DOMAIN DECOMPOSITION FOR A NON-SMOOTH CONVEX MINIMIZATION PROBLEM AND ITS APPLICATION TO PLASTICITY

CARSTEN CARSTENSEN

Abstract. P.L. Lions's work on the Schwarz alternating method for convex minimization problems is generalized to a certain non-smooth situation where the non-differentiable part of the functionals is additive and independent with respect to the decomposition. Such functionals arise naturally in plasticity where the material law is a variational inequality formulated in $L^2(\Omega)$. The application to plasticity with hardening is sketched and gives contraction numbers which are independent of the discretization parameter h and of a possible regularization of the non-smooth material law.

1. Introduction

Domain decomposition is proved to be an important tool for parallel computations in the numerical analysis of partial differential equations. *The Schwarz method does converge for many different types of problems* because of *two reasons: the first is that it has a variational interpretation and* the *second one is its interpretation in terms of the maximum principle* [L, p.2]. Motivated by the study of elastoplasticity (see §4), we consider a minimization problem for $f = \phi + \psi$ where $\phi : X \to \mathbb{R}$ is smooth and uniformly convex, but $\psi : X \to (-\infty, +\infty]$ is just convex and may be some discontinuous indicator function (of a convex closed subset). We stress that the minimization problem can also be recast as a variational inequality where we seek $u \in X$ such that $\psi(u) < \infty$ and

$$D\phi(u; u - v) \leq \psi(v) - \psi(u) \qquad (v \in X). \tag{1.1}$$

P.L. Lions proved geometric convergence of the Schwarz alternating method, a domain decomposition method, in case $\psi = 0$ by variational arguments and for a special case for $\psi \neq 0$ by using the maximum principle [L]. Therefore, an extension of his proof to plasticity problems seems non-obvious.

In this paper, we make use of the observation that, typically in plasticity, the material law is local, i.e., the non-smooth contributions are formulated pointwise in the domain Ω. Then, a direct decomposition is possible leading to a *non-overlapping* partition of internal and stress variables. Loosely speaking, the goal of this paper is to prove that the following philosophy is actually true: if there is no overlapping in the decomposition of some components, then the functional is allowed to be non-smooth in these components. A precise statement will be given in §2 where Theorem 2.1 guarantees linear convergence of the domain decomposition method (with a contraction number which is independent of discretization parameters). An outline of the rest of the paper is as follows: All hypotheses are introduced in §2 where the convergence result is stated in Theorem 2.1 and proved in §3. After a brief introduction of elastoplasticity with hardening in §4, examples of application of domain decomposition methods in plasticity conclude the paper with §5.

2. General result

Let $X_1, \ldots, X_J$ be $J > 1$ subspaces of the normed vector space X with dual X^* satisfying

$$X = \sum_{j=1}^{J} X_j. \tag{2.1}$$

We assume that there are linear mappings $P_j : X \to X_j$, $j \in \{1, \ldots, J\}$, with

$$P_1 + \cdots + P_J = \mathrm{Id} \tag{2.2}$$

1991 *Mathematics Subject Classification.* 65N30, 65R20, 73C50.

Key words and phrases. domain decomposition, convex minimization, variational inequality, elastoplasticity, plasticity with hardening.

and that there exists a constant $C_X > 0$ such that for all subsets $\Lambda \subseteq \{1, \ldots, J\}$ and all $x \in X$ there holds

$$\| \sum_{\lambda \in \Lambda} P_\lambda x \|_X \leq C_X \cdot \| x \|_X . \tag{2.3}$$

We consider the minimization of a functional $f : X \to (-\infty, +\infty]$ which can be decomposed as

$$f = \phi + \psi \tag{2.4}$$

where $\phi : X \to \mathbb{R}$ is Fréchet-differentiable and $D\phi$ is uniformly elliptic and Lipschitz continuous, i.e., there are two constants α and L $(0 < \alpha, L < \infty)$ such that the Fréchet-derivative satisfies, for all $u, v \in X$,

$$\alpha \cdot \| u - v \|_X^2 + D\phi(u; v - u) \leq \phi(v) - \phi(u) \tag{2.5}$$

$$\| D\phi(u; \cdot) - D\phi(v; \cdot) \|_{X^*} \leq L \cdot \| u - v \|_X . \tag{2.6}$$

The functional $\psi : X \to (-\infty, +\infty]$ is convex, lower semi-continuous and ψ is assumed to be additive and independent with respect to the partition (2.1) in the sense that, for all $(x_1, \ldots, x_J) \in X_1 \times \cdots \times X_J$,

$$\psi(\sum_{j=1}^J x_j) = \sum_{j=1}^J \psi(x_j) \tag{2.7}$$

and, for all $j \in \{1, \ldots, J\}$, for all $x_j \in X_j$ and for all $y_j \in \sum_{k=1, k\neq j}^J X_k$, there holds

$$\psi(x_j + P_j y_j) = \psi(x_j). \tag{2.8}$$

Under the preceding assumptions, we consider the following domain decomposition method.

Algorithm 2.1. *Given u_0, compute a sequence $(u_\nu)_{\nu \in I} \in X^I$, $I := \{0, 1/J,\ 2/J,\ 3/J,\ \ldots\ \}$: For all $n = 0, 1, 2, 3, \ldots$ and all $j = 1, \ldots, J$ determine $u_{n+j/J} \in u_{n+(j-1)/J} + X_j$ that minimizes the functional f on the affine subspace $u_{n+(j-1)/J} + X_j$.*

Remark 2.1. The assumptions guarantee that f is coercive and strictly convex. Hence there exists a unique minimizer u of f in X and each minimization problem in Algorithm 2.1 has a unique solution [ET, GLT, Z]; in other words, Algorithm 2.1 is feasible.

