SIAM J. NUMER. ANAL.
Vol. 32, No. 6, pp. 1778–1807, December 1995

ON THE FINITE ELEMENT METHOD FOR MIXED VARIATIONAL INEQUALITIES ARISING IN ELASTOPLASTICITY*

WEIMIN HAN† AND B. DAYA REDDY‡

Abstract. We analyze the finite-element method for a class of mixed variational inequalities of the second kind, which arises in elastoplastic problems. An abstract variational inequality, of which the elastoplastic problems are special cases, has been previously introduced and analyzed [B. D. Reddy, *Nonlinear Anal.*, 19 (1992), pp. 1071–1089], and existence and uniqueness results for this problem have been given there. In this contribution the same approach is taken; that is, finite-element approximations of the abstract variational inequality are analyzed, and the results are then discussed in further detail in the context of the concrete problems. Results on convergence are presented, as are error estimates. Regularization methods are commonly employed in variational inequalities of this kind, in both theoretical and computational investigations. We derive a posteriori error estimates which enable us to determine whether the solution of a regularized problem can be taken as a sufficiently accurate approximation of the solution of the original problem.

Key words. elastoplastic problems, mixed variational inequalities, finite-element method, convergence, error estimates, regularization method, a posteriori error estimates

AMS subject classifications. 65N30, 73V05

1. Introduction.

Mixed finite-element approximations play a central role in a variety of problems of continuum mechanics. Perhaps the most prominent example concerns finite-element methods for problems with the constraint of incompressibility, such as the Stokes or Navier–Stokes equations for viscous Newtonian fluids, and analogous problems in elasticity. There exists a large body of literature devoted to the development of stable and convergent finite-element schemes for this class of problems; comprehensive surveys may be found in the monographs by Brezzi and Fortin [7] and by Girault and Raviart [14].

Another popular mixed problem which arises particularly in the context of elasticity is that obtained from the Hellinger–Reissner variational principle. In this problem it is not a constraint which produces a mixed or saddle-point problem; rather, such a problem arises from the fact that both the stress and the displacement are treated as unknown variables. Details of finite-element approximations of this class of problems may be found in the work by Brezzi and Fortin [7], as well as in a number of papers devoted to this subject (see, for example, Johnson and Mercier [25], Arnold and coworkers [1], [2], Arnold and Falk [3], Pitkäranta and Stenberg [29], and Stenberg [38], [39]).

There is now a sizeable literature on the numerical approximation of variational inequalities (see, for example, the works by Glowinski et al. [16] and by Hlaváček et al. [21]), which includes investigations of variational inequalities arising in plasticity. Analyses of finite-element approximations of the elastoplastic problem have enjoyed limited but steady attention, in contrast to the voluminous literature devoted to computational and algorithmic aspects of this problem. Havner and Patel [19] and Jiang [22] analysed approximations of the so-called rate problem; this is an elliptic

---

* Received by the editors September 24, 1993; accepted for publication (in revised form) March 15, 1994.

† Department of Mathematics, University of Iowa, Iowa City, Iowa 52242 (whan@math.uiowa.edu).
The work of this author was supported by the Old Gold Fellowship of the University of Iowa.

‡ Department of Mathematics and Applied Mathematics, University of Cape Town, 7700 Rondebosch, South Africa (bdr@maths.uct.ac.za).
The work of this author was supported by the Foundation for Research Development.

variational inequality in which the primary unknowns are the velocity, rather than the displacement, and the plastic multiplier. Reddy and Griffin [32] and Han [17] have considered finite-element approximations of the holonomic or time-independent problem which arises as a typical step when time discretization is introduced in the full problem. These works are all displacement based. Mercier [26] and Oden and Whiteman [27] have both studied finite-element approximations of versions of the closely related Hencky problem of plasticity. Johnson [23] has considered a formulation of the elastoplasticity problem in which stress is the primary variable and has derived error estimates for the fully discrete (that is, time-and-space) problem (see also related work by Hlaváček [20] and a summary account in [21]).

With regard to work on *mixed* variational problems in the form of variational inequalities, little has appeared. Johnson [24] has considered fully discrete finite-element approximations in the context of plasticity, and Brezzi et al. [9] have treated finite-element approximations of the time-independent Hencky problem for elastoplastic plates. Brezzi et al. [8] studied problems which arise in the context of the obstacle problem for a membrane and the unilateral contact problem. The variational inequalities in all these works arise as a result of the problems being posed on convex subsets; that is, these are variational inequalities of the first kind.

In an earlier work, Reddy [30] has considered the problem of mixed variational problems which take the form of variational inequalities. This work was motivated by mixed variational problems which arise in elastoplasticity; these are mixed problems either because of the constraint of plastic incompressibility or because the problem is of Hellinger–Reissner type, so that the stress is treated explicitly along with the displacement and plastic strain. Furthermore, the problems take the form of variational inequalities of the second kind; that is, these are inequalities because of the presence of a nondifferentiable functional. The issues of existence and uniqueness of solutions to these problems have been addressed in [30].

The aim of this contribution is to return to that work and to consider finite-element approximations. The problem is of some importance in engineering applications, and the particular model treated here forms the basis for one approach to large-scale finite element codes for the simulation of elastoplastic behavior [12], [33]. The variational inequality considered in the present work is time independent and arises typically either when time discretization is introduced into the time-dependent problem or alternatively when the applied forces vary linearly with time—the problem of proportional loading—so that the problem reduces to one which is time independent. To place matters in proper perspective, it may be worth mentioning that the problem considered here is a more general version of the Hencky problem, which has been the subject of much investigation in recent times (see [40] and references therein), and which differs from the present problem in that it applies to isotropically elastic, perfectly plastic materials; neither of these restrictions are present in the problem considered here.

While the problem treated here is motivated by applications in elastoplasticity, there are clearly other areas in which it would be of interest, for example, problems involving frictional contact [11]. Furthermore, as indicated earlier, a study of finite-element approximations of mixed variational inequalities of the second kind, that is, those involving a nondifferentiable functional, appears to be lacking.

This work is organized as follows. In §2 we give full details of the elastoplastic problem. In §3, we present some of the mixed formulations for the elastoplastic problems; these come about due to the constraint of plastic incompressibility and/or as a result of inclusion of the stress as a variable (the so-called Hellinger–Reissner formulation, in the context of elasticity). These problems are special cases of an abstract problem, which is then formulated in §4; the main results of [30], which concern existence and uniqueness of solutions to these problems, are reviewed here for use in subsequent sections.

We consider finite-element approximations for the abstract mixed variational inequality problem in §5. We prove some convergence results, together with error estimates. In §6 we apply the results of §5 to the elastoplastic problems formulated in §4. We elaborate on the error estimates of finite-element approximations.

The main result on the solvability of the abstract mixed variational inequality is proved in [30] by introducing a regularizing sequence. In applications, the regularizing sequence technique is also used for numerical computations. A regularization method depends on a small parameter $\varepsilon > 0$, and convergence is obtained when ε goes to 0. However, as $\varepsilon \to 0$, the conditioning of a regularized problem deteriorates. So there is a tradeoff in the selection of the regularization parameter. Theoretically, to obtain more accurate approximations, we need to use smaller values of ε. On the other hand, if ε is too small, the numerical solution of the regularized problem cannot be computed accurately. Thus, it is highly desirable to have a posteriori error estimates that can provide computable error bounds once solutions of the regularized problems have been found. We derive such a posteriori error estimates in §7.

2. The elastoplastic problem. We consider the problem of quasistatic behaviour of an elastoplastic body which occupies a bounded domain $\Omega \subset R^d$ with Lipschitz boundary Γ. The plastic behaviour of the material is assumed to be describable within the classical framework of a convex yield surface coupled with the normality law. We adopt the equivalent form of the flow law in which the dissipation function, rather than the yield function, is employed. This formulation has been studied in some detail both theoretically and computationally in the works [31], [33].

Full details of the formulation considered here are presented in [30]. As indicated in that work, this formulation may be arrived at by approximating rates, for example, by an Euler backward difference. The effect of this assumption would be that the rate problem is approximated by a sequence of incremental problems, in the sense that it is required to determine the response of the body to forces at time $t_0 + \Delta t$, given the complete state of the body at time t_0. The boundary-value problem which we consider arises in a typical time step.

Alternatively, the problem considered here arises when there is proportional loading; that is, the applied forces vary linearly with time, so that a time discretization is rendered unnecessary. This is in fact similar to the formulation adopted in [17], [32].

The elastoplastic material under consideration is assumed to undergo nonlinear kinematic hardening; the nonlinear term takes the form of an exponential decay, and is one which is in current use in numerical treatments of this class of problems (see, for example, [36]). The assumption of a hardening material, apart from the fact that it represents realistic material behaviour, serves also to allow for a complete analysis within a Sobolev space framework, the special case of perfect plasticity requiring that the displacements be sought in the space $BD(\Omega)$ of functions of bounded deformation (see, for example, the text [40] and references therein).

Of special interest here is the classical assumption of no volume change accompanying plastic deformation. This is an assumption which is conventionally accommodated by expressing the yield condition in terms of the stress deviator. We treat this constraint explicitly through the introduction of a Lagrange multiplier.

Under these circumstances, the following equations govern the problem [30]:
(1) the equilibrium equation

$$\operatorname{div}\sigma + b = 0; \tag{2.1}$$

(2) the constitutive equations

$$\sigma = C(\epsilon - p), \tag{2.2}$$
$$\chi \equiv \sigma^D - \sigma_0^D \in \partial D(p); \tag{2.3}$$

(3) the strain-displacement relation

$$\epsilon(u) = \frac{1}{2}(\nabla u + (\nabla u)^T); \tag{2.4}$$

(4) and the condition of plastic incompressibility

$$\operatorname{tr} p := I \cdot p = 0 \tag{2.5}$$

or $p_{kk} = 0$. Here and henceforth summation is implied on repeated indices, unless otherwise stated. Equations (2.1)–(2.5) are required to hold on Ω; we take the boundary condition to be

$$u = 0 \text{ on } \Gamma. \tag{2.6}$$

In the above, σ denotes the stress tensor, $\sigma^D := \sigma - \frac{1}{d}(\operatorname{tr}\sigma)I$ denotes the stress deviator, b is the body force, ϵ is the strain tensor, u is the displacement vector, and p is the plastic strain tensor. The subdifferential ∂D of D is defined to be the set

$$\partial D(p) = \{\chi \in M^d : D(q) \geq D(p) + \chi \cdot (q - p) \quad \forall q \in M^d\},$$

where M^d is the set of all real symmetric $d \times d$ matrices, and $\chi \cdot q = \chi_{ij} q_{ij}$.

The quantity C is a fourth-order tensor of elastic coefficients, which has the symmetry properties

$$C_{ijkl} = C_{jikl} = C_{klij}, \tag{2.7}$$

and we assume that

$$C_{ijkl} \in L^\infty(\Omega), \tag{2.8}$$

and that C is pointwise stable; that is, there exists a constant $c_0 > 0$ such that

$$C_{ijkl}(x)\,\zeta_{ij}\zeta_{kl} \geq c_0 \zeta_{ij}\zeta_{ij}, \quad \forall \zeta \in M^d, \text{ a.e. in } \Omega. \tag{2.9}$$

Equation (2.3) characterises nonlinear kinematic hardening, as mentioned earlier, and this is represented by the term σ_0 appearing there. This term is the back stress, which we assume to be given by

$$\sigma_0(p) = h(|p|)p, \tag{2.10}$$

where $h(\cdot)$ is a scalar-valued hardening function (see, for example, [36]). The function $h(\cdot)$ is assumed to be of the form

$$h(\alpha) = h_0 + h_1 e^{-\nu\alpha}, \tag{2.11}$$

where h_0, h_1, and ν are scalars whose values depend on the material constitution. Here ν is a constant with

$$\nu > 0; \tag{2.12}$$

furthermore, it is assumed that $h_0,\ h_1 \in L^\infty(\Omega)$, and that

$$h_0(x) \geq \eta_0 > 0, \quad h_1(x) \geq 0, \quad \text{a.e. in } \Omega, \tag{2.13}$$

for some constant η_0. We will also require the assumption that a constant $\theta \in (0,1)$ exists such that

$$h_1(x) < \theta h_0(x) e^2, \quad \text{a.e. in } \Omega; \tag{2.14}$$

this is a reasonable approximation for a wide range of materials. The consequences of the approximations embodied in this hardening law are discussed in [30].

The function $D : M^d \to [0, \infty]$, which is known as the dissipation function, is a gauge, that is,

$$D(q) \geq 0,\ D(0) = 0, \tag{2.15}$$

$$D \text{ is convex and positively homogeneous.} \tag{2.16}$$

For realistic models of plasticity it is necessary to assume further that

$$D(q) = 0 \quad \text{if and only if} \quad q = 0, \tag{2.17}$$

$$D \text{ is continuous,} \tag{2.18}$$

and that

$$D(q) < \infty, \quad \text{for all } q \in M^d. \tag{2.19}$$

The properties (2.15)–(2.19) ensure that D is a norm on M^d. Furthermore, the properties of convexity and positive homogeneity imply that

$$|D(p) - D(q)| \leq D(p - q),$$

and since all norms on M^d are equivalent, it follows that

$$D \text{ is Lipschitz continuous on } \operatorname{dom} D. \tag{2.20}$$

REMARK 2.1. *The property* (2.20) *is true under rather weak assumptions. Indeed, let*

$$\partial' D(p) = \left\{ \chi \in M^d : \liminf_{q \to p} \frac{D(q) - D(p) - \chi \cdot (q - p)}{|q - p|^2} > -\infty \right\}.$$

One has $\partial' D(p) \subset \partial D(p)$. It is proved in [34] *that for a proper, lower semicontinuous function D, if $|\chi| \leq k$, $\chi \in \partial' D(p)$ at any point $p \in M^d$ where $\partial' D(p)$ exists, and if $D(p_0) < \infty$ at some $p_0 \in M^d$, then*

$$\begin{aligned} &D(p) < \infty, \quad \forall\, p \in M^d, \\ &D \text{ is Lipschitz continuous on } M^d. \end{aligned}$$

The relationship between the dissipation function and the possibly more familiar yield function may be summarised as follows (see [35, §15], for further details, and [13] for a discussion and further developments in the context of plasticity).

