximum over all triangles (we used $\varrho = 0.5$). Closely related adaptive strategies were given in [15] and, in an abstract Least Squares setting, in [1].

5.5 Discussion of the numerical results

We present the results obtained when the initial grid was 9 by 9 with 100 nodes; the initial guess was simply set to zero at all nodes, except for those where the functions has values prescribed by the boundary conditions.

For the "4 by 4" reformulation, the numerical experiments showed rather poor decrease (or even increase) of the FE error as compared to the decrease of the residual error. This clearly contradicts the theory presented above and thus indicates that the Regularity condition does not hold ($\sigma = \infty$). The latter claim was checked numerically by estimating the minimum eigenvalue of a generalized eigenvalue problem associated with the corresponding stiffness and mass matrices, which showed that even for linear FE bases, the stability constant σ approaches 100 and is growing further along with the refinement of the FE space. Thus, the residual norm appears as a not good error indicator when trying to solve a problem that does not satisfy the Regularity property.

On the other hand, the "9 by 7" reformulation demonstrated very good convergence behavior even with linear FE bases, and the numerical experiments were in nice agreement with the theory. The numerical estimates obtained for σ were not much larger than 10 for all grids. However, the iteration costs (especially the computer memory space requirements) are considerably higher when one uses seven unknown functions instead of only four ones.

Similar results were obtained for the quadratic elements. Typically, the adaptive quadratic elements perform much better than the linear ones and sometimes give a possibility to solve the problem even using an insufficiently regular first order reformulation. In our case, it was possible to obtain a satisfactory approximation to the solution using the "4 by 4" reformulation, but the use of more regular "9 by 7" one made it possible to attain the same precision with much less computational cost.

The convergence histories for both of these formulations are given in Figures 1-4, where the logarithms of error and residual norms are given versus the logarithm of flops needed to compute the FE solution. It is seen that for both types of elements the error norm is larger than the residual norm for the "4 by 4" reformulation and, on the contrary, the reverse relation holds for the "9 by 7" one. However, for both reformulations, the residuals behave virtually alike. Generally (in cases when the exact solution is unknown), it appears expedient to check the numerical estimates of the regularity constant σ in order to recognise the lack of regularity of the first order reformulation used.

6 Conclusions

As suggested by both theory and the numerical tests presented above, one of the key points in the analysis of First Order System Least Squares FE techniques is to find a proper (regular) reformulation of the original problem. Otherwise, any improvement in the Finite Element or nonlinear and linear solution techniques will be rather useless. An alternative to (or a preliminary stage for) the theoretical analysis of different settings of the problem may be a series of carefully planned numerical tests