Theorem 2.1. *Let u be the minimizer of f in X and let the sequence $(u_\nu)_{\nu \in I}$ be generated by Algorithm 2.1 with initial data u_0. Define $q := \gamma/(1+\gamma)$, $\gamma := J \cdot C_X^2 \cdot L^2 \cdot \alpha^{-2}/2$ and $C_0 := 2(1+\gamma) \cdot \alpha^{-1} \cdot (f(u_0) - f(u))$. Then, for any $\nu \geq 1$ there holds*

$$\| u - u_\nu \|_X^2 \leq C_0 \cdot q^{[\nu]}. \tag{2.9}$$

Remark 2.2. In case $\psi = 0$, Theorem 2.1 was proved by P.L. Lions in [L, Theorem I.3].

Remark 2.3. The philosophy mentioned in the introduction is hidden in the assumptions (2.7) and (2.8): No smoothness on ψ is required because the minimization of ψ globally (i.e., on X) can be performed locally (i.e., on each X_j separately).

3. Proof of the theorem

Let $\nu = n + j/J \geq 1$. Since $f(\cdot + u_\nu)|_{X_j}$ has a minimum at 0, we have from $0 \in \partial f(u_\nu)|_{X_\nu}$ and the sum rule $\partial f(\cdot + u_\nu)|_{X_j} = \partial\phi(\cdot + u_\nu)|_{X_j} + \partial\psi(\cdot + u_\nu)|_{X_j}$ [ET, Z], for all $\eta \in X_j$,

$$D\phi(u_\nu; \eta) \leq \psi(u_\nu - \eta) - \psi(u_\nu). \tag{3.1}$$

Combining (2.5) and (3.1) for $\eta = u_\nu - u_{\nu-1/J}$ we conclude

$$\alpha \cdot \| u_\nu - u_{\nu-1/J} \|_X^2 \leq f(u_{\nu-1/J}) - f(u_\nu). \tag{3.2}$$

According to (2.5),

$$\alpha \cdot \| u - u_\nu \|_X^2 + \phi(u_\nu) - \phi(u) \leq D\phi(u_\nu; u_\nu - u). \tag{3.3}$$

Utilizing (2.2) and understanding all lower indices in P_ℓ and X_ℓ modulo J (i.e., e.g., $P_\ell = P_m$ iff $\ell \equiv m \pmod{J}$), we write

$$u_\nu - u = \sum_{k=j-J+1}^{j} P_k(u_\nu - u). \tag{3.4}$$

As in (3.1) we obtain for all $k \in \{j-J+1, \ldots, j\}$ and $\eta = P_k(u_\nu - u)$

$$\begin{aligned} D\phi\big(u_{n+k/J}; \sum_{m=j-J+1}^{k} P_m(u_\nu - u)\big) \quad &\leq \quad D\phi\big(u_{n+k/J}; \sum_{m=j-J+1}^{k-1} P_m(u_\nu - u)\big) \\ &\quad + \psi(u_{n+k/J} - P_k(u_\nu - u)) - \psi(u_{n+k/J}). \end{aligned}$$

By using these identities successively for $k = j, j-1, \ldots, j-J+1$, we infer from (3.4)

$$\begin{aligned} D\phi(u_\nu; u_\nu - u) &\leq \sum_{k=j-J+1}^{j} \big(\psi(u_{n+k/J} - P_k(u_\nu - u)) - \psi(u_{n+k/J})\big) \\ &+ \sum_{k=j-J+1}^{j} \big(D\phi(u_{n+k/J};\cdot) - D\phi(u_{n+(k-1)/J};\cdot)\big)\big(\sum_{m=j-J+1}^{k-1} P_m(u_\nu - u)\big). \end{aligned}$$

According to (2.3) and (2.6), this leads to

$$\begin{aligned} D\phi(u_\nu; u_\mu - u) \quad &\leq \quad C_X \cdot L \cdot \| u - u_\nu \|_X \cdot \sum_{k=j-J+1}^{j} \| u_{n+k/J} - u_{n+(k-1)/J} \|_X \\ &\quad + \sum_{k=j-J+1}^{j} \big(\psi(u_{n+k/J} - P_k(u_\nu - u)) - \psi(u_{n+k/J})\big). \end{aligned} \tag{3.5}$$

The last sum on the right-hand side of (3.5) will be expressed by utilizing (2.7) and (2.8). By construction of Algorithm 2.1, we have

$$u_\nu - u_{n+k/J} \in \sum_{m=k+1}^{j} X_m \qquad (k \in \{j-J+1, \ldots, j\}) \tag{3.6}$$

which, according to (2.8), yields

$$\psi(P_k(u_{n+k/J} - u_\nu) + P_k u) = \psi(P_k u) \quad \text{and} \quad \psi(P_k u_{n+k/J}) = \psi(P_k u_\nu). \tag{3.7}$$

Furthermore, (2.2), (2.7), and (2.8) lead to

$$\begin{aligned} &\psi(u_{n+k/J} - P_k(u_\nu - u)) - \psi(u_{n+k/J}) \\ &= \sum_{m=j-J+1}^{j} \psi(P_m u_{n+k/J} - P_m P_k(u_\nu - u)) - \sum_{m=j-J+1}^{j} \psi(P_m u_{n+k/J}) \\ &= \psi(P_k u_{n+k/J} - P_k P_k(u_\nu - u)) - \psi(P_k u_{n+k/J}) \end{aligned}$$

and verify

$$\begin{aligned} &\psi(P_k(u_{n+k/J} - u_\nu + u)) = \psi(P_k u_{n+k/J} - P_k(u_\nu - u)) \\ &= \psi(P_k u_{n+k/J} - \sum_{\ell=j-J+1}^{j} P_\ell P_k(u_\nu - u)) \\ &= \psi\big(P_k(u_{n+k/J} - P_k(u_\nu - u)) - \sum_{\ell=k+1}^{k+J-1} P_\ell P_k(u_\nu - u)\big) \\ &= \psi\big(P_k u_{n+k/J} - P_k P_k(u_\nu - u)\big). \end{aligned}$$

