Least squares finite element methods for systems of nonlinear partial differential equations of first order: A posteriori error estimates

O. Axelsson[*] I.E. Kaporin[†]

Abstract

In recent years there has been much renewed interest in least squares finite element methods for systems of first order partial differential equations.
In the present paper nonlinear such systems are considered and an a posteriori error estimate in L_2-norm based on the L_2-norm of the residual, taking into account the inexactness of the solution of the discretized (and linearized) system, is derived. The error estimates can be used for adaptive refinements. Some aspects of the solution of the arising nonlinear systems are also discussed.

1 Introduction

Recent years have shown a much renewed interest in least-squares finite element methods for systems of first order partial differential equations, see [5], [9], [10] and [7], for instance. Since any (system of) partial differential equations can be rewritten as a system of first order equations such methods are generally applicable. The disadvantage with them is that many different unknown (physical) variables are introduced, such as all the components of the gradient vectors, for instance.
However, as we shall see, there are many advantages with least-squares methods,

[*]Faculty of Mathematics and Informatics, University of Nijmegen, The Netherlands
[†]High-Speed Rolling Stocks Research Center, Moscow State University of Railway Communication, Moscow, Russia

The research for the first author was supported by a travel grant from the Copernicus CP 940820 of the European Commission, and for the second author by a grant from the Netherlands Organisation of Scientific Research, NWO B61 - 298

which may outweigh this disadvantage. (Eventually, perhaps some combination of classical methods for higher order systems and first order systems may turn out to be the most efficient. For instance, one may use the more classical methods in the major part of the domain and least squares only in some minor, but critical parts.)
The main advantages are:
(i) The choice of finite elements is much simplified. For instance, for (divergence) constrained problems there is no need to find stable finite element pairs (for the pressure and velocity components, for instance). It is known that finding such stable pairs, satisfying the so-called LBB-condition is difficult, in particular in 3-space dimensional problems. On the other hand, if one accepts that not all variables will always be approximated to full accuracy, in least squares methods one may use the same finite element space for all variables involved. Hence, also one may use a single finite element mesh.
(ii) The finite element matrices can be kept on 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. This simplifies the a posteriori error analysis and adaptive refinement of the underlying finite element mesh. (In practice, it may turn out to be most efficient to use adaptive refinement for different groups of unknowns separately. However, this will not be discussed in the present paper.)

While previous papers have considered mostly linear systems, in the present paper we consider nonlinear systems. This adds the complication of solving the arising nonlinear algebraic systems.
We show how the L_2-norm of the errors can be estimated based on the L_2-norm of the residuals. This includes all error sources, among them an estimate of the error arising from the inexact solution of the (linearized) discretized system of nonlinear algebraic equations.
As the above mentioned error estimates assumes, in particular, that the approximate solution is sufficiently close to the exact solution, this discussion includes methods which converge from an (in principle) arbitrary initial point. For more details on the topic of robust Newton type methods, see [1], [2], [3], and [13], and references quoted therein.

2 Least squares discretization error estimate in L_2-norm for linear problems

As a preliminary for the analysis of the discretization error for nonlinear problems, consider first a linear problem

$$\mathcal{L}u = f,$$

where $f \in L^2(\Omega)$ and $\mathcal{L}$ is a first order linear differential operator, i.e., involving no derivatives of higher order than one and defined on a domain Ω in $\mathbb{R}^n$.
Hence $\mathcal{L} : V \to V'$, where $V = H^1(\Omega)$, $V' = L^2(\Omega)$. 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.
We make then two basic assumptions. The first is a regularity assumption consisting of two parts. Part *a)* is intended for use in practical error estimators, while part *b)* is intended to show the rate of convergence. The second is an approximation property assumption on the finite element space involving the operator and consists also of two similar parts.

(i) Regularity assumptions:

For any $\varphi \in V$, there exists a ψ which is a solution of

$$(\mathcal{L}\psi, \mathcal{L}v) = (\varphi, v) \quad \forall v \in \mathring{V} \subset V. \tag{2.1}$$

Here $\mathring{V}$ contains typically functions in V satisfying homogeneous essential boundary conditions.
The solution is assumed to be sufficiently regular, in the sense that

$$a) \qquad \|\mathcal{L}\psi\| \le c_1\|\varphi\| \tag{2.2}$$

or

$$b) \qquad \|\psi\|_{1+\alpha} \le C_\alpha\|\varphi\|, \text{ for some } \alpha,\ \ 0 < \alpha \le 1, \tag{2.3}$$

where $\|\cdot\|_{1+\alpha}$ denotes a (possibly fractional order) Sobolev space.
Note that part *a)* holds if $\psi \in \mathring{V}$, and $\|\mathcal{L}\psi\| \ge \text{const}\|\psi\|$ (or, equivalently, the inverse operator $\mathcal{L}^{-1}$ is bounded), because then (2.1) shows that

$$\|\mathcal{L}\psi\| \le \text{const}\|\varphi\|.$$

For more general functions ψ, it holds if $\|\mathcal{L}\psi\| \leq \text{const}\|\psi\|$ and $\psi = \psi_0 + \widetilde{\psi}$, where $\widetilde{\psi} \in \mathring{V}$ and ψ_0 is sufficiently smooth and satisfies the inhomogeneous boundary conditions. Clearly part *a)* holds if part *b)* holds with $\alpha = 1$.

