SPLITTING ALGORITHMS FOR THE SUM OF TWO NONLINEAR OPERATORS*

P. L. LIONS† AND B. MERCIER‡

Abstract. Splitting algorithms for the sum of two monotone operators.

We study two splitting algorithms for (stationary and evolution) problems involving the sum of two monotone operators. These algorithms are well known in the linear case and are here extended to the case of multivalued monotone operators. We prove the convergence of these algorithms, we give some applications to the obstacle problem and to minimization problems; and finally we present numerical computations comparing these algorithms to some other classical methods.

We consider the nonlinear multivalued evolution equation

$$\frac{\partial u}{\partial t} + C(u) \ni 0,$$

(1)

$$u(0) = u^0,$$

and the associated stationary equation

(2)
$$C(\bar{u}) \ni 0,$$

where C is a monotone operator on a Hilbert space H, which is *multivalued* that is, for $u \in H$, $C(u)$ is a (possibly empty) subset of H.

We shall consider the case where $C = A + B$ and A and B are maximal monotone. We study the following algorithms:

(3)
$$u^{n+1} = (I + \lambda B)^{-1}(I - \lambda A)(I + \lambda A)^{-1}(I - \lambda B)u^n,$$

which was introduced, in the case of linear operators, by Peaceman–Rachford [16], [18], and

(4)
$$u^{n+1} = (I + \lambda B)^{-1}[(I + \lambda A)^{-1}(I - \lambda B) + \lambda B]u^n,$$

which was introduced by Douglas–Rachford [8].

In this paper, we show that these algorithms can be used to solve the evolution equation (1) by letting $\lambda \to 0$, and λ appears then as a time step. We show also that they have the interesting property that, *for fixed λ, $u^n \to u_\infty$ as $n \to \infty$*, where u_∞ is a solution of the stationary problem (2).

Last, but not least, they are both *unconditionally stable*, i.e. u^n remains bounded independently of n for any λ.

This set of properties is remarkable if we compare them to that we get for more standard algorithms.

The first one is

(5)
$$u^{n+1} = (I + \lambda A)^{-1}(I - \lambda B)u^n,$$

which is *not* unconditionally stable, but converges to the solution of the stationary problem for λ sufficiently small if B is Lipschitz continuous (see Goldstein [12], Bruck [6]).

The second one is

(6) $$u^{n+1} = (I + \lambda A)^{-1}(I + \lambda B)^{-1} u^n,$$

which is unconditionally stable but does not converge to the solution of the stationary problem for any λ, except with some special modification (see Lions [14]).

All these are called "splitting" algorithms since, up to the introduction of a fractionary step, they can be interpreted as the combination of a step for A and a step for B.

As an example, (5) can be written

$$\frac{u^{n+1/2} - u^n}{\lambda} + Bu^n = 0,$$

$$\frac{u^{n+1} - u^{n+1/2}}{\lambda} + Au^{n+1} = 0,$$

which shows that (5) results from the combination of a forward step on B and a backward step on A.

An outline of this paper is as follows: in § 1, we study the convergence of (3), (4), in the stationary case; in § 2, we study consistency of these algorithms and the convergence to the solution of the evolution equation (as $\lambda \to 0$); in § 3, we give some numerical results. Finally, § 4 contains some comments.

1. The stationary case.
1.1. Assumptions and notations
We shall assume that A and B are maximal monotone. We denote by $D(A)$ the domain of A. We recall that A monotone means

$$(y - z, u - v) \geqq 0, \quad \forall u \in H, y \in A(u), v \in H, z \in A(v),$$

where $(\cdot, \cdot)$ denotes the scalar product of H and $|\cdot|$ the associated norm; a maximal monotone is equivalent to saying the resolvent $J_A^\lambda = (I + \lambda A)^{-1}$ is a contraction defined on H. When A is single valued, we have the identity

(7) $$(I + \lambda A) J_A^\lambda = I.$$

In this section, we shall assume that the stationary problem (2) has at least one solution. Hence

(8) *there exists $u \in H$, $a \in A(u)$, $b \in B(u)$ such that $a + b = 0$.*

In the case here considered, where A and B are multivalued, we need make precise the definition of algorithms (3) and (4). For both, $u^0 \in D(B)$ is given, and we *choose* $b^0 \in B(u^0)$ and set $v^0 = u^0 + \lambda b^0$ in such a way that $u^0 = J_B^\lambda v^0$. We then define by induction the sequence v^n in the following way.

ALGORITHM I.

(9) $$v^{n+1} = (2J_A^\lambda - I)(2J_B^\lambda - I)v^n;$$

ALGORITHM II.