Hence, we obtain

$$\psi(u_{n+k/J} - P_k(u_\nu - u)) - \psi(u_{n+k/J}) = \psi(P_k(u_{n+k/J} - u_\nu + u)) - \psi(P_k u_{n+k/J})$$

and by (3.7)

$$\psi(u_{n+k/J} - P_k(u_\nu - u)) - \psi(u_{n+k/J}) = \psi(P_k u) - \psi(P_k u_\nu).$$

Combining this with (3.2) and (3.5), we conclude

$$\begin{aligned} D\phi(u_\nu; u_\nu - u) \quad &\leq \quad \psi(u) - \psi(u_\nu) + C_X \cdot \alpha^{-1/2} \cdot L \cdot \| u - u_\nu \|_X \\ &\quad \cdot \sum_{k=j-J+1}^{j} \left(f(u_{n+(k-1)/J}) - f(u_{n+k/J})\right)^{1/2}. \end{aligned}$$

In the following we write $\delta_\mu := f(u_\mu) - f(u)$ and $\epsilon_\mu := \| u - u_\mu \|_X$ so that (3.3) yields

$$\begin{aligned} \alpha \cdot \epsilon_\nu^2 + \delta_\nu \quad &\leq \quad C_X \cdot L \cdot \alpha^{-1/2} \cdot \epsilon_\nu \cdot \sum_{\mu=\nu-1+1/J}^{j} \left(\delta_{\mu-1/J} - \delta_\mu\right)^{1/2} \\ &\leq \quad \gamma \cdot (\delta_{\nu-1} - \delta_\nu) + 2^{-1} \cdot \alpha \cdot \epsilon_\nu^2 \end{aligned}$$

which implies

$$\frac{\alpha}{2(1+\gamma)} \cdot \epsilon_\nu^2 + \delta_\nu \leq q \cdot \delta_{\nu-1}. \tag{3.8}$$

First, (3.8) verifies $\delta_\nu \leq \delta_0 \cdot q^{[\nu]}$ and, secondly, the assertion (2.9). The proof is finished. □

Remark 3.1. At first glance, it seems that is was not used that u is a minimizer. Indeed, in the proof, we merely use that $u \in X$ satisfies $f(u) \leq f(u_\nu)$ for all $\nu \in I$. From (3.2) and from the convergence of $f(u_\nu)$ (because the sequence is bounded and monotone decreasing) we infer that u_ν is a Cauchy sequence and hence convergent towards some u^*. Is u^* a minimizer of f? At this stage, the standard technique [L] requires some continuity of f to guarantee that u^* is a minimizer of f; but here, this continuity is missing. As a modification, we fix u arbitrarily such that δ_μ remains non-negative: every u with $f(u) \leq \liminf_{\nu\to\infty} f(u_\nu)$ will be sufficient. At the end, the above estimates show $\lim_{\nu\to\infty} u_\nu = u$ which, then, proves that $u = u^*$ is indeed the unique minimizer of f.

4. Elastoplasticity

For convenient reading, we outline in this section the strong formulation of small strain elastoplasticity with kinematic or isotropic hardening as used in engineering literature and recast the problem to a dual form which fits into the frame of our main result in §2.

The body under consideration occupies a bounded Lipschitz domain Ω in $\mathbb{R}^n$. Local equilibrium for the stress field $\sigma \in L^2(\Omega; \mathbb{R}^{n\times n})$, requires

$$\sigma = \sigma^T \qquad \text{in} \quad \Omega, \tag{4.1}$$

$$\operatorname{div} \sigma + b = 0 \qquad \text{in} \quad \Omega, \tag{4.2}$$

where b is the vector field of given body forces and we understand (4.2) in the sense of distributions.

With the displacement field $u \in H_0^1(\Omega)^n$ ($H_0^1(\Omega)$ being the Sobolev space of functions in $L^2(\Omega)$ with a weak derivative in $L^2(\Omega)^n$ and vanishing traces on the boundary, i.e., $u|_{\partial\Omega} = 0$) we associate the linear Green strain tensor

$$\epsilon(u) := \left(\nabla u + (\nabla u)^T\right)/2 \qquad \text{a.e. in} \quad \Omega, \tag{4.3}$$

that is the symmetric part of the gradient of u. In the context of small strain elastoplasticity, the strain $\epsilon(u)$ is split additively into two contributions

$$\epsilon(u) = \mathbb{A}\sigma + p \qquad \text{a.e. in} \quad \Omega, \tag{4.4}$$

where the linear mapping $\mathbb{A} : \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n}$ describes linear elasticity and is symmetric and positive definite. The elasticity operator is $\mathbb{C} := \mathbb{A}^{-1} : \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n}$ and $\mathbb{A}\sigma$ is the elastic and p is the plastic part of the total strain $\epsilon(u)$. Note that the elastic material behaviour is characterized by $p = 0$ and, here, we need another material law to determine p. Moreover, there are restrictions on the stress variables prescribed by a dissipation-functional φ which is convex and non-negative but may be $+\infty$. The first restriction is that

$$\varphi(\sigma,\alpha) < \infty \qquad \text{a.e. in} \quad \Omega. \tag{4.5}$$

In this way, the hardening parameter α steers the set of admissible stresses; sometimes, the couple (σ,α) is called *generalized stresses* and they are called *admissible* if $\varphi(\sigma,\alpha) < \infty$.

The evolution of p and α is given by the so-called Prandtl-Reuß normality law which states that for all other generalized stresses (τ,β) there holds

$$\dot{p} : (\tau - \sigma) - \dot{\alpha} : (\beta - \alpha) \leq \varphi(\tau,\beta) - \varphi(\sigma,\alpha) \quad \text{a.e. in } \Omega. \tag{4.6}$$

Here, the dot denotes time derivatives, e.g., $\dot{p} := \frac{\partial p}{\partial t}$, and $:$ is a scalar product of matrices such that $A : B = \sum_{i,j=1}^{n} A_{ij} \cdot B_{ij}$ if $A, B \in \mathbb{R}^{n\times n}$. In the same way, we define the Euclidean length $|\cdot|$ by $|A| := \sqrt{A : A}$ if A is a matrix.