(ii) Approximation property of the F.E. space:

Let $V_h \subset V$ be a finite element subspace and assume that

$$a) \qquad \inf_{v_h \in V_h} \|\mathcal{L}(\psi - v_h)\| \leq \varepsilon_h \|\varphi\|$$

or

$$b) \qquad \inf \|\mathcal{L}(\psi - v_h)\| \leq \widetilde{\varepsilon}_h \|\psi\|_{1+\alpha},$$

where $\varepsilon_h, \widetilde{\varepsilon}_h = o(1)$, $h \to 0$.
Typically, in case (b), $\widetilde{\varepsilon}_h = Ch^\alpha$ for some constant C, independent of h.
Note that if part *b)* and (2.3) both holds, then part *a)* holds with $\varepsilon_h = C_\alpha \widetilde{\varepsilon}_h$.
Let u_h be the least squares F.E. method solution for $\mathcal{L}u = f$, i.e., let u_h satisfy the essential boundary conditions and let

$$(\mathcal{L}u_h, \mathcal{L}v_h) = (f, \mathcal{L}v_h) \quad \forall v_h \in V_h \cap \mathring{V},$$

The error $u - u_h$ satisfies then

$$(\mathcal{L}(u - u_h), \mathcal{L}v_h) = 0 \quad \forall v_h \in V_h. \tag{2.4}$$

For the derivation of the error estimates, let w, an auxiliary function, be the solution of

$$(\mathcal{L}w, \mathcal{L}v) = (u - u_h, v) \quad \forall v \in \mathring{V}. \tag{2.5}$$

Then, since $u - u_h \in \mathring{V}$, with $v = u - u_h$ it follows

$$(u - u_h, u - u_h) = (\mathcal{L}w, \mathcal{L}(u - u_h))$$

and, using the orthogonality property (2.4), we find for any $v_h \in V_h$,

$$\begin{aligned} \|u - u_h\|^2 &= (\mathcal{L}(w - v_h), \mathcal{L}(u - u_h)) \\ &\leq \|\mathcal{L}(w - v_h)\| \, \|\mathcal{L}(u - u_h)\|, \end{aligned}$$

so

$$\|u - u_h\|^2 \leq \inf_{v_h \in V_h} \|\mathcal{L}(w - v_h)\| \, \|f - \mathcal{L}u_h\|.$$

Together with (2.5), the approximation assumption, part *a)* yields now

$$\|u - u_h\|^2 \le \varepsilon_h \|u - u_h\| \, \|f - \mathcal{L}u_h\| = \varepsilon_h \|u - u_h\| \, \|r\|,$$

or

$$\|u - u_h\| \le \varepsilon_h \|r\|, \quad r = f - \mathcal{L}u_h. \tag{2.6}$$

Note that $\mathcal{L}u_h \in V' = L^2(\Omega)$, since $\mathcal{L}$ is a first order operator, so $r \in L^2(\Omega)$.
Similarly, (2.5) with the approximation and regularity assumption, part *b)* yield

$$\|u - u_h\|^2 \le \tilde{\varepsilon}_h \|w\|_{1+\alpha} \|f - \mathcal{L}u_h\| \le \tilde{\varepsilon}_h C_\alpha \|u - u_h\| \, \|r\|,$$

or

$$\|u - u_h\| \le C_\alpha \tilde{\varepsilon}_h \|r\|, \tag{2.7}$$

where $\tilde{\varepsilon} = Ch^\alpha$.
The above show that the discretization error in L_2-norm converges faster than the L_2-norm of the residuals. The latter corresponds to a H^1-norm of the error and (2.4) 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 [12].

3 An L_2-norm estimate of the full error for the inexact solution of the linearized discrete equations

We consider now the error estimate for a nonlinear equation

$$F(u) = 0, \quad F : V \to V', \tag{3.1}$$

where F is differentiable, with Frechet derivative F' and where $u \in V$ is a solution. Thereby we will include also the effect of inexact solution of the nonlinear discretized and linearized equations. As has been shown in previous papers by the authors, see [13], [2] and [3], the proper condition to measure the effect of the inexactness of the solution of the linearized equation is the cosine of the acute angle between the residual and the finite element space, mapped by the Frechet derivative. More precisely, we assume that $u_h \in V_h$ is a given approximation to the exact solution u of $F(u) = 0$, such that

$$|(r, Av_h)| \le \delta \|r\| \, \|Av_h\| \quad \forall v_h \in V_h, \tag{3.2}$$

where $r = F(u_h)$, $A = F'(u_h)$ and δ, $\delta < 1$ is a sufficiently small parameter. Note that $\delta \geq \max_{v_h \in V_h} \cos(\angle[r, Av_h])$. From now on, A replaces the linear operator notation $\mathcal{L}$, used in Section 2.

