

A POSTERIORI ERROR ESTIMATION FOR VARIATIONAL PROBLEMS WITH UNIFORMLY CONVEX FUNCTIONALS

SERGEY I. REPIN

ABSTRACT. The objective of this paper is to introduce a general scheme for deriving a posteriori error estimates by using duality theory of the calculus of variations. We consider variational problems of the form

$$\inf_{v \in V} \{F(v) + G(\Lambda v)\},$$

where $F : V \rightarrow \mathbb{R}$ is a convex lower semicontinuous functional, $G : Y \rightarrow \mathbb{R}$ is a uniformly convex functional, V and Y are reflexive Banach spaces, and $\Lambda : V \rightarrow Y$ is a bounded linear operator. We show that the main classes of a posteriori error estimates known in the literature follow from the *duality error estimate* obtained and, thus, can be justified via the duality theory.

1. INTRODUCTION

In this paper, we consider methods of a posteriori error estimation for a class of variational problems with convex functionals. The basic problem, in its general form, is to find u in a Banach space V such that

$$(1.1) \quad J(u, \Lambda u) = \inf_{v \in V} J(v, \Lambda v),$$

where $J(v) = F(v) + G(\Lambda v)$, F is a convex, lower semicontinuous functional, G is a uniformly convex functional and $\Lambda : V \rightarrow Y$ is a bounded linear operator. We assume that V and Y are reflexive Banach spaces endowed with the norms $\|\cdot\|_V$ and $\|\cdot\|$, respectively. Let $v \in V$ be an approximation of u , then $e = v - u$ is the approximation error. The aim of a posteriori error analysis is to obtain a computable error majorant $\mathcal{M} = \mathcal{M}(v; \mathcal{D})$ which depends only on v and the given data set \mathcal{D} . This majorant must possess the following two basic properties:

$$(1.2) \quad \|e\|_V \leq \mathcal{M}(v; \mathcal{D}) \quad \forall v \in V,$$

$$(1.3) \quad \mathcal{M}(v_k; \mathcal{D}) \xrightarrow[k \rightarrow +\infty]{} 0 \quad \text{if } v_k \rightarrow u \text{ in } V.$$

Methods of a posteriori error estimation for partial differential equations received attention more than two decades ago (see [2, 3, 4, 5, 20]). Nowadays the literature on this subject is vast (see, e.g., [1, 15, 21, 23, 36] and the references therein). However, almost all methods can be collected into three main groups:

Received by the editor April 1, 1997.

1991 *Mathematics Subject Classification*. Primary 65N30.

Key words and phrases. A posteriori error estimates, duality theory, nonlinear variational problems.

This research was supported by INTAS Grant N 96-0835.

- (A) residual methods,
- (B) methods based on gradient recovery,
- (C) equilibrated data methods.

In the residual method (see, e.g., [1, 3, 4, 13, 36]) a weak norm of the residual function is taken for \mathcal{M} . Methods (B) (see, e.g., [6, 37, 38]) are based on averaging (smoothing) approximate solutions obtained by the finite element method. These types of post-processing procedures give new approximations, which often are much more accurate. For this reason, the difference between the direct approximation and the averaged one can be used as an error indicator. Complementary energy principles were applied for getting error estimates in [7, 17, 18, 19] and in other papers. They formed the basis of methods (C) which apply special numerical procedures designed for getting the so-called “equilibrated functions” in complementary energy principles.

In this paper, we present a unified approach to a posteriori error estimation that follows from the *duality theory* of the calculus of variations. In earlier papers, we used this theory to obtain a priori error estimates for variational problems with linear growth functionals [24, 25, 32] and a posteriori estimates for some classes of nonlinear variational problems [26, 27, 28, 29, 33, 34]. The aim of the analysis below is to introduce a *general scheme* for deriving a posteriori error estimates and to show that methods (A)–(C) can be identified with particular forms of the duality error estimate.

The paper is organized as follows. In Section 2 we obtain the general a posteriori estimate (2.12). The right-hand side of (2.12) is a sum of two nonnegative functionals M_F and M_G which are equal to zero if and only if the *duality relations* (2.9)(i)–(ii) are satisfied. In the remainder of Section 2, we pay special attention to the frequently encountered case when F is a linear functional. In this case, the estimate (2.12) should be replaced by a modified one (2.25). The modified error majorant is a sum of two nonnegative functionals M_R and M_D , which represent a generalized measure of the *residual* and the error in the *duality relations*, respectively. In Section 3 we apply abstract results of Section 2 to several classes of variational problems. The goal of Section 4 is to compare the duality method with the aforementioned methods (A)–(C) and to show that they can be uniformly justified via the duality theory.

2. DUALITY A POSTERIORI ERROR ESTIMATES

2.1. Preliminaries. We begin by recalling some definitions. Let X be a reflexive Banach space. We consider functionals defined on elements of X with values in $\overline{\mathbb{R}} := \mathbb{R} \cup \{\pm\infty\}$. For a convex functional \mathcal{F} one can define its domain $dom \mathcal{F} := \{x \in X \mid \mathcal{F}(x) < +\infty\}$ and its epigraph $epi \mathcal{F} := \{(x, \alpha) \in X \times \mathbb{R} \mid \mathcal{F}(x) \leq \alpha\}$. We say that \mathcal{F} is a *proper* functional if $dom \mathcal{F} \neq \emptyset$ and $\mathcal{F}(x) > -\infty$ for any $x \in X$. The functional \mathcal{F} is said to be *lower semicontinuous* (l.s.c.) if $epi \mathcal{F} \in X \times \mathbb{R}$ is a closed set. For the set of all proper, convex, l.s.c functionals we use the notation $\Gamma_0(X)$. Let X^* be the space topologically dual to X with duality pairing $\langle \cdot, \cdot \rangle$ and $\mathcal{F} \in \Gamma_0(X)$. The function $\mathcal{F}^* : X^* \rightarrow \overline{\mathbb{R}}$ defined by

$$\mathcal{F}^*(x^*) := \sup_{x \in X} \{\langle x^*, x \rangle - \mathcal{F}(x)\}$$

is called the *Fenchel conjugate* of \mathcal{F} . Directly from this definition it follows that

$$(2.1) \quad \mathcal{F}(x) + \mathcal{F}^*(x^*) - \langle x^*, x \rangle \geq 0 \quad \forall x \in X, \quad x^* \in X^*.$$

An element $x^* \in X^*$ satisfying the equality $\mathcal{F}(x) + \mathcal{F}^*(x^*) - \langle x^*, x \rangle = 0$ is called a *subgradient* of \mathcal{F} at x . The set of all subgradients of \mathcal{F} at x is denoted by $\partial\mathcal{F}(x)$. If $\partial\mathcal{F}(x)$ consists of the unique element x^* , then \mathcal{F} is said to be *Gâteaux-differentiable* at x and we write $x^* = \mathcal{F}'(x)$.

Throughout this paper we use two pairs of dual Banach spaces (V, V^*) and (Y, Y^*) with duality pairings $\langle \cdot, \cdot \rangle$ and $\langle\langle \cdot, \cdot \rangle\rangle$, respectively. The spaces Y and Y^* are endowed with the norms $\|\cdot\|$ and $\|\cdot\|_*$. Let Λ be an element of the space $\mathcal{B}(V, Y)$ of all bounded linear operators from V to Y . We assume that

$$(2.2) \quad \|\Lambda w\| \geq c_0 \|w\|_V \quad \forall w \in V,$$

where c_0 is a positive constant independent of w . In addition to Λ we introduce its conjugate $\Lambda^* \in \mathcal{B}(Y^*, V^*)$ as the operator satisfying the identity

$$(2.3) \quad \langle\langle y^*, \Lambda v \rangle\rangle = \langle \Lambda^* y^*, v \rangle \quad \forall y^* \in Y^*, v \in V.$$

Besides, we introduce two convex functionals

$$(2.4) \quad F \in \Gamma_0(V) \quad \text{and} \quad G \in \Gamma_0(Y)$$

which compose the functional

$$J(v, \Lambda v) := F(v) + G(\Lambda v).$$

The latter functional is assumed to be coercive on V , i.e.,

$$(2.5) \quad J(v, \Lambda v) \rightarrow +\infty \quad \text{if} \quad \|v\|_V \rightarrow +\infty.$$

Lastly, we note that \mathbb{R}_+ denotes the set of all positive real numbers, the abbreviation “iff” is used instead of the words “if and only if” and the symbol “:=” means “equal by definition”.

2.2. Primal and dual problems. Let us start by giving the formal statement of the considered variational problem.

Problem \mathcal{P} . Find $u \in V$ such that

$$(2.6) \quad J(u, \Lambda u) = \inf \mathcal{P} := \inf_{v \in V} J(v, \Lambda v).$$

The problem dual to (2.6) is (see, e.g., [12])

Problem \mathcal{P}^* . Find $p^* \in Y^*$ such that

$$(2.7) \quad \begin{aligned} -J^*(\Lambda^* p^*, -p^*) &= \sup \mathcal{P}^* := \sup_{y^* \in Y^*} -J^*(\Lambda^* y^*, -y^*), \\ J^*(\Lambda^* y^*, -y^*) &:= F^*(\Lambda^* y^*) + G^*(-y^*), \end{aligned}$$

where F^* and G^* are the functionals conjugate of F and G , respectively.