Remark 4.1. The minus-sign in front of $\dot{\alpha} : (\beta - \alpha)$ is not a misprint. It corresponds to the plus sign in front of $\dot{p} : (\tau - \sigma) = \dot{\epsilon}(u) : (\tau - \sigma) - A^{-1}\dot{\sigma} : (\tau - \sigma)$ because, here, $\dot{\sigma}$ has a minus sign as well.

According to the Prandtl-Reuß normality law (4.6) the problem is time-dependent and so we are seeking variables in a time interval $[0,T]$. Given a Hilbert space H, let $W^{1,p}(0,T;H)$ be the space of functions $f : [0,T] \to H$ that are absolutely continuous such that $\dot{f}$ exists and satisfies $\|\dot{f}\| \in L^p(0,T;\mathbb{R})$. Given an exterior load $b \in W^{2,2}(0,T;L^2(\Omega)^n)$ (i.e., b and $\dot{b}$ belong to $W^{1,2}(0,T;L^2(\Omega)^n)$ with $b(0) = 0$, the time-dependent problem reads as follows.

Definition 4.1 (Problem (P)). *Find* $(u,p,\sigma,\alpha) \equiv (u,(p,\sigma,\alpha)) \in W^{1,2}\Big(0,T;H_0^1(\Omega)^n \times L^2(\Omega;\mathbb{R}^{n\times n}_{sym} \times \mathbb{R}^{n\times n}_{sym} \times \mathbb{R}^m)\Big)$ *with* $(u,p,\sigma,\alpha)(0) = 0$ *and satisfying, almost everywhere in time,* (4.1)—(4.6).

In the examples for a dissipation functional φ there is a given yield function Φ such that the admissible generalized stresses are characterized by

$$\Phi(\sigma,\alpha) \leq 0 \qquad \text{in} \quad \Omega. \tag{4.7}$$

(To guarantee that the set of admissible generalized stresses is convex, we require that Φ is convex.) Then, φ is the characteristic functional of the set of admissible generalized stresses, i.e.,

$$\varphi(\sigma,\alpha) := \begin{cases} 0 & \text{if } \Phi(\sigma,\alpha) \leq 0, \\ \infty & \text{if } \Phi(\sigma,\alpha) > 0. \end{cases} \tag{4.8}$$

In this case, (4.5) reads $\Phi(\sigma,\alpha) \leq 0$ and for all (τ,β) with $\Phi(\tau,\beta) \leq 0$ there holds

$$\dot{p} : (\tau - \sigma) - \dot{\alpha} : (\beta - \alpha) \leq 0 \qquad \text{a.e. in} \quad \Omega. \tag{4.9}$$

Example 4.1 (Isotropic hardening). *Let $m = 1$, i.e., α is a scalar, and define*

$$\Phi(\sigma,\alpha) := |dev\,\sigma| - \sigma_y \cdot (1 + H \cdot \alpha) \tag{4.10}$$

in case $\alpha \geq 0$ (and $\Phi(\sigma,\alpha) = \infty$ if $\alpha < 0$ which, thereby, is not allowed). Here,

$$dev\,A := A - (tr\,A) \cdot I_n \qquad (A \in \mathbb{R}^{n\times n}),$$

where $tr\,A := \sum_{i=1}^{n} A_{ii}$ and I_n is the $n \times n$-unit matrix. The material constant $\sigma_y > 0$ is the yield stress and the constant $H > 0$ is the modulus of hardening. Then, there exists a unique solution of Problem (P) provided the exterior load b is slightly more regular (and then there holds Johnson's so-called safe-load assumption) [C1, HR, J1, J3].

Example 4.2 (Kinematic hardening). *Let $m = n(n+1)/2$ and identify $\mathbb{R}^m \equiv \mathbb{R}^{n\times n}_{sym} := \{A \in \mathbb{R}^{n\times n} : A = A^T\}$. As the stress σ we consider α (pointwise) in $\mathbb{R}^{n\times n}_{sym}$ and define*

$$\Phi(\sigma,\alpha) := |dev\,\sigma - dev\,\alpha| - \sigma_y. \tag{4.11}$$

Then, there exists a unique solution of Problem (P) provided the exterior load b is slightly more regular (and then there holds Johnson's so-called safe-load assumption) [C1, HR, J1, J3].

Example 4.3 (Perfect plasticity). *The case $m = 0$ of no hardening, i.e., the internal variables are absent, the von-Mises yield condition is given by*

$$\Phi(\sigma) := |dev\,\sigma| - \sigma_y. \tag{4.12}$$

Although the resulting problem is covered in this section, the missing hardening results in a different functional analytical frame. There exist solutions of Problem (P) in a much weaker sense (spaces of bounded deformation) if the load is limited (here, the safe-load assumption is a very restrictive assumption) [C1, S, T]. *This excludes the so-called perfect plasticity from our study of domain decomposition.*

Example 4.4 (Viscoplasticity). *In Examples 4.1, 4.2, and 4.3 the dissipation functional* (4.7) *is non-smooth and can be approximated by a smoother functional. The Yosida-regularization leads to viscoplastic material description in the sense of Perzyna where, given a viscosity $\mu > 0$, for all preceding examples of Φ we define*

$$\varphi(\sigma,\alpha) := \frac{1}{\mu}\inf\big\{|(\sigma-\tau,\alpha-\beta)| : (\tau,\beta)\in\mathbb{R}^{n\times n}_{sym}\times\mathbb{R}^m \text{ with } \Phi(\tau,\beta)\le 0\big\}. \tag{4.13}$$

For $\mu > 0$ there exists a unique solution of Problem (P) [J1, S]. *The dissipation functional* (4.11) *is, in some sense, converging towards* (4.7) *as $\mu \to 0$* [S].

The time-dependent problem is solved with an implicit time discretization such as generalized mid-point rules like Crank-Nicolson or backward Euler schemes. The schemes converge under some regularity assumptions on the solutions [C2, HR, J3] so we focus on the numerical analysis of one typical time step written in the standard weak form.