To get some further insight in the size and role played by δ we indicate first that when u_h is close to the exact least squares Galerkin solution to a linearized equation, which is linearized at some approximate solution $\tilde{u}_h$ near the solution of $F(u) = 0$, then δ can be taken close to zero.

This follows, since then

$$r = F(u_h) = F(u_h) - F(u) = A(u - u_h) + o(\|u - u_h\| + \|u - \tilde{u}_h\|),$$

where $A = F'(\tilde{u}_h)$.

Neglecting the last (nonlinear) term we have then

$$(r, Av_h) = (A(u - u_h), Av_h) = 0,$$

so δ would be exactly zero. Therefore, when the nonlinear term above is small, δ can normally be taken small.

On the other hand, when δ is small for some u_h, then it can be seen that, in a neighbourhood of u_h (which can be naturally defined as $\{y = \|F(y)\| \leq \|F(u_h)\|\}$), there does not exist any other solution having a significantly smaller residual than $r = F(u_h)$. This follows from the next Theorem.

Theorem 3.1 *Assume that (3.2) and the nonlinearity condition*

$$\begin{aligned} &\|F(y) - F(u_h) - F'(u_h)(y - u_h)\| \leq \\ &\quad \tfrac{1}{2} K_0 \|F'(u_h)(y - u_h)\| \, \|F(y) - F(u_h)\| \end{aligned}$$

are satisfied, where K_0 is a non-negative constant. Assume also that $\|F(u_h)\| \leq \frac{1}{3K_0}$. Then for any $y \in V_h$ such that $\|F(y)\| \leq \frac{1}{3K_0}$, it holds

$$\|F(y)\| \geq (1 - 4.5\delta^2)^{\frac{1}{2}} \|F(u_h)\|.$$

Proof. Let $\tau = \|F(y) - F(u_h)\|$. Then, for any $y \in V_h$ we have

$$\|Q\| \leq \frac{1}{2} K_0 \tau \|A(y - u_h)\|,$$

where $Q = F(y) - F(u_h) - A(y - u_h)$.

Note that, if $\|F(y)\| \leq \frac{1}{3K_0}$, it follows that $\tau \leq \frac{2}{3K_0}$. Hence, by

$$\begin{aligned} \|A(y - u_h)\| &\leq \|F(y) - F(u_h)\| + \|Q\| \\ &\leq \tau + \tfrac{1}{2} K_0 \tau \|A(y - u_h)\|, \end{aligned}$$

the estimate

$$\|A(y-u_h)\| \leq \frac{2\tau}{2-K_0\tau} \leq \frac{3}{2}\tau,$$

and therefore

$$\|Q\| \leq \frac{3}{4}K_0\tau^2,$$

follow. Thus using the last two estimates with the identity

$$\|F(y)\|^2 \equiv \|r\|^2 + \tau^2 + 2(r, A(y-u_h)+Q),$$

where $r = F(u_h)$, (3.2) shows the following estimate:

$$\begin{aligned} \|F(y)\|^2 &\geq \|r\|^2 + \tau^2 - 2\delta\|r\|\,\|A(y-u_h)\| - 2\|r\|\,\|Q\| \\ &\geq \|r\|^2 + \tau^2 - 3\delta\|r\|\tau - \tfrac{3}{2}K_0\|r\|\tau^2 \\ &\geq \|r\|^2 - 3\delta\|r\|\tau + \tfrac{\tau^2}{2}. \end{aligned}$$

Hence

$$\begin{aligned} \|F(y)\|^2 &\geq \min_{\tau\geq 0}(\|r\|^2 - 3\delta\|r\|\tau + \tfrac{\tau^2}{2}) \\ &= \|r\|^2(1-4.5\delta^2). \end{aligned}$$

□

Remark 3.1 The nonlinearity condition in Theorem 3.1 is a special case ($\nu = 1$) of condition (3.12) to follow. A similar result as in Theorem 3.1 can be shown to hold for the general condition (3.12). For the proof one can then make use of the inequality $\tau^{1+\nu} \leq (1-\nu)\tau + \nu\tau^2$, which follows directly from (3.13).