LEMMA 2.2. *Set*

$$K = \{\chi \in M^d : \chi \cdot q \le D(q) \ \ \textit{for all } q \in M^d\}.$$

Then

$$D(p) = \sup_{\chi \in K} \chi \cdot p; \tag{2.21}$$

$$K = \partial D(0); \tag{2.22}$$

$$\chi \in \partial D(p) \Leftrightarrow p \in N_K(\chi). \tag{2.23}$$

The set K is convex and is referred to as the region of admissible (generalised) stresses; its boundary is known as the *yield surface.* Here $N_K(\chi)$ denotes the normal cone to K at χ.

The function D, by virtue of its properties as a gauge and the property (2.18), admits a *polar* function $f : M^d \to [0, \infty]$, defined by

$$f(\chi) = \sup_{q \neq 0} \frac{\chi \cdot q}{D(q)}. \tag{2.24}$$

This function is known in the present context as the *canonical yield function.* That is, it is a gauge and also has the property

$$f(\chi) = 0 \quad \text{if and only if} \quad \chi = 0. \tag{2.25}$$

It derives its name as a yield function from the fact that its level set at unity describes K

$$K = \{\chi : f(\chi) \le 1\}. \tag{2.26}$$

For $\chi \in K$ and $\chi \in \partial D(p)$,

$$\chi \cdot p = f(\chi) D(p). \tag{2.27}$$

We also observe from (2.25) and (2.26) that

$$\chi \in \operatorname{int} K \ \Rightarrow \ N_K(\chi) = \{0\} \ \Rightarrow \ f(\chi) < 1. \tag{2.28}$$

The polar function f is also a norm on M^d. Since all norms on M^d are equivalent, there then exists a constant $c > 0$ such that

$$|\chi| \le c f(\chi) \le c, \quad \forall \chi \in K, \tag{2.29}$$

where $|\cdot|$ denotes the Euclidean norm on M^d. Thus it follows in particular from (2.3) and Lemma 2.2 that

$$|(\sigma - \sigma_0)^D| \le c. \tag{2.30}$$

Example. A simple and popular example is that corresponding to the von Mises yield condition, for which

$$K = \{\chi \in M^d : \ |\chi| \le k\},$$

where $\chi = \sigma^D - \sigma_0^D$ and k is a positive scalar; then

$$D(q) = k|q| = k\sqrt{q_{ij}q_{ij}}\ , \tag{2.31}$$

and the canonical yield function is given by

$$f(\chi) = \frac{|\chi|}{k}.$$

3. The variational problems. As a prelude to presenting the variational formulations of the elastoplastic problem of the last section, we define the spaces

$$V = [H_0^1(\Omega)]^d,$$
$$Q = \{q = (q_{ij}) : q_{ij} \in L^2(\Omega),\ q_{ji} = q_{ij}\}, \quad \text{and} \quad Q_0 = \{q \in Q : \operatorname{tr} q = 0\},$$

where $H_0^1(\Omega)$ is the space of distributions which together with their first derivatives are in $L^2(\Omega)$, and whose traces on Γ vanish. Both V and Q are Hilbert spaces with inner products

$$(u, v)_V = \int_\Omega \frac{\partial u_i}{\partial x_j}\frac{\partial v_i}{\partial x_j}\, dx \quad \text{and} \quad (p, q)_Q = \int_\Omega p \cdot q\ dx = \int_\Omega p_{ij}q_{ij}\, dx,$$

and norms $||v||_V = (v, v)^{1/2}$, $||q||_Q = (q, q)^{1/2}$. Furthermore, Q_0 is a closed subspace of Q.

We define the product space $\overline{V} = V \times Q$, which is a Hilbert space, with the inner product

$$(\overline{u}, \overline{v})_{\overline{V}} := (u, v)_V + (p, q)_Q$$

and norm $||\overline{u}||_{\overline{V}} = (\overline{u}, \overline{u})_{\overline{V}}^{1/2}$, where $\overline{u} = (u, p)$ and $\overline{v} = (v, q)$. We also define $\overline{V}_0 = V \times Q_0$, a closed subspace of $\overline{V}$. The topological dual of a Hilbert space X is denoted by X^*.

We define the operator $A_1 : \bar{V} \to \bar{V}^*$ by

$$\begin{aligned}\langle A_1\bar{u}, \bar{v}\rangle &= \int_\Omega \left[C\left(\epsilon(u) - p\right) \cdot (\epsilon(v) - q) + \sigma_0(p) \cdot q\right]\, dx \\ &= \int_\Omega \left[C_{ijkl}(\epsilon_{ij}(u) - p_{ij})(\epsilon_{kl}(v) - q_{kl}) + (\sigma_0(p))_{ij}q_{ij}\right]\, dx,\end{aligned} \tag{3.1}$$

the linear functional

$$l : \overline{V} \to R, \qquad \langle l, \overline{v}\rangle = \int_\Omega b \cdot v\, dx$$

and the functional

$$j : \overline{V} \to R, \qquad j(\overline{v}) = \int_\Omega D(q(x))\, dx, \tag{3.2}$$

where as before $\overline{u} = (u, p)$ and $\overline{v} = (v, q)$. The functionals $l(\cdot)$ and $j(\cdot)$ are easily shown to be bounded and, from the properties of D, $j(\cdot)$ is a convex, positively homogeneous, nonnegative continuous functional. Note that, however, in general,

j is not differentiable. As with the dissipation function D, convexity and positive homogeneity of j imply that

$$j \text{ is Lipschitz continuous on } \overline{V}. \tag{3.3}$$

We remark also that the Lipschitz continuity of j is assured under weaker assumptions, as implied by Remark 2.1.

The classical problem defined by (2.1)–(2.5) is formally equivalent [30] to the following Problem P_1.

Problem P_1. Find $\overline{u} = (u,p) \in \overline{V}_0$ such that

$$\langle A_1\overline{u}, \overline{v} - \overline{u}\rangle + j(\overline{v}) - j(\overline{u}) - \langle l, \overline{v} - \overline{u}\rangle \geq 0, \quad \forall \overline{v} \in \overline{V}_0.$$

We also have the following result.

THEOREM 3.1 (see [30]). *Problem* P_1 *has a unique solution.*

This result depends in particular on the fact that the operator A_1 is Lipschitz continuous and strongly monotone. That is, there are constants $\alpha_0, \alpha_1 > 0$, such that

$$||A_1\overline{u} - A_1\overline{v}||_{\overline{V}^*} \leq \alpha_1 ||\overline{u} - \overline{v}||_{\overline{V}}, \tag{3.4}$$

$$\langle A_1\overline{u} - A_1\overline{v}, \overline{u} - \overline{v}\rangle \geq \alpha_0 ||\overline{u} - \overline{v}||_{\overline{V}}^2, \tag{3.5}$$

for all $\overline{u}, \overline{v} \in \overline{V}$.

The main concern here is not with variational inequalities of the form of Problem P_1, but rather with *mixed* variational problems associated with P_1. Two such mixed variational problems are considered in [30]. The first is as follows.

Problem P_2. Find $\overline{u} = (u,p) \in \overline{V}$ and $\lambda \in \Lambda$ such that

$$\begin{cases} \langle A_1\overline{u}, \overline{v} - \overline{u}\rangle + j(\overline{v}) - j(\overline{u}) + b_1(\overline{v} - \overline{u}, \lambda) \geq \langle l, \overline{v} - \overline{u}\rangle, & \forall \overline{v} = (v,q) \in \overline{V}, \\ b_1(\overline{u}, \mu) = 0, \quad \forall \mu \in \Lambda, \end{cases} \tag{3.6}$$

where $\Lambda = L^2(\Omega)$ and the bilinear form

$$b_1 : \overline{V} \times \Lambda \to R, \qquad b_1(\overline{v}, \mu) = -\int_\Omega \mu \,\mathrm{tr}\, q \, dx. \tag{3.7}$$

The second mixed variational inequality considered in [30] may be obtained by extending to the case of elastoplastic materials the classical Hellinger–Reissner variational problem [7] for elasticity, and at the same time relaxing the constraint of plastic incompressibility. Let E be the elastic compliance tensor, inverse to C. The tensor E is bounded, symmetric, and pointwise stable. We set $\Sigma = Q$, $\overline{\Sigma} = \Sigma \times Q$, and for $\overline{\sigma} = (\sigma, p)$ and $\overline{\tau} = (\tau, q)$ in $\overline{\Sigma}$ define the operator A_2 by

$$A_2 : \overline{\Sigma} \to \overline{\Sigma}^*, \quad \langle A_2\overline{\sigma}, \overline{\tau}\rangle = \int_\Omega [(E\sigma + p)\cdot\tau + (\sigma_0(p) - \sigma)\cdot q] \, dx. \tag{3.8}$$

Define $N = V \times \Lambda$, the new space of Lagrange multipliers, where $\Lambda = L^2(\Omega)$, and the bilinear form $b_2 : \overline{\Sigma} \times N \to R$,

$$b_2(\overline{\tau}, n) = -\int_\Omega \tau \cdot \epsilon(v) \, dx - \int_\Omega \mu \,\mathrm{tr} q \, dx, \quad n = (v, \mu), \ \overline{\tau} = (\tau, q). \tag{3.9}$$

Problem P_3. Find $\overline{\sigma} = (\sigma, p) \in \overline{\Sigma}$ and $m = (u, \lambda) \in N$ such that

$$\begin{cases} \langle A_2\overline{\sigma}, \overline{\tau} - \overline{\sigma}\rangle + j(\overline{\tau}) - j(\overline{\sigma}) + b_2(\overline{\tau} - \overline{\sigma}, m) \geq 0, & \forall \overline{\tau} = (\tau, q) \in \overline{\Sigma}, \\ b_2(\overline{\sigma}, n) = \langle g, n\rangle, \quad \forall n = (v, \mu) \in N, \end{cases} \tag{3.10}$$

where the linear functional $g : N \to R$ is defined by $\langle g, n\rangle = \int_\Omega b \cdot v\, dx$.

There are many mixed finite-element methods in elasticity which take as a variational basis not the appropriate reduction of Problem P_3, which is (after setting $p = q = 0$ throughout): find $(\sigma, u) \in \Sigma \times V$ such that

$$\int_\Omega E\sigma \cdot \tau\, dx - \int_\Omega \epsilon(u) \cdot \tau\, dx = 0, \qquad \forall \tau \in \Sigma, \tag{3.11}$$

$$\int_\Omega \epsilon(v) \cdot \sigma\, dx = \int_\Omega b \cdot v\, dx, \qquad \forall v \in V, \tag{3.12}$$

but instead the version of the Hellinger–Reissner formulation in which the differentiability condition on the displacement is transferred to the stress. For this purpose it is necessary to define the spaces

$$W = [L^2(\Omega)]^d \quad \text{and} \quad H = \{\tau = (\tau_{ij}) : \tau_{ji} = \tau_{ij}, \tau_{ij} \in L^2(\Omega),\ \operatorname{div} \tau \in W\},$$

where $\operatorname{div} \tau$ is the vector with components $\partial \tau_{ij}/\partial x_j$ (with summation implied on j). The space W has the standard L^2-product norm, while H is endowed with the norm

$$||\tau||_H^2 = ||\tau||_\Sigma^2 + ||\operatorname{div} \tau||_W^2.$$

It is straightforward to derive the alternative Hellinger–Reissner formulation for elasticity in the following form: find $(\sigma, u) \in H \times W$ such that

$$\int_\Omega E\sigma \cdot \tau\, dx + \int_\Omega u \cdot \operatorname{div} \tau\, dx = 0, \qquad \forall \tau \in H, \tag{3.13}$$

$$\int_\Omega v \cdot \operatorname{div} \sigma\, dx = -\int_\Omega b \cdot v\, dx, \qquad \forall v \in W. \tag{3.14}$$

This problem has formed the basis of most investigations of mixed finite-element methods for elasticity problems [1], [2], [25], [29], [38], [39], though the formulation (3.11)–(3.12) is favoured in many engineering applications (see, for example, [28]). An exception is the work [37], in which the abstract conditions for stability in the works cited above are given an interesting mechanical interpretation. Numerical examples are also given in [37].

The appropriate generalisation to elastoplasticity is as follows.

Problem P_4. Find $\overline{\sigma} = (\sigma, p) \in \overline{H} = H \times Q$ and $m = (u, \lambda) \in L = W \times \Lambda$ such that

$$\left\{ \begin{array}{l} \langle A_2\overline{\sigma}, \overline{\tau} - \overline{\sigma}\rangle + j(\overline{\tau}) - j(\overline{\sigma}) + b_3(\overline{\tau} - \overline{\sigma}, m) \geq 0, \quad \forall \overline{\tau} = (\tau, q) \in \overline{H}, \\ b_3(\overline{\sigma}, n) = \langle g, n\rangle\ , \quad \forall n = (v, \mu) \in L, \end{array} \right. \tag{3.15}$$

where

$$A_2 : \overline{H} \to \overline{H}^*, \quad \langle A_2\overline{\sigma}, \overline{\tau}\rangle = \int_\Omega [(E\sigma + p) \cdot \tau + (\sigma_0(p) - \sigma) \cdot q]\, dx, \tag{3.16}$$

and

$$b_3(\overline{\tau}, n) = \int_\Omega (v \operatorname{div} \tau - \mu \operatorname{tr} q)\, dx. \tag{3.17}$$

Problems P_2–P_4 can be conveniently studied as special cases of an abstract mixed variational inequality, which we now formulate.