Definition 4.2 (Problem $(P^*_{\Delta t})$). *Let $H \subseteq H^1_0(\Omega)^n$, $L^{n\times n}_{sym} \subseteq L^2(\Omega;\mathbb{R}^{n\times n}_{sym})$, $L^m \subseteq L^2(\Omega;\mathbb{R}^m)$ be closed subspaces. Given $u_0 \in H$, $p_0, \sigma_0 \in L^{n\times n}_{sym}$, and, $\alpha_0 \in L^m$ at some time t_0 one step of a backward Euler scheme consists of seeking $(u, p, \sigma, \alpha) \in H \times L^{n\times n}_{sym} \times L^{n\times n}_{sym} \times L^m)$, which approximate the variables at time $t_1 = t_0 + \Delta t$, and satisfies*

$$\int_\Omega \sigma : \epsilon(v)\,dx = \int_\Omega v\cdot b\,dx \tag{4.14}$$

for all $v \in H$, where $b := b(t_1) \in L^2(\Omega)^n$, and, for all $(\tau,\beta) \in L^{n\times n}_{sym} \times L^m$ there holds

$$\int_\Omega \big\{(p-p_0):(\tau-\sigma) - (\alpha-\alpha_0):(\beta-\alpha)\big\}\,dx \le \Delta t\int_\Omega \varphi(\tau,\beta)\,dx - \Delta t\int_\Omega \varphi(\sigma,\alpha)\,dx \tag{4.15}$$

where it is understood that $\int_\Omega \varphi(\sigma,\alpha)\,dx < \infty$ and we set

$$p = \epsilon(u) - \mathbb{C}^{-1}\sigma. \tag{4.16}$$

Remark 4.2. For the backward Euler scheme, the derivatives $(\dot p, \dot\alpha)$ are approximated by differences $(p - p_0, \alpha - \alpha_0)/\Delta t$ and the conditions are satisfied at t_1. The generalized mid-point rule leads to a problem with a different update which is essentially the same. It is known there exists a unique solution to Problem $(P^*_{\Delta t})$ [C2, HR, J1, J2, J3, J4].

Remark 4.3. Definition 4.2 includes the continuous case (i.e., no space discretization) for $H := H^1_0(\Omega)^n$, $L^{n\times n}_{sym} := L^2(\Omega;\mathbb{R}^{n\times n}_{sym})$, $L^m := L^2(\Omega;\mathbb{R}^m)$. Furthermore, a space discretization is included in Definition 4.2 as well. For example, if $\mathcal{T}$ is a regular triangulation in the sense of Ciarlet [BS], let $S_1(\mathcal{T})$ be the set of continuous functions with vanishing boundary values which are affine on each element $T \in \mathcal{T}$, and $S_0(\mathcal{T})$ the space of functions that are constant on each element $T \in \mathcal{T}$. Then, let $H := S_1(\mathcal{T})^n$, $L^m := (S_0(\mathcal{T}))^m$ etc.. Further details are given in the next section in the context of domain decomposition.

In [HR], the problem in Definition 4.2 is called the dual formulation. The primary problem is obtained utilizing notions in convex analysis as follows. The condition (4.15) reads

$$\frac{1}{\Delta t}(p - p_0, -\alpha + \alpha_0) \in \partial\varphi(\sigma,\alpha), \tag{4.17}$$

when we use the notion of a subdifferential

$$a \in \partial\varphi(b) \quad \text{if and only if for all } c \quad \int_\Omega a : (c-b)\,dx \le \int_\Omega \varphi(c)\,dx - \int_\Omega \varphi(b)\,dx. \tag{4.18}$$

Since φ is convex, (4.17) is equivalent to

$$(\sigma,\alpha)\in\partial\varphi^*\Big(\frac{1}{\Delta t}(p-p_0,\alpha_0-\alpha_0)\Big), \tag{4.19}$$

where φ^* is the dual functional (also called Fenchel-transformed) to φ, given by

$$\varphi^*(b):=\sup_c\{b:c-\varphi(c)\}. \tag{4.20}$$

We refer to standard references on convex analysis [ET, Z].

Let us postpone the computation of φ^* to the end of this section and recast $(P^*_{\Delta t})$ to the primal problem (4.14), (4.16), (4.19). If we use (4.16) to eliminate σ, we finally obtain the following equivalent problem (recall $\mathbb{A}^{-1}=\mathbb{C}$).

Definition 4.3 (Problem $(P_{\Delta t})$). *Under the same conditions as in Definition* (4.2) *we seek* $(u,p,\alpha)\in H\times L^{n\times n}_{sym}\times L^m$ *satisfying, for all* $(v,q,\beta)\in H\times L^{n\times n}_{sym}\times L^m$,

$$\textstyle\int_\Omega \mathbb{C}[\epsilon(u)-p]:\epsilon(v)\,dx=\int_\Omega v\cdot b\,dx, \tag{4.21}$$

$$\begin{aligned}&\int_\Omega\big\{\mathbb{C}[\epsilon(u)-p]:(\Delta t q-p+p_0)+\alpha:(-\Delta t\beta+\alpha_0-\alpha)\big\}\,dx\\&\le\Delta t\int_\Omega\varphi^*(q,\beta)\,dx-\Delta t\int_\Omega\varphi^*((p-p_0)/\Delta t,(\alpha_0-\alpha)/\Delta t)\,dx.\end{aligned} \tag{4.22}$$

With standard arguments one proves that Problem $(P_{\Delta t})$ is the stationary condition of a minimizer in the minimization problem (M), i.e., a minimizer of f is a solution of $(P_{\Delta t})$ and vice versa.

Definition 4.4 (Problem (M)). *Under the same conditions as in Definition 4.3 we seek a minimizer* (u,p,α) *of*

$$f(u,p,\alpha)\;:=\;\frac12\int_\Omega\mathbb{C}[\epsilon(u)-p]:(\epsilon(u)-p)\,dx+\frac12\int_\Omega|\alpha|^2\,dx \tag{4.23}$$

$$+\Delta t\int_\Omega\varphi^*((p-p_0)/\Delta t,(\alpha_0-\alpha)/\Delta t)\,dx-\int_\Omega b\cdot u\,dx \tag{4.24}$$

in $H\times L^{n\times n}_{sym}\times L^m$.