(10) $$v^{n+1} = J_A^\lambda (2J_B^\lambda - I)v^n + (I - J_B^\lambda)v^n.$$

One can check that in the case where A and B are single valued, Algorithm I is equivalent to the Peaceman–Rachford algorithm (3) up to the change of variable $u^n = J_B^\lambda v^n$ (with the use of (7)). Likewise, Algorithm II equivalent to the Douglas–Rachford (4). But (9) and (10) have a meaning even for A and B multivalued.

Remark 1. Algorithm I has been considered by Brezis [4] and Lions [14], but with a severe restriction: B is assumed to be coercive, that is $B - \alpha I$ monotone for some $\alpha > 0$.

1.2. Convergence of Algorithm I. We shall use the following notation

$$v = u + \lambda b, \qquad w = u + \lambda a,$$

$$w^n = 2u^n - v^n, \qquad b^n = \frac{v^n - u^n}{\lambda}, \qquad a^n = \frac{w^n - v^{n+1}}{2\lambda}.$$

We prove the following result.

PROPOSITION 1. *Under the assumption* (8), *the sequences* u^n, v^n, w^n, a^n, b^n, *remain bounded. Moreover*

(11a)
$$\lim_{n \to +\infty} (b^n - b, u^n - u) = 0;$$

(11b)
$$\lim_{n \to +\infty} \left(a^n - a, \frac{v^{n+1} + w^n}{2} - u \right) = 0.$$

Proof. From the definition of b^n, we have $v^n = u^n + \lambda b^n$. As $u^n = J_B^\lambda v^n$, we have $v^n \in u^n + \lambda B u^n$, hence $b^n \in B u^n$. From the monotonicity of B, we get

(12)
$$0 \leq (b^n - b, u^n - u) = \frac{1}{4\lambda} (|v^n - v|^2 - |w^n - w|^2),$$

where we have used

$$u^n = \frac{v^n + w^n}{2}, \quad u = \frac{v + w}{2}, \quad b^n = \frac{v^n - w^n}{2\lambda}, \quad b = \frac{v - w}{2\lambda}.$$

On the other hand, from (9), we get

$$v^{n+1} = (2J_A^\lambda - I)w^n \quad \Rightarrow \quad \frac{v^{n+1} + w^n}{2} = J_A^\lambda w^n.$$

Hence

$$w^n \in \frac{v^{n+1} + w^n}{2} + \lambda A \left(\frac{v^{n+1} + w^n}{2} \right),$$

$$a^n = \frac{w^n - v^{n+1}}{2\lambda} \in A \left(\frac{v^{n+1} + w^n}{2} \right).$$

From the monotonicity of A, we get

(13)
$$0 \leq \left(a^n - a, \frac{v^{n+1} + w^n}{2} - u \right) = \left(\frac{w^n - v^{n+1}}{2\lambda} - \frac{w - v}{2\lambda}, \frac{v^{n+1} + w^n}{2} - \frac{v + w}{2} \right)$$

$$= \frac{1}{4\lambda} (|w^n - w|^2 - |v^{n+1} - v|^2),$$

which shows, with (12), that

$$|v^{n+1} - v|^2 \leq |w^n - w|^2 \leq |v^n - v|^2,$$

and the sequences v^n, w^n are bounded, thus proving the first part. Finally, as $|v^n - v|^2 - |v^{n+1} - v|^2 \to 0$ as $n \to +\infty$, (12), (13) imply (11) thereby completing the proof. □

COROLLARY 1. *If B is single valued and satisfies the following property*:
(14) *For all x_n, $x \in D(B)$ such that Bx_n is bounded, $x_n \rightharpoonup \bar{x}$ weakly, and*

$$(Bx_n - Bx, x_n - x) \to 0$$

as $n \to +\infty$, then one has $x = \bar{x}$; then u^n converges weakly to u, solution of (2), which is unique.

Proof. Let us prove first the uniqueness. Let u_1, u_2 be two solutions of (2); we have

$$0 \leqq (Bu_1 - Bu_2, u_1 - u_2) = -(Au_1 - Au_2, u_1 - u_2) \leqq 0,$$

hence $(Bu_1 - Bu_2, u_1 - u_2) = 0$ which, together with property (14) implies $u_1 = u_2$.

Let $u_{n'}$ denote a subsequence of the bounded sequence u_n converging weakly to $\bar{u}$; from (11a) and (14), one gets $u = \bar{u}$, and the whole sequence u_n converges weakly to u. $\square$

Remark 2. We give some cases where property (14) is satisfied:

1. *B is coercive*: there exists some $\alpha > 0$ such that $B - \alpha I$ is monotone. In this case, the sequence u^n converges even strongly to u.