The following existence theorem is known in the calculus of variations (see [12]).

Theorem 2.1. *Let the assumptions (2.4) and (2.5) hold. If the functional F is finite at some $u_0 \in V$ and the functional G is continuous and finite at $\Lambda u_0 \in Y$, then there exists a minimizer u to Problem \mathcal{P} and a maximizer p^* to Problem \mathcal{P}^* . Besides,*

$$(2.8) \quad \inf \mathcal{P} = \sup \mathcal{P}^*$$

and the following duality relations hold

$$(2.9) \quad \begin{aligned} (i) \quad &F(u) + F^*(\Lambda^* p^*) - \langle \Lambda^* p^*, u \rangle = 0, \\ (ii) \quad &G(\Lambda u) + G^*(-p^*) + \langle\langle p^*, \Lambda u \rangle\rangle = 0. \end{aligned}$$

Remark 2.1. The above relations are equivalent to the generalized differential equations

$$(2.10) \quad \begin{aligned} (i) \quad & \Lambda^* p^* \in \partial F(u), \\ (ii) \quad & -p^* \in \partial G(\Lambda u). \end{aligned}$$

2.3. Problems with uniformly convex functionals. Henceforth we assume that the functional G possesses the additional convexity properties formulated below.

Definition 1. We say that a continuous functional $G : Y \rightarrow \mathbb{R}$ is uniformly convex on a ball $B(0, \delta) := \{y \in Y \mid \|y\| < \delta\}$ if there exists a continuous functional $\Phi_\delta : Y \rightarrow \mathbb{R}_+$ such that $\Phi_\delta(y) = 0$ iff y is zero element of Y and

$$(2.11) \quad G\left(\frac{y_1+y_2}{2}\right) + \Phi_\delta(y_2 - y_1) \leq \frac{1}{2} (G(y_1) + G(y_2)) \quad \forall y_1, y_2 \in B(0, \delta).$$

It follows directly from (2.11) that any continuous uniformly convex functional is convex. Moreover, the functional Φ_δ (*forcing functional* [14]) reinforces the usual convexity inequality. Several examples of uniformly convex functionals are presented in Section 3.

Remark 2.2. Typically, the functional Φ_δ is given by a continuous strictly increasing function of the norm $\|y\|$. One can find the corresponding definitions of uniformly convex functionals in [14], [22].

Now, we are in a position to present a *general form* of a posteriori error estimate for variational problems with uniformly convex functionals.

Theorem 2.2. *Assume that the conditions of Theorem 2.1 hold and*

- (i) G is uniformly convex on a ball $B(0, \delta)$,
- (ii) the solution u of Problem \mathcal{P} and an element $v \in V$ are such that $\Lambda u, \Lambda v \in B(0, \delta)$.

Then

$$(2.12) \quad \Phi_\delta(\Lambda e) \leq M(v, y^*) := M_F(\Lambda^* y^*, v) + M_G(y^*, \Lambda v) \quad \forall y^* \in Y^*,$$

where $e = v - u$ and

$$\begin{aligned} M_F(\Lambda^* y^*, v) &:= \frac{1}{2} (F(v) + F^*(\Lambda^* y^*) - \langle \Lambda^* y^*, v \rangle), \\ M_G(y^*, \Lambda v) &:= \frac{1}{2} (G(\Lambda v) + G^*(-y^*) + \langle y^*, \Lambda v \rangle). \end{aligned}$$

Proof. Since $F \in \Gamma_0(V)$ and G is uniformly convex, we obtain

$$\Phi_\delta(\Lambda e) \leq \frac{1}{2} \left[(F(v) + G(\Lambda v)) + (F(u) + G(\Lambda u)) \right] - G(\Lambda(\frac{v+u}{2})) - F(\frac{v+u}{2}).$$

The element u is a minimizer, therefore,

$$G(\Lambda u) + F(u) = J(u) \leq G(\Lambda(\frac{u+v}{2})) + F(\frac{u+v}{2}),$$

and we arrive at the basic estimate

$$(2.13) \quad \Phi_\delta(\Lambda e) \leq \frac{1}{2} (J(v, \Lambda v) - J(u, \Lambda u)) \quad \forall v \in B(0, \delta).$$

In view of Theorem 2.1

$$(2.14) \quad F(u) + G(\Lambda u) = \inf \mathcal{P} = \sup \mathcal{P}^* = -F^*(\Lambda^* p^*) - G^*(-p^*).$$

Now, (2.13), (2.14) and the inequality

$$-J^*(\Lambda^*p^*, -p^*) \geq -J^*(\Lambda^*y^*, -y^*) \quad \forall y^* \in Y^*$$

imply

$$\begin{aligned} \Phi_\delta(\Lambda e) &\leq \frac{1}{2} (F(v) + G(\Lambda v) + F^*(\Lambda^*p^*) + G^*(-p^*)) \\ &\leq \frac{1}{2} (F(v) + G(\Lambda v) + F^*(\Lambda^*y^*) + G^*(-y^*)). \end{aligned}$$

The above inequality, together with (2.3), results in the desired (2.12). □

Theorem 2.2 deserves some more detailed comments. The right-hand side of (2.12) is the sum of two functionals $M_F : V^* \times V \rightarrow \mathbb{R}$ and $M_G : Y^* \times Y \rightarrow \mathbb{R}$. These functionals are nonnegative (see (2.1)) and vanishes iff v and y^* satisfy the relations (2.9)(i)–(ii) (i.e., iff $v = u$ and $y^* = p^*$). Therefore, the majorant $\mathbf{M}(v, y^*)$ is, in fact, a measure of the error in the *duality relations* for the pair (v, y^*) .

Remark 2.3. Let the functional F be uniformly convex on V with a forcing functional φ_δ . Then instead of (2.13) we have

$$(2.15) \quad \Phi_\delta(\Lambda e) + \varphi_\delta(e) \leq \frac{1}{2}(J(v, \Lambda v) - J(u, \Lambda u))$$

and, as a result, (2.12) is replaced by the strengthened estimate

$$(2.16) \quad \Phi_\delta(\Lambda e) + \varphi_\delta(e) \leq \mathbf{M}(v, y^*) \quad \forall y^* \in Y^*.$$

Remark 2.4. Some practically interesting variational problems (e.g., elasticity problems with nonconvex energy) are related to functionals which do not satisfy the condition (2.11). Nevertheless, the duality approach can be successfully extended to this case if the key equality $\inf \mathcal{P} = \sup \mathcal{P}^*$ holds. For these problems a posteriori error estimates are obtained in terms of the dual problem (see [30]).

It is not difficult to verify that

$$\mathbf{M}(v, y^*) - \mathbf{M}(v, p^*) = J^*(\Lambda^*y^*, -y^*) - J^*(\Lambda^*p^*, -p^*) \geq 0.$$

Therefore, for any v the right-hand side of (2.12) is minimal if $y^* = p^*$. Consequently, to make the estimate effective we have to find some y^* close to p^* in Y^* . In principle, this can be done by solving Problem \mathcal{P}^* numerically. Regrettably, very often the latter problem is more complicated than Problem \mathcal{P} and, for this reason, it is more effective to use duality relations (2.10) for getting a suitable approximation of p^* . To this end, we set $y^* = \sigma^*(v)$, where

$$-\sigma^*(v) \in \partial G(\Lambda v) \subset Y^*.$$

Hence, $M_G(\sigma^*(v), \Lambda v) = 0$ and we get the estimate

$$(2.17) \quad \Phi_\delta(\Lambda e) \leq M_F(\Lambda^*\sigma^*(v), v)$$

whose right-hand side depends on v only.

However, the estimate (2.17) cannot be directly applied in one practically important case which we consider below.

2.4. Problems with linear functional F . Let

$$(2.18) \quad F(v) = \langle l^*, v \rangle, \quad l^* \in V^*.$$

Since

$$F^*(v^*) = \sup_{v \in V} \langle v^* - l^*, v \rangle = \begin{cases} 0 & \text{if } v^* = l^*, \\ +\infty & \text{if } v^* \neq l^*, \end{cases}$$

we see that

$$M_F(\Lambda^* y^*, v) = \frac{1}{2}(F^*(\Lambda^* y^*) + \langle l^* - \Lambda^* y^*, v \rangle) = \begin{cases} 0 & \text{if } y^* \in Q_l^*, \\ +\infty & \text{if } y^* \notin Q_l^*, \end{cases}$$

where

$$Q_l^* := \{y^* \in Y^* \mid \langle \Lambda^* y^*, w \rangle = \langle l^*, w \rangle \quad \forall w \in V\}.$$

Notice that, in general, σ^* and $\Pi\sigma^*$ do not belong to the set Q_l^* , so that the right-hand side of (2.17) can become infinite. Therefore, the aim of our subsequent analysis is to obtain a modified error majorant $\widetilde{M}(v, y^*)$ which is finite for all $v \in V$ and all $y^* \in Y^*$. The first step on this way is provided by the following