Altogether, we indicated in this section that a time-discretization of the elastoplastic evolution leads typically to the Problem (M) which is of the form considered in §2. Further details on possible subspaces are discussed in the next section.

A computation of the duals of the dissipation functionals as described above concludes this section.

Example 4.5 (Isotropic hardening). *Let* φ *be defined by* (4.8) *with* Φ *given in* (4.10), *we have, for all* $A\in\mathbb{R}^{n\times n}_{sym}$ *and* $B\in\mathbb{R}$,

$$\varphi^*(A,B)=\sup_{|dev\,\sigma|\le\sigma_y(1+H\alpha)}(\sigma:A+\alpha B)$$

where $|dev\,\sigma|\le\sigma_y(1+H\alpha)$ *indicates that in the supremum* (σ,α) *is an arbitrary element satisfying* $\sigma\in\mathbb{R}^{n\times n}_{sym}$, $\alpha\ge0$ *and* $|dev\,\sigma|\le\sigma_y(1+H\alpha)$. *Assume* $\varphi^*(A,B)<\infty$. *Since* $(\sigma,\alpha)=s\cdot(I_{n\times n},0)$ *satisfies those conditions for any* $s\in\mathbb{R}$, $\varphi^*(A,B)<\infty$ *implies* $tr\,A=0$. *Furthermore, for all* $\alpha\ge0$, $\Phi(\lambda A,\alpha)\le0$ *if* $\lambda:=\sigma_y(1+H\alpha)/|A|$. *Hence,*

$$\lambda|A|^2+\alpha B=|A|\sigma_y+\alpha(B+H\sigma_y|A|)\le\varphi^*(A,B)<\infty$$

for all $q\ge0$. *Thus, we obtain* $B+H\sigma_y|A|\le0$. *According to Cauchy's inequality, recall* $A=dev\,A$ *by* $tr\,A=0$,

$$dev\,\sigma:dev\,A+qB\le\sigma_y(1+H\alpha)|dev A|+\alpha B=\sigma_y|A|+\alpha(\sigma_yH|A|+B),$$

for all admissible (σ,α). *The above calculations prove*

$$\varphi^*(A,B)=\begin{cases}\sigma_y|A| & \textit{if } tr\,A=0\wedge B+H\sigma_y|A|\le0,\\ \infty & \textit{if } tr\,A\ne0\vee B+H\sigma_y|A|>0.\end{cases} \tag{4.25}$$

Given (u,p) in Problem (M), the minimization with respect to α concerns only the term $\frac{1}{2}\int_\Omega |\alpha|^2\,dx$ and the side restriction $B + H\sigma_y|A| \le 0$ where $A = (p-p_0)/\Delta t$ and $B = (\alpha_0 - \alpha)/\Delta t$. The side restriction reads $\alpha \ge \alpha_0 + H\sigma_y|p-p_0| \ge 0$ (because $\alpha_0, \alpha \ge 0$) and so the minimization with respect to α has the unique solution $\alpha = \alpha_0 + H\sigma_y|p-p_0|$. Putting this into the functional f, we get a simpler minimization problem where we seek a minimizer (u,p) of

$$
\begin{aligned}
f(u,p) \quad := \quad & \frac{1}{2}\int_\Omega \mathbb{C}[\epsilon(u)-p] : (\epsilon(u)-p)\,dx + \frac{1}{2}\int_\Omega (\alpha_0 + H\sigma_y|p-p_0|)^2\,dx \\
& + \int_\Omega \sigma_y|p-p_0|\,dx - \int_\Omega b\cdot u\,dx
\end{aligned}
\tag{4.26}
$$

in $H \times L^{n\times n}_{sym}$ under the side restriction $tr\,(p-p_0) = 0$.

Remark 4.4. Note that the side restriction reduces the $n(n+1)/2$ unknown variables of p (pointwise in the continuous case or on each element in the space-discrete situation) to $n(n+1)/2-1$ and is easily implemented in numerical calculations.

Example 4.6 (Kinematic hardening). *Let φ be defined by (4.8) with Φ given in (4.11), we have, for all $A, B \in \mathbb{R}^{n\times n}_{sym}$*

$$\varphi^*(A,B) = \sup_{|dev\,\sigma - dev\,\alpha| \le \sigma_y} (\sigma : A + \alpha : B),$$

where $\sigma, \alpha \in \mathbb{R}^{n\times n}_{sym}$ are arbitrary under the given restriction. Arguing as in Example 4.5 assume $\varphi^*(A,B) < \infty$ and infer $tr\,A = 0 = tr\,B$. Since (σ,σ) is admissible for any $\sigma \in \mathbb{R}^{n\times n}_{sym}$, we further conclude $B = -A$. Then, again with $A = dev\,A$,

$$\varphi^*(A,B) = \sup_{|dev\,\sigma - dev\,\alpha| \le \sigma_y} dev\,(\sigma - \alpha) : A = \sup_{|dev\,\sigma| \le \sigma_y} dev\,\sigma : A = \sigma_y|A|,$$

which finally shows

$$
\varphi^*(A,B) = \begin{cases} \sigma_y|A| & \text{if } tr\,A = 0 \wedge B = -A, \\ \infty & \text{if } tr\,A \ne 0 \vee B \ne -A. \end{cases}
\tag{4.27}
$$

The restriction $B = -A$ for $\varphi^*(A,B) < \infty$ allows us to reduce the set of variables because $\alpha = p - p_0 - \alpha_0$. Indeed, in Problem (M) we seek a minimizer (u,p) of

$$
\begin{aligned}
f(u,p) \quad := \quad & \frac{1}{2}\int_\Omega \mathbb{C}[\epsilon(u)-p] : (\epsilon(u)-p)\,dx + \frac{1}{2}\int_\Omega |p-p_0-\alpha_0|^2\,dx \\
& + \int_\Omega \sigma_y|p-p_0|\,dx - \int_\Omega b\cdot u\,dx
\end{aligned}
\tag{4.28}
$$

in $H \times L^{n\times n}_{sym}$ under the side restriction $tr\,(p-p_0) = 0$.