4. The abstract problem. Let Ψ and N be two Hilbert spaces, A an operator from Ψ to its dual Ψ^*, $b : \Psi \times N \to R$ a bilinear form, and $j : \Psi \to R$ a functional. The bilinear form b is assumed to have the following two properties.

(i) $b(\cdot,\cdot)$ is bounded, so that there exist bounded linear operators B, B^T defined by

$$B : \Psi \to N^*, \qquad \langle B\psi, n\rangle = b(\psi, n) \quad \text{and} \quad B^T : N \to \Psi^*, \qquad \langle B^T n, \psi\rangle = b(\psi, n)$$

for all $\psi \in \Psi$ and $n \in N$. B^T is thus the adjoint operator of B. The kernels of B and B^T are defined by

$$\operatorname{Ker} B = \{\psi \in \Psi : B\psi = 0\} \text{ and } \operatorname{Ker} B^T = \{n \in N : B^T n = 0\}.$$

(ii) There exists a constant $\beta > 0$ such that

$$\sup_{\psi \in \Psi} \frac{b(\psi, n)}{||\psi||_\Psi} \geq \beta \, ||n||_{N/\operatorname{Ker} B^T}, \quad \forall n \in N. \tag{4.1}$$

The operator A is assumed to be Lipschitz continuous on Ψ; that is, there exists a constant $\alpha_1 > 0$ such that

$$||A\phi - A\psi||_{\Psi^*} \leq \alpha_1 ||\phi - \psi||_\Psi, \quad \forall \phi, \psi \in \Psi. \tag{4.2}$$

For any given $\phi_1 \in (\operatorname{Ker} B)^\perp$, let $\tilde{A} : \Psi \to \Psi^*$ be the operator defined by

$$\tilde{A}\chi = A(\chi + \phi_1) \quad \text{for all } \chi \in \operatorname{Ker} B. \tag{4.3}$$

The operator $\tilde{A}$ is required to be strongly monotone on Ker B; that is, there exists a constant $\alpha_0 > 0$ such that

$$\langle \tilde{A}\phi - \tilde{A}\psi, \phi - \psi\rangle \geq \alpha_0 ||\phi - \psi||_\Psi^2, \quad \forall \phi, \psi \in \operatorname{Ker} B. \tag{4.4}$$

We assume also that

(4.5) j is convex, nonnegative, and continuous, but not differentiable.

REMARK 4.1. *The assumption* (4.4) *replaces that of strong monotonicity of A on* Ker B, *which was assumed in* [30] (*see* (3.3) *in that work*). *The strong monotonicity of A on* Ker B *is an insufficient assumption, since it is readily established, as in the analysis of the auxiliary problem* (3.8) *in* [30], *that A is actually required to be strongly monotone on the affine set $\phi_1 + \operatorname{Ker} B$. The assumption* (4.4) *guarantees this. Note that the condition of strong monotonicity of A on Ψ implies* (4.4).

The mixed problems defined in §3 are all particular cases of the following general problem.

Problem P. Given $f \in \Psi^*$ and $g \in N^*$, find $(\phi, m) \in \Psi \times N$ such that

$$\begin{cases} \langle A\phi, \ \psi - \phi\rangle + j(\psi) - j(\phi) + b(\psi - \phi, m) \geq \langle f, \ \psi - \phi\rangle, & \forall \psi \in \Psi, \\ b(\phi, n) = \langle g, n\rangle, \quad \forall n \in N. \end{cases} \tag{4.6}$$

This problem is approached in [30] by introducing a regularized version of the problem, which reduces the inequality $(4.6)_1$ to an equation. We introduce the family of functionals $j_\varepsilon : \Psi \to R$ parametrized by $\varepsilon \in (0, 1]$ and with the following properties:

(4.7) j_ε is convex and differentiable, with Gâteaux derivative $j'_\varepsilon : \Psi \to \Psi^*$;

(4.8) $j_\varepsilon(\psi) \to j(\psi)$ as $\varepsilon \to 0$, uniformly with respect to $\psi \in \Psi$;

(4.9) $||j'_\epsilon(\psi)||_{\Psi^*} \leq c, \ \forall \psi \in \Psi$.

The constant c is required to be independent of ε and $\psi \in \Psi$.

The regularized problem then takes the following form.

Problem P_ε. Find $\phi_\varepsilon \in \Psi$ and $m_\varepsilon \in N$ such that

$$(4.10)\quad \begin{cases} \langle A\phi_\varepsilon, \psi - \phi_\varepsilon\rangle + j_\varepsilon(\psi) - j_\varepsilon(\phi_\varepsilon) + b(\psi - \phi_\varepsilon, m_\varepsilon) \geq \langle f, \psi - \phi_\varepsilon\rangle, & \forall \psi \in \Psi, \\ b(\phi_\varepsilon, n) = \langle g, n\rangle, \quad \forall n \in N. \end{cases}$$

Since j_ε is differentiable, the inequality $(4.10)_1$ reduces to a variational equation

$$(4.11)\qquad \langle A\phi_\varepsilon, \psi\rangle + \langle j_\varepsilon'(\phi_\varepsilon), \psi\rangle + b(\psi, m_\varepsilon) = \langle f, \psi\rangle, \quad \forall \psi \in \Psi.$$

Then we have the following theorem.

THEOREM 4.2. *Assume that conditions* (4.1)–(4.5) *and* (4.7)–(4.9) *hold, and that* $g \in \operatorname{Im} B$, *the range of* B. *Then*

(a) *there exists a unique solution* $(\phi_\varepsilon, m_\varepsilon) \in \Psi \times (N/\mathrm{Ker} B^T)$ *to Problem* P_ε;

(b) *there exists a solution* $(\phi, m) \in \Psi \times N$ *to Problem* P. *Furthermore,* ϕ *is unique, and is the strong limit of* ϕ_ε *as* $\varepsilon \to 0$.

In [30], this theorem is proved for the case $\operatorname{Ker} B = \{0\}$, and when the condition (4.8) is replaced by the less general condition

$$0 \leq j_\varepsilon(\psi) - j(\psi) \leq c_1\varepsilon, \quad \forall \psi \in \Psi.$$

It follows readily from the proof in [30] that the result holds with the somewhat more general assumptions (4.1) and (4.8). The weaker assumption (4.8), in particular, allows the possibility of treating more general types of yield conditions, such as nonsmooth yield functions (see, for example, [5]).

In [30], it is shown that Problem P_2 satisfies the conditions of Theorem 4.2, that it has a solution $(u, p, \lambda) \in \overline{V} \times \Lambda$, and that (u, p) is unique. It is also shown that Problem P_3 has a solution $(\sigma, p, u, \lambda) \in \overline{\Sigma} \times N$, with (σ, p) being unique. Furthermore, it may be proved that the Lagrange multiplier u is unique, though not the multiplier λ.

To obtain similar results for Problem P_4, which was not treated in [30], we need to show that A_2 is Lipschitz continuous on $\overline{H} = H \times Q$ and that $\tilde{A}_2$ is strongly monotone on $\operatorname{Ker} B_3$, defined by

$$\begin{aligned} \operatorname{Ker} B_3 &= \left\{ (\tau, q) \in \overline{H} : \int_\Omega v \cdot \operatorname{div} \tau \, dx = 0, \ \forall v \in W \text{ and } \int_\Omega \mu \operatorname{tr} q \, dx = 0, \ \forall \mu \in \Lambda \right\} \\ (4.12)\quad &= \{(\tau, q) \in \overline{H} : \ \operatorname{div} \tau = 0 \text{ and } \operatorname{tr} q = 0, \quad \text{a.e. in } \Omega\}. \end{aligned}$$

We also have to show that the Babuška–Brezzi condition holds. That is, that a constant $\beta > 0$ exists such that

$$(4.13)\qquad \sup_{\overline{\tau} \in \overline{H}} \frac{b_3(\overline{\tau}, n)}{\| \overline{\tau} \|_{\overline{H}}} \geq \beta \| n \|_{L/\operatorname{Ker} B_3^T} .$$

The functional j is as in Problems P_2 and P_3, and its fulfilment of properties (4.7)–(4.9) in the case of the von Mises condition has been established in [30].

Lipschitz continuity of A_2 on $\overline{H}$ follows from the fact that

$$(4.14)\qquad \langle A_2\overline{\sigma} - A_2\overline{\tau}, \overline{p}\rangle \leq c(\| \sigma - \tau \|_Q + \| p - q \|_Q) \| \overline{p} \|_{Q \times Q},$$

as has been established in [30], and using also the fact that $\| \sigma \|_Q \leq \| \sigma \|_H$ for $\sigma \in H$.

Whereas A_2 is strongly monotone on $\overline{\Sigma} \times \overline{\Sigma}$ (for Problem P_3), it is *not* strongly monotone on $\overline{H} = H \times Q$. Instead, we show that (4.4) holds. We have, given $\overline{\sigma}_1 = (\sigma_1, p_1) \in (\operatorname{Ker} B_3)^\perp$, with $\tilde{A}_2\overline{\tau} = A_2(\overline{\sigma}_1 + \overline{\tau})$, and for $\overline{\sigma}$, $\overline{\tau} \in \operatorname{Ker} B_3$,

$$\begin{aligned}
\langle \tilde{A}_2\overline{\sigma} - \tilde{A}_2\overline{\tau}, \overline{\sigma} - \overline{\tau}\rangle &= \langle A_2(\overline{\sigma} + \overline{\sigma}_1) - A_2(\overline{\tau} + \overline{\sigma}_1), \overline{\sigma} - \overline{\tau}\rangle \\
&= \int_\Omega \{E(\sigma - \tau)\cdot(\sigma - \tau) + [\sigma_0(p + p_1) - \sigma_0(q + p_1)]\cdot(p - q)\}\, dx \\
&\geq \alpha_0 \left(\| \sigma - \tau \|_\Sigma^2 + \| p - q \|_Q^2\right) \\
&= \alpha_0 \left(\| \sigma - \tau \|_H^2 + \| p - q \|_Q^2\right) \\
&= \alpha_0 \| \overline{\sigma} - \overline{\tau} \|_{\overline{H}}^2,
\end{aligned} \tag{4.15}$$

for some constant $\alpha_0 > 0$. Here we have used the fact that $\operatorname{div}\sigma = 0$ and $\operatorname{div}\tau = 0$, since $\overline{\sigma}$, $\overline{\tau} \in \operatorname{Ker} B_3$, and also the inequality

$$\begin{aligned}
&[\sigma_0(q) - \sigma_0(r)]\cdot(q - r) \\
&\quad\geq \left[(\theta h_0 + h_1 e^{-\nu|q|})q - (\theta h_0 + h_1 e^{-\nu|r|})r\right]\cdot(q - r) + (1-\theta)h_0|q - r|^2 \\
&\quad\geq (1-\theta)\eta_0|q - r|^2
\end{aligned}$$

(see also Lemma 2 in [30]).

Finally, it is straightforward to show [38] that the Babuška–Brezzi condition (4.13) holds. Thus Problem P_4 has a solution $\overline{\sigma} \in \overline{H}$, $n = (u, \lambda) \in L$, and $\overline{\sigma}$ is unique. It may furthermore be established that u is unique.

For the convergence analysis of finite-element solutions later, we introduce a formulation equivalent to (4.6). For this, we need the notion of the subdifferential $\partial j : \Psi \to 2^{\Psi^*}$, defined by

$$\partial j(\phi) = \{\phi^* \in \Psi^* : j(\psi) \geq j(\phi) + \langle \phi^*, \psi - \phi\rangle,\ \forall \psi \in \Psi\}.$$

We observe that the problem (4.6) is equivalent to the problem of finding $\phi \in \Psi$, $m \in N$ and $\phi^* \in \Psi^*$ such that

$$\left\{\begin{array}{l}
\langle A\phi, \psi\rangle + \langle \phi^*, \psi\rangle + b(\psi, m) = \langle f, \psi\rangle, \quad \forall \psi \in \Psi, \\
b(\phi, n) = \langle g, n\rangle, \quad \forall n \in N, \\
\phi^* \in \partial j(\phi).
\end{array}\right. \tag{4.16}$$

REMARK 4.3. *From the equivalent formulation* (4.16), *it is easily seen that if* $\phi^* \in \Psi^*$ *is unique, then* $m \in N$ *is unique in* $N/\mathrm{Ker}B^T$. *In the case when there is no constraint, such a* $\phi^* \in \Psi^*$ *is indeed unique* (*see* [17, Prop. 4.2]). *For the mixed problem* P, *however, the uniqueness of* ϕ^* (*and hence that of* $m \in N/\mathrm{Ker}B^T$) *depends on both* b *and* j. *It is not difficult to see that we have uniqueness for* ϕ^* *if and only if the following condition is satisfied:*

(4.17) *If* $\phi^* \in \partial j(\phi)$, *then* $\partial j(\phi)$ *does not contain elements of the form* $\phi^* + B^T m_0$ *with* $m_0 \in N/\mathrm{Ker}B^T$.

Usually, however, the condition (4.17) *cannot be verified easily.*

For a mechanical interpretation of the nonuniqueness of m, see [30].