2. *B^{-1} is coercive*: there exists $\beta > 0$ such that

$$(Bu - Bv, u - v) \geqq \beta |Bu - Bv|^2, \quad \forall u, v \in D(B);$$

in such a case, (11a) implies that $Bu^n \to Bu$ strongly. Moreover, if B^{-1} is univoque, (14) is satisfied since $Bx_n \to Bx$; as $x_n \rightharpoonup \bar{x}$, $\bar{x} \in D(B)$ and $B\bar{x} = Bx$; thus $\bar{x} = x$.

3. *B is strictly monotone and weakly closed*: that is if $u, v \in D(B)$ and $(Bu - Bv, u - v) = 0$, one has $u = v$; if $x_n \in D(B)$, with $x_n \rightharpoonup \bar{x} \in H$ weakly and such that there exists $y_n \in Bx_n$ with $y_n \rightharpoonup y \in H$, then $\bar{x} \in D(B)$ and $y = B\bar{x}$.

Let us prove that in this case again (14) is satisfied. Indeed, let x_n be as in (14) and $x_{n'}$ denote a subsequence such that $Bx_{n'} \rightharpoonup y$ weakly. From the weak closure of B, we have $\bar{x} \in D(B)$ and $y \in B\bar{x}$. The assumption on x_n implies

$$\lim_{n' \to \infty} (Bx_{n'}, x_{n'}) = \lim_{n' \to \infty} (Bx, x_{n'} - x) + (Bx_{n'}, x)$$

$$= (Bx, \bar{x} - x) + (B\bar{x}, x)$$

$$\leqq (B\bar{x}, \bar{x}),$$

where we have applied the monotonicity of B. A result of Brezis [3, Prop. 2.5] proves then that the inequality is actually an equality, which implies $(Bx - B\bar{x}, x - \bar{x}) = 0$; then $x = \bar{x}$, which proves that (14) holds.

Remark 3. If B is linear, or if J_B^λ is compact, then B is weakly closed.

Remark 4. In Corollary 1, if we assume that A instead of B satisfies (14), then the sequence $(v^{n+1} + w^n)/2$ converges weakly to the unique solution of the stationary problem (2).

Remark 5. We can prove that, if a subsequence $(v^{n'})$ of (v^n) is bounded, then the stationary problem (2) has one solution u. Indeed, let $T = (2J_A^\lambda - I)(2J_B^\lambda - I)$, we have $v^{n+1} = Tv^n$. As $2J_A^\lambda - I$ and $2J_B^\lambda - I$ are nonexpansive, T itself is nonexpansive. The boundedness of a subsequence $v^{n'}$ implies then that T has a fixed point v, with $Tv = v$.

Let $u = J_B^\lambda v$, we have $u \in D(B)$ and

$$v = (2J_A^\lambda - I)(2u - v),$$

$$u = J_A^\lambda (2u - v);$$

hence $u \in D(A)$. Let $t \in B(u)$ satisfy $v = u + \lambda t$; we have

$$u = J_A^\lambda (u - \lambda t) \quad \Rightarrow \quad u - \lambda t \in u + \lambda Au,$$

that is $-t \in A(u)$, hence u is a solution of the stationary problem (2).

Remark 6. *A counter-example.* Let $H = \mathbb{C}$, $A = B = i : Az = iz$. Note that A is maximal monotone on $\mathbb{R}^2$; moreover, A is linear compact and one to one; however, $2J_A^\lambda - I = (1 - \lambda i)/(1 + \lambda i)$ is a rotation and v^n does not converge strongly to the solution of (2) which is 0.

1.3. Convergence of Algorithm II. We shall see that the convergence for Algorithm II occurs for more general operators A and B than those considered previously in Remark 2. Indeed, we shall prove the following result.

THEOREM 1. *Under the assumption* (8), *the sequence* v^n *generated by the algorithm* (10), *converges weakly to* $v \in H$, *as* $n \to \infty$, *such that* $u = J_B^\lambda v$ *is a solution of* (2). *Furthermore, let* $u^n = J_B^\lambda v^n$,

  (i) *if B is linear, u^n converges weakly to a solution of* (2);

  (ii) *if A and B are odd, u^n converges strongly to a solution of* (3);

  (iii) *If $A + B$ is maximal monotone, any subsequence of (u^n) converges weakly to a solution of* (3).