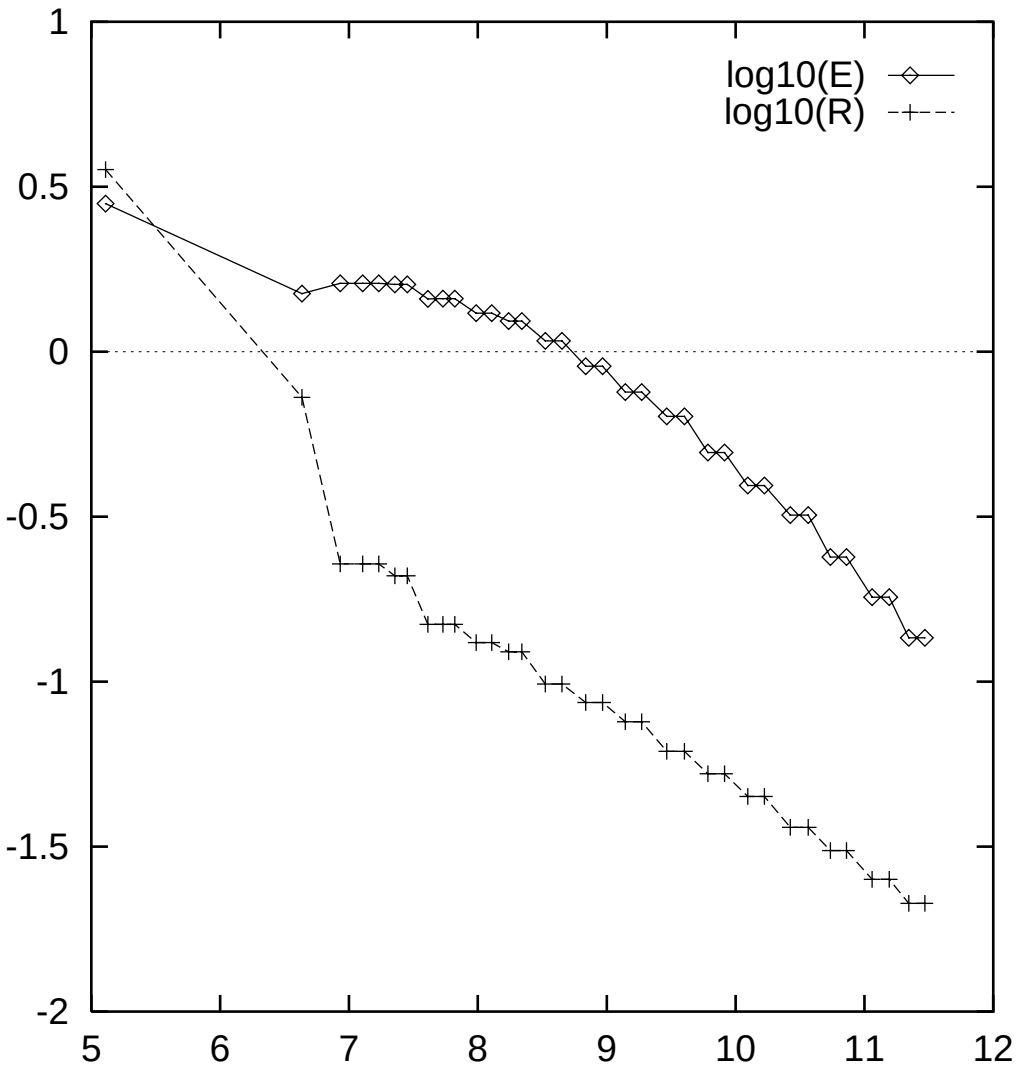

Figure 1: Residual and error norms vrs. flops count: "4 by 4" reformulation; linear elements