Given the residual r, we want now to estimate the L_2-error $\|u-u_h\|$, based on $\|r\|$. (In this respect, note that r is computable when F is a first order operator, even when V_h contains piecewise polynomial continuous basis functions.)
Let then $\widehat{u}_0$, an auxiliary function−only to be used for the derivation, be the exact least-squares solution of

$$(r + A(\widehat{u}_0 - u_h), Av) = 0 \quad \forall v \in \mathring{V}. \tag{3.3}$$

Also, let $w \in \mathring{V}$ be the solution of

$$(Aw, Av) = (\widehat{u}_0 - u_h, v) \quad \forall v \in \mathring{V}. \tag{3.4}$$

Letting $v = \widehat{u}_0 - u_h$, we find

$$\|\widehat{u}_0 - u_h\|^2 = (Aw, A(\widehat{u}_0 - u_h))$$

so, by (3.3),

$$\|\widehat{u}_0 - u_h\|^2 = -(r, Aw).$$

Hence, for any $v_h \in V_h$,

$$\begin{aligned} \|\widehat{u}_0 - u_h\|^2 &= -(r, A(w - v_h)) - (r, Av_h) \\ &\leq \|r\| \, \|A(w - v_h)\| + \delta \|r\| \, \|Av_h\|. \end{aligned} \tag{3.5}$$

Let now $v_h \equiv \arg \min_{v_h \in V_h} \|A(w - v_h)\|$. Then

$$(A(w - v_h), Av_h) = 0 \quad \forall \widetilde{v}_h \in V_h,$$

so

$$(Av_h, A\widetilde{v}_h) = (Aw, A\widetilde{v}_h) \leq \|Aw\| \, \|A\widetilde{v}_h\|,$$

i.e.,

$$\|Av_h\| \leq \|Aw\|.$$

Hence (3.5), the approximation assumption and the regularity assumption *a)* yield

$$\begin{aligned} \|\widehat{u}_0 - u_h\|^2 &\leq \|r\| \|(\inf_{v_h \in V_h} \|A(w - v_h)\| + \delta \|Aw\|) \\ &\leq \|r\|(\varepsilon_h + \delta)\|Aw\|. \end{aligned}$$

Here, by (3.4) and the regularity assumption *a)* for $A = F'(u_h)$, it follows that

$$\|Aw\| \leq c_1 \|\widehat{u}_0 - u_h\|$$

so

$$\|\widehat{u}_0 - u_h\| \leq c_1(\varepsilon_h + \delta)\|r\|. \tag{3.6}$$

If the regularity and approximation assumptions *b)* hold, we find similarly

$$\begin{aligned} \|\widehat{u}_0 - u_h\|^2 &\leq \|r\|(\widetilde{\varepsilon}_h \|w\|_{1+\alpha} + \delta \|Aw\|) \\ &\leq \|r\|(\widetilde{\varepsilon} C_\alpha + \delta \text{const }) \|\widehat{u}_0 - u_h\|, \end{aligned}$$

or

$$\|\widehat{u}_0 - u_h\| \leq (C_\alpha \widetilde{\varepsilon}_h + \text{const} \cdot \delta)\|r\|.$$

We can now state the following Lemma.

Lemma 3.2 *Let u be a solution of (3.1), let u_h be an approximation for which (3.2) holds and such that $F'(u_h)$ satisfies the regularity assumption a) in Section 2. Then the L_2-error is bounded by*

$$\|u - u_h\| \leq \|u - \widehat{u}_0\| + c_1(\varepsilon_h + \delta)\|r\| \tag{3.7}$$

Here

(i) $\|u-\widehat{u}_0\|$ *is the term arising from the nonlinearity of the mapping* F*, (i.e. arising from the linearization)*

(ii) $c_1\varepsilon_h\|r\|$ *is the FE approximation error arising from the* $H^1 \to L^2$*-lifting (2.6).*

(iii) $c_1\delta\|r\|$ *is error due to inexact solution of the linearized FE equation.*

Proof. Use the triangle inequality and (3.6). □

Remark 3.2 A similar estimate holds if the regularity assumption *b)* is satisfied.

It remains to estimate the first term in the right hand side of (3.7). This will be done via the error $\|A(u-\widehat{u}_0)\|$ where $A = F'(u_h)$. We will then assume that A is nonsingular in the direction $u - \widehat{u}_0$ with a boundedness constant $\widehat{C}_1$, i.e.,

$$\|u-\widehat{u}_0\| \leq \widehat{C}_1\|A(u-\widehat{u}_0)\| \tag{3.8}$$

and we shall estimate $\|A(u-\widehat{u}_0)\|$ from $\|r\|$.
Consider then (3.3),

$$(r + A(\widehat{u}_0 - u_h), Av) = 0 \quad \forall v \in \mathring{V}. \tag{3.9}$$