1.3.1. Preliminary results. We need first to recall a definition and a result.

DEFINITION. Let $K \subset H$ be a closed convex subset of H. An operator $T : K \to K$ is said to be *firmly nonexpansive* if, for all $x, y \in K$,

$$(Tx - Ty, x - y) \geqq |Tx - Ty|^2.$$

(Note that firmly nonexpansive $\Rightarrow$ nonexpansive).

RESULT (Browder [5]). *If T is nonexpansive from K into K, and if T has at least one fixed point $\xi \in K$, then, for all $x \in K$, $T^{n+1}x - T^n x \to 0$ strongly, as $n \to \infty$, and $T^n x \to \bar{x} \in K$ weakly, where $\bar{x} = T\bar{x}$ is a fixed point of T.*

We shall prove the following result.

LEMMA 1. *Let T_1 and T_2 be two firmly nonexpansive operators from K into K, then $S = T_1(2T_2 - I) + I - T_2$ is firmly nonexpansive. More precisely, one has*

$$(15) \qquad (Sx - Sy, x - y) \geqq |Sx - Sy|^2 + (T_2 x - T_2 y, (I - T_2)x - (I - T_2)y).$$

Proof. We have

$$|Sx - Sy|^2 = |T_1(2T_2 - I)x - T_1(2T_2 - I)y|^2 + |(I - T_2)x - (I - T_2)y|^2$$
$$+ 2(T_1(2T_2 - I)x - T_1(2T_2 - I)y, (I - T_2)x - (I - T_2)y).$$

As T_1 is firmly nonexpansive, we get

$$|Sx - Sy|^2 \leqq (T_1(2T_2 - I)x - T_1(2T_2 - I)y, x - y) + |(I - T_2)x - (I - T_2)y|^2$$
$$= (Sx - Sy, x - y) - (T_2 x - T_2 y, (I - T_2)x - (I - T_2)y)$$

which proves (15). The positivity of the last term results from the fact that T_2 is firmly nonexpansive. $\square$

1.3.2. Convergence. We shall apply this result to

$$(16) \qquad G(\lambda) = J_A^\lambda (2J_B^\lambda - I) + I - J_B^\lambda,$$

and we note that the algorithm (10) can be written

$$v^{n+1} = G(\lambda)v^n.$$

PROPOSITION 2. *The operator $G(\lambda)$ defined in (16) is firmly nonexpansive and satisfies*

(17)
$$(G(\lambda)x - G(\lambda)y, x - y) \geq |G(\lambda)x - G(\lambda)y|^2$$
$$+ ((I - J_B^\lambda)x - (I - J_B^\lambda)y, J_B^\lambda x - J_B^\lambda y).$$

Then, under the assumption (8), we have

(18) $v^n \to v$ *weakly, where* $v = G(\lambda)v$,
        u^n *is bounded and* $u^{n+1} - u^n \to 0$.

Let $b^n = (v^n - u^n)/\lambda$; b^n *is bounded,* $b^{n+1} - b^n \to 0$, *and*

(19)
$$(b^n - b, u^n - u) \to 0, \quad as \ n \to \infty.$$

Proof. As J_B^λ is firmly nonexpansive, we may apply Lemma 1 to prove (17). From the "result" above, we have $v^{n+1} - v^n \to 0$ and $v^n \rightharpoonup v$ weakly, where v is a fixed point of $G(\lambda)$. As $u^n = J_B^\lambda v^n$, (18) results from the nonexpansiveness of J_B^λ. Finally, applying (17) with $x = v^n$ and $y = v^n$, we get

(20)
$$(v^{n+1} - v, v^n - v) \geq |v^{n+1} - v|^2 + \lambda(b^n - b, u^n - u),$$

which proves that $|v^n - v|^2$ decreases and (19). □

COROLLARY 2. *Assume that J_B^λ is weakly closed; then under the assumption (8), u^n converges weakly to a solution $u = J_B^\lambda v$ of (2).*

Proof. From Proposition 2, $v^n \rightharpoonup v$ and $u = J_B^\lambda v$ is a solution of (2). As the sequence u^n is bounded, for any subsequence $u^{n'}$ converging weakly to some $\bar{u} \in H$, we have $\bar{u} = J_B^\lambda v = u$, hence the weak convergence of the whole sequence to u. □

COROLLARY 3. *If J_A^λ is compact, then $u^n \to u$ the solution of (2), strongly.*

Proof. From (10), we have

(21)
$$u^n = v^n - v^{n+1} + J_A^\lambda(2u^n - v^n),$$

which shows that the sequence u^n remains in a compact set. But, for any convergent subsequence $u^{n'} \to u$, as J_B^λ is maximal monotone, $u = J_B^\lambda v$. □

Remark 7. If J_B^λ is compact, J_B^λ is also weakly closed and Corollary 2 applies. However, the convergence is obviously strong.