REMARK 4.4. *In the case of plasticity with the von Mises yield condition* (*see* (2.31) *and* (2.3)), *we have* $\Psi = \overline{V} = V \times Q$, *and* j *is given by*

$$j(\overline{v}) = \int_\Omega k\,|q|\,dx$$

where k is bounded, measurable, and nonnegative. Then $(4.16)_3$ is equivalent to the two conditions (with $\phi^* = (u^*, p^*)$):

$$p^*(x) = k\,\frac{p(x)}{|p(x)|} \quad \text{if } p(x) \neq 0 \quad \text{and} \quad |p^*(x)| \leq k \quad \text{if } p(x) = 0, \tag{4.18}$$

which hold a.e. in Ω. Whereas we have uniqueness when $p(x) \neq 0$, $\forall\, x \in \Omega$, in general we have no a priori control over the solution p, and p^* fails to be unique. This is easy to appreciate if it is observed, from (2.3), that p^* is equivalent to the quantity $\sigma^D - \sigma_0^D(p)$ ($= \sigma^D$ when $p = 0$), and $(4.18)_2$ expresses the fact that this lies in the set of admissible stresses, which on its own does not determine σ^D or p^* uniquely.

Note also that by application of the arguments leading to (2.29) we have uniform boundedness of $\{p^*\}$

$$|p^*(x)| \leq k \quad \text{a.e. in } \Omega. \tag{4.19}$$

REMARK 4.5. *In certain other applications (cf. [15], [18]), $\Psi = [H^1(\Omega)]^d$ or $[H_0^1(\Omega)]^d$, and j is of the form*

$$j(\psi) = \int_\Omega k\,|\nabla\psi|\,dx \tag{4.20}$$

where k is bounded, measurable, and nonnegative. It can be proved that $\phi^* \in \partial j(\phi)$ is equivalent to the two conditions

$$\phi^*(x) = k\,\frac{\nabla\phi(x)}{|\nabla\phi(x)|} \quad \text{if } \nabla\phi(x) \neq 0 \quad \text{and} \quad |\phi^*(x)| \leq k \quad \text{if } \nabla\phi(x) = 0.$$

Once again, the uniqueness of ϕ^* depends on the form of b (cf. the condition (4.17) above). We still have uniform boundedness of $\{\phi^*\}$, though.

5. Finite-element approximation of the abstract mixed variational inequality. To study finite-element approximations of the mixed problems of §3, we first consider that of the abstract Problem P. Let $\Psi_h \subset \Psi$ and $N_h \subset N$ be finite-dimensional subspaces, $h > 0$ being a discretization parameter. We assume that

$$\lim_{h\to 0}\inf_{\psi_h\in\Psi_h} \|\psi - \psi_h\|_\Psi = 0, \quad \forall\,\psi \in \Psi,$$
$$\lim_{h\to 0}\inf_{n_h\in N_h} \|n - n_h\|_N = 0, \quad \forall\, n \in N.$$

The finite-element approximation to the abstract Problem P is as follows.

Problem P_h. Find $\phi_h \in \Psi_h$ and $m_h \in N_h$, such that

$$\begin{cases} \langle A\phi_h, \psi_h - \phi_h\rangle + j(\psi_h) - j(\phi_h) + b(\psi_h - \phi_h, m_h) \geq \langle f, \psi_h - \phi_h\rangle, & \forall\,\psi_h \in \Psi_h, \\ b(\phi_h, n_h) = \langle g, n_h\rangle, \quad \forall\, n_h \in N_h. \end{cases} \tag{5.1}$$

Let

$$Z_h(g) = \{\xi_h \in \Psi_h : \ b(\xi_h, n_h) = \langle g, n_h\rangle, \ \forall\, n_h \in N_h\}$$

and introduce the discrete operators B_h and B_h^T through the relation

$$b(\psi_h, n_h) = \langle B_h\psi_h, n_h\rangle = \langle B_h^T n_h, \psi_h\rangle, \quad \forall\,\psi_h \in \Psi_h,\ \forall\, n_h \in N_h.$$

The kernels of B_h and B_h^T are defined by

$$\operatorname{Ker} B_h = \{\psi_h \in \Psi_h : \; b(\psi_h, n_h) = 0 \;\; \forall n_h \in N_h\}, \tag{5.2}$$

$$\operatorname{Ker} B_h^T = \{n_h \in N_h : \; b(\psi_h, n_h) = 0 \;\; \forall \psi_h \in \Psi_h\}. \tag{5.3}$$

For any given $\phi_{1h} \in (\operatorname{Ker} B_h)^\perp$, we will also require the operator $\tilde{A}_h : \Psi_h \to \Psi_h^*$ defined by

$$\tilde{A}_h \chi_h = A(\chi_h + \phi_{1h}) \quad \text{for all } \chi_h \in \operatorname{Ker} B_h. \tag{5.4}$$

Applying Theorem 4.2 to Problem P_h we have Theorem 5.1.

THEOREM 5.1. *Assume that the conditions of Theorem* 4.2 *hold, with the exception that the condition* (4.1) *is replaced by its discrete counterpart*

$$\sup_{\psi_h \in \Psi_h} \frac{b(\psi_h, n_h)}{\|\psi_h\|_\Psi} \geq k_h \|n_h\|_{N/\operatorname{Ker} B_h^T}, \quad \forall n_h \in N_h, \quad \textit{for some } k_h > 0, \tag{5.5}$$

and the condition (4.4) *is replaced by the condition that a constant $\alpha_0 > 0$, independent of h and independent of the function ϕ_{1h} used in defining $\tilde{A}_h$ in* (5.4), *exists such that*

$$\langle \tilde{A}_h \phi_h - \tilde{A}_h \psi_h, \phi_h - \psi_h \rangle \geq \alpha_0 \|\psi_h - \phi_h\|_\Psi^2 \quad \textit{for all } \phi_h, \psi_h \in \operatorname{Ker} B_h. \tag{5.6}$$

Then, if $Z_h(g) \neq \emptyset$, Problem P_h has a solution $(\phi_h, m_h) \in \Psi_h \times N_h$, ϕ_h being unique. Furthermore, $(\phi_h, m_h) \in \Psi_h \times N_h$ is a solution of Problem P_h if and only if there exists a $\phi_h^* \in \Psi_h^*$, such that

$$\begin{gathered} \langle A\phi_h, \psi_h \rangle + \langle \phi_h^*, \psi_h \rangle + b(\psi_h, m_h) = \langle f, \psi_h \rangle, \quad \forall \psi_h \in \Psi_h, \\ b(\phi_h, n_h) = \langle g, n_h \rangle, \quad \forall n_h \in N_h, \\ \phi_h^* \in \partial_h j(\phi_h), \end{gathered} \tag{5.7}$$

where

$$\partial_h j(\phi_h) = \{\phi_h^* \in \Psi_h^* : j(\psi_h) \geq j(\phi_h) + \langle \phi_h^*, \psi_h - \phi_h \rangle, \; \forall \psi_h \in \Psi_h\}. \tag{5.8}$$

To study the convergence of the finite-element solution, we first derive an error estimate for $\phi - \phi_h$. We will denote a solution of (4.16) by $\{\phi, \phi^*, m\}$.

THEOREM 5.2. *Assume the conditions of Theorems* 4.2 *and* 5.1 *hold. For any $\xi_h \in \Psi_h$, set*

$$I(\xi_h, \phi, \phi^*) = j(\xi_h) - j(\phi) - \langle \phi^*, \xi_h - \phi \rangle.$$

Then there is a constant $c > 0$, independent of h, such that

$$\|\phi - \phi_h\|_\Psi \leq c \left[\inf_{\xi_h \in Z_h(g)} \left(\|\phi - \xi_h\|_\Psi + |I(\xi_h, \phi, \phi^*)|^{1/2} \right) + \inf_{n_h \in N_h} \|m - n_h\|_N \right]. \tag{5.9}$$

Proof. For any $\xi_h \in Z_h(g)$,

$$\|\phi - \phi_h\|_\Psi \leq \|\phi - \xi_h\|_\Psi + \|\phi_h - \xi_h\|_\Psi. \tag{5.10}$$

Now write $\phi_h = \phi_{0h} + \phi_{1h}$, where $\phi_{0h} \in \operatorname{Ker} B_h$, $\phi_{1h} \in Z_h(g) \cap (\operatorname{Ker} B_h)^\perp$, and define $\tilde{A}_h$ as in (5.4) by this ϕ_{1h}. Observe that $\xi_h - \phi_{1h} \in \operatorname{Ker} B_h$. Thus, by property (5.6) we have

$$\begin{aligned}
\alpha_0 \|\phi_h - \xi_h\|_\Psi^2 &= \alpha_0 \|\phi_{0h} - (\xi_h - \phi_{1h})\|_\Psi^2 \\
&\le \langle \tilde{A}_h \phi_{0h} - \tilde{A}_h(\xi_h - \phi_{1h}), \phi_h - \xi_h \rangle \\
&= \langle A\phi_h - A\xi_h, \phi_h - \xi_h \rangle \\
&= \langle A\phi_h, \phi_h - \xi_h \rangle + \langle A\phi - A\xi_h, \phi_h - \xi_h \rangle + \langle A\phi, \xi_h - \phi_h \rangle.
\end{aligned}$$

We add the two inequalities

$$\begin{gathered}
\langle A\phi_h, \xi_h - \phi_h \rangle + j(\xi_h) - j(\phi_h) \ge \langle f, \xi_h - \phi_h \rangle, \\
\langle A\phi, \phi_h - \phi \rangle + j(\phi_h) - j(\phi) + b(\phi_h - \phi, m) \ge \langle f, \phi_h - \phi \rangle
\end{gathered}$$

to obtain

$$\langle A\phi_h, \xi_h - \phi_h \rangle + \langle A\phi, \phi_h - \phi \rangle + j(\xi_h) - j(\phi) + b(\phi_h - \phi, m) \ge \langle f, \xi_h - \phi \rangle;$$

that is,

$$\langle A\phi_h, \phi_h - \xi_h \rangle \le \langle A\phi, \phi_h - \phi \rangle + j(\xi_h) - j(\phi) + b(\phi_h - \phi, m) - \langle f, \xi_h - \phi \rangle.$$

Hence

$$\begin{aligned}
\alpha_0 \|\phi_h - \xi_h\|_\Psi^2 &\le \langle A\phi, \xi_h - \phi \rangle + j(\xi_h) - j(\phi) + b(\xi_h - \phi, m) - \langle f, \xi_h - \phi \rangle \\
&\quad + \langle A\phi - A\xi_h, \phi_h - \xi_h \rangle + b(\phi_h - \xi_h, m) \\
&= j(\xi_h) - j(\phi) - \langle \phi^*, \xi_h - \phi \rangle \\
&\quad + \langle A\phi - A\xi_h, \phi_h - \xi_h \rangle + b(\phi_h - \xi_h, m - n_h)
\end{aligned}$$

for any $n_h \in N_h$, and so

$$\begin{aligned}
&\alpha_0 \|\phi_h - \xi_h\|_\Psi^2 \\
&\quad \le |I(\xi_h, \phi, \phi^*)| + \alpha_1 \|\phi - \xi_h\|_\Psi \|\phi_h - \xi_h\|_\Psi + \|b\| \, \|\phi_h - \xi_h\|_\Psi \|m - n_h\|_N.
\end{aligned}$$

Thus

$$\|\phi_h - \xi_h\|_\Psi \le c \left[\|\phi - \xi_h\|_\Psi + |I(\xi_h, \phi, \phi^*)|^{1/2} + \|m - n_h\|_N \right], \quad \forall n_h \in N_h. \tag{5.11}$$

Combining (5.10) and (5.11), we get the desired error estimate. ☐

In the special case when $\operatorname{Ker} B_h \subset \operatorname{Ker} B$, $b(\phi_h - \xi_h, m - n_h) = 0$ in the proof above. Then we have Proposition 5.3.

PROPOSITION 5.3. *If we further assume that* $\operatorname{Ker} B_h \subset \operatorname{Ker} B$, *then*

$$\|\phi - \phi_h\|_\Psi \le c \inf_{\xi_h \in Z_h(g)} \left(\|\phi - \xi_h\|_\Psi + |I(\xi_h, \phi, \phi^*)|^{1/2} \right).$$

The nature of the bound on the term $|I(\xi_h, \phi, \phi_h)|$ will depend on the particular form taken by the functional j. In elastoplasticity j is Lipschitz continuous (see (3.3)), and so

$$|I(\xi_h, \phi, \phi^*)| \le c \|\phi - \xi_h\|_\Psi. \tag{5.12}$$

To bound $\inf_{\xi_h \in Z_h(g)} \|\phi - \xi_h\|_\Psi$ by the more standard approximation quantity $\inf_{\psi_h \in \Psi_h} \|\phi - \psi_h\|_\Psi$, we need the following result ([7, p. 55]).

LEMMA 5.4. *Assume that the discrete* inf-sup *condition* (5.5) *holds*; *then*

$$\inf_{\xi_h \in Z_h(g)} \|\phi - \xi_h\|_\Psi \le \left(1 + \frac{\|b\|}{k_h}\right) \inf_{\psi_h \in \Psi_h} \|\phi - \psi_h\|_\Psi .$$

An obvious consequence of Theorem 5.2 and Lemma 5.4 is the following convergence result for ϕ_h.

THEOREM 5.5. *Under the assumptions of Theorem* 5.2, *if* (5.12) *holds and if*

$$\frac{1}{k_h} \inf_{\psi_h \in \Psi_h} \|\phi - \psi_h\|_\Psi \to 0 \quad \text{as } h \to 0, \tag{5.13}$$

then

$$\phi_h \to \phi \text{ in } \Psi \quad \text{as } h \to 0.$$

If the pair $\{\Psi_h, N_h\}$ of finite-element spaces is such that k_h is bounded away from 0, independently of h, then the condition (5.13) is automatically satisfied. To have convergence, however, we do not require k_h to be bounded away from 0, as long as k_h does not tend to 0 too fast (in the sense that (5.13) holds).