Setting here $v = \widehat{u}_0 - u_h$, we find

$$\|A(\widehat{u}_0 - u_h)\|^2 = -(r, A(\widehat{u}_0 - u_h)) \leq \|r\| \, \|A(\widehat{u}_0 - u_h)\|,$$

i.e.,

$$\|A(\widehat{u}_0 - u_h)\| \leq \|r\|. \tag{3.10}$$

(Recall that $r = F(u_h)$.)
Setting now $v = u - \widehat{u}_0$ in (3.9) we find

$$(r + A(\widehat{u}_0 - u_h), A(u-\widehat{u}_0)) = 0,$$

so

$$\begin{aligned} & \|r + A(\widehat{u}_0 - u_h)\|^2 + \|A(u-\widehat{u}_0)\|^2 \\ = \; & \|r + A(\widehat{u}_0 - u_h) + A(u-\widehat{u}_0)\|^2 \\ = \; & \|r + A(u-u_h)\|^2 = \|F(u) - F(u_h) - F'(u_h)(u-u_h)\|^2, \end{aligned}$$

and, in particular,

$$\|A(u-\widehat{u}_0)\| \leq \|F(u) - F(u_h) - F'(u_h)(u-u_h)\|. \tag{3.11}$$

The next theorem shows that $\|A(u-\widehat{u}_0\|$ can be estimated by a constant times $\|A(u-u_h)\|^{1+\nu}$ if a certain nonlinearity condition holds.

Theorem 3.3 *Assume that a nonlinearity condition in the form*

$$\begin{aligned} &\|F(u) - F(u_h) - F'(u_h)(u - u_h)\| \leq \\ &\tfrac{1}{2}K_0\|F'(u_h)(u - u_h)\| \, \|F(u) - F(u_h)\|^\nu \end{aligned} \tag{3.12}$$

for some constants $K_0 > 0$ and ν, $0 < \nu \leq 1$, holds. Then, if

$$K_0\nu\|F'(u_h)(u - u_h)\|^\nu \leq 1,$$

the estimate

$$\|F(u) - F(u_h) - F'(u_h)(u - u_h)\| \leq K_0\|F'(u_h)(u - u_h)\|^{1+\nu}$$

holds.

Proof. With $A = F'(u_h)$, using the nonlinearity condition and the triangle inequality we have

$$\begin{aligned} &\|F(u) - F(u_h) - A(u - u_h)\| \leq \tfrac{1}{2}K_0\|A(u - u_h)\| \, \|F(u) - F(u_h)\|^\nu \\ \leq \;& \tfrac{1}{2}K_0\|A(u - u_h)\|(\|F(u) - F(u_h) - A(u - u_h)\| + \|A(u - u_h)\|)^\nu. \end{aligned}$$

Using the inequality

$$(1 + t)^\nu \leq 1 + \nu t \tag{3.13}$$

(which follows from the concavity of the mapping $f(t) = (1 + t)^\nu$, $0 < \nu \leq 1$), one obtains

$$\begin{aligned} &\|F(u) - F(u_h) - A(u - u_h)\| \\ &\leq \tfrac{1}{2}K_0\|A(u - u_h)\|^{1+\nu}(1 + \nu\tfrac{\|F(u)-F(u_h)-A(u-u_h)\|}{\|A(u-u_h)\|}) \\ &= \tfrac{1}{2}K_0\|A(u - u_h)\|^{1+\nu} + \tfrac{1}{2}K_0\nu\|A(u - u_h)\|^\nu \cdot \|F(u) - F(u_h) - A(u - u_h)\|, \end{aligned}$$

so, by the assumption made,

$$\|F(u) - F(u_h) - A(u - u_h)\| \leq K_0\|A(u - u_h)\|^{1+\nu}. \qquad \square$$

Remark 3.3 Theorem 3.3 shows that the nonlinearity condition (3.12) is of a quasi-Hölder type.

We find now from (3.11) and Theorem 3.3,

$$\begin{aligned} \|A(u - \widehat{u}_0)\| &\leq \|Fu - Fu_h - F'(u_h)(u - u_h)\| \\ &\leq K_0\|F'(u_h)(u - u_h)\|^{1+\nu} \\ &= K_0\|A(u - u_h)\|^{1+\nu}, \end{aligned} \tag{3.14}$$

when

$$K_0\nu\|F'(u_h)(u - u_h)\|^\nu \leq 1 \tag{3.15}$$

holds. Clearly, condition (3.15) holds if u_h is a sufficiently close approximation to u. We can now formulate the first main theorem.

Theorem 3.4 *Let u be a solution of problem (3.1), where F is differentiable and F and F' satisfies the regularity assumption (2.2) and the nonlinearity condition (3.12). Then, if u_h is sufficiently close to the solution u of (3.1) so that*

$$K_0\nu\|F'(u_h)(u-u_h)\|^\nu \leq 1$$

holds, the following error estimate in L_2-norm follows:

$$\|u-u_h\| \leq 2^{2+\nu}\widehat{C}_1 K_0\|r\|^{1+\nu} + c_1(\varepsilon_h+\delta)\|r\|.$$

Here $r = F(u_h)$, K_0 is the constant in the nonlinearity condition, $\widehat{C}_1$ is the boundedness constant of the inverse of the Frechet derivative $F'(u_h)$ in the direction $u - \widehat{u}_0$ and c_1 is the constant in the regularity condition a) for the operator $F'(u_h)$.