Remark 8. If A and B are odd, then v^n converges strongly to a fixed point of $G(\lambda)$, and u^n to a solution of (2). This results from a result by Baillon [1], for odd firmly nonexpansive operators, which $G\{\lambda)$ is, in this case.

The following result gives some additional information about the convergence of the sequences u^n and v^n.

PROPOSITION 3. *Under the assumption (8), there exists a sequence ε^n such that*

(22) $\varepsilon^n + a^n + b^n = 0$ *where* $b^n = \dfrac{v^n - u^n}{\lambda}$, $a^n \in A(u^n + \lambda\varepsilon^n)$ *and* $\varepsilon^n \to 0$ *strongly.*

Furthermore

(23a)
$$(b^n - b, u^n - u) \to 0,$$

(23b)
$$(a^n - (-b), u^n + \varepsilon^n - u) \to 0, \quad as \ n \to +\infty.$$

Proof. Let a^n be the element of $A(u^n + v^{n+1} - v^n)$ such that (see (21))

$$u^n + v^n - v^{n+1} + \lambda a^n = 2u^n - v^n = u^n - \lambda b^n,$$

which implies (22) with the choice $\varepsilon^n = (v^n - v^{n+1})/\lambda$. To complete the proof, we notice that

$$(\varepsilon^n + a^n + b^n, u^n - u) = 0,$$

$$(\varepsilon^n, u^n - u) + (a^n + b, u^n + \lambda\varepsilon^n - u) + (b^n - b, u^n - u) = \lambda(a^n + b, \varepsilon^n).$$

As b^n is bounded (Proposition 2), (22) shows that a^n is bounded. Hence,

$$(a^n + b, u^n + \lambda\varepsilon^n - u) + (b^n - b, u^n - u) \to 0.$$

As each term is positive by monotonicity of A and B, we get (23). □

Remark 9. If A or B satisfies property (14), the sequence u^n converges weakly to a solution u of the stationary problem (2). If A or B is coercive, the convergence is strong, and one has $u = J_B^\lambda v$. In any case, if $u^n \rightharpoonup u = J_B^\lambda v$ weakly, then $b^n = (v^n - u^n)/\lambda$ converges weakly to an element b satisfying (8).

1.3.3. Speed of convergence. We shall give an evaluation of the speed of convergence of Algorithm II, in the special case where B is both coercive and Lipschitz. Then, there exists α and $M > 0$ such that

(24a) $$|Bx_1 - Bx_2| \leqq M|x_1 - x_2|,$$

(24b) $$(Bx_1 - Bx_2, x_1 - x_2) \geqq \alpha|x_1 - x_2|^2,$$

for all $x_1, x_2 \in H$.

PROPOSITION 4. *Under the assumption* (24), *there exists a constant* C_1 *such that*

$$|v^n - v| \leqq C_1 k^n; \qquad |u^n - u| \leqq C_1 k^n,$$

where $u = J_B^\lambda v$ *is the unique solution of* (2) *and* $k = (1 - 2\lambda\alpha/(1 + \lambda M)^2)^{1/2}$.

Proof. In view of (20) and (24b), we have

$$|v^{n+1} - v|^2 + 2\lambda\alpha|u^n - u|^2 \leqq |v^n - v|^2,$$

where we have applied that $v^n = u^n + \lambda b^n$, $v = u + \lambda b$, $b^n \in Bu^n$, $b \in Bu$, which gives also with (24a)

$$|v^n - v|^2 \leqq (1 + \lambda M)^2 |u^n - u|^2.$$

Hence

$$|v^{n+1} - v|^2 \leqq \left(1 - \frac{2\lambda\alpha}{(1 + \lambda M)^2}\right)|v^n - v|^2,$$

which gives the first estimate. The second one results from the fact that J_B^λ is a contraction. □

Remark 10. In view of (12) and (13), we can derive a similar estimate for Algorithm I. Both estimates show that there is an optimal value for λ, which is confirmed by the experiments. The best estimate for k above corresponds to $\lambda = 1/M$ which gives $k = (1 - \alpha/(2M))^{1/2}$. The experiments show that this estimate is rough. Notice that for the algorithm (5), one gets only $(1 - \alpha^2/(M^2))^{1/2}$.

1.3.4. The general case. Finally, in the general case, the algorithm is "almost convergent."

PROPOSITION 5. *Assume that* $A + B$ *is maximal monotone and* $0 \in R(A + B)$, *then* $v^n \rightharpoonup v$ *weakly and* u *is solution of* (3) *with* $u = J_B^\lambda u$.