Lemma 2.1. *Let the assumptions of Theorem 2.2 hold and F be given by (2.18). Then*

$$(2.19) \quad \Phi_\delta(\Lambda e) \leq \left[M_G(y^*, \Lambda v) + \frac{1}{2} \inf_{q^* \in Q_l^*} \inf_{\xi \in \partial G^*(-q^*)} \langle \langle y^* - q^*, \xi - \Lambda v \rangle \rangle \right].$$

Proof. Since

$$\begin{aligned} 2(M_F(\Lambda^* q^*, v) + M_G(q^*, \Lambda v)) &= G(\Lambda v) + G^*(-y^*) \\ &\quad + \langle \langle y^*, \Lambda v \rangle \rangle + \langle \langle q^* - y^*, \Lambda v \rangle \rangle \\ &\quad + G^*(-q^*) - G^*(-y^*) \quad \forall q^* \in Q_l^*, y^* \in Y^* \end{aligned}$$

and

$$G^*(-q^*) - G^*(-y^*) \leq \langle \langle y^* - q^*, \xi \rangle \rangle \quad \forall \xi \in \partial G^*(-q^*)$$

we obtain

$$(2.20) \quad \begin{aligned} &2(M_F(\Lambda^* q^*, v) + M_G(q^*, \Lambda v)) \\ &\leq G(\Lambda v) + G^*(-y^*) + \langle \langle y^*, \Lambda v \rangle \rangle + \langle \langle y^* - q^*, \xi - \Lambda v \rangle \rangle \\ &= 2M_G(y^*, \Lambda v) + \langle \langle y^* - q^*, \xi - \Lambda v \rangle \rangle \quad \forall q^* \in Q_l^*. \end{aligned}$$

Now (2.12) and (2.20) imply

$$(2.21) \quad \Phi_\delta(\Lambda e) \leq M_G(y^*, \Lambda v) + \frac{1}{2} \langle \langle y^* - q^*, \xi - \Lambda v \rangle \rangle.$$

Taking the infimum over q^* and ξ we end up with (2.19). □

Corollary 2.1. *If G^* is Gâteaux-differentiable, then from (2.21) we derive the estimate*

$$(2.22) \quad \begin{aligned} \Phi_\delta(\Lambda e) &\leq M_G(y^*, \Lambda v) + \frac{1}{2} \inf_{q^* \in Q_l^*} \langle \langle y^* - q^*, G^{*'}(-q^*) - \Lambda v \rangle \rangle \\ &= M_G(y^*, \Lambda v) + \frac{1}{2} \inf_{q^* \in Q_l^*} \left(\langle \langle y^* - q^*, G^{*'}(-q^*) - G^{*'}(-y^*) \rangle \rangle \right. \\ &\quad \left. + \langle \langle y^* - q^*, G^{*'}(-y^*) - \Lambda v \rangle \rangle \right). \quad \square \end{aligned}$$

Let $H \in \Gamma_0(Y)$, $H(y) \geq 0$ for all $y \in \text{dom } H$ and $H(0) = 0$. By $H^* : Y^* \rightarrow \mathbb{R}_+$ we denote the functional conjugate of H . Then, in virtue of the Joung–Fenchel inequality

$$(2.23) \quad \langle\langle \xi^*, \xi \rangle\rangle \leq H^*(\xi^*) + H(\xi) \quad \forall \xi \in Y, \xi^* \in Y^*,$$

we obtain

$$(2.24) \quad \langle\langle y^* - q^*, G^{*'}(y^*) - \Lambda v \rangle\rangle \leq H(G^{*'}(-y^*) - \Lambda v) + H^*(y^* - q^*).$$

Now, from (2.22) and (2.24) we deduce a modified majorant \widetilde{M} :

$$(2.25) \quad \Phi_\delta(\Lambda e) \leq \widetilde{M}(v, y^*) := M_D(y^*, \Lambda v) + M_R(y^*).$$

Here

$$(2.26) \quad M_D(y^*, \Lambda v) := M_G(y^*, \Lambda v) + \frac{1}{2}H(G^{*'}(-y^*) - \Lambda v)$$

and

$$(2.27) \quad M_R(y^*) := \frac{1}{2} \inf_{q^* \in Q_l^*} \left[\langle\langle y^* - q^*, G^{*'}(-q^*) - G^{*'}(-y^*) \rangle\rangle + H^*(y^* - q^*) \right].$$

We note that both summands of \widetilde{M} depend on the functional H whose form is rather arbitrary, e.g., in the simplest case, one can take

$$H(y) = \frac{1}{2} \|y\|^2, \quad H^*(y^*) = \frac{1}{2} \|y^*\|_*^2.$$

Thus, we see that the relations (2.25)–(2.27) give the general form of various a posteriori estimates. In practice, this freedom can be utilized to get the most rigorous error majorant.

We also note that these two summands have different, but clear sense. The first term $M_D(y^*, \Lambda v)$ is nonnegative and equal to zero iff v and y^* satisfy *the duality relation*

$$\Lambda v = G^{*'}(-y^*),$$

which is true for exact solutions p^* and u (cf. (2.9)(ii)). Hence, $M_D(y^*, \Lambda v)$ should be considered to be a measure of the error in these relations.

The term $M_R(y^*)$ is nonnegative and finite (unlike M_F !) for all $y^* \in Y^*$. It is equal to zero iff $y^* \in Q_l^*$, i.e., iff the equation $\Lambda^* y^* - l^* = 0$ holds. Consequently, $M_R(y^*)$ is a generalized measure of *the residual*

$$\mathcal{R}(y^*) = \Lambda^* y^* - l^*$$

expressed via the dual variable y^* .

The functional $M_D(y^*, \Lambda v)$ can be directly computed if v and y^* are given. However, finding the value of M_R necessitates solving an auxiliary minimization problem on the set Q_l^* . Below we consider the case when computing M_R is reduced to an unconstrained minimization problem for a convex functional J_0 . For the sake of simplicity we prove this assertion under some additional assumptions which, however, are not very restrictive and can be verified in concrete examples.

Assumption. Suppose that there exist two convex continuous functions $h : \mathbb{R} \rightarrow \mathbb{R}_+$ and $h^* : \mathbb{R} \rightarrow \mathbb{R}_+$ which are mutually conjugate and satisfy the inequalities

$$(2.28) \quad c_1 |t|^{\alpha_1} \leq h(t) \leq c_2 |t|^{\alpha_2},$$

$$(2.29) \quad \langle\langle \eta^* - y^*, G^{*'}(\eta^*) - G^{*'}(y^*) \rangle\rangle \leq h^*(\|\eta^* - y^*\|_*) \quad \forall \eta^*, y^* \in Y^*,$$

where $c_2 \geq c_1 > 0$ and $\alpha_2 \geq \alpha_1 > 1$.

Theorem 2.3. *Let the conditions of Lemma 2.1 and the foregoing Assumption be satisfied. Then*

$$(2.30) \quad \Phi_\delta(\Lambda e) \leq M_D(y^*, \Lambda v) + M_R(y^*),$$

where

$$(2.31) \quad M_D(y^*, \Lambda v) = M_G(y^*, \Lambda v) + \frac{1}{2}h(\|G^{*'}(-y^*) - \Lambda v\|),$$

$$(2.32) \quad M_R(y^*) = - \inf_{w \in V} \{h(\|\Lambda w\|) + \langle y^*, \Lambda w \rangle - \langle l^*, w \rangle\}.$$

Proof. Let us set $H(y) = h(\|y\|)$ and $H^*(y^*) = h^*(\|y^*\|_*)$. Making use of (2.27) and (2.29) we represent the function M_R as

$$(2.33) \quad M_R(y^*) = \inf_{q^* \in Q_l^*} h^*(\|q^* - y^*\|_*).$$

Hence, we obtain

$$\begin{aligned} M_R(y^*) &= - \sup_{q^* \in Q_l^*} \{-h^*(\|q^* - y^*\|_*)\} \\ &= - \sup_{q^* \in Y^*} \inf_{w \in V} \{\langle \Lambda^* q^* - l^*, w \rangle - h^*(\|q^* - y^*\|_*)\} \\ &= - \sup_{\eta^* \in Y^*} \inf_{w \in V} L(w, \eta^*), \end{aligned}$$

where

$$L(w, \eta^*) := \langle \Lambda^* \eta^* - l^*, w \rangle + \langle \Lambda^* y^*, w \rangle - h^*(\|\eta^*\|_*).$$

The function $w \mapsto L(w, \eta^*)$ is convex and continuous for any $\eta^* \in Y^*$. The function $\eta^* \mapsto L(w, \eta^*)$ is concave and continuous for any $w \in V$. Besides, from (2.28) it follows that $h^*(t) \geq c_2^* t^{\alpha_2^*}$, where $c_2^* = (c_2 \alpha_2)^{1-\alpha_2^*} (\alpha_2^*)^{-1}$ and $\alpha_2^* = \frac{\alpha_2}{\alpha_2-1} > 1$, so that $L(0, \eta^*) \rightarrow -\infty$ if $\|\eta^*\|_* \rightarrow +\infty$. Therefore, $\inf \sup L = \sup \inf L$ and

$$(2.34) \quad M_R(y^*) = - \inf_{w \in V} \sup_{\eta^* \in Y^*} L(w, \eta^*) = - \inf_{w \in V} J_0(w),$$

where

$$J_0(w) = h(\|\Lambda w\|) + \langle \mathcal{R}(y^*), w \rangle.$$

In view of (2.28) and (2.2) the functional J_0 is coercive on V . Thus, by standard technique, we establish the existence of $\hat{w} \in V$ such that $J_0(\hat{w}) \leq J_0(w) \forall w \in V$. Now, (2.26) comes in the form (2.31) and (2.34) yields (2.32). \square

At the end of this section we prove the consistency of the duality error majorant given by the estimate (2.30).