Proof. Using Theorem 3.3 and the triangle inequality, we find

$$\|A(u-u_h)\| \leq \|r\| + K_0\|A(u-u_h)\|^{1+\nu},$$

or

$$\|A(u-u_h)\| \leq 2\|r\|.$$

Further, Lemma 3.2, (3.8) and (3.12) show that

$$\begin{aligned}\|u-u_h\| &\leq 2\widehat{C}_1 K_0\|A(u-u_h)\|^{1+\nu} + c_1(\varepsilon_h+\delta)\|r\| \\ &\leq 2^{2+\nu}\widehat{C}_1 K_0\|r\|^{1+\nu} + c_1(\varepsilon_h+\delta)\|r\|. \qquad \square\end{aligned}$$

Condition (3.15) is somewhat inappropriate as it can normally not be checked in practice. We replace it now with a condition on the smallness of the residual, which can readily be checked, if K_0 can be estimated, because the residual is explicitly available.

Theorem 3.5 *Let u be a solution of problem (3.1) where F is differentiable and F and F' satisfies the regularly assumption (2.2) and the nonlinearity condition (3.12). Then, if the residual is sufficiently small so that*

$$\|r\| \leq (1/K_0)^{1/\nu} \tag{3.16}$$

holds, the following error estimate in L_2-norm follows:

$$\|u-u_h\| \leq \widehat{C}_1 K_0\|r\|^{1+\nu} + c_1(\varepsilon_h+\delta)\|r\|.$$

Proof. Using (3.11), the nonlinearity assumption, and (3.10) we find

$$\begin{aligned}\|A(u-\widehat{u}_0\| &\leq \tfrac{1}{2}K_0\|A(u-u_h)\|\,\|r\|^{\nu}\\ &\leq \tfrac{1}{2}K_0\|r\|^{\nu}(\|A(u-\widehat{u}_0)\|+\|A(\widehat{u}_0-u_h)\|)\\ &\leq \tfrac{1}{2}K_0\|r\|^{\nu}(\|A(u-\widehat{u}_0)\|+\|r\|)\\ &\leq \tfrac{1}{2}\|A(u-\widehat{u}_0)\|+\tfrac{1}{2}K_0\|r\|^{1+\nu},\end{aligned}$$

where we have also used the assumption (3.15). This yields

$$\|A(u-\widehat{u}_0)\| \leq K_0\|r\|^{1+\nu}$$

and use of (3.8) completes the proof. □

From Theorems 3.4 and 3.5 it can be seen that, under the stated conditions, we can control the total error in L_2-norm by choosing a sufficiently fine finite element mesh and by solving the linearized systems sufficiently accurately. This holds at least if the approximate solution is sufficiently close to the exact solution. In general, the constants involved in the error estimates are not known. However, c_1 can be approximated, using a power iteration method. Clearly δ and $\|r\|$ are computable. Due to the superconvergence of the nonlinear term, it can be expected to be much smaller than the other terms, when $\|r\|$ is small. The finite element error factor ε_h can be estimated at some stage in the nonlinear iteration process, using extrapolation for some approximations for a sequence of meshes and the property

$$\inf_{v_h}\|A(\psi-v_h)\| \leq \|A(\psi-\widehat{u}_h)\|,$$

where $\widehat{u}_h$ is the exact finite element solution of

$$A(\psi-u_h) = -F(u_h),$$

where $A = F'(u_h)$. The error term arising from the linearization converges eventually superlinearly to zero with the current residual. The latter depends on the current finite element mesh and the accuracy of the solution of the linearized system. Hence, there is an interplay between the terms involved. The theorems show also that it makes sense to modify the mesh in such a way that the local values of the residuals get about equal size. These residuals are explicitly available and can hence readily be checked. Since the local regularity constants can vary much, in some problems it may be more efficient to include the regularity (stability) constant in the norm. For a similar approach (for linear problems), see [8], [11].

As the theorems hold under the assumption that u_h is already sufficiently close to the solution u for (3.15) to be valid, it remains to discuss methods to find such solutions,

when no further information of the solution is available.
There exists at least two such methods which can be recommended. The first is based on approximate Newton type methods using extended subspaces . Under very general conditions, such methods converge to the solution starting from an arbitrary initial approximation. See [1], [13], [2] and [3] for further details and for references to similar publications by other authors.
The second method is based on a two-level mesh method and will now be considered.