Now we consider the convergence of m_h. Here it is necessary to turn again to the motivating problems for further information about the behaviour of the sequence $\{\phi_h^*\}$. We identify Ψ_h^* with Ψ_h and view Ψ_h^* as a subspace of Ψ^*: for any $\psi_h^* \in \Psi_h^*$, we extend ψ_h^* from Ψ_h^* to Ψ^* by setting $\langle \psi_h^*, \psi \rangle = 0$, $\forall\, \psi \in \Psi_h^\perp$. Now in Problems P_2–P_4, ϕ_h is the ordered pair (u_h, p_h) or (σ_h, p_h), and we find that, from the Lipschitz continuity of D,

$$\phi_h^* \in \partial_h j(\phi_h) \;\Rightarrow\; \|p_h^*\|_{\Psi^*} \le c, \quad \text{for a constant } c \text{ independent of } h. \tag{5.14}$$

Hence,

$$\{\phi_h^*\} \text{ is weakly precompact in } \Psi^*. \tag{5.15}$$

Thus every subsequence of $\{\phi_h^*\}$ contains a subsequence weakly converging in Ψ^*.

Using the discrete inf-sup condition (5.5), the relation (5.7), and the boundedness of the sequence $\{\phi_h^*\}$, we find that

$$\begin{aligned} k_h \|m_h\|_{N/\mathrm{Ker} B_h^T} &\le \sup_{\psi_h \in \Psi_h} \frac{b(\psi_h, m_h)}{\|\psi_h\|_\Psi} \\ &= \sup_{\psi_h \in \Psi_h} \frac{1}{\|\psi_h\|_\Psi} \{\langle f, \psi_h\rangle - \langle A\phi_h, \psi_h\rangle - \langle \phi_h^*, \psi_h\rangle\} \\ &\le c\,(\|f\| + \|\phi_h\|_\Psi + \|\phi_h^*\|_{\Psi^*}). \end{aligned} \tag{5.16}$$

Since $\phi_h \to \phi$ in Ψ, $\{\|\phi_h\|_\Psi\}$ is uniformly bounded with respect to h. Thus, under the Babuška–Brezzi condition [4], [6]

$$k_h \ge k_1 > 0, \tag{5.17}$$

we can modify m_h by elements in Ker B_h^T, the modified multiplier being denoted once again by m_h, such that $\{\|m_h\|_N\}$ is uniformly bounded with respect to h. Therefore we can find a subsequence $\{m_{h'}\}$ and an element $\tilde{m} \in Q$, such that

$$m_{h'} \to \tilde{m} \text{ weakly in } N, \quad \text{as } h' \to 0. \tag{5.18}$$

Since $\{\phi_h^*\}$ is weakly precompact in Ψ^*, we can find a further subsequence of the subsequence $\{\phi_{h'}^*\}$, still denoted by $\{\phi_{h'}^*\}$, and a $\tilde{\phi}^* \in \Psi^*$, such that

$$\phi_{h'}^* \to \tilde{\phi}^* \text{ weakly in } \Psi^*. \tag{5.19}$$

From (5.19), the strong convergence $\phi_h \to \phi$, the approximability of any $\psi \in \Psi$ by finite-element functions, and the continuity of j, we get

$$\tilde{\phi}^* \in \partial j(\phi). \tag{5.20}$$

Now fixing a finite-element test function $\psi_h \in \cup_{h'} \Psi_{h'}$, taking the limit in the first relation of (5.7) along the subsequence h', and then approximating an arbitrary test function $\psi \in \Psi$ by ψ_h, we obtain

$$\langle A\phi, \psi\rangle + \langle \tilde{\phi}^*, \psi\rangle + b(\psi, \tilde{m}) = \langle f, \psi\rangle, \quad \forall\, \psi \in \Psi. \tag{5.21}$$

From (5.20), (5.21), and Theorem 5.5, we then know that $\{\phi, \tilde{\phi}^*, \tilde{m}\}$ is a solution of (4.16); in other words, $\{\phi, \tilde{m}\}$ is a solution of Problem P.

So far, we have proved the following.

THEOREM 5.6. *Under the assumptions of Theorem* 5.5 *together with the condition* (5.17), *we have*

$$m_{h'} \to \tilde{m} \quad \textit{weakly in } N,$$

where $m_{h'}$ is a suitably chosen solution of (5.1).

Usually, we can say more about the convergence of the multipliers of the discrete problems for Problems P_2–P_4. From the assumption that D is positively homogeneous (cf. (2.16)), we find that

$$\int_\Omega D(q_h(x))\, dx \ge \int_\Omega D(p_h(x))\, dx + \langle p_h^*, q_h - p_h\rangle, \quad \forall\, q_h \in Q_h$$

is equivalent to the two relations

$$\langle p_h^*, p_h\rangle = \int_\Omega D(p_h(x))\, dx \quad \text{and} \quad \langle p_h^*, q_h\rangle \le \int_\Omega D(q_h(x))\, dx, \quad \forall\, q_h \in Q_h. \tag{5.22}$$

Here Ψ_h will be a product space of the form $\Psi_h = X_h \times Q_h$, where Q_h is the space of discrete plastic strains. Since D is a norm on M^d (see §2) we have

$$\int_\Omega D(q(x))\, dx \le c\, \|q\|_{(L^1(\Omega))^{d\times d}}, \quad \forall\, q \in (L^1(\Omega))^{d\times d},$$

so that

$$\|p_h^*\|_{L^\infty(\Omega)} \le c, \quad \text{for a constant } c \text{ independent of } h. \tag{5.23}$$

Indeed, if (5.23) is not true, we can find a subsequence $\{p_{h'}^*\}$ and a $q \in (L^2(\Omega))^{d\times d}$ such that

$$\|q\|_{(L^1(\Omega))^{d\times d}} = 1 \quad \text{and} \quad \langle p_{h'}^*, q\rangle \to \infty.$$

Let $\Pi_{h'} q \in Q_{h'}$ be the $(L^2(\Omega))^{d\times d}$-projection of q to $Q_{h'}$, then since $\langle p_{h'}^*, q - \Pi_{h'} q\rangle = 0$, we have

$$\langle p_{h'}^*, \Pi_{h'} q\rangle = \langle p_{h'}^*, q\rangle \to \infty. \tag{5.24}$$

On the other hand, since

$$\|q - \Pi_{h'}q\|_{(L^2(\Omega))^{d\times d}} \to 0 \quad \text{as } h' \to 0,$$

the sequence $\{\Pi_{h'}q\}$ is bounded in $(L^1(\Omega))^{d\times d}$. But then from (5.22), we get

$$\langle p^*_{h'}, \Pi_{h'}q\rangle \le \int_\Omega D(\Pi_{h'}q(x))\,dx \le c\,\|\Pi_{h'}q\|_{(L^1(\Omega))^{d\times d}} \le c,$$

which contradicts (5.24).

We incorporate the property (5.23) of elastoplasticity solutions in a more general assumption, namely, that

$$\{\phi^*_h\} \text{ is precompact in } \Psi^*. \tag{5.25}$$

Also assume

$$\mathrm{Ker} B_h^T = \{0\}, \tag{5.26}$$

as is the case for the applications in the next section. We can then further show that $m_h \to \tilde{m}$ strongly in N, for a subsequence $\{m_h\}$. From now on, we will use $\{m_h\}$ and $\{\phi^*_h\}$ to denote the convergent subsequences $\{m_{h'}\}$ and $\{\phi^*_{h'}\}$.

To prove the strong convergence of m_h, we write

$$\|\tilde{m} - m_h\|_N \le \|\tilde{m} - n_h\|_N + \|n_h - m_h\|_N, \quad \forall n_h \in N_h. \tag{5.27}$$

By the condition (5.17),

$$\|n_h - m_h\|_N \le \frac{1}{k_1} \sup_{\psi_h\in\Psi_h} \frac{b(\psi_h, n_h - m_h)}{\|\psi_h\|_\Psi}.$$

Now we have

$$b(\psi_h, n_h - m_h) = b(\psi_h, \tilde{m} - m_h) + b(\psi_h, n_h - \tilde{m})$$

and so, from (4.16) with ϕ^* and m being replaced by $\tilde{\phi}^*$ and $\tilde{m}$, and (5.7), we get

$$b(\psi_h, \tilde{m} - m_h) = -\langle A\phi - A\phi_h, \psi_h\rangle - \langle\tilde{\phi}^* - \phi^*_h, \psi_h\rangle.$$

Thus

$$\begin{aligned}\|n_h - m_h\|_N &\le \frac{1}{k_1} \sup_{\psi_h\in\Psi_h} \frac{1}{\|\psi_h\|_\Psi} \left\{-\langle A\phi - A\phi_h, \psi_h\rangle - \langle\tilde{\phi}^* - \phi^*_h, \psi_h\rangle + b(\psi_h, n_h - \tilde{m})\right\} \\ &\le c\left[\|\phi - \phi_h\|_\Psi + \|\tilde{\phi}^* - \phi^*_h\|_{\Psi^*} + \|\tilde{m} - n_h\|_N\right].\end{aligned}$$

Combining this result with (5.27), we now have

$$\|\tilde{m} - m_h\|_N \le c\left(\|\phi - \phi_h\|_\Psi + \|\tilde{\phi}^* - \phi^*_h\|_{\Psi^*} + \|\tilde{m} - n_h\|_N\right), \quad \forall n_h \in N_h. \tag{5.28}$$

We summarize this result in the following theorem.

THEOREM 5.7. *Under the assumptions made in Theorem* 5.6, *together with* (5.25) *and* (5.26), *for a subsequence* $\{m_h\}$,

$$m_h \to \tilde{m} \text{ strongly in } N.$$

REMARK 5.8. *Theorem* 5.2 *provides an error estimate for* $\phi - \phi_h$. *To estimate the convergence order, for some applications, it is inappropriate to use* (5.12) *to bound* $|I(\xi_h, \phi, \phi^*)|$. *Such is the situation when* $\Psi = [L^2(\Omega)]^d$ *and* $j(\psi) = \int_\Omega k\,|\psi|\,dx$. *One needs to dig into the special structure of the finite-element space* Ψ_h *and try to construct an interpolant* ξ_h *from the set* $Z_h(g)$ *in such a way that no loss in the order of convergence is introduced. For some other applications, however,* (5.12) *readily leads to an optimal error estimate. As an example, when* $\Psi = [H^1(\Omega)]^3$ *and the nondifferentiable functional* j *is of the form* (4.20), *then* $I(\xi_h, \phi, \phi^*)$ *becomes*

$$I(\xi_h, \phi, \phi^*) = j(\xi_h) - j(\phi) - \int_\Omega k\lambda \cdot \nabla(\xi_h - \phi)\,dx$$

for some measurable vector function λ *satisfying* $|\lambda(x)| \le 1$ *a.e. in* Ω. *In this case, the estimate*

$$|I(\xi_h, \phi, \phi^*)| \le c\,\|\nabla(\xi_h - \phi)\|_{L^1(\Omega)} \le c\,\|\xi_h - \phi\|_\Psi$$

does not cause loss in the order of convergence, and the optimal error estimate is

$$\begin{aligned}\|\phi - \phi_h\|_\Psi &\le c\left[\inf_{\xi_h \in Z_h(g)} \|\phi - \xi_h\|_\Psi^{1/2} + \inf_{n_h \in N_h} \|m - n_h\|_N\right]\\ &\le c\left[\left(1 + \frac{1}{\sqrt{k_h}}\right) \inf_{\psi_h \in \Psi_h} \|\phi - \psi_h\|_\Psi^{1/2} + \inf_{n_h \in N_h} \|m - n_h\|_N\right].\end{aligned}$$

6. Application to the elastoplastic problems. We return now to the mixed problems of §3, and apply the results of §5. We discuss in detail finite-element approximations of Problem P_4 only, since the corresponding treatments for Problems P_2 and P_3 follow in a similar way (and are in fact more straightforward).

The condition (4.17) takes a common form for all problems. Since in all cases

$$j(\phi) = \int_\Omega D(p(x))\,dx, \tag{6.1}$$

condition (4.17) states that there is no $\lambda_0 \in \Lambda/\mathrm{Ker}\,B^T$ such that

$$\phi^* \in \partial j(\phi) \iff \int_\Omega [D(q) - D(p) - (p^* + \lambda_0 I) \cdot (q - p)]\;dx \ge 0 \quad \forall\, q \in Q. \tag{6.2}$$

By setting $q = 0$ and $q = 2p$, and by using the fact (see (2.27)) that $p^* \cdot p = D(p)$, we obtain the condition

$$\int_\Omega \lambda_0 \mathrm{tr}\, p\;dx = 0. \tag{6.3}$$

The inequality (6.2) takes a slightly different form in the elastic domain, which is defined by $\Omega^e = \{x \in \Omega : p(x) = 0 \text{ a.e.}\}$. From (6.2) it follows that

$$\int_{\Omega^e} D(q)\,dx \ge \int_{\Omega^e} p^* \cdot q\;dx + \int_{\Omega^e} \lambda_0 \mathrm{tr}\, q\;dx. \tag{6.4}$$

It is not easy to verify that there is no $\lambda_0 \ne 0$ satisfying (6.3) and (6.4). On the other hand, in the fully plastic case, that is, when $\Omega^e = \emptyset$, it is a straightforward matter to verify (4.17).

Also common to all the example problems is the question of the existence of a regularizing sequence j_ε satisfying (4.7)–(4.9). For the case of the von Mises yield condition (see (2.31)) one may set

$$j_\varepsilon(\bar{v}) = \int_\Omega D_\varepsilon(q)\,dx,$$

where

$$D_\varepsilon(q) = k\sqrt{|q|^2 + \varepsilon^2}$$

or

$$D_\varepsilon(q) = \begin{cases} k\,(|q| - \varepsilon/2), & \text{if } |q| \geq \varepsilon, \\ k\,|q|^2/(2\varepsilon), & \text{if } |q| \leq \varepsilon, \end{cases}$$

for example. We recall also from (3.3) that j is Lipschitz continuous.