Theorem 2.4. *Suppose that the functionals $G(y)$ and $G^*(-y^*)$ are continuous at $y = \Lambda u$ and $y^* = p^*$, respectively. Let $\{v_k\}$ and $\{y_k^*\}$ be such sequences that*

$$\|v_k - u\|_V \xrightarrow{k \rightarrow +\infty} 0 \quad \text{and} \quad \|y_k^* - p^*\|_* \xrightarrow{k \rightarrow +\infty} 0.$$

Then the right-hand side of (2.19) tends to zero as $k \rightarrow +\infty$.

If, in addition, the functional G^ has a continuous Gâteaux derivative, then*

$$(2.35) \quad \widetilde{M}(v_k, y_k^*) \rightarrow 0 \quad \text{as } k \rightarrow 0.$$

Proof. It is straightforward to show that, under the continuity conditions imposed on Λ , G and G^* , the following limit relations hold

$$\begin{aligned} \Lambda v_k &\rightarrow \Lambda u \text{ in } Y, \\ G(\Lambda v_k) &\rightarrow G(\Lambda u), \\ G^*(-y_k^*) &\rightarrow G^*(-p^*). \end{aligned}$$

Thus, we obtain

$$(2.36) \quad M_G(y_k^*, \Lambda v_k) \rightarrow \frac{1}{2}[G(\Lambda u) + G^*(-p^*) + \langle\langle p^*, \Lambda u \rangle\rangle] = 0.$$

By virtue of (2.9)(ii), $\Lambda u \in \partial G^*(-p^*)$. Therefore,

$$(2.37) \quad \inf_{q^* \in Q^*} \inf_{\xi \in \partial G^*(-q^*)} \langle\langle y_k^* - q^*, \xi - \Lambda v_k \rangle\rangle \leq \langle\langle y_k^* - p^*, \Lambda u - \Lambda v_k \rangle\rangle \leq \|y_k^* - p^*\|_* \|\Lambda u - \Lambda v_k\| \rightarrow 0.$$

Based on (2.36) and (2.37), it is concluded that the estimate (2.19) tends to zero.

Let us prove (2.35). We have

$$(2.38) \quad \begin{aligned} M_D(y_k^*, \Lambda v_k) &= M_G(y_k^*, \Lambda v_k) + \frac{1}{2}h \left(\|G^{*'}(-y_k^*) - \Lambda v_k\| \right) \\ &\leq M_G(y_k^*, \Lambda v_k) + \frac{1}{2}h \left(\|G^{*'}(-y_k^*) - G^{*'}(-p^*)\| + \|\Lambda(u - v_k)\| \right). \end{aligned}$$

Due to (2.36) and the continuity of $G^{*'}$, the right-hand side of (2.38) tends to zero so that

$$(2.39) \quad M_D(y_k^*, \Lambda v_k) \rightarrow 0 \text{ as } k \rightarrow +\infty.$$

By setting $q^* = p^*$ in the right-hand side of (2.27), we obtain the inequality

$$M_R(y_k^*) \leq \frac{1}{2} \left[\langle\langle y_k^* - p^*, G^{*'}(-p^*) - G^{*'}(-y_k^*) \rangle\rangle + h^*(\|y_k^* - p^*\|_*) \right] \rightarrow 0$$

which together with (2.39) yields (2.35). This completes the proof. □

2.5. Particular cases of the estimate (2.25). The majorant $\widetilde{\mathcal{M}}$ depends on $v \in V$ and $y^* \in Y^*$. Since v is known, the question of how to define y^* arises. We explore this important question below.

Let us assume that $p^* \in Q^* \subset Y^*$ and let $\Pi : Y^* \rightarrow Q^*$ be a continuous operator such that $\Pi q^* = q^*$ for all $q^* \in Q^*$. Typically, the form of Q^* is dictated by a priori differentiability properties of the exact solution and the operator Π is defined by some post-processing procedure. If v is known, then one can define its counterpart in the space Y^* via the duality mapping (2.10)(ii):

$$(2.40) \quad \sigma^*(v) = -G'(\Lambda v).$$

By setting $y^* = \Pi \sigma^*(v)$ in (2.25) we obtain a common form for an a posteriori error majorant \mathcal{M} (cf. (1.2)–(1.3)):

$$(2.41) \quad \Phi_\delta(\Lambda e) \leq \mathcal{M}(v) := \widetilde{\mathcal{M}}(v, \Pi \sigma^*(v)).$$

From Theorem 2.4 and the foregoing assumptions it follows that

$$\mathcal{M}(v_k) \rightarrow 0 \text{ if } v_k \rightarrow u \text{ in } V.$$

Three main forms of the estimate (2.41) arise when the set Q^* is defined in accordance with (a), (b) or (c) below:

$$(2.42) \quad \begin{aligned} (a) \quad & Q^* \equiv Y^*; \\ (b) \quad & Q^* \subset Y^*, \quad Q^* \neq Y^*; \quad Q^* \neq Q_l^*; \\ (c) \quad & Q^* = Q_l^*. \end{aligned}$$

Case (a). If $Q^* \equiv Y^*$, then Π is the identity operator, so that

$$M_D(\Pi\sigma^*(v), \Lambda v) = M_D(\sigma^*(v), \Lambda v) = 0$$

and (2.41) yields the estimate

$$(2.43) \quad \Phi_\delta(\Lambda e) \leq M_R(\sigma^*(v))$$

whose right-hand side consists of the term M_R only.

Case (b). Let Q^* be a proper subset of Y^* which contains p^* , and let $Q^* \neq Q_l^*$. Then both terms M_D and M_R can be positive and the corresponding error estimate has the form

$$(2.44) \quad \Phi_\delta(\Lambda e) \leq M_D(\Pi\sigma^*(v), \Lambda v) + M_R(\Pi\sigma^*(v)).$$

Case (c). If $Q^* = Q_l^*$, then $\Pi\sigma^*(v)$ is an element of the set Q_l^* . Therefore,

$$M_R(\Pi\sigma^*(v)) = 0$$

and (2.41) becomes

$$(2.45) \quad \Phi_\delta(\Lambda e) \leq M_D(\Pi\sigma^*(v), \Lambda v).$$

Thus, we have described the basic principles by which one can handle efficiently the construction of various a posteriori error estimates. In subsequent sections we use them on several concrete problems.

3. EXAMPLES

Let Ω be a bounded connected domain in the Euclidean space \mathbb{R}^d with Lipschitz continuous boundary $\partial\Omega$ and let V denote a subspace of the Sobolev space $W^{1,\alpha}(\Omega)$ formed by functions vanishing on $\partial\Omega$ in the usual sense of traces. We set $\Lambda v := \nabla v$ and consider variational problems for the functional

$$J(v, \nabla v) = \int_{\Omega} (g(\nabla v) + f(v)) \, dx.$$

Now G and F are integral functionals whose integrands $g : \mathbb{R}^d \rightarrow \mathbb{R}$ and $f : \mathbb{R} \rightarrow \mathbb{R}$ are convex differentiable functions. As usual, we denote their conjugate functions g^* and f^* , respectively. We identify the spaces Y and Y^* with the Lebesgue spaces $L^\alpha(\Omega, \mathbb{R}^d)$ and $L^{\alpha^*}(\Omega, \mathbb{R}^d)$, $\alpha^* = \frac{\alpha}{\alpha-1}$ and the number $\alpha > 1$ is taken such that the above integral has sense. Lastly, in the considered case

$$\langle y^*, y \rangle := \int_{\Omega} y^* \cdot y \, dx \quad \text{and} \quad \Lambda^* y^* := -\operatorname{div} y^* \in V^*.$$

3.1. **Example 1.** Let $g(y) = \frac{1}{2}Ay \cdot y$, where A is a symmetric real matrix satisfying the conditions

$$(3.1) \quad \nu_1 |\eta|^2 \leq A\eta \cdot \eta \leq \nu_2 |\eta|^2 \quad \forall \eta \in \mathbb{R}^d,$$

for some $\nu_2 \geq \nu_1 > 0$. It is straightforward to check that the functional G is uniformly convex on any ball. The two parts of the error majorant M (cf. (2.12)) are given by the relations

$$(3.2) \quad M_G(y^*, \Lambda v) = \frac{1}{4} \int_{\Omega} (\bar{A}y^* + \nabla v) \cdot (y^* + A\nabla v) \, dx,$$

$$(3.3) \quad M_F(\Lambda^*y^*, v) = \frac{1}{2} \int_{\Omega} (f(v) - y^* \cdot \nabla v) \, dx + \frac{1}{2} \sup_{w \in V} \int_{\Omega} (y^* \cdot \nabla w - f(w)) \, dx,$$

where \bar{A} is the matrix inverse of A . If the function $f^*(-\operatorname{div}y^*)$ is summable, then (3.3) can be estimated by a more symmetric expression

$$(3.4) \quad M_F(v, y^*) \leq \frac{1}{2} \int_{\Omega} (f(v) + f^*(-\operatorname{div}y^*) - y^* \cdot \nabla v) \, dx.$$