Proof. In view of (22), we deduce that, for any $x \in D(A + B)$, $y \in Ax$, $z \in Bx$,

$$(a^n + b^n - (y + z), u^n - x) = (a^n - y, u^n + \lambda \varepsilon^n - x) + (b^n - z, u^n - x) - \lambda(a^n - y, \varepsilon^n)$$

$$\geqq -\lambda(a^n - y, \varepsilon^n).$$

Thus, if a subsequence $u^{n'}$ converges weakly to $\tilde{u}$, we have

$$(-(y + z), \tilde{u} - x) \geqq 0 \quad \forall x \in D(A + B),$$

hence $0 \in (A + B)\tilde{u}$, since $A + B$ is maximal monotone. $\quad\square$

Remark 11. In the special case where $A = B$, we claim that the whole sequence u^n converges weakly to $u \in A^{-1}(0)$. Indeed, in this case, $u = J_A^\lambda u = v$, hence $|u^n + \lambda b^n - u|^2 = |v^n - v|^2$ decreases. From (23a), $2\lambda(b^n, u^n - u)$ tends to zero, hence the sequence $|u^n - u|^2 + \lambda^2|b^n|^2$ converges to a real number depending on u, say $p(u)$. Let us show that this implies the uniqueness of the weak limit of any convergent subsequence, hence the weak convergence of the whole sequence u^n. Indeed, let $u^{n'} \to u_1 \in A^{-1}(0)$ and $u^{m'} \to u_2 \in A^{-1}(0)$; one has

$$|b^n|^2 + |u^n - u_1|^2 = |u^n - u_2|^2 + 2(u^n - u_2, u_2 - u_1) + |u_2 - u_1|^2 + |b^n|^2.$$

By taking the limit with respect to the subsequences n' or m', we get $p(u_1) = p(u_2) \pm |u_2 - u_1|^2$, hence $u_2 = u_1$.

As a bibliographical comment, we would like to mention that Algorithm II has been studied by Lieutaud [13, Chap V, § 2] in the finite dimensional and univoque case.

For additional bibliography, we refer the interested reader to Varga [18] and Marchouk [15].

A generalization of Algorithm II to the sum of n operators has been given by Douglas–Gunn [7]. However, the convergence seems difficult to prove in this general framework, even in the case of a sum of 3 operators. Varga [18, p. 240] suggests to combine the Peaceman–Rachford and the Douglas–Rachford algorithm in a single algorithm depending on a parameter ω, which gives Peaceman–Rachford for $\omega = 0$ and Douglas–Rachford for $\omega = 1$.

Let $F(\lambda) = (2J_A^\lambda - I)(2J_B^\lambda - I)$, the generalized algorithm suggested by Varga is nothing but

$$v^{n+1} = [(1 - \omega)F(\lambda) + \omega G(\lambda)]v^n$$

that is a convex combination of Algorithms I and II.

As the Browder result above extends to the convex combination of a nonexpansive operator $(F(\lambda))$ and a firmly nonexpansive operator $(G(\lambda))$, the new algorithm has the same properties as Algorithm II in the case $0 < \omega < 1$.

1.4. Application to a class of optimization problems. We shall consider the following abstract optimization problem introduced by Rockafellar [17]: Find $\bar{x} \in X$, solution of

$$(25) \qquad\qquad \inf_{x \in X} [f(\Lambda x) + g(x)],$$

where $\Lambda: X \to Y$ is a continuous linear operator, with closed range, X and Y are two Hilbert spaces, $f: Y \to (-\infty, +\infty]$ and $g: Y \to (-\infty, +\infty]$ are two convex lower semicontinuous functions.

This problem occurs in many situations in mechanics and economics. We shall explain how the previous algorithms can be applied to solve this problem.

1.4.1. The case where g is linear. Let $g(x) = \langle b, x \rangle$ where $b \in X'$ dual of X, and $\langle \cdot, \cdot \rangle$ denote the duality pairing between X and X'. We call $\Lambda^*: Y' \to X'$ the adjoint of Λ. As the range of Λ is assumed to be closed, there exists $\beta \in Y'$ such that $\Lambda^*\beta = b$.

Let $h(y) = f(y) + \langle \beta, y \rangle$, one can check that (25) is equivalent to

$$(26) \qquad \qquad \inf_{y \in K} h(y),$$

where $K = R(\Lambda)$ denotes the range of Λ.