Example 4.7 (Perfect plasticity). *If φ is defined by (4.8) with Φ given in (4.12), then a simple calculation shows, for any $A \in \mathbb{R}^{n\times n}_{sym}$,*

$$
\varphi^*(A) = \begin{cases} \sigma_y|dev\,A| & \text{if } tr\,A = 0, \\ \infty & \text{if } tr\,A \ne 0. \end{cases}
\tag{4.29}
$$

5. Domain decomposition in plasticity

In this section, we consider Algorithm 2.1 in the minimization problem (M) of elastoplasticity described in §4. In the case of isotropic and kinematic hardening, we seek the minimizer (u,p) of f in $H \times L$ where $H \subseteq H^1_0(\Omega)$ and $L \subseteq \{r \in L^2(\Omega; \mathbb{R}^{n\times n}_{sym}) : \mathrm{tr}\,r = 0\}$ are closed subspaces (we shifted $p := p - p_0$ so that L is a vector space). The function f has the form (2.4), i.e., $f = \phi + \psi$, with

$$\phi(u,p) \quad := \quad \frac{1}{2}\int_\Omega \mathbb{C}[\epsilon(u)-p] : (\epsilon(u)-p)\,dx + \frac{1}{2}\int_\Omega k^2|p|^2\,dx + \ell(u,p) \tag{5.1}$$

$$\psi(u,p) \quad := \quad \int_\Omega \kappa\cdot\sigma_y|p|\,dx \tag{5.2}$$

where ℓ is a linear functional (resulting from f as from shifting the variable p) and k and κ are positive constants ($k = 1$ or $k = H\sigma_y$ and $\kappa = 1$ or $\kappa = 1 + 2H\alpha_0$).

It is known that the continuous quadratic functional $\phi(u,p) - \ell(u,p)$ is $H \times L$-elliptic and so the conditions (2.5) and (2.6) are satisfied.

Since the non-differentiable part ψ only affects internal variables in L, Theorem 2.1 guarantees that Algorithm 2.1 yields a linear convergent procedure in each of the examples below.

In our model situation, we consider a regular triangulation $\mathcal{T}$ in the sense of Ciarlet (see, e.g., [BS] for a definition) and consider

$$\mathcal{V} := \{r \in L : \forall T \in \mathcal{T}\; r|_T \in \mathcal{P}_0(T; \mathbb{R}^{n\times n}_{sym})\} \tag{5.3}$$

$$\mathcal{W} := \{u \in H : \forall T \in \mathcal{T}\; u|_T \in \mathcal{P}_1(T; \mathbb{R}^{n})\} \tag{5.4}$$

where $\mathcal{P}_k(T;\mathbb{R}^m)$ is the space of polynomial with total degree $\leq k$, $k = 0, 1$, on T in each of the m components. This defines $X = \mathcal{V} \times \mathcal{W}$.

5.1. **Alternating u-p-minimization.** Let $J = 2$, $X_1 = \mathcal{V} \times \{0\}$ and $X_2 = \{0\} \times \mathcal{W}$. Then, in any step of Algorithm 2.1 we solve one linear equation and one nonlinear minimization problem,

$$u_{\nu+1} := \operatorname{argmin}\ \phi(\cdot, p_\nu) \tag{5.5}$$

$$p_{\nu+1} := \operatorname{argmin}\ \phi(u_\nu, \cdot) + \psi, \tag{5.6}$$

where argmin g is the minimizer of a minimization problem with the function g. The problem (5.6) is local, i.e., it can be solved for each point x in Ω or for any element $T \in \mathcal{T}$ separately. Moreover, this minimization problem may be solved analytically and hence we assume that it is solved exactly.

The conditions on X_1, X_2 are obviously satisfied with $C_X = 1$, $J = 2$, and (2.5) and (2.6) are trivially true. According to Theorem 2.1 the alternating scheme (5.5)-(5.6) is linearly convergent with a convergence rate q independent of $\mathcal{T}$.

5.2. **Multiplicative additive Schwarz method for the displacements.** As a typical example in domain decomposition [L, X], we assume that the finite element space $\mathcal{V}$ is decomposed into subspaces

$$\mathcal{V}_i := \{v \in \mathcal{V} : v = 0 \text{ on } \Omega \setminus \Omega_i\} \quad (i = 1, \dots, I) \tag{5.7}$$

as follows. We assume that we are given a partition $\Omega_1^0, \dots, \Omega_I^0$ of Ω which are quasi-uniform of size h_0. Then, the subdomain Ω_i is defined to be a mesh subdomain containing Ω_i^0 with a distance from $\partial\Omega \cap \Omega$ to Ω_i^0 greater than or equal to ηh_0 with some prescribed fixed constant $\eta > 0$. Finally, let $\mathcal{V}_0 \subset \mathcal{V}$ be a coarse finite element space with respect to quasi-uniform triangulation of size h_0.

Then, there exist bounded linear mappings $Q_i : \mathcal{V} \to \mathcal{V}_i$, $i = 0, 1, \dots, I$, such that

$$\sum_{i=0}^{I} \|Q_i v\|^2_{H_0^1(\Omega)^n} \leq C_X \|v\|^2_{H_0^1(\Omega)^n}. \tag{5.8}$$

(See, e.g., [X, Lemma 7.1] for a proof and the explicit construction of Q_i.) Then, $X_i := \mathcal{V}_{i-1} \times \{0\}$ for $i = 1, \dots, I+1$, $X_{I+2} := \{0\} \times \mathcal{W}$, $J := I + 2$ satisfy (2.1)-(2.3) while (2.5) and (2.6) are trivial. According to Theorem 2.1 the Algorithm 2.1 is linearly convergent with a convergence rate q independent of $\mathcal{T}$.