In particular, if $f(v) = \frac{\lambda}{2}v^2 + \mu v$, where $\mu \in \mathbb{R}$ and $\lambda \in \mathbb{R}_+$, then $f^*(v^*) = \frac{1}{2\lambda}(v^* - \mu)^2$. In this case, $\alpha = 2$ and for any

$$y^* \in H(\Omega; \operatorname{div}) := \{ \eta^* \in Y^* \mid \operatorname{div}\eta^* \in L^2(\Omega) \}$$

we obtain

$$(3.5) \quad M_F(v, y^*) \leq \frac{1}{4\lambda} \|\lambda v + \operatorname{div}y^* + \mu\|_{\Omega}^2,$$

where $\|\cdot\|_{\Omega}$ denotes the norm in $L^2(\Omega)$. Now both functionals G and F are uniformly convex, and the relation (2.15) holds for

$$\Phi(\nabla e) = \frac{1}{4} \int_{\Omega} A\nabla e \cdot \nabla e \, dx, \quad \varphi(e) = \frac{\lambda}{4} \int_{\Omega} |e|^2 \, dx.$$

As a consequence, we get (2.16) in the form

$$(3.6) \quad \int_{\Omega} A\nabla(v - u) \cdot \nabla(v - u) \, dx + \lambda \|v - u\|_{\Omega}^2 \leq \int_{\Omega} (\bar{A}y^* + \nabla v) \cdot (y^* + A\nabla v) \, dx + \frac{1}{\lambda} \|\lambda v + \operatorname{div}y^* + \mu\|_{\Omega}^2.$$

This estimate deteriorates if f is a linear function, so that for $\lambda = 0$ we should use the majorant \widetilde{M} (see 3.4).

3.2. **Example 2.** Let $g(y) = \frac{1}{\alpha} |y|^\alpha$ and $f(v) = l^*v$, where $\alpha > 1$. Then Problem \mathcal{P} is to minimize

$$I_{\alpha}(v) := \int_{\Omega} \left(\frac{1}{\alpha} |\nabla v|^\alpha + l^*v \right) \, dx$$

over the space V , and Problem \mathcal{P}^* is to maximize

$$I_{\alpha^*}^*(y^*) = -\frac{1}{\alpha^*} \int_{\Omega} |y^*|^{\alpha^*} \, dx$$

over the set

$$Q_l^* = \left\{ y^* \in Y^* := L_{\alpha^*}(\Omega, \mathbb{R}^d) \mid \int_{\Omega} y^* \cdot \nabla w \, dx = \int_{\Omega} l^*w \, dx \quad \forall w \in V \right\}.$$

The functional $G(y) = \int_{\Omega} g(y) dx$ is uniformly convex. For $\alpha \geq 2$ this fact follows from the inequality (see [35])

$$(3.7) \quad \int_{\Omega} \left| \frac{y_1 + y_2}{2} \right|^{\alpha} dx + \int_{\Omega} \left| \frac{y_1 - y_2}{2} \right|^{\alpha} dx \leq \frac{1}{2} \int_{\Omega} (|y_1|^{\alpha} + |y_2|^{\alpha}) dx,$$

which is valid for all $y_1, y_2 \in Y$. Hence, (3.7) implies (2.11) with

$$\Phi(y) = \frac{1}{2^{\alpha}} \int_{\Omega} |y|^{\alpha} dx.$$

One can prove that for $1 < \alpha \leq 2$ this functional is uniformly convex also [22].

By virtue of (3.7) we derive the basic duality estimate

$$(3.8) \quad \frac{1}{\alpha 2^{\alpha}} \int_{\Omega} |\nabla e|^{\alpha} dx \leq \frac{1}{2} (I_{\alpha}(v) - I_{\alpha}^{*}(q^{*})) \quad \forall v \in V, \quad q^{*} \in Q_i^{*},$$

which is, in fact, a particular form of (2.13) for power growth functionals. Further analysis of the duality error estimates for this class of variational problems can be found in [27].

3.3. Example 3. Let $g(\nabla y) = \frac{1}{2} A \nabla y \cdot \nabla y + \psi(|\nabla y|)$, where $\psi : \mathbb{R} \rightarrow \mathbb{R}$ is a convex differentiable function, $f(v) = l^{*}v$ and $l^{*} \in L^2(\Omega)$. In this case, the choice of functional spaces depends on the growth condition for ψ . We assume that this growth is less than quadratic. Then $V = \{v \in W^{1,2}(\Omega) \mid v = 0 \text{ on } \partial\Omega\}$, $\|v\|_V^2 = \int_{\Omega} |\nabla v|^2 dx$ and Y can be identified with the space $L^2(\Omega, \mathbb{R}^d)$. Now the relation (2.40) reads

$$(3.9) \quad \begin{aligned} \text{(i)} \quad & \sigma^{*}(v)(x) = -g'(\nabla v)(x), \\ \text{(ii)} \quad & \nabla v(x) = g^{*'}(-\sigma^{*})(x) \text{ for a.e. } x \in \Omega, \end{aligned}$$

where $g'(y) = Ay + \psi'(|y|)\frac{y}{|y|}$. It is straightforward to prove that the functional G is uniformly convex and that

$$\frac{1}{2} \int_{\Omega} A \nabla(v - u) \cdot \nabla(v - u) dx \leq J(v, \nabla v) - J(u, \nabla u) \quad \forall v \in V.$$

Therefore, (cf. (2.13)) the basic estimate (2.25) holds with $\Phi(\Lambda e) = \frac{1}{4} \|e\|^2$, where

$$\|e\|^2 := \int_{\Omega} A \nabla(v - u) \cdot \nabla(v - u) dx.$$

One can prove that under the above assumptions, the functional G^{*} is Gâteaux differentiable and

$$(3.10) \quad \|G^{*'}(y^{*}) - G^{*'}(\eta^{*})\| \leq c_3 \|y^{*} - \eta^{*}\|_{*},$$

$$(3.11) \quad \langle\langle y^{*} - \eta^{*}, G^{*'}(y^{*}) - G^{*'}(\eta^{*}) \rangle\rangle \leq c_3 \| \eta^{*} - y^{*} \|_{*}^2 \quad c_3 = \nu_1^{-1}.$$

Thus, setting

$$(3.12) \quad h(t) = \frac{\nu_1}{4} t^2, \quad h^{*}(t) = c_3 t^2$$

in (2.31)–(2.32) we deduce the estimate

$$(3.13) \quad \frac{1}{4} \|e\|^2 \leq M_D(y^{*}, \nabla v) + M_R(y^{*}) \quad \forall y^{*} \in Y^{*},$$

where

$$(3.14) \quad M_D(y^*, \nabla v) = \frac{1}{2} \left[\int_{\Omega} \left(g(\nabla v) + g^*(-y^*) + y^* \cdot \nabla v \right) dx + c_4 \int_{\Omega} \left(g^{*'}(-y^*) - \nabla v \right)^2 dx, \right]$$

$$(3.15) \quad M_R(y^*) = - \inf_{w \in V} \int_{\Omega} (c_4 |\nabla w|^2 - \mathcal{R}(y^*)w) dx,$$

$c_4 = \frac{\nu_1}{4}$ and

$$\mathcal{R}(y^*) := \operatorname{div} y^* + l^*.$$

Since

$$(3.16) \quad g(y) + g^*(-y^*) + y^* \cdot y \leq (y - g^{*'}(-y^*)) \cdot (y^* + g'(y)),$$

we see that the term $M_D(y^*, \nabla v)$ vanishes if the duality relations (3.9) hold.

Now, we focus our attention on $M_R(y^*)$. First of all we note that for arbitrary $y^* \in Y^*$, the term $\mathcal{R}(y^*)$ should be understood in the sense of distributions. Therefore, an adequate measure of $M_R(y^*)$ is given by the quantity

$$\|\mathcal{R}(y^*)\|_{(-1)} := \sup_{\substack{w \in V \\ w \neq 0}} \frac{\int_{\Omega} (l^*w - y^* \cdot \nabla w) dx}{\|\nabla w\|_{\Omega}},$$

which is nonnegative and finite for any $y^* \in Y^*$. Indeed, we can estimate the term $M_R(y^*)$ as follows

$$(3.17) \quad \begin{aligned} M_R(y^*) &= \sup_{w \in V} \int_{\Omega} (\mathcal{R}(y^*)w - c_4 |\nabla w|^2) dx \\ &\leq \sup_{t \in \mathbb{R}_+} \left(\|\mathcal{R}(y^*)\|_{(-1)} t - c_4 t^2 \right) \leq \nu_1^{-1} \|\mathcal{R}(y^*)\|_{(-1)}^2. \end{aligned}$$

If $y^* \in H(\Omega; \operatorname{div})$, then $\mathcal{R}(y^*) \in L^2(\Omega)$ and

$$(3.18) \quad \|\mathcal{R}(y^*)\|_{(-1)} \leq C(\Omega) \|\mathcal{R}(y^*)\|_{\Omega},$$

where $C(\Omega)$ is a constant in the Poincaré–Friedrichs inequality

$$\|w\|_{\Omega} \leq C(\Omega) \|\nabla w\|_{\Omega} \quad \forall w \in V.$$

Whence, in this case we can estimate $M_R(y^*)$ via the L_2 -norm of the residual

$$(3.19) \quad M_R(y^*) \leq C^2(\Omega) \nu_1^{-1} \|\mathcal{R}(y^*)\|_{\Omega}^2.$$

3.4. Example 4. Let $g(y) = \frac{1}{2}Ay \cdot y$ and $f(v) = l^*v$. This simple and at the same time important example deserves special consideration. We use it to show the performance of our method in a more transparent form.