We call I_K the indicator[1] function of K; assuming that

$$(27) \qquad \qquad \partial(h + I_K) = \partial h + \partial I_K,$$

which requires some qualification hypothesis on h (see Ekeland–Temam [9]), we have a special case of the stationary problem (2) with $H = Y$, $A = \partial I_K$, and $B = \partial h$.

The resolvent operator J_A^λ is then the projection operator onto K, which is easy to compute since K is a linear subspace.

Algorithms I and II can be applied for solving (26). Indeed, II is equivalent to one of the penalty-duality algorithms considered in [10], [11], where many numerical results can be found.

The application of I and II is easy when $I + \lambda\, \partial h$ is easy to invert. That is easy in the following situation which covers most of the examples considered in [10] and [11]:

$$X = H_0^1(\Omega), \qquad Y = (L^2(\Omega))^n, \qquad \Lambda = \operatorname{grad},$$

where Ω is an open bounded set of $\mathbb{R}^n$, and

$$h(y) = \int_\Omega \psi(y(\xi))\, d\xi,$$

where $\psi: \mathbb{R}^n \to \mathbb{R}$ is a convex lower semi-continuous function.

To solve $(I + \lambda\, \partial h)(y) \ni z$ is then equivalent to solving almost everywhere (on each element in the discrete case) the nonlinear equation in n variables

$$(28) \qquad \qquad (I + \lambda\, \partial\psi)y(\xi) \ni z(\xi).$$

Note that I and II can also be applied to the dual

$$(29) \qquad \qquad \inf_{p \in K^*} h^*(p),$$

where K^* is the kernel of Λ^* (again a linear subspace) and h^* is the conjugate of h.

1.4.2. The general case. In the general case where g is nonlinear, following an idea of J. P. Aubin, the problem (25) can be considered as a particular case of the stationary problem (2), but with $H = X \times Y$. Indeed, (25) is equivalent to

$$(30) \qquad \qquad \inf_{v \in K} \phi(v),$$

where $K = \{\{x, y\} \in X \times Y : y - \Lambda x = 0\}$ and $\phi(v) = f(y) + g(x)$ for $v = \{x, y\}$. Provided that $\partial\phi + \partial I_K = \partial(\phi + I_K)$, which requires again a qualification hypothesis, (25) is then a particular case of (2) with $A = \partial I_K$ and $B = \partial\phi = \begin{pmatrix} \partial g & 0 \\ 0 & \partial f \end{pmatrix}$.

Again K is a linear subspace, hence $J_A^\lambda = P_K$ is easy to compute, which makes Algorithms I and II very attractive to solve (25). They can also be applied to the dual of (25) which has the same structure (cf. [9]).

---

[1] $I_K(y) = 0$ if $y \in K$, and $+\infty$ otherwise.

2. Application of Algorithms I and II to the evolution equations.

2.1. Introduction. We recall the definition (16) of $G(\lambda)$ and write

$$(31) \qquad\qquad F(\lambda) = (2J_A^\lambda - I)(2J_B^\lambda - I).$$

Let $u^0 \in D(A) \cap D(B)$ and $b^0 \in B(u^0)$ be given; from now on we choose $\lambda = t/n$, where $t > 0$ is given. We let $v_n^0 = u^0 + (t/n)b^0$ and

ALGORITHM I.

$$v^n = F\left(\frac{t}{n}\right)v_n^0, \quad u^n = \left(I + \frac{t}{n}B\right)^{-1} v^n.$$

ALGORITHM II.

$$v^n = G\left(\frac{t}{n}\right)v_n^0, \quad u^n = \left(I + \frac{t}{n}B\right)^{-1} v^n.$$

In this section, we shall prove the sequence (u^n), generated by these algorithms, converges to the solution of the evolution equation (1) in the following sense:

$$(32a) \qquad\qquad u^n \to u(2t), \quad \text{as } n \to +\infty, \quad \text{in case I;}$$

$$(32b) \qquad\qquad u^n \to u(t), \quad \text{as } n \to \infty, \quad \text{in case II.}$$

We shall denote by E^0 the minimal section of an operator E, such that $E^0 x$ is the projection of the origin on the set Ex. Our main tool for this purpose is the following result of Brezis–Pazy (see e.g. Brezis [3]).

LEMMA 2. *Let C be a maximal monotone operator and $u(t)$ denote the solution of the evolution equation* (1). *Let $T(\lambda)$ be a family of contractions from $\overline{D(C)}$ into $\overline{D(C)}$ such that for any $\rho > 0$*

$$(33) \qquad \lim_{\lambda \to 0}\left(I + \rho\left(\frac{I - T(\lambda)}{\lambda}\right)\right)^{-1} x \to (I + \rho C)^{-1}x \quad \text{for all } x \in D(C);$$

then,

$$T\left(\frac{t}{n}\right)^n u^0 \to u(t),$$

uniformly for t in any compact set of $[0, +\infty)$.