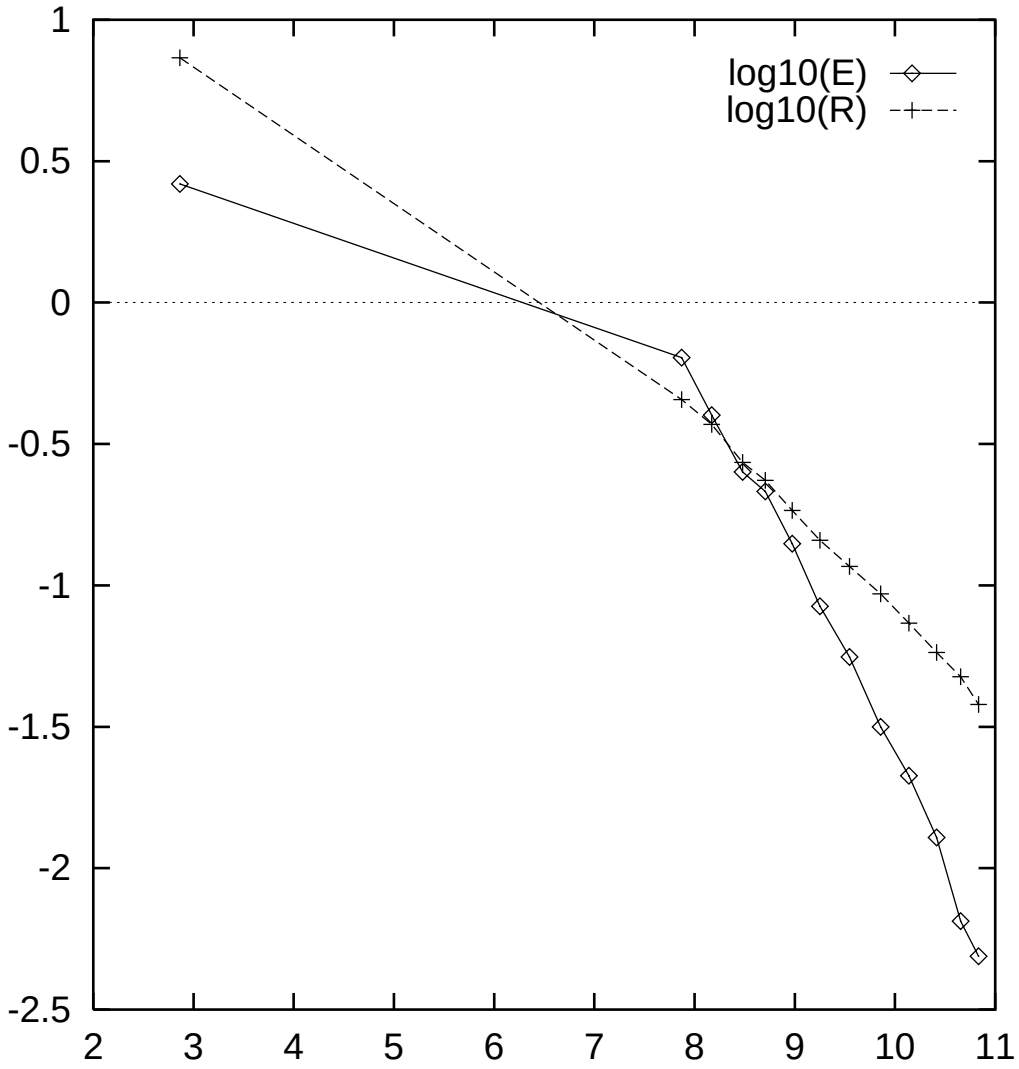

Figure 2: Residual and error norms vrs. flops count: "9 by 7" reformulation; linear elements