In the considered case, \bar{A} is a symmetric matrix defined as in Example 1, V , Y and Y^* are defined as in Example 3, and $g^*(y^*) = \frac{1}{2}\bar{A}y^* \cdot y^*$. It is easily verified using elementary manipulations that the basic duality inequality yields the following

estimate

(3.20)

$$\begin{aligned} & \frac{1}{2} \int_{\Omega} A \nabla(v - u) \cdot \nabla(v - u) \, dx \\ &= J(v, \nabla v) - J(u, \nabla u) = J(v, \nabla v) - \sup \mathcal{P}^* \\ &\leq \int_{\Omega} \left(\frac{1}{2} (A \nabla v \cdot \nabla v + \bar{A} y^* \cdot y^*) + l^* v + \frac{1}{2} (\bar{A} q^* \cdot q^* - \bar{A} y^* \cdot y^*) \right) dx, \end{aligned}$$

where $q^* \in Q_i^*$ and $y^* \in Y^*$. Since

$$\bar{A} q^* \cdot q^* - \bar{A} y^* \cdot y^* = \bar{A} (q^* - y^*) \cdot (q^* - y^*) - 2 \bar{A} y^* \cdot (y^* - q^*)$$

and q^* meets the integral identity

$$\int_{\Omega} q^* \cdot \nabla w \, dx = \int_{\Omega} l^* w \, dx \quad \forall w \in V,$$

we rewrite (3.20) as

$$\begin{aligned} (3.21) \quad \frac{1}{2} \| e \|^2 &\leq \int_{\Omega} \left(\frac{1}{2} A \nabla v \cdot \nabla v + \frac{1}{2} \bar{A} y^* \cdot y^* + \nabla v \cdot y^* \right) dx \\ &+ \int_{\Omega} (\nabla v + \bar{A} y^*) \cdot (q^* - y^*) \, dx \\ &+ \frac{1}{2} \int_{\Omega} \bar{A} (q^* - y^*) \cdot (q^* - y^*) \, dx. \end{aligned}$$

Now we apply the inequality (2.23) with

$$H(y) = \frac{\beta}{2} \int_{\Omega} A y \cdot y \, dx, \quad H^*(y) = \frac{1}{2\beta} \int_{\Omega} \bar{A} y^* \cdot y^* \, dx, \quad \beta > 0$$

to the second integral in the right-hand side of (3.21). This results in the estimate

$$(3.22) \quad \| e \|^2 \leq (1 + \beta) m_D^2(y^*, \nabla v) + (1 + \beta^{-1}) m_R^2(y^*),$$

where for the sake of convenience we have introduced the terms

$$m_D^2(y^*, \nabla v) = \int_{\Omega} (A \nabla v + y^*) \cdot (\nabla v + \bar{A} y^*) \, dx$$

and

$$m_R^2(y^*) = \inf_{q^* \in Q_i^*} \int_{\Omega} \bar{A} (y^* - q^*) \cdot (y^* - q^*) \, dx.$$

By taking the infimum in the right-hand side of (3.22) over the parameter β , we arrive at the final estimate

$$(3.23) \quad \| e \| \leq m_D(y^*, \nabla v) + m_R(y^*).$$

To obtain computationally more attractive forms of $m_R(y^*)$, we note that

$$(3.24) \quad \frac{1}{2} m_R^2(y^*) = - \inf_{w \in V} \int_{\Omega} \left(\frac{1}{2} A \nabla w \cdot \nabla w - \mathcal{R}(y^*) w \right) dx$$

(cf. (2.32)). From (3.24) by analogy with (3.17)–(3.18) we obtain

$$\begin{aligned} m_R(y^*) &\leq \nu_1^{-1/2} \|\mathcal{R}(y^*)\|_{(-1)} && \forall y^* \in Y^*, \\ m_R(y^*) &\leq C(\Omega, A) \|\mathcal{R}(y^*)\|_\Omega && \forall y^* \in H(\Omega; \text{div}), \end{aligned}$$

where $C(\Omega, A)$ is a constant in the inequality

$$\int_\Omega |w|^2 dx \leq C^2(\Omega, A) \int_\Omega A \nabla w \cdot \nabla w dx \quad \forall w \in V.$$

It is worth remarking that (3.23) holds for any $y^* \in Y^*$. This freedom can be utilized to get the most rigorous error bound (see [31]).

We conclude this consideration with comments about the relationship between duality and projection error estimates. Let V_h be a set of finite-dimensional spaces embedded in V which satisfy the usual conditions (see, e.g., [8, 20]) required to guarantee that the corresponding Galerkin approximations u_h tend to u as $h \rightarrow 0$.

Since u and u_h are minimizers of Problem \mathcal{P} and of its discrete analog, respectively, we have

$$\begin{aligned} J(u_h, \nabla u_h) - J(u, \nabla u) &\leq J(v_h, \nabla v_h) - J(u, \nabla u) \\ &= \frac{1}{2} \int_\Omega A \nabla(v_h - u) \cdot \nabla(v_h - u) dx \quad \forall v_h \in V_h. \end{aligned}$$

Therefore, (3.20) yields the inequality

$$\begin{aligned} (3.26) \quad &\int_\Omega A \nabla(u_h - u) \cdot \nabla(u_h - u) dx \\ &\leq \int_\Omega A \nabla(v_h - u) \cdot \nabla(v_h - u) dx \quad \forall v_h \in V_h, \end{aligned}$$

which gives

$$(3.27) \quad \|u - u_h\|_V \leq c_5 \inf_{v_h \in V_h} \|u - v_h\|_V, \quad c_5 = \sqrt{\nu_2/\nu_1}.$$

This inequality (also known as Cea Lemma—see, e.g., [8]) means that an upper bound of the error is given by the distance (in the space V) between the exact solution u of Problem \mathcal{P} and the set V_h containing the Galerkin approximation u_h .

Let us set $v = u_h$, $y^* = y_h^* := -\bar{A} \nabla u_h$ and apply (3.23). Since $m_D(y_h^*, \nabla u_h) = 0$ we obtain

$$\begin{aligned} (3.28) \quad &\int_\Omega A \nabla(u_h - u) \cdot \nabla(u_h - u) dx \leq m_R^2(y_h^*) \\ &\leq \int_\Omega \bar{A}(y_h^* - q^*) \cdot (y_h^* - q^*) dx \quad \forall q^* \in Q_l^*. \end{aligned}$$

This inequality yields the estimate

$$(3.29) \quad \|u - u_h\|_V \leq c_6 \inf_{q^* \in Q_l^*} \|y_h^* - q^*\|_*,$$

where $c_6 = \nu_1^{-1}$. The estimate (3.29) is, in a sense, dual to (3.27). It shows that an upper bound of the error is given by the distance (in the dual space Y^*) between y_h^* (which is a dual counterpart of the Galerkin approximation u_h) and the set Q_l^* containing the exact solution p^* of Problem \mathcal{P}^* .

4. CONNECTION WITH OTHER METHODS

In this section, we apply the general scheme presented in subsection 2.5 to a class of variational problems, and we show that the a posteriori error estimates (A), (B) and (C) can be derived from the duality error estimate (see also [26, 28]). For this purpose we take the problem

$$(4.1) \quad \inf_{v \in V} J(v, \nabla v), \quad J(v, \nabla v) = \int_{\Omega} (g(\nabla v) + l^*v) \, dx,$$

that has been analyzed in Example 3.

4.1. Residual error estimates. The Euler-Lagrange equation associated with problem (4.1) is

$$\operatorname{div} p^* + l^* = 0, \quad p^* = -g'(\nabla u).$$

Hence, for any $v \in V$ the function

$$R(v) := \mathcal{R}(\sigma^*(v)) = \operatorname{div} \sigma^*(v) + l^*$$

is the residual of this equation if $\sigma^*(v)$ is defined in accordance with the duality relations

$$(4.2) \quad \sigma^*(v) = -g'(\nabla v).$$

By setting $y^* = \sigma^*(v)$ in (3.13) (cf. (2.42)(a)) we obtain

$$(4.3) \quad \begin{aligned} \frac{1}{4} \| e \|^2 &\leq M_R(\sigma^*(v)) \\ &= - \inf_{w \in V} \int_{\Omega} (c_4 | \nabla w |^2 + \sigma^*(v) \cdot \nabla w + l^*w) \, dx \\ &\leq \frac{1}{\nu_1} \| R(v) \|_{(-1)}^2. \end{aligned}$$

Thus, we see that such choice of y^* leads to an a posteriori error estimate whose right-hand side is given by the negative norm of the residual. If v is a finite element approximation of u , then the explicit calculation of $\| R(v) \|_{(-1)}$ is based on the Clement’s interpolation theorem [9] and the corresponding error estimate is a sum of element residuals and interelement jumps (see, e.g., [36]).