Going on now to finite-element approximations of Problem P_4, we will assume for simplicity that the domain Ω is polygonal (resp., polyhedral) so that Ω is completely covered by triangular (resp., tetrahedral) elements. We make the identification $\Psi = \overline{H} = H \times Q$, $N = L = W \times \Lambda$, $\phi = \overline{\sigma} = (\sigma, p)$, $\psi = \overline{\tau} = (\tau, q)$, $m = (u, \lambda)$, $n = (v, \mu)$, $A = A_2$, and $b(\cdot,\cdot) = b_3(\cdot,\cdot)$.

Suppose that we choose $H_h \subset H$, $W_h \subset W$, $Q_h \subset Q$, and $\Lambda_h \subset \Lambda$; then $\overline{H}_h = H_h \times Q_h \subset \overline{H}$ and $L_h = W_h \times \Lambda_h \subset L$. We define Problem $\mathrm{P}_{4,h}$.

Problem $\mathrm{P}_{4,h}$. Find $\overline{\sigma}_h = (\sigma_h, p_h) \in \overline{H}_h$ and $m_h = (u_h, \lambda_h) \in L_h$ such that

$$\begin{cases} \langle A_2\overline{\sigma}_h, \overline{\tau}_h - \overline{\sigma}_h\rangle + j(\overline{\tau}_h) - j(\overline{\sigma}_h) + b_3(\overline{\tau}_h - \overline{\sigma}_h, m_h) \geq 0, & \forall \overline{\tau}_h = (\tau_h, q_h) \in \overline{H}_h, \\ b_3(\overline{\sigma}_h, n_h) = \langle g, n_h\rangle, \quad \forall n_h = (v_h, \mu_h) \in L_h. \end{cases} \tag{6.5}$$

Various finite-element spaces have been constructed for the purpose of obtaining stable and convergent approximations for the purely elastic case (see [7]). For the purpose of illustration we consider here the element introduced by Johnson and Mercier [25], in the context of the 2-dimensional problem and assuming isotropic elasticity. In this case the operator A_2 takes the form

$$\langle A_2\overline{\sigma}, \overline{\tau}\rangle = \int_\Omega \left[\left(\frac{1}{2\mu}\sigma^D + p\right)\cdot\tau^D + \frac{1}{\lambda+\mu}(\operatorname{tr}\sigma)(\operatorname{tr}\tau) + (\sigma_0(p) - \sigma)^D \cdot q\right] dx, \tag{6.6}$$

where λ and μ are Lamé's constants. The polygonal domain Ω is partitioned into triangular elements, and the Johnson–Mercier element is constructed as follows. A generic element K is subdivided into three subtriangles K_j, $j = 1, 2, 3$, these having a common vertex at the centroid of K. We then define the space H_K by

$$H_K = \{\tau \in H : \tau|_{K_j} \in [P_1(K_j)]^{2\times 2},\ j = 1, 2, 3\},$$

where $P_1(K_j)$ is the space of the polynomials of degree ≤ 1 on K_j, and the space H_h by

$$H_h = \left\{\tau_h \in H : \tau_h|_K \in H_K,\ \int_\Omega \operatorname{tr}\tau_h\,dx = 0\right\}.$$

The space W_h is simply defined by

$$W_h = \{v_h \in W : v_h|_K \in [P_1(K)]^2\}, \tag{6.7}$$

and we define Q_h and Λ_h by

$$Q_h = \{q_h \in Q : \; q_h|_K \in [P_1(K)]^{2\times 2}\}, \quad \Lambda_h = P_1', \tag{6.8}$$

where $P_1' = \{v \in L^2(\Omega) : \; v|_K$ is a polynomial of degree $1\}$. Then with this choice of spaces it can be shown [25] that the *elastic* version of Problem $\mathrm{P}_{4,h}$ (which is obtained by setting $p_h = q_h = 0$) has a unique solution, and that, if $\sigma \in (H^2(\Omega))^{2\times 2}$ and $u \in (H^2(\Omega))^2$, then

$$||\sigma - \sigma_h||_0 \le Ch^2, \qquad ||u - u_h||_0 \le Ch^2, \tag{6.9}$$

where $||\cdot||_0$ denotes the product L^2-norm.

The proof relies on the fact that the elastic version of A_2 is $\operatorname{Ker} B_h^e$-elliptic, that $\operatorname{Ker} B_h^e \subset \operatorname{Ker} B^e$, and that the discrete condition (5.5) holds, with a constant k_h independent of h. Here, B^e and B_h^e are defined through the bilinear form

$$b^e(\tau, v) = \int_\Omega v \operatorname{div}\tau \, dx$$

by

$$b^e(\tau, v) = \langle B^e\tau, v\rangle = \langle B^{eT}\tau, v\rangle, \quad \forall \tau \in H, v \in W,$$

and

$$b^e(\tau_h, v_h) = \langle B_h^{eT} v_h, \tau_h\rangle = \langle B_h^e \tau_h, v_h\rangle, \quad \forall \tau_h \in H_h, v_h \in W_h.$$

The operator $\tilde{A}_{2,h}$, defined by $\tilde{A}_{2,h}\overline{\tau}_h = A_2(\overline{\tau}_h + \overline{\sigma}_{1h})$ for all $\overline{\tau}_h \in \overline{H}_h$, where $\overline{\sigma}_{1h} \in (\mathrm{Ker}B_h)^\perp$ satisfies $b(\overline{\sigma}_{1h}, n_h) = \langle g, n_h\rangle$ for all $n_h \in N_h$, is shown to be strongly monotone on Ker B_h in the same way as the corresponding result is derived for $\tilde{A}_2$ (see (4.15)).

Properties of the operators B_h and B_h^T follow also by exploiting the properties of the elastic problem. That is, it follows readily from the definition (3.17) of $b = b_3$ and the properties of its elastic part, that $\operatorname{Ker} B_h \subset \operatorname{Ker} B$ and that the bilinear form satisfies the discrete Babuška–Brezzi condition, with $\operatorname{Ker} B_h^T = \{0\}$. The property $\operatorname{Ker} B_h \subset \operatorname{Ker} B$ follows first from

$$\left\{\sigma_h \in H_h : \int_\Omega \operatorname{div}\sigma_h \cdot v_h \, dx = 0 \;\text{ for all } v_h \in W_h\right\} \subset \{\sigma \in H : \operatorname{div}\sigma = 0\}$$

as in the elastic case, and second from

$$\left\{p_h \in Q_h : \int_\Omega \mu_h \operatorname{tr} p_h \, dx = 0, \; \forall \mu_h \in \Lambda_h\right\} \subset \{p \in Q : \operatorname{tr} p = 0\}.$$

Thus Problem $\mathrm{P}_{4,h}$ has a solution $(\sigma_h, p_h) \in \overline{H}_h$ and $(u_h, \lambda_h) \in L_h$, and (σ_h, p_h) is unique. Furthermore, it is possible to show that the multiplier u_h is unique; setting $q_h = p_h$ in (6.5), this problem reduces to

(6.10)
$$\int_\Omega \left[\frac{\sigma_h^D \cdot \tau_h^D}{2\mu} + \frac{(\operatorname{tr}\sigma_h)\,(\operatorname{tr}\tau_h)}{\lambda + \mu}\right] dx + \int_\Omega u_h \cdot \operatorname{div}\tau_h \, dx = -\int_\Omega p_h \cdot \tau_h \, dx, \; \forall \tau_h \in H_h,$$

$$\int_\Omega v_h \cdot \operatorname{div}\sigma_h \, dx = -\int_\Omega b \cdot v_h \, dx, \quad \forall v_h \in W_h, \tag{6.11}$$

and this problem, which is a minor variation of the elastic problem, has a unique solution.

To obtain an error estimate that extends to the present problem, the estimate (6.9) which is valid for the elastic case, we return to Proposition 5.3, and set $\Psi = \overline{\Sigma}$ where $\overline{\Sigma}$ has the same definition as in the previous problem. We also note that A_2 is strongly monotone on $\overline{\Sigma}$, and set

$$Z_h(g) = \{\overline{\tau}_h \in \overline{H}_h : \ b_3(\overline{\tau}_h, n_h) = \langle g, n_h \rangle \ \ \forall n_h \in N_h\}.$$

Then by following the steps taken in the proof of Theorem 5.2, using the inequality

$$\alpha_1 \|\overline{\sigma}_h - \overline{\tau}_h\|_{\overline{\Sigma}}^2 \leq \langle A_2\overline{\sigma}_h - A_2\overline{\tau}_h, \overline{\sigma}_h - \overline{\tau}_h\rangle, \quad \forall \overline{\sigma}_h, \overline{\tau}_h \in \overline{H}_h$$

and noticing that $\mathrm{Ker}B_h \subset \mathrm{Ker}B$, we find that

$$\|\overline{\sigma} - \overline{\sigma}_h\|_{\overline{\Sigma}} \leq c \inf_{\overline{\tau}_h \in Z_h(g)} \left(\|\overline{\sigma} - \overline{\tau}_h\|_{\overline{\Sigma}} + \|p - q_h\|_Q^{1/2} \right).$$

Applying Lemma 5.4, we then find that if $\sigma \in (H^1(\Omega))^{2\times 2}$, $p \in (H^2(\Omega))^{2\times 2}$,

$$\|\sigma - \sigma_h\|_0 \leq ch, \qquad \|p - p_h\|_0 \leq ch.$$

Note in particular the reduction in order; the elastic problem yields an error estimate of $O(h^2)$. This reduction is due to the presence of the nondifferentiable term.

We may also obtain an error estimate for $\|u - u_h\|_0$. We first write down the continuous analogue of (6.10), that is,

$$\int_\Omega \left[\frac{1}{2\mu} \sigma^D \cdot \tau^D + \frac{1}{\lambda + \mu} (\mathrm{tr}\, \sigma)\, (\mathrm{tr}\, \tau) \right] dx + \int_\Omega u \cdot \mathrm{div}\, \tau \, dx = - \int_\Omega p \cdot \tau \, dx, \quad \forall \tau \in H$$

which, together with (6.10), implies the relation

$$\begin{aligned} &\int_\Omega (v_h - u_h) \cdot \mathrm{div}\tau_h \, dx \\ &= - \int_\Omega \left[(p - p_h) \cdot \tau_h + \frac{1}{2\mu} (\sigma^D - \sigma_h^D) \cdot \tau_h^D + \frac{\mathrm{tr}(\sigma - \sigma_h)\, \mathrm{tr}\tau_h}{\lambda + \mu} + (u - v_h) \cdot \mathrm{div}\tau_h \right] dx, \\ &\quad \forall \tau_h \in H_h. \end{aligned}$$

Now from (5.2) of [25],

$$\sup_{\tau_h \in H_h} \frac{(v_h, \mathrm{div}\tau_h)}{\|\tau_h\|_H} \geq \beta \, \|v_h\|_0.$$

Hence we have, for all $v_h \in W_h$,

$$\begin{aligned} &\beta \, \|v_h - u_h\|_0 \\ &\quad \leq \sup_{\tau_h \in H_h} \frac{1}{\|\tau_h\|_H} (v_h - u_h, \mathrm{div}\tau_h) \\ &\quad = \sup_{\tau_h \in H_h} \frac{-1}{\|\tau_h\|_H} \int_\Omega \Big[(p - p_h) \cdot \tau_h + \frac{(\sigma^D - \sigma_h^D) \cdot \tau_h^D}{2\mu} + \frac{\mathrm{tr}(\sigma - \sigma_h)\, \mathrm{tr}\tau_h}{\lambda + \mu} \\ &\qquad\qquad + (u - v_h) \cdot \mathrm{div}\tau_h \Big] dx \\ &\quad \leq c \left[\|p - p_h\|_0 + \|\sigma - \sigma_h\|_0 + \|u - v_h\|_0 \right]. \end{aligned}$$

As a result, from the triangle inequality we get

$$\|u - u_h\|_0 \leq c \left[\|p - p_h\|_0 + \|\sigma - \sigma_h\|_0 + \inf_{v_h \in W_h} \|u - v_h\|_0 \right] \leq ch.$$

7. A posteriori error analysis of regularizing sequences. Because of the difficulty in dealing with the nondifferentiable term j, one rarely solves the finite-element system (5.1) directly. In practice, there are several approaches to circumvent the difficulty caused by the nondifferentiability. One approach is to introduce a Lagrange multiplier for the nondifferentiable term, and the problem (5.1) is solved by an iterative procedure; for details, see, for example, [15]. Here, we concentrate on another approach, namely, the regularization method. The idea of the regularization method is to approximate the nondifferentiable term by a sequence of differentiable ones. The regularizing sequence technique has been used in proving Theorem 4.2 (see [30]). Here, we use the technique as a numerical method to solve the mixed variational inequality. It is easy to give an a priori error estimate which implies convergence of the regularization method (cf. [30]). Our main concern in this section is to derive a posteriori error estimates for solutions of the regularized problems. We will derive such an a posteriori error estimate for solving Problem P_2. For Problems P_3 and P_4, the same techniques presented here can be employed to give similar a posteriori error estimates.

As in [17], we need a result from convex analysis (cf. [10]).