4 Estimate of the error in L_2-norm of the nonlinear two-level method

The two-level method to solve nonlinear partial differential equation problems has been considered previously in [14], [15], [1], [4] and [6] and others.. Given a nonlinear partial differential equation $Fu = 0$, with solution u in a function space V, the method consists of the following steps:

(i) Solve $F(u) = 0$ on a finite element space $V_H \subset V$ with relatively few degrees of freedom, corresponding to a coarse mesh, i.e., solve the nonlinear problem (exactly),

$$(F(u_H), v_H) = 0 \quad \forall v_H \in V_H \cap \mathring{V} \tag{4.1}$$

(ii) Correct the solution once on a much finer space (mesh) $V_h \subset V$ based on the linearized problem, i.e., solve (for u_h)

$$(F'(u_H)(u_h - u_H), v_h) = -(F(u_H), v_h) \quad \forall v_h \in V_h.$$

The major concern of the previous papers was finding an asymptotic relation between the mesh parameters h and H such that the rate of convergence $\|u - u_h\|_V$ has the same order as $\|u - \widehat{u}_h\|_V$, where $\widehat{u}_h$ is the Galerkin (finite element) solution of $(F(\widehat{u}_h), v_h) = 0 \ \forall v_h \in V_h$. A typical relation found was $H = O(h^{\frac{1}{2}})$, or, for quasi linear problems, even $H = O(h^{\frac{1}{4}})$. This means that the coarse mesh problem can be extremely coarse relative to the fine mesh problem.
However, the above results assume a globally sufficient smoothness of the solution but when the solution has interior and/or boundary layers, such a smoothness is not seen. In such a case, the coarse mesh may be unable to find a sufficiently accurate solution on which the linearized problem can be based, if the coarse mesh is taken so coarse as indicated above. It will then not suffice with just one Newton (linear equation) solution on the finer space. Still, the use of a two-level method can be efficient

in providing at least a somewhat reasonable initial approximation for starting the nonlinear iterations on the finer mesh using an approximate Newton type solver. The modified two-level method will then have the following form:

(i) Solve $F(u) = 0$ on a coarse mesh V_H (fairly accurately) to find a solution u_H.

(ii) With u_H as initial approximation, $u_h^{(0)} = u_H$, for $k = 0, 1, \ldots$ iterate until convergence, i.e., solve approximately

$$(F'(u_h^{(k)})\xi_h^{(k)}, v_h) = -(F(u_h^{(k)}), v_h) \quad \forall v_h \in V_h \cap \mathring{V}$$

and let

$$u_h^{(k+1)} = u_h^{(k)} + \xi_h^{(k)}.$$

Thereby, the linear systems are solved such that the condition

$$|(r(u_h^{(k+1)}), Av_h)| \leq \delta \|r(u_h^{(k+1)})\| \, \|Av_h\|$$

holds for some $\delta < 1$, where

$$r(u_h^{(k+1)}) = F(u_h^{(k+1)}).$$

For further details, see [1], [13]. Should it happen that u_H is not a sufficiently accurate approximation, one may have to use damping of the stepsize, at least for the first iteration steps.

We consider now the accuracy of the solution u_H and show that if the solution u is sufficiently smooth and if H is sufficiently small, the solution u_H can be computed to a sufficient accuracy for (3.14) to hold for u_H.

Therefore, the nonlinear solution process in the modified two level method will converge with at step k, a solution $u_h^{(k)}$ for which the bound in Theorem 3.1 holds.

Let then u_H be computed to solve (4.1) approximately by some method where

$$|(r_H, Av_H)| \leq \delta_H \|r_H\| \, \|Av_H\|$$

holds for some $\delta_H < 1$, to be specified below, where

$$r_H = F(u_H), \; A = F'(u_H).$$

We want to estimate $\|u - u_H\|$ asymptotically as a function of H.

We let u_H be an approximate solution of (4.1) computed by an approximate Newton type method where at each linearization step, the linearized equations are solved (approximately) by a least-squares finite element method. Hence, the solution u_H satisfies approximately

$$(A(u_H - u_H^*), Av_H) = -(F(u_H^*), Av_H) \quad \forall v_H \in V_H \cap \overset{\circ}{V}, \tag{4.2}$$

where $A = F'(u_H^*)$ and u_H^* is the approximation computed before the final linearization step.
By the remark below Lemme 3.1 it follows that

$$\|u - u_H\| \le K_1 \|r_H\|^{1+\nu} + K_2(\widetilde{\varepsilon}_H + \delta)\|r_H\| \tag{4.3}$$

where $r_H = F(u_H)$, $\widetilde{\varepsilon}_H = K_3 H^\alpha$, and K_1, K_2, K_3 are constants.
We choose $\delta = O(H^\alpha)$ (or in practice $\delta = O(H)$ since, in general, α is not known). It remains to estimate $\|r_H\|$ as a function of H.
We have by (3.13)

$$\begin{aligned} \|r\| &= \|F(u_H) - F(u)\| \le \\ &\quad \|A(u_H - \widehat{u}_H)\| + \|A(\widehat{u}_H - u)\| + O(\|A(u - u_H)\|^{1+\nu}), \end{aligned} \tag{4.4}$$

where $A = F'(u_H^*)$ and $\widehat{u}_H$ is the exact least-squares finite element solution of (4.2). Since u_H satisfies