4.2. Error estimates based on gradient averaging. In many cases, it is known a priori that p^* possesses additional differentiability properties and, therefore, belongs to a proper subset Q^* of Y^* . Then, it is natural to expect that any “good” approximation of p^* should also be an element of Q^* . An obvious way to get the one is to find a regularized function $\tilde{\sigma}^*$ by mapping $\sigma^*(v)$ on Q^* . For this purpose we need a computationally inexpensive continuous operator $\Pi : Y^* \rightarrow Q^*$. Operators with these properties are known in the literature as *gradient (stress) averaging* or *recovery* operators (see, e.g., [5, 6, 37]). Below we justify recovery based methods via the duality theory and show that the duality error majorant (3.13) yields the estimate (4.4) whose main part coincides with the estimate of the group (B).

Proposition 4.1. *If $\tilde{\sigma}^* = \Pi \sigma^*(v)$, where $\Pi : Y^* \rightarrow Q^*$ is a recovery operator, then*

$$(4.4) \quad \frac{1}{4} \| e \|^2 \leq c_7 \| \sigma^*(v) - \tilde{\sigma}^* \|_{\Omega}^2 + M_R(\tilde{\sigma}^*).$$

Proof. By virtue of (3.16) we present the term M_D as

$$M_D(\tilde{\sigma}^*, \nabla v) = \frac{1}{2} \left[\int_{\Omega} (\nabla v - g^{*'}(-\tilde{\sigma}^*)) \cdot (\tilde{\sigma}^* + g'(\nabla v)) \, dx + c_4 \int_{\Omega} (g^{*'}(-\tilde{\sigma}^*) - \nabla v)^2 \, dx \right].$$

Now we recall that

$$-\sigma^*(v) = g'(\nabla v) \quad \text{and} \quad \nabla v = g^{*'}(-\sigma^*(v)),$$

so that

$$(4.5) \quad M_D(\tilde{\sigma}^*, \nabla v) = \frac{1}{2} \left[\int_{\Omega} (g^{*'}(-\sigma^*(v)) - g^{*'}(-\tilde{\sigma}^*)) \cdot (\tilde{\sigma}^* - \sigma^*(v)) \, dx + c_4 \int_{\Omega} (g^{*'}(-\tilde{\sigma}^*) - g^{*'}(\sigma^*(v)))^2 \, dx \right].$$

Making use of (3.10), (3.11) and (4.5) we obtain

$$(4.6) \quad M_D(\tilde{\sigma}^*, \nabla v) \leq \frac{c_3}{2} (1 + c_3 c_4) \|\sigma^*(v) - \tilde{\sigma}^*\|_{\Omega}^2.$$

Now (4.4) follows from (3.13) and (4.6) if set $c_7 = \frac{5c_3}{8}$. □

Remark 4.1. The second term in the right-hand side of (4.4) can be estimated as

$$M_R(\tilde{\sigma}^*) = - \inf_{w \in V} \int_{\Omega} (c_4 |\nabla w|^2 + (\tilde{\sigma}^* - p^*) \cdot \nabla w) \, dx \leq c_3 \|\tilde{\sigma}^* - p^*\|_{\Omega}^2.$$

Hence, under the usual assumption that $\|\tilde{\sigma}^* - p^*\|_{\Omega}$ is negligible with respect to $\|\tilde{\sigma}^* - \sigma^*(v)\|_{\Omega}$, we arrive at the recovery based error estimate

$$\|e\|^2 \leq c_8 \|\sigma^*(v) - \tilde{\sigma}^*\|_{\Omega}^2,$$

where $c_8 = \frac{5}{2}c_3$. However, it is quite possible that in some cases the above assumption is not true. Therefore, a *rigorous* a posteriori estimate for averaged approximations has the form (4.4) and must include the term M_R (see [31] for a more detailed discussion of this subject and for numerical examples).

Remark 4.2. The efficiency of the above estimates strongly depends on the choice of an operator Π that must be mathematically stable and computationally inexpensive. A study of concrete operators and their mathematical properties suggests an important but separate problem of approximation theory which, however, is beyond the frame of the present paper. At this point we refer to, e.g., [1, 6, 10, 11, 37], where these questions are addressed for finite element methods.

4.3. Error estimates based on data equilibration. Since $p^* \in Q_i^*$, it is possible to set $Q^* = Q_i^*$ (cf. (2.42)(c)). The set Q_i^* consists of functions q^* satisfying the equation $\text{div} q^* + f = 0$ that often appears in applications as the *equilibrium* equation. Hence, a mapping $\Pi : Y^* \rightarrow Q_i^*$ is naturally called an *equilibration* operator. Let us define a function

$$\hat{\sigma}^* = \Pi \sigma^*(v) \in Q_i^*.$$

Then $M_R(\hat{\sigma}^*) = 0$ and by analogy with (4.5) we obtain

$$\begin{aligned}
 M_D(\hat{\sigma}^*, \nabla v) &= \frac{1}{2} \left[\int_{\Omega} (g^{*'}(-\sigma^*(v)) - g^{*'}(-\hat{\sigma}^*)) \cdot (\hat{\sigma}^* - \sigma^*(v)) \, dx \right. \\
 &\quad \left. + c_4 \int_{\Omega} (g^{*'}(-\hat{\sigma}^*) - g^{*'}(\sigma^*(v)))^2 \, dx \right] \\
 &\leq \frac{1}{2} \int_{\Omega} \left(c_3 |\hat{\sigma}^* - \sigma^*(v)|^2 + c_4 |g^{*'}(-\hat{\sigma}^*) - \nabla v|^2 \right) \, dx
 \end{aligned}$$

which together with (3.13) implies the estimate

$$(4.7) \quad \frac{1}{2} \| e \|^2 \leq \int_{\Omega} \left(c_3 |\hat{\sigma}^* + g'(\nabla v)|^2 + c_4 |g^{*'}(-\hat{\sigma}^*) - \nabla v|^2 \right) \, dx.$$

The latter gives the general form of the equilibration type error estimator for the considered class of problems. It should be emphasized that the right-hand side of (4.7) is, in fact, a measure of the error in the duality relations (3.9).

The function $\hat{\sigma}^*$ can be also used in a different way. Indeed, let us substitute $y^* = \sigma^*(v)$ in (3.13). Then,

$$\begin{aligned}
 M_R(\sigma^*(v)) &= - \inf_{w \in V} \int_{\Omega} (c_4 |\nabla w|^2 - \mathcal{R}(\sigma^*(v))w) \, dx \\
 &= c_3 \inf_{\eta^* \in Q_l^*} \int_{\Omega} |\sigma^*(v) - \eta^*|^2 \, dx
 \end{aligned}$$

and we obtain a simple estimate

$$(4.8) \quad \| e \|^2 \leq c_9 \int_{\Omega} |\sigma^*(v) - \hat{\sigma}^*|^2 \, dx,$$

where $c_9 = 4c_3$. Note also that for quadratic functionals a similar estimate (but with a smaller constant) follows straightforwardly from (3.29).

It should be remarked that, in general, an operator $\Pi : Y^* \rightarrow Q_l^*$ is difficult to construct. However, there is one particular case when the estimate (4.8) can be applied fairly easily. Suppose that $n = 2$ and that we know a function σ_0^* satisfying the equation $\text{div} \sigma_0^* + l^* = 0$ exactly (the latter assumption may be rather obligatory if $l^* \neq \text{const}$). Then any function $y^* \in Q_l^*$ can be presented via $\sigma_0^* = (\sigma_{01}^*, \sigma_{02}^*)$ and a stream function $\theta \in V$:

$$(4.9) \quad y_1^* = \sigma_{01}^* - \frac{\partial \theta}{\partial x_2}, \quad y_2^* = \sigma_{02}^* + \frac{\partial \theta}{\partial x_1}.$$

By substituting (4.9) into (4.8) we get the estimate

$$\| e \|^2 \leq c_9 \inf_{\theta \in V} \int_{\Omega} \left(\left| \sigma_1^* - \sigma_{01}^* + \frac{\partial \theta}{\partial x_2} \right|^2 + \left| \sigma_2^* - \sigma_{02}^* - \frac{\partial \theta}{\partial x_1} \right|^2 \right) \, dx.$$

For a class of nonlinear 2D-problems in continuum mechanics, this type of a posteriori error estimate was used in, e.g., [22].