Let V, Λ be two normed spaces, with V^*, Λ^* their dual spaces. Assume there exists a linear continuous operator $F \in \mathcal{L}(V, \Lambda)$, with transpose $F^* \in \mathcal{L}(\Lambda^*, V^*)$. Let J be a function mapping $V \times \Lambda$ into $R \cup \{+\infty\}$. Consider the minimization problem

$$\inf_{v \in V} J(v, Fv). \tag{7.1}$$

Define the conjugate function of J by

$$J^*(v^*, \mu^*) = \sup_{v \in V, \mu \in \Lambda} \left[\langle v, v^* \rangle + \langle \mu, \mu^* \rangle - J(v, \mu) \right]. \tag{7.2}$$

THEOREM 7.1. *Assume that*

(i) V *is a reflexive Banach space,* Λ *is a normed space.*

(ii) $J : V \times \Lambda \to R \cup \{+\infty\}$ *is a proper, lower semicontinuous, strictly convex function.*

(iii) $\exists\, u_0 \in V$, *such that* $J(u_0, Fu_0) < \infty$ *and* $\mu \mapsto J(u_0, \mu)$ *is continuous at* Fu_0.

(iv) $J(v, Fv) \to +\infty$, *as* $\|v\| \to \infty$, $v \in V$.

Then problem (7.1) *has a unique solution* $u \in V$, *and*

$$-J(u, Fu) \le J^*(F^*\mu^*, -\mu^*), \quad \forall\, \mu^* \in \Lambda^*. \tag{7.3}$$

We will apply Theorem 7.1 to derive an a posteriori error estimate for the regularizing technique for solving (3.6), that is, Problem P_2, and its discrete version, in the context of the von Mises yield condition. Instead of the Problem P_2, we consider a slightly more general problem, namely, the constraint $b_1(\overline{u}, \mu) = 0$, $\forall\, \mu \in \Lambda$ is replaced by

$$b_1(\overline{u}, \mu) = \langle g, \mu \rangle, \quad \forall\, \mu \in \Lambda. \tag{7.4}$$

In this way, one will see more clearly how to employ the techniques presented here to derive a posteriori error estimates for Problems P_3 and P_4. We choose the following regularizing function for the dissipation function:

$$D_\varepsilon(q) = k\sqrt{|q|^2 + \varepsilon^2}. \tag{7.5}$$

First, we need to rewrite the problem (3.6) in the form of (7.1). To do this, set

$$S = \{s = (s_{ij}) : s_{ij} = s_{ji} \in L^2(\Omega),\ 1 \le i,j \le d\}$$

and identify S^* with S. We make use of the spaces V, Q, Λ and $\overline{V} = V \times Q$ used earlier in Problem P_2, and define the operator $F : \bar{V} \to S$ by

$$F\bar{v} = \epsilon(v), \quad \forall\, \bar{v} \in \bar{V}.$$

Let

$$Z(g) = \{\bar{v} \in \overline{V} : b_1(\bar{v}, \mu) = \langle g, \mu\rangle,\ \forall\, \mu \in \Lambda\}.$$

We now define the energy function on $\bar{V} \times S$ by

(7.6)

$$J(\bar{v}, s) = \begin{cases} \displaystyle\int_\Omega \left[\tfrac{1}{2}\, C\,(s-q)\cdot(s-q) + H(|q|) + k\,|q| - b\cdot v\right] dx, & \text{if } \bar{v} \in Z(g), \\ +\infty, \quad \text{otherwise}, \end{cases}$$

where

$$H(\alpha) = \frac{1}{2}\, h_0 \alpha^2 + \frac{1}{\nu^2}\, h_1(1 - e^{-\nu\alpha}) - \frac{1}{\nu}\, h_1 \alpha\, e^{-\nu\alpha}$$

(cf. (2.10) and (2.11)). Then it can be shown that the problem (3.6) with the more general constraint (7.4), is equivalent to the minimization problem

$$\bar{u} \in \bar{V}, \quad J(\bar{u}, F\bar{u}) = \inf_{\bar{v} \in \bar{V}} J(\bar{v}, F\bar{v}).$$

To use Theorem 7.1, we need to compute $J^*(F^*s^*, -s^*)$, for $s^* \in S^*$. We have

$$\begin{aligned} J^*(F^*s^*, -s^*) &= \sup_{\bar{v} \in \bar{V}, s \in S} \left[\langle \bar{v}, F^*s^*\rangle - \langle s, s^*\rangle - J(\bar{v}, s)\right] \\ &= \sup_{\bar{v} \in \bar{V}, s \in S} \left[\langle F\bar{v}, s^*\rangle - \langle s, s^*\rangle - J(\bar{v}, s)\right] \\ &= \sup_{\bar{v} \in Z(g), s \in S} \int_\Omega \Big[\epsilon(v)\cdot s^* - s\cdot s^* - \tfrac{1}{2}\, C\,(s-q)\cdot(s-q) \\ &\qquad\qquad - H(|q|) - k\,|q| + bv\Big]\, dx \\ &= \tfrac{1}{2} \int_\Omega C^{-1} s^* \cdot s^* dx + \sup_{\bar{v} \in Z(g)} \int_\Omega \left[\epsilon(v)\cdot s^* + b\cdot v\right] dx + \int_\Omega K(|s^*|)\, dx, \end{aligned}$$

where

(7.7)
$$K(|s^*|) = T(t(|s^*|)),$$

with

$$T(t) = (|s^*| - k)\, t - H(t),$$

and $t(|s^*|) = 0$ if $|s^*| \le k$, $t(|s^*|) > 0$ being the unique solution of the equation (the unique solvability is guaranteed by the assumption $h_0 > 0$)

$$(h_0 + h_1 e^{-\nu t})\, t = |s^*| - k \quad \text{if } |s^*| > k.$$

Next, we deal with the term

$$\sup_{\overline{v}\in Z(g)} \int_\Omega [\epsilon(v)\cdot s^* + b\cdot v]\,dx.$$

We have

$$\sup_{\overline{v}\in Z(g)} \int_\Omega [\epsilon(v)\cdot s^* + b\cdot v]\,dx = \int_\Omega [\epsilon(u_\varepsilon)\cdot s^* + b\cdot u_\varepsilon]\,dx + \sup_{\overline{v}\in Z(0)} \int_\Omega [\epsilon(v)\cdot s^* + b\cdot v]\,dx$$

$$= \begin{cases} \displaystyle\int_\Omega [\epsilon(u_\varepsilon)\cdot s^* + b\cdot u_\varepsilon]\,dx \\ \qquad \text{if } \displaystyle\int_\Omega [\epsilon(v)\cdot s^* + b\cdot v]\,dx = 0,\ \forall \overline{v}\in Z(0), \\ +\infty \quad \text{otherwise.} \end{cases}$$

Applying (4.11) to the problem (3.6), we find that

$$\langle A\bar{u}_\varepsilon, \overline{v}\rangle + \langle j_\varepsilon'(\bar{u}_\varepsilon), \overline{v}\rangle = \langle b, \overline{v}\rangle, \quad \forall \overline{v}\in Z(0),$$

that is, $\forall \overline{v}\in Z(0)$,

$$\int_\Omega \left[C\left(\epsilon(u_\varepsilon) - p_\varepsilon\right)\cdot(\epsilon(v) - q) + h(|p_\varepsilon|)\,p_\varepsilon\cdot q + \frac{k p_\varepsilon\cdot q}{\sqrt{|p_\varepsilon|^2+\varepsilon^2}}\right] dx = \int_\Omega b\cdot v\,dx.$$

Hence,

$$-\,C\left(\epsilon(u_\varepsilon) - p_\varepsilon\right) + h(|p_\varepsilon|)\,p_\varepsilon + \frac{k p_\varepsilon}{\sqrt{|p_\varepsilon|^2+\varepsilon^2}} = 0, \tag{7.8}$$

$$\int_\Omega C\left(\epsilon(u_\varepsilon) - p_\varepsilon\right)\cdot\epsilon(v)\,dx = \int_\Omega b\cdot v\,dx, \quad \forall v\in \mathrm{Ker}B. \tag{7.9}$$

With (7.9), we choose

$$s^* = -C\left(\epsilon(u_\varepsilon) - p_\varepsilon\right); \tag{7.10}$$

then

$$\sup_{v\in Z(g)} \int_\Omega [\epsilon(v)\cdot s^* + b\cdot v]\,dx = \int_\Omega \left[-C\left(\epsilon(u_\varepsilon) - p_\varepsilon\right)\cdot\epsilon(u_\varepsilon) + b\cdot u_\varepsilon\right]\,dx.$$

Therefore, with the choice (7.10) for the dual variable s^*, we have

$$\begin{aligned} J^*(F^*s^*, -s^*) = \int_\Omega \Big[&\tfrac{1}{2} C\left(\epsilon(u_\varepsilon) - p_\varepsilon\right)\cdot\left(\epsilon(u_\varepsilon) - p_\varepsilon\right) \\ &- C\left(\epsilon(u_\varepsilon) - p_\varepsilon\right)\cdot\epsilon(u_\varepsilon) + b\cdot u_\varepsilon + K(|C\left(\epsilon(u_\varepsilon) - p_\varepsilon\right)|)\Big]\,dx. \end{aligned} \tag{7.11}$$

Now consider the difference

$$J(\bar{u}_\varepsilon, F\bar{u}_\varepsilon) - J(\bar{u}, F\bar{u}).$$

By Theorem 7.1, an upper bound for the difference, with s^* given by (7.10), is

$$\begin{aligned} &J(\bar{u}_\varepsilon, F\bar{u}_\varepsilon) - J(\bar{u}, F\bar{u}) \\ &\quad \le J(\bar{u}_\varepsilon, F\bar{u}_\varepsilon) + J^*(F^*s^*, -s^*) \\ &\quad = \int_\Omega \left[k\,|p_\varepsilon|\,\frac{\sqrt{|p_\varepsilon|^2+\varepsilon^2} - |p_\varepsilon|}{\sqrt{|p_\varepsilon|^2+\varepsilon^2}} + H(|p_\varepsilon|) - h(|p_\varepsilon|)\,|p_\varepsilon|^2 + K(|C(\epsilon(u_\varepsilon) - p_\varepsilon)|)\right] dx. \end{aligned}$$

In the derivation above, we used the relation (7.8). We then turn to a lower bound of the difference. Taking $\bar{v} = \bar{u}_\varepsilon$ in $(3.6)_1$, we obtain

$$\begin{aligned}(7.12)\quad &\int_\Omega [k\,|p_\varepsilon| - k\,|p| - b\cdot(u_\varepsilon - u)]\,dx \\ &\geq \int_\Omega [-C\,(\epsilon(u)-p)\cdot((\epsilon(u_\varepsilon)-p_\varepsilon)-(\epsilon(u)-p)) - h(|p|)\,p\cdot(p_\varepsilon - p)]\,dx.\end{aligned}$$

Thus

$$\begin{aligned}&J(\bar{u}_\varepsilon, F\bar{u}_\varepsilon) - J(\bar{u}, F\bar{u}) \\ &\quad = \int_\Omega \Big[\tfrac{1}{2}C\,(\epsilon(u_\varepsilon)-p_\varepsilon)\cdot(\epsilon(u_\varepsilon)-p_\varepsilon) + H(|p_\varepsilon|) + k\,|p_\varepsilon| - b\cdot u_\varepsilon \\ &\qquad\qquad - \tfrac{1}{2}C\,(\epsilon(u)-p)\cdot(\epsilon(u)-p) - H(|p|) - k\,|p| + b\cdot u\Big]\,dx \\ &\quad \geq \int_\Omega \Big\{\tfrac{1}{2}c_0\,|(\epsilon(u)-p)-(\epsilon(u_\varepsilon)-p_\varepsilon)|^2 + H(|p_\varepsilon|) - H(|p|) - h(|p|)\,p\cdot(p_\varepsilon - p)\Big\}\,dx\end{aligned}$$

where we have made use of (7.12).

Now define the function

$$H_1(\alpha) = H(\alpha) - \frac{1}{2}\,h_0\alpha^2 = h_1\left[\frac{1}{\nu^2}(1-e^{-\nu\alpha}) - \frac{1}{\nu}\alpha\,e^{-\nu\alpha}\right],$$

the part of H related to h_1. Then

$$H_1''(\alpha) = h_1(1-\nu\alpha)\,e^{-\nu\alpha} \geq -e^{-2}h_1 > -\theta h_0,$$

using (2.14). Hence,

$$\begin{aligned}&H(|p_\varepsilon|) - H(|p|) - h(|p|)\,p\cdot(p_\varepsilon - p) \\ &\quad = \tfrac{1}{2}h_0(|p_\varepsilon|^2 - |p|^2 - 2p\cdot(p_\varepsilon - p)) + H_1(|p_\varepsilon|) - H_1(|p|) - h_1 e^{-\nu|p|}p\cdot(p_\varepsilon - p) \\ &\quad \geq \frac{h_0}{2}\,|p_\varepsilon - p|^2 + H_1(|p_\varepsilon|) - H_1(|p|) - H_1'(|p_\varepsilon| - |p|) \\ &\quad \geq \frac{h_0}{2}\,|p_\varepsilon - p|^2 - \frac{\theta h_0}{2}\,|\,|p_\varepsilon| - |p|\,|^2 \\ &\quad \geq \frac{(1-\theta)h_0}{2}\,|p_\varepsilon - p|^2.\end{aligned}$$

Thus

$$\begin{aligned}&J(\bar{u}_\varepsilon, F\bar{u}_\varepsilon) - J(\bar{u}, F\bar{u}) \\ &\quad \geq \int_\Omega \Big\{\tfrac{1}{2}c_0\,|(\epsilon(u)-\epsilon(u_\varepsilon)) - (p-p_\varepsilon)|^2 + \tfrac{1}{2}\,(1-\theta)\,h_0|p-p_\varepsilon|^2\Big\}\,dx \\ &\quad \geq \bar{\alpha}\,(||u-u_\varepsilon||_V^2 + ||p-p_\varepsilon||_Q^2)\end{aligned}$$

where

$$\bar{\alpha} = \frac{1}{2}\,\eta_0(1-\theta)\,\min\left\{1, \frac{Kc_0}{c_0 + \eta_0(1-\theta)/2}\right\};$$

the last inequality is obtained using the trick employed in proving Lemma 2.1 in [32].