5.3. **General situation.** We consider a set of overlapping subdomains $\Omega_1, \dots, \Omega_I$ and a partition $\omega_1, \dots, \omega_I$ of Ω,

$$\Omega = \Omega_1 \cup \dots \cup \Omega_I = \omega_1 \cup \dots \cup \omega_I, \tag{5.9}$$

whose boundaries align the mesh $\mathcal{T}$ which describes $\mathcal{V}$ and $\mathcal{W}$. Define $\mathcal{V}_i$ as in (5.7) and consider $\mathcal{V}_0$ empty (i.e., no coarse grid correction) or as in Subsection 5.2 (with coarse grid correction). Assume (5.8) (e.g., as in the situation of Subsection 5.2) and define

$$\mathcal{W}_i := \{w \in \mathcal{W} : w = 0 \text{ on } \Omega \setminus \omega_i\} \quad (i = 1, \dots, I). \tag{5.10}$$

To be more flexible, we allow $\Omega_i = \emptyset$ (resp. $\omega_i = \emptyset$) corresponding to $\mathcal{V}_i = \{0\}$ (resp. $\mathcal{W}_i = \{0\}$) for some $i \in \{0, \dots, I\}$.

Then, define $X_1 := \mathcal{V}_0 \times \{0\}$, $\chi_0 := 0$, $X_{i+1} = \mathcal{V}_i \times \mathcal{W}_i$ and $\chi_i(x) = 1$ for $x \in \omega_i$ and $\chi_i(x) = 0$ otherwise for $i = 1, \ldots, I$, $J := I + 1$, and

$$P_j : X \to X_j, \binom{u}{p} \mapsto \binom{Q_{j-1}u}{\chi_{j-1} \cdot p} \qquad (j \in \{1, \ldots, J\}). \tag{5.11}$$

It is important to notice that, necessarily, $\Omega_1, \ldots, \Omega_I$ is overlapping and $\omega_1, \ldots, \omega_I$ is a non-overlapping decomposition of Ω. Then, (2.1)–(2.3) are satisfied according to (5.8) as (2.7)-(2.8).

According to Theorem 2.1 the Algorithm 2.1 is linearly convergent with a convergence rate q independent of $\mathcal{T}$.

The natural choice $\omega_i = \Omega_i^0$ (Ω_i^0 as in Subsection 5.2) has the disadvantage that a coupled nonlinear minimization problem has to be solved numerically and the only advantage of reducing the dimension of the original problem.

5.4. **Remarks.** The linear convergent procedures in Subsection 5.1 and 5.2 avoid a minimization of a (high-dimensional) non-smooth nonlinear problem. Alternatively, regularization techniques with a posteriori error control were suggested, cf., e.g., [HR].

The analysis of this paper suggests that the appearance of the non-smooth nonlinear contribution ψ does not affect the quality of the convergence behaviour (compared to the linear elastic situation $\psi = 0$). For example, Theorem 2.1 guarantees linear convergence of Algorithm 2.1 also for regularized problems proposed in the literature which approximate ψ by some smooth functional [HR] with the same convergence properties.

The above notation is linked to the finite element discretization because of its practical importance. However, we stress that the continuous case is included as well.

References

[BS] S.C. Brenner, L.R. Scott: *The mathematical theory of finite element methods.* Texts in Applied Mathematics **15** Springer New York 1994.

[C1] C. Carstensen: Interface problems in viscoplasticity and plasticity *SIAM J. Math. Anal.* **25** (1994) 1468—1487.

[C2] ———: Coupling of FEM and BEM for interface problems in viscoplasticity and plasticity with hardening. *SIAM J. Numer. Anal.* **33** (1996) 171—207.

[ET] I. Ekeland, R. Temam: *Convex analysis and variational problems.* North-Holland, Amsterdam 1976.

[GLT] R. Glowinski, J.L. Lions, R. Trémolières: *Numerical analysis of variational inequalities.* North-Holland, Amsterdam 1981.

[HR] W. Han, B.D. Reddy: Computational plasticity: the variational basis and numerical analysis. *Computational Mechanics Advances* **2** (1995) 285—400.

[J1] C. Johnson: On Finite Element Methods for Plasticity Problems. *Numer. Math.* **26** (1976) 79—84.

[J2] ———: Existence Theorems for Plasticity Problems. *J. Math. pures et appl.* **55** (1976) 431—444.

[J3] ———: A mixed finite element method for plasticity problems with hardening. *SIAM J. Numer. Anal.* **14** (1977) 575—583.

[J4] ———: On Plasticity with Hardening. *Journ. of Math. Anal. and Appl.* **62** (1978) 325—336.

[L] P.L. Lions: On the Schwarz alternating method. In *Proceedings of the first international symposium of domain decomposition methods for partial differential equations.* R. Glowinski, G. Golub, G. Meurant, and J. Periaux, eds., 1987. pp. 1–42.

[S] P.-M. Suquet: Dicontinuities and Plasticity. In *Nonsmooth mechanics and applications,* CISM Courses **302**, Eds. J.J. Moreau, P.D. Panagiotopoulos eds., Springer-Verlag New York, 1988, 279–341.

[T] R. Temam: *Mathematical problems in plasticity.* Gauthier-Villars, Paris, 1985.

[X] J. Xu: Iterative methods by space decomposition and subspace correction. *SIAM Review* **34** (1992) 581–613.

[Y] H. Yserentant: Old and new convergence proofs for multigrid methods. *Acta Numerica* (1993) 285–326.

[Z] E. Zeidler: *Nonlinear functional analysis and its applications.* Springer-Verlag, New York, II Vol. A and B. 1990; III 1985; IV 1988.

Mathematisches Seminar, Christian-Albrechts-Universität zu Kiel, Ludewig-Meyn-Str. 4, D-24098 Kiel, FRG.
E-mail address: `cc@numerik.uni-kiel.de`