4.4. Concluding remarks. Finally, we add some remarks on the scope of the methods analyzed in this section.

The general estimate (3.13) is valid for all pairs of functions $(v, y^*) \in V \times Y^*$. Various estimates can be derived from (3.13) if the dual variable y^* is defined by means of the approximate solution v , duality relations (DR) and a post-processing operator $\Pi : Y^* \rightarrow Q^* \in Y^*$. This procedure can be presented diagrammatically as

$$v \xrightarrow{\text{DR}} \sigma^*(v) \xrightarrow{\Pi} y^* \Rightarrow \widetilde{M}(v, y^*).$$

If Π is the identity operator (i.e., if no post-processing is used), then $y^* = \sigma^*(v)$ and $\widetilde{M}(v, y^*)$ yields the residual error estimate. Other methods are related to some post-processing of $\sigma^*(v)$. If Π is an averaging (smoothing) operator, then this way leads to error estimators of the group (B) (e.g., to the so-called ZZ-estimator [37]). If Π is a procedure that puts $\sigma^*(v)$ in equilibrium, then $\widetilde{M}(v, \Pi\sigma^*(v))$ coincides with an estimator of the group (C).

It should be emphasized that the above scheme is very flexible. It can be applied to a wide variety of *ad hoc* operators Π and, thus, provides a simple way for taking into account any a priori information on such properties of the exact solution as higher differentiability, boundedness, localization of singularities, etc.

5. ACKNOWLEDGMENT

The author wishes to thank the reviewer for his constructive remarks.

REFERENCES

- [1] M. Ainsworth and J. T. Oden, *A unified approach to a posteriori error estimation using element residual methods*, Numer. Math., **65** (1993), 23–50. MR **95a**:65185
- [2] J. P. Aubin and H. G. Burchard, *Some aspects of the method of hypercircle applied to elliptic variational problems*, In B. Hubbard, editor, Numerical Solutions of Partial Differential Equations – II, SYNSPADE 1970. Academic Press, New York, London, 1971. MR **44**:2359
- [3] I. Babuška and W. C. Rheinboldt, *A-posteriori error estimates for the finite element method*, Internat. J. Numer. Meth. Engrg., **12** (1978), 1597–1615.
- [4] I. Babuška and W. C. Rheinboldt, *Error estimates for adaptive finite element computations*, SIAM J. Numer. Anal., **15**(4) (1978), 736–754. MR **58**:3400
- [5] I. Babuška and W. C. Rheinboldt, *A posteriori error analysis of finite element solutions for one-dimensional problems*, SIAM J. Numer. Anal., **18**(3) (1981), 565–589. MR **82j**:65082
- [6] I. Babuška and R. Rodriguez, *The problem of the selection of an a posteriori error indicator based on smoothing techniques*, Internat. J. Numer. Meth. Engrg., **36** (1993), 539–567.
- [7] R. E. Bank and A. Weiser, *Some a posteriori error estimators for elliptic partial differential equations*, Math. Comp., **44** (1985), 283–301. MR **86g**:65207
- [8] P. Ciarlet, *The finite element method for elliptic problems*, North-Holland, Amsterdam, 1978. MR **58**:25001
- [9] Ph. Clément, *Approximations by finite element functions using local regularization*, RIAO Anal. Numer., **2** (1975), 77–84.
- [10] M. Crouzeix and V. Thomée, *The stability in L_p and W_p^1 of the L_2 -projection onto finite element function spaces*, Math. Comp., **48** (1987), 521–532. MR **88f**:41016
- [11] J. Douglas, Jr. T. Dupont and L. Wahlbin, *The stability in L^q of the L^2 projection into finite element function spaces*, Numer. Math., **23** (1975), 193–197. MR **52**:4669
- [12] I. Ekeland and R. Temam, *Convex analysis and variational problems*, North-Holland, Amsterdam, Oxford, New-York, 1976. MR **57**:3931b
- [13] K. Eriksson and C. Johnson, *An adaptive finite element method for linear elliptic problems*, Math. Comp., **50** (1988), 361–383. MR **89c**:65119
- [14] R. Glowinski, *Numerical methods for nonlinear variational problems*. Springer-Verlag, New-York, Berlin, Heidelberg, Tokyo, 1984. MR **86c**:65004
- [15] C. Johnson and P. Hansbo, *Adaptive finite element methods in computational mechanics*, Comput. Methods Appl. Mech. Engrg, **101** (1992), 143–181. MR **93m**:65157
- [16] C. Johnson, Yi-Yong Nie, and V. Thomée, *An a posteriori error estimate and adaptive timestep control for a backward Euler discretization of a parabolic problem*, SIAM J. Numer. Anal., **27** (1990), 277–291. MR **91g**:65199
- [17] D. W. Kelly, *The self-equilibration of residuals and complementary a posteriori error estimates in the finite element method*, Internat. J. Numer. Meth. Engrg., **20** (1984), 1491–1506. MR **85h**:73039
- [18] P. Ladevèze and D. Leguillon, *Error estimate procedure in the finite element method and applications*, SIAM J. Numer. Anal., **20**(3) (1983), 485–509. MR **84g**:65150

- [19] P. Ladevèze, J.-P. Pelle, and Ph. Rougeot, *Error estimation and mesh optimization for classical finite elements*, Engineering Computations, **8** (1991), 69–80. MR **93f**:73100
- [20] S. G. Mikhailin, *Variational methods in mathematical physics*. Pergamon, Oxford, 1964. MR **30**:2712
- [21] S. G. Mikhailin, *Error Analysis in Numerical Processes*. Wiley, New York, 1991. MR **92j**:65002
- [22] P. P. Mosolov and V. P. Myasnikov, *Mechanics of rigid plastic media*. Nauka, Moscow, 1981 (in Russian). MR **84e**:73029
- [23] J. T. Oden, L. Demkowicz, W. Rachowicz, and T. A. Westermann, *Towards a universal h-p adaptive finite element strategy. Part 2. A posteriori error estimation*, Comput. Methods Appl. Mech. Engrg., **77** (1989), 113–180. MR **91m**:76073
- [24] S. I. Repin, *A priori error estimates of variational-difference methods for Hencky plasticity problems*, Zapiski Nauchnykh Seminarov POMI, **221** (1995), 226–234. (in Russian, English translation in the Journal of Mathematical Sciences 1995). MR **97c**:73029
- [25] S. I. Repin, *Errors of finite element methods for perfectly elasto-plastic problems*, Math. Models Meth. Appl. Sci., **6**(5) (1996), 587–604. MR **97f**:73049
- [26] S. I. Repin, *A posteriori error estimation for nonlinear variational problems by duality theory*, Zapiski Nauchnykh Seminarov POMI, **242** (1997), 196–208. MR **99e**:49049
- [27] S. I. Repin, *A posteriori error estimates for approximate solutions of variational problems with power growth functionals*, Zapiski Nauchnykh Seminarov POMI, **249** (1997), 244–255.
- [28] S. I. Repin, *A posteriori error estimates for approximate solutions of variational problems*, In the Proceedings of the Second European Conference on Numerical Mathematics and Advanced Applications, Heidelberg 1997. VSP, Utrecht (1999).
- [29] S. I. Repin, *A posteriori error estimates for approximate solutions of variational problems with strongly convex functionals*, Problems of Mathematical Analysis, **17** (1997), 199–226. (in Russian, English translation in the Journal of Mathematical Sciences 1998).
- [30] S. I. Repin, *A posteriori estimates of the accuracy of variational methods in problems with nonconvex functionals*, Algebra i Analiz, **11** (1999) (in Russian).
- [31] S. I. Repin, *A unified approach to a posteriori error estimation by duality error majorants*, Mathematics and Computers in Simulation (1999).
- [32] S. I. Repin and G. A. Seregin, *Error estimates for stresses in the finite element analysis of the two-dimensional elasto-plastic problems*, Internat. J. Engrg. Sci., **33**(2), (1995), 255–268. MR **96c**:73068
- [33] S. I. Repin and L. S. Xanthis, *A posteriori error estimation for elasto-plastic problems based on duality theory*, Comput. Methods in Appl. Mech. Engrg., **138** (1996), 317–339. MR **97k**:73032
- [34] S. I. Repin and L. S. Xanthis, *A posteriori error estimation for nonlinear variational problems*, C. R. Acad. Sci. Paris, Serie I., **324** (1997), 1169–1174. MR **98d**:49043
- [35] S. L. Sobolev, *Some Applications of Functional Analysis in Mathematical Physics*. Izdt. Leningrad. Gos. Univ., Leningrad, 1955. in Russian. English translation in Translation of Mathematical Monographs, Volume 90 American Mathematical Society, Providence, Rhode Island. MR **92e**:46067
- [36] R. Verfürth, *A review of a posteriori error estimation and adaptive mesh-refinement techniques*. Wiley, Teubner, 1996.
- [37] O. C. Zienkiewicz and J. Z. Zhu, *A simple error estimator and adaptive procedure for practical engineering analysis*, Internat. J. Numer. Meth. Engrg., **24** (1987), 337–357. MR **87m**:73055
- [38] O. C. Zienkiewicz and J. Z. Zhu, *The superconvergent patch recovery (SPR) and adaptive finite element refinement*, Comput. Methods Appl. Mech. Engrg., **101** (1992) 207–224. MR **93h**:73042

DEPARTMENT OF APPLIED MATHEMATICS ST. PETERSBURG STATE TECHNICAL UNIVERSITY
195251, ST. PETERSBURG, RUSSIA

E-mail address: repin@mat.amd.stu.neva.ru