$$|(r, Av_H)| \le \delta \|r\| \, \|Av_H\| \quad \forall v_H \in V_H \cap \overset{\circ}{V},$$

it follows from elementary geometry that

$$\|A(u_H - \widehat{u}_H)\| = O(\delta).$$

Further,

$$(A(\widetilde{u}_H - \widehat{u}_H), Av_H) = (A(\widetilde{u}_H - u), Av_H) \quad \forall v_H \in V_H \cap \overset{\circ}{V}, \tag{4.5}$$

where $\widetilde{u}_H \in V_H$ is arbitrary and where we have used the orthogonality property of the finite element error $A(u - \widehat{u}_H)$.
Letting $v_H = \widetilde{u}_H - \widehat{u}_H$ in (4.5), we find

$$\|A(\widetilde{u}_H - \widehat{u}_H)\| \le \|A(\widetilde{u}_H - u)\|.$$

Hence $\|A(u - \widehat{u}_H)\| \le 2\|A(u - \widetilde{u}_H)\|$ so, by (4.4),

$$\begin{aligned} \|r\| &= \|A(u_H - \widehat{u}_H)\| + \|A(\widehat{u}_H - u)\| + O(\|A(u - u_H)\|^{1+\nu}) \\ &\le O(\delta) + 2 \inf_{\widetilde{u}_H \in V_H} \|A(u - \widetilde{u}_H)\| + O(\|A(u - u_H)\|^{1+\nu}). \end{aligned}$$

Here, assuming that the regularity assumption b) holds, it follows by elementary interpolation theory,

$$\inf_{\widetilde{u}_H \in V_H} \|A(u - \widetilde{u}_H)\| = O(H^\alpha)\|u\|_{1+\alpha},$$

$0 < \alpha \leq 1$. Hence, asymptotically, the above implies that if $\delta = O(H^\alpha)$, then

$$\|r\| = O(H^\alpha), \quad H \to 0,$$

and, by (4.3),

$$\|u - u_H\| = O(H^{\alpha(1+\nu)}) + O(H^{2\alpha}).$$

Under the stated assumption on smoothness of the solution, this shows a sufficient convergence which enables choosing V_H sufficiently fine such that the initial approximation in (ii) is sufficiently accurate.
If we have full regularity, $\alpha = 1$ and if the nonlinearity condition holds with $\nu = 1$, then

$$\|u - u_H\| = O(H^2).$$

For some numerical results, using the above methods, see [2], [3], and [13].

References

[1] O. Axelsson, *On mesh independence and Newton-type methods*, Report 9217, Department of Mathematics, University of Nijmegen, The Netherlands, August 1992; appeared later in: Applications of Mathematics, 38 (1995), 249-265.

[2] O. Axelsson and I.E. Kaporin, *On the solution of nonlinear equations for nondifferentiable mappings*, in: Fast Solvers for Flow Problems, (W. Hackbusch, G. Wittum, editors), Vieweg-Verlag, Braunschweig/Wiesbaden, 1994, pp. 38-51.

[3] O. Axelsson and I.E. Kaporin, *Minimum residual adaptive multilevel procedure for the finite element solution of nonlinear stationary problems*, Report No. 9522, Department of Mathematics, University of Nijmegen, The Netherlands, May 1995: submitted to SIAM J. Numer. Anal.

[4] O. Axelsson and B. Layton, *A two-level method for the discretization of nonlinear boundary value problems*, Report 9314, Department of Mathematics, University of Nijmegen, The Netherlands, March 1993: SIAM J. Numer. Anal. (1996), to appear.

[5] A. Aziz, R. Kellogg and A. Stephens, *Least-squares methods for elliptic systems*, Math. Comp. 44 (1985), 53-70.

[6] R.E. Bank and R.K. Smith, *A posteriori error estimates based on hierarchical bases*, SIAM J. Numer. Anal. 30 (1993), 921-935.

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

[8] K. Ericsson and C. Johnson, *Adaptive streamline diffusion finite element methods for stationary convection-diffusion problems*, Math. Comput., 60 (1993), 167-188.

[9] 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.

[10] 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.

[11] C. Johnson, *Adaptive finite element methods for diffusion and convection problems*, Computer Methods Appl. Mech. Engrg., 82 (1990), 301-322.

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

[13] I.E. Kaporin and O. Axelsson, *On a class of nonlinear equation solvers based on the residual norm reduction over a sequence of affine subspaces*, SIAM J. Sci. Comput., 16 (1995), 228-249.

[14] J. Xu, *Two grid finite element discretization for linear and nonlinear elliptic equations*, Techn. Report AM105, Dept. Math., Pennsylvania State Univ., University Park, July 1992.

[15] J. Xu, *A novel two-grid method for semilinear elliptic equations*, SIAM J. Sci. Comput., 15 (1994), 231-237.

[16] M. Zlamal, *On the finite element method*, Numer. Math., 12 (1968), 394-409.