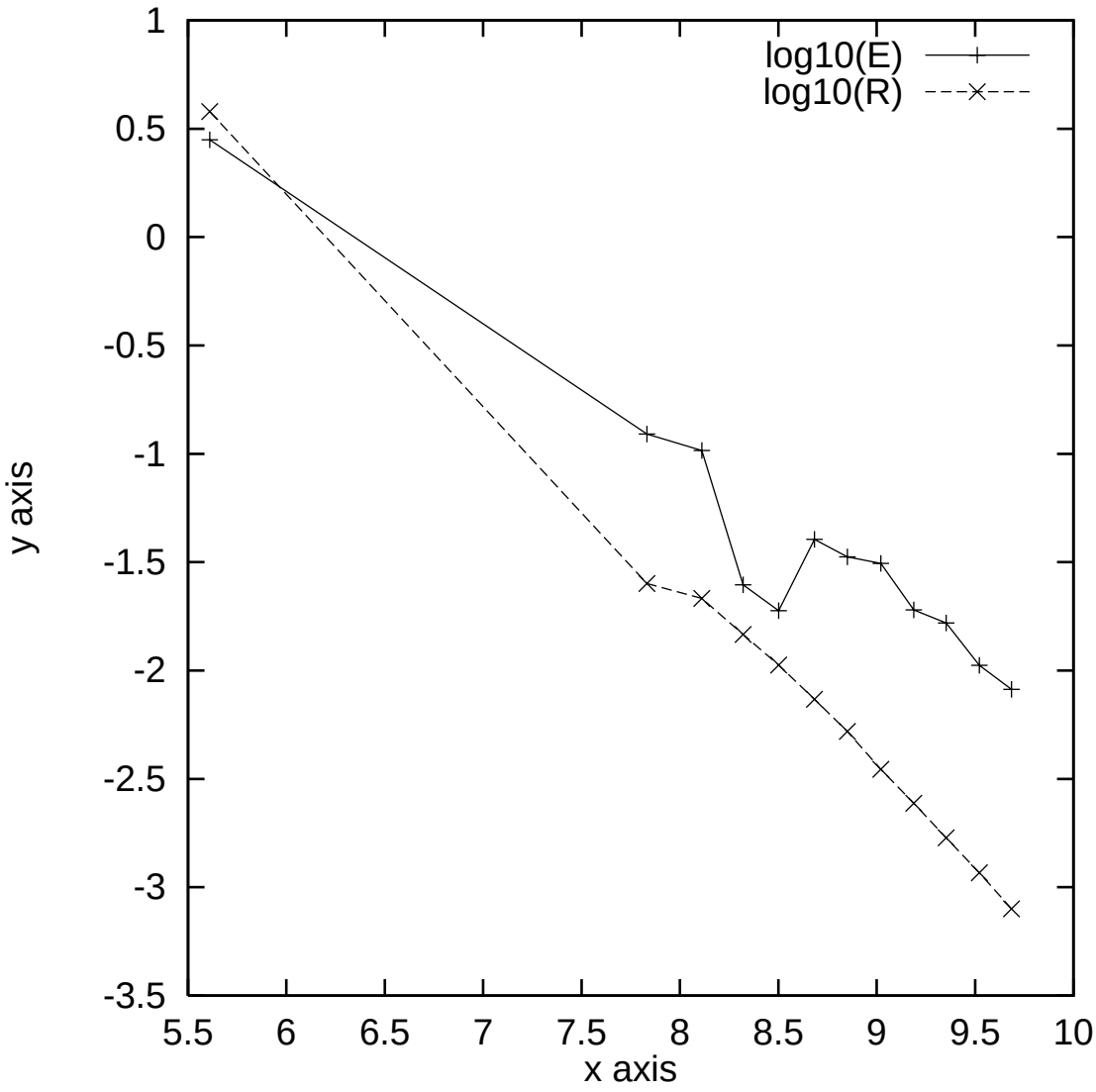

Figure 3: Residual and error norms vrs. flops count: "4 by 4" reformulation; quadratic elements

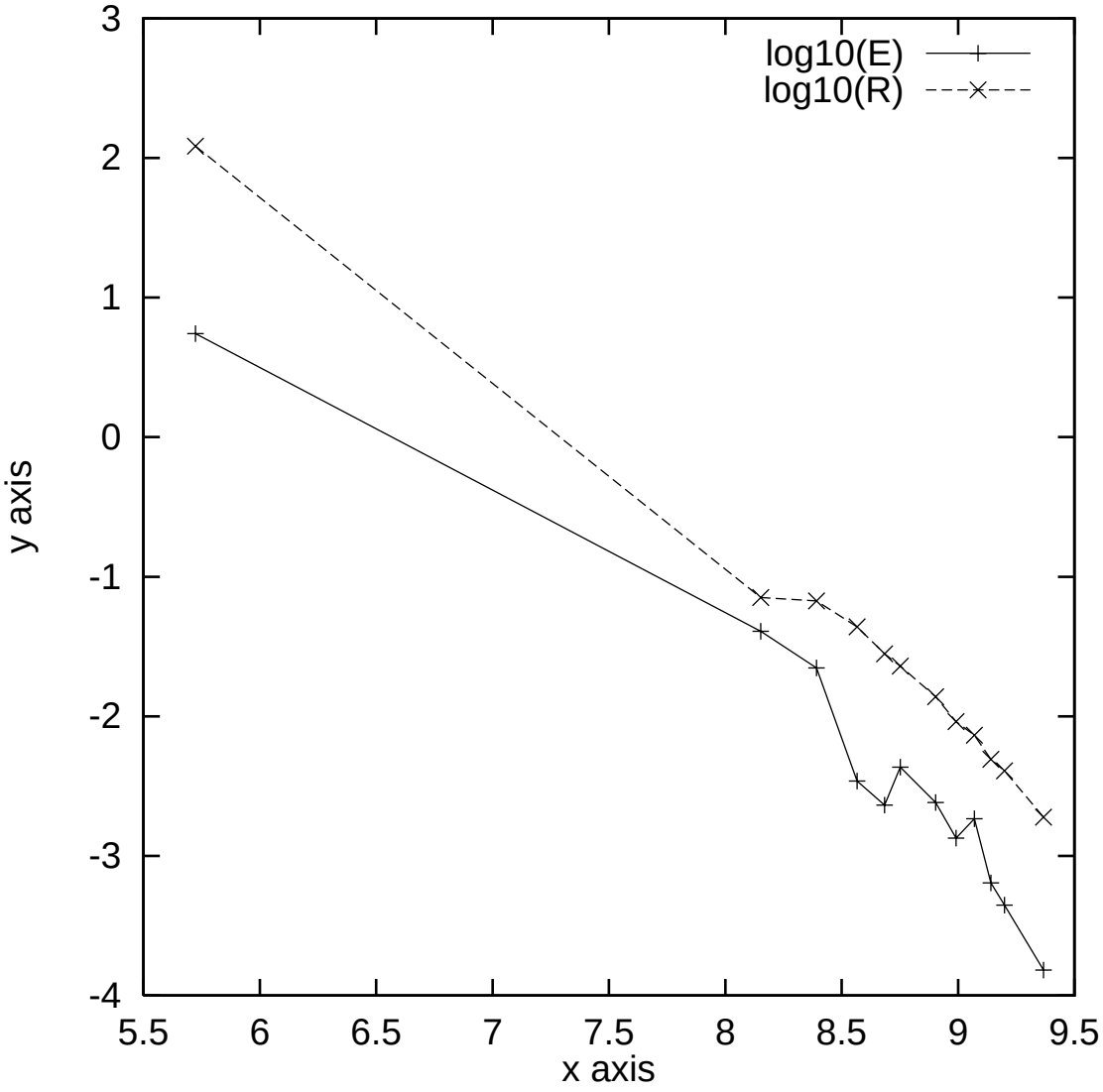

Figure 4: Residual and error norms vrs. flops count: "9 by 7" reformulation; quadratic elements

controlling the estimates of the Regularity constant σ and/or using a properly chosen analytical solution.