Combining the two bounds on the difference $J(\bar{u}_\varepsilon, F\bar{u}_\varepsilon) - J(\bar{u}, F\bar{u})$, we then have the a posteriori error estimate for the regularizing technique for solving the problem (3.6).

THEOREM 7.2. *Under the assumptions made on the problem* (3.6), *the following inequality holds:*

$$\begin{aligned} &\bar{\alpha}\left(\|u-u_\varepsilon\|_V^2 + \|p-p_\varepsilon\|_Q^2\right) \\ &\quad \le \int_\Omega \Bigg[\frac{k\,|p_\varepsilon|\,\varepsilon^2}{\sqrt{|p_\varepsilon|^2+\varepsilon^2}\left(\sqrt{|p_\varepsilon|^2+\varepsilon^2}+|p_\varepsilon|\right)} \\ &\qquad + H(|p_\varepsilon|) - h(|p_\varepsilon|)\,|p_\varepsilon|^2 + K(|C(\epsilon(u_\varepsilon)-p_\varepsilon)|)\Bigg]\,dx. \end{aligned} \tag{7.13}$$

To see more clearly the effectiveness of the a posteriori error estimate (7.13), we consider the simpler case when the material undergoes linear hardening, that is, when the function $h(\alpha)$ in (2.11) is of the form

$$h(\alpha) = h_0,$$

and $h_0(x) \ge \eta_0 > 0$, a.e. in Ω. We can compute

$$K(|s^*|) = \frac{1}{2h_0}\left[(|s^*|-k)_+\right]^2,$$

where

$$t_+ = \begin{cases} t, & \text{if } t \ge 0, \\ 0, & \text{if } t < 0. \end{cases}$$

In this special case of linear hardening, the a posteriori error estimate assumes the simpler form

$$\begin{aligned} &\bar{\alpha}\left(\|u-u_\varepsilon\|_V^2 + \|p-p_\varepsilon\|_Q^2\right) \\ &\quad \le \int_\Omega \Bigg[\frac{k\,|p_\varepsilon|\,\varepsilon^2}{\sqrt{|p_\varepsilon|^2+\varepsilon^2}\left(\sqrt{|p_\varepsilon|^2+\varepsilon^2}+|p_\varepsilon|\right)} \\ &\qquad - \tfrac{1}{2}\,h_0|p_\varepsilon|^2 + \frac{1}{2h_0}\left[(|C(\epsilon(u_\varepsilon)-p_\varepsilon)|-k)_+\right]^2\Bigg]\,dx, \end{aligned} \tag{7.14}$$

where

$$\bar{\alpha} = \frac{1}{2}\,\eta_0 \min\left\{1, \frac{Kc_0}{c_0+\eta_0/2}\right\}.$$

If

$$(|C(\epsilon(u_\varepsilon)-p_\varepsilon)|-k)_+ \le h_0|p_\varepsilon|, \tag{7.15}$$

then from (7.14) we have

$$\bar{\alpha}\left(\|u-u_\varepsilon\|_V^2 + \|p-p_\varepsilon\|_Q^2\right) \le \int_\Omega \frac{k\,|p_\varepsilon|\,\varepsilon^2}{\sqrt{|p_\varepsilon|^2+\varepsilon^2}\left(\sqrt{|p_\varepsilon|^2+\varepsilon^2}+|p_\varepsilon|\right)}\,dx$$

which indicates that (7.14) (and (7.13), at least when h_1 is small) is a useful a posteriori error estimate. To prove (7.15), we notice that from (7.8),

$$C(\epsilon(u_\varepsilon) - p_\varepsilon) = \left(h_0 + \frac{k}{\sqrt{|p_\varepsilon|^2 + \varepsilon^2}}\right) p_\varepsilon.$$

Thus

$$|C(\epsilon(u_\varepsilon) - p_\varepsilon)| = \left(h_0 + \frac{k}{\sqrt{|p_\varepsilon|^2 + \varepsilon^2}}\right) |p_\varepsilon|,$$

and so

$$|C(\epsilon(u_\varepsilon) - p_\varepsilon)| - k = h_0|p_\varepsilon| - \frac{k\,\varepsilon^2}{\sqrt{|p_\varepsilon|^2 + \varepsilon^2}\left(\sqrt{|p_\varepsilon|^2 + \varepsilon^2} + |p_\varepsilon|\right)}.$$

Obviously,

$$\left(h_0|p_\varepsilon| - \frac{k\,\varepsilon^2}{\sqrt{|p_\varepsilon|^2 + \varepsilon^2}\left(\sqrt{|p_\varepsilon|^2 + \varepsilon^2} + |p_\varepsilon|\right)}\right)_+ \le h_0|p_\varepsilon|.$$

Therefore, (7.15) follows.

For the finite-element system (5.1), we can also use the regularization technique. So instead of solving (5.1), which is difficult because of the presence of the nondifferentiable term, we solve a sequence of regularized problems: find $u_{h,\varepsilon} \in V_h$ and $p_{h,\varepsilon} \in Q_h$, such that

$$\begin{cases} \langle Au_{h,\varepsilon}, v_h - u_{h,\varepsilon}\rangle + j_\varepsilon(v_h) - j_\varepsilon(u_{h,\varepsilon}) + b(v_h - u_{h,\varepsilon}, p_{h,\varepsilon}) \ge \langle b, v_h - u_{h,\varepsilon}\rangle, \ \forall\, v_h \in V_h, \\ b(u_{h,\varepsilon}, q_h) = \langle g, q_h\rangle, \ \forall\, q_h \in Q_h. \end{cases} \tag{7.16}$$

We can apply the results in Theorem 7.2 to the discrete problems, (5.1) and (7.16), to obtain the a posteriori error estimate

$$\begin{aligned} &\bar{\alpha}\left(||u - u_\varepsilon||_V^2 + ||p - p_\varepsilon||_Q^2\right) \\ &\quad \le \int_\Omega \left[\frac{k\,|p_{h,\varepsilon}|\,\varepsilon^2}{\sqrt{|p_{h,\varepsilon}|^2 + \varepsilon^2}\left(\sqrt{|p_{h,\varepsilon}|^2 + \varepsilon^2} + |p_{h,\varepsilon}|\right)} \right. \\ &\qquad \left. + H(|p_{h,\varepsilon}|) - h_0|p_{h,\varepsilon}|^2 + K(|C(\epsilon(u_{h,\varepsilon}) - p_{h,\varepsilon})|)\right] dx. \end{aligned} \tag{7.17}$$

Note that the computable error estimate (7.17) can help one to determine whether a solution of the regularized problem can be accepted as the solution of the original finite-element problem.

Acknowledgment. The authors thank an anonymous referee for comments that led to an improvement in the presentation of this paper.

REFERENCES

[1] D. N. Arnold, F. Brezzi, and J. Douglas, *PEERS: a new mixed finite element for plane elasticity*, Japan J. Appl. Math., 1 (1984), pp. 347–367.

[2] D. N. Arnold, J. Douglas, and C. P. Gupta, *A family of higher order mixed finite element methods for plane elasticity*, Numer. Math., 45 (1984), pp. 1–22.

[3] D. N. Arnold and R. Falk, *A new mixed formulation for elasticity*, Numer. Math., 53 (1988), pp. 13–30.

[4] I. Babuška and A. K. Aziz, *Survey lectures on the mathematical foundations of the finite element method*, in The Mathematical Foundations of the Finite Element Method with Applications to Partial Differential Equations, A. K. Aziz, ed., Academic Press, New York, 1972, pp. 3–359.

[5] F. Barlat and J. Lian, *Plastic behavior and stretchability of sheet metals. Part I: A yield function for orthotropic sheets under plane stress conditions*, Internat. J. Plasticity, 5 (1989), pp. 51–66.

[6] F. Brezzi, *On the existence, uniqueness and approximation of saddle-point problems arising from Lagrange multipliers*, RAIRO Modél Math. Anal. Numér., 8 (1974), pp. 129–151.

[7] F. Brezzi and M. Fortin, *Mixed and Hybrid Finite Element Methods*, Springer-Verlag, Berlin, 1991.

[8] F. Brezzi, W. W. Hager, and P. A. Raviart, *Error estimates for the finite element solution of variational inequalities. Part II. Mixed methods*, Numer. Math., 31 (1978), pp. 1-16.

[9] F. Brezzi, C. Johnson, and B. Mercier, *Analysis of a mixed finite element method for elasto-plastic plates*, Math. Comp. 31 (1977), pp. 809–817.

[10] I. Ekeland and R. Temam, *Convex Analysis and Variational Problems*, North-Holland, Amsterdam, 1976.

[11] C. M. Elliott, A. Mikelič, and M. Shillor, *Constrained anisotropic elastic materials in unilateral contact with or without friction*, Nonlinear Anal., 16 (1991), pp. 155–182.

[12] R. A. Eve and B. D. Reddy, *The variational formulation and solution of problems of finite-strain elastoplasticity based on the use of a dissipation function*, Internat. J. Numer. Methods Engrg., 37 (1994), pp. 1673–1695.

[13] R. A. Eve, B. D. Reddy, and R. T. Rockafellar, *An internal variable theory of elastoplasticity based on the maximum plastic work inequality*, Quart. Appl. Math., 48 (1990), pp. 59–84.

[14] V. Girault and P. A. Raviart, *Finite Element Methods for Navier-Stokes Equations*, Springer, Berlin, 1986.

[15] R. Glowinski, *Numerical Methods for Nonlinear Variational Problems*, Springer-Verlag, New York, 1984.

[16] R. Glowinski, J. L. Lions, and R. Trémolières, *Numerical Analysis of Variational Inequalities*, North-Holland, Amsterdam, 1981.

[17] W. Han, *Finite element analysis of a holonomic elastic-plastic problem*, Numer. Math., 60 (1992), pp. 493–508.

[18] ———, *Quantitative error estimates in modeling the laminar stationary flow of a Bingham fluid*, Appl. Math. Comput., 47 (1992), pp. 15–24.

[19] K. S. Havner and H. P. Patel, *On convergence of the finite element method for a class of elastic-plastic solids*, Quart. Appl. Math., 34 (1976), pp. 59–68.

[20] I. Hlaváček, *A finite element solution for plasticity with strain-hardening*, RAIRO Modél Math. Anal. Numér., 14 (1980), pp. 347–368.

[21] I. Hlaváček, J. Haslinger, J. Nečas, and J. Lovíšek, *Solution of Variational Inequalities in Mechanics*, Springer-Verlag, New York, 1988.

[22] L. Jiang, *On an elastic-plastic problem*, J. Differential Equations, 51 (1984), pp. 97–115.

[23] C. Johnson, *On finite element methods for plasticity problems*, Numer. Math., 26 (1976), pp. 79–84.

[24] C. Johnson, *A mixed finite element method for plasticity problems with hardening*, SIAM J. Numer. Anal., 14 (1977), pp. 575–583.

[25] C. Johnson and B. Mercier, *Some equilibrium finite elements for two-dimensional elasticity problems*, Numer. Math., 30 (1978), pp. 103–116.

[26] B. Mercier, *On the finite element approximation and the solution by a penalty-duality algorithm of an elasto-plastic problem*, Rapport IRIA Laboria, 149, 1975.

[27] J. T. Oden and J. R. Whiteman, *Analysis of some finite element methods for a class of problems in elasto-plasticity*, Internat. J. Sci. Engrg., 20 (1982), pp. 977–988.

[28] T. H. H. Pian, *Finite elements based on consistently assumed stresses and displacements*, Finite Elem. Anal. Des., 1 (1985), pp. 131–140.

[29] J. PITKÄRANTA AND R. STENBERG, *Analysis of some mixed finite element methods for plane elasticity equations*, Math. Comp., 41 (1983), pp. 399–423.

[30] B. D. REDDY, *Mixed variational inequalities arising in elastoplasticity*, Nonlinear Anal., 19 (1992), pp. 1071–1089.

[31] ———, *Existence of solutions to quasistatic problems in plasticity*, in Progress in Partial Differential Equations: Calculus of Variations, Applications, C. Bandle et al., eds., Longman, London, 1992, pp. 299–311.

[32] B. D. REDDY AND T. B. GRIFFIN, *Variational principles and convergence of finite element approximations of a holonomic elastic-plastic problem*, Numer. Math., 52 (1988), pp. 101–117.

[33] B. D. REDDY AND J. B. MARTIN, *Algorithms for the solution of internal variable problems in plasticity*, Comput. Methods Appl. Mech. Engrg., 93 (1991), pp. 253–273.

[34] R. REDHEFFER AND W. WALTER, *The subgradient in R^n*, Nonlinear Anal., 20 (1993), pp. 1345–1348.

[35] R. T. ROCKAFELLAR, *Convex Analysis*, Princeton University Press, New Jersey, 1970.

[36] J. C. SIMO AND R. L. TAYLOR, *Consistent tangent operators for rate-independent elastoplasticity*, Comput. Methods Appl. Mech. Engrg., 48 (1985), pp. 101–118.

[37] E. STEIN AND R. ROLFES, *Mechanical conditions for stability and optimal convergence of mixed finite elements for linear plane elasticity*, Comp. Methods. Appl. Mech. Engrg., 84 (1990), pp. 77–95.

[38] R. STENBERG, *On the construction of optimal mixed finite element methods for the linear elasticity problem*, Numer. Math., 48 (1986), pp. 447–462.

[39] ———, *A family of mixed finite elements for the elasticity problem*, Numer. Math., 53 (1988), pp. 513–538.

[40] R. TEMAM, *Mathematical Problems in Plasticity*, Gauthier–Villars, Paris, 1985.