However, some formulation of a discrete analogue of the Regularity condition with $\sigma = \sigma(h)$ moderately growing as $h \to 0$, as well as the further development of the theory may lead to a reasonable compromise between strict regularity and simplest possible first order reformulation. This is of special importance for the nonlinear case since the nonlinearity constant c_0 (as was defined in Section 4) may often grow as the FE space refinement progresses.

References

[1] O. Axelsson and I.E. Kaporin, *A posteriori error estimates and adaptive refinement strategies for a minimum residual finite element solution of nonlinear problems*, Report No. 9560, Department of Mathematics, University of Nijmegen, The Netherlands, December 1995;

[2] O. Axelsson and I.E. Kaporin, *Least squares finite element methods for systems of nonlinear partial differential equations of first order: a posteriori error estimates*, Report No. 9612, Department of Mathematics, University of Nijmegen, The Netherlands, March 1996.

[3] O. Axelsson and I.E. Kaporin, *Minimum residual adaptive multilevel finite element procedure for the solution of nonlinear stationary problems*, SIAM J. Numer. Anal. 35 (1998), 1213-1229.

[4] M. Berndt, T.A.Manteuffel, and S.F.McCormick, *Local error estimates and adaptive refinement for first-order system least squares (FOSLS)*, Electronic Trans. Numer. Anal, 6 (1997), 35-43.

[5] P. Bochev, Z. Cai, T. A. Manteuffel, and S. F. McCormick, *Analysis of velocity-flux first-order system least-squares principles for the Navier-Stokes equations: Part I*, SIAM J. Numer. Anal. 35 (1998), 990-1009.

[6] Pavel B. Bochev and Max D. Gunzburger, *Analysis of least squares finite element methods for the Stokes equations*, Math. Comp. 63 (1994), 479-506.

[7] C.L.Chang and B.-N.Jiang, *An error analysis of least-squares finite element method of velocity-pressure-vorticity formulation for Stokes problem*, Comput. Methods Appl. Mech. Engrg. 84 (1990), 247-255.

[8] D.A. Dunavant, *High degree efficient symmetrical Gaussian quadrature rules for the triangle*, Int. J. Numer. Methods Engrg. 21 (1985), 1129–1148.

[9] K.O.Friedrichs, *Symmetric positive differential equations*, Comm. Pure Appl. Math., 11 (1958) 333-418.

[10] Handbook on Numerical Analysis, vol. 2, Finite Element Methods (Part I) (P.G. Ciarlet, J.L. Lions, eds.), North-Holland etc., 1991.

[11] B.-N. Jiang, T.L. Lin, and L.A. Povinelli, *Least-squares finite element method for fluid dynamics*, Comput. Methods Appl. Mech. Engrg. 81 (1990), 13-37.

[12] B.-N. Jiang, T.L. Lin and L.A. Povinelli, *Large-scale computation of incompressible viscous flow by least-squares finite element method*, Comput. Methods Appl. Mech. Engrg. 114 (1994), 213-231.

[13] I.E. Kaporin, *High quality preconditioning of a general symmetric positive definite matrix based on its* $U^TU+U^TR+R^TU$*-decomposition*, Numer. Linear Algebra Appl., 5, (1998), 484-509.

[14] T. Manteuffel, S. McCormick, and C. Pflaum, *Improved Discretization Error Estimates for First-Order System Least Squares (FOSLS)*, Math. Inst. Univ. Würzburg Prepr. No. 227 (Sept. 1998), 16 p.

[15] W.F.Mitchell, *Optimal multilevel iterative methods for adaptive grids*, SIAM J. Sci. Comput. 13 (1992) 146-167.

[16] S.-Y.Yang and J.-L.Liu, *A unified analysis of a weighted least squares method for first-order systems*, Appl. Math. Comput., 92 (1998), 9-27.