2.2. Convergence. We shall prove the following result.

THEOREM 2. *The families of nonexpansive operators $F(\lambda)$ and $G(\lambda)$, defined in* (31) *and* (16), *satisfy the consistency property* (33). *As a corollary, if $C = A + B$ is maximal monotone, the sequences u^n generated by Algorithms I and II satisfy the convergence property* (32) *to the solution of the evolution equation* (1).

The proof of this theorem relies on the following lemma, originally proved by Baillon-Mercier [2], of which we give a simpler proof.

LEMMA 3. *Let A be a maximal monotone operator on H; $u \in D(A)$ and $u_\lambda \in H$ satisfy $(u_\lambda - u)/\lambda \to y$ as $\lambda \to 0$, then*

$$\frac{u_\lambda - J_A^\lambda u_\lambda}{\lambda} \to P_{Au}(y) \quad \text{as } \lambda \to 0,$$

where $P_{Au}(\cdot)$ denotes the projection operator onto Au.

Proof. Let E denote the maximal monotone operator such that $Eu = Au - y$ for all $u \in D(A)$; we have

$$J_E^\lambda (u_\lambda - \lambda y) = J_A^\lambda u_\lambda.$$

Let $v_\lambda = u_\lambda - \lambda y$; we have

$$\frac{1}{\lambda} (v_\lambda - J_E^\lambda v_\lambda) = \frac{1}{\lambda} (v_\lambda - u) + \frac{1}{\lambda} (u - J_E^\lambda u) + \frac{1}{\lambda} (J_E^\lambda u - J_E^\lambda v_\lambda);$$

we notice that $\lim_{\lambda \to 0} (v_\lambda - u)/\lambda = 0$, by assumption, that $(1/\lambda)(u - J_E^\lambda u) \to E^0 u = P_{Eu}(0)$ from a result of Brezis [3, p. 28]; finally, as J_E^λ is a contraction,

$$\left| \frac{1}{\lambda} (J_E^\lambda u - J_E^\lambda v_\lambda) \right| \le \left| \frac{1}{\lambda} (u - v_\lambda) \right| \to 0 \quad \text{as } \lambda \to 0,$$

which proves that

$$\frac{1}{\lambda} (v_\lambda - J_E^\lambda v_\lambda) \to P_{Eu}(0),$$

which is equivalent to the desired result. $\square$

Proof of Theorem 2. For $b \in Bx$, we define $x_\lambda = x + \lambda b$ and we notice that

$$C^0 x = b + P_{Ax}(-b).$$

Let us consider

$$\frac{1}{\lambda} (x_\lambda - F(\lambda)x_\lambda) = \frac{1}{\lambda} (2J_B^\lambda x_\lambda - 2J_A^\lambda (2J_B^\lambda x_\lambda - x_\lambda))$$

$$= \frac{1}{\lambda} (2x - 2J_A^\lambda (x - \lambda b))$$

$$= 2b + \frac{2}{\lambda} (u_\lambda - J_A^\lambda u_\lambda),$$

where we have introduced $u_\lambda = x - \lambda b$. We can then apply Lemma 3 for $x = u$ and $y = -b$ which shows that

$$\lim_{\lambda \to 0} \frac{1}{\lambda} (x_\lambda - F(\lambda)x_\lambda) = 2(b + P_{Au}(-b)) = 2C^0 u.$$

In other words, $C_\lambda = (1/\lambda)(I - F(\lambda))$ is such that there exists $y_\lambda \in C_\lambda x_\lambda$ with $y_\lambda \to 2C^0 x$. By a slight extension of a result of Brezis [3, Prop. 2.8, p. 29] this proves that

$$(I + \rho C_\lambda)^{-1} x \to (I + \rho 2C)^{-1} x \quad \text{for } x \in D(C) \text{ and } \lambda \to 0.$$

Then $F(\lambda)$ is consistent with $2C$ in the sense of (33); this shows, with Lemma 2, that $F(t/n)^n u^0 \to u(2t)$. Finally, to show that v^n, and therefore $u^n = (I + (t/n)B)^{-1} v^n \to u(2t)$, we notice that

$$\left| v^n - F\left(\frac{t}{n}\right)^n u^0 \right| \le \frac{t}{n} |b^0|$$

since $F(\lambda)$ is a contraction. In the same way, we have, from (16),

$$\frac{1}{\lambda}(x_\lambda - G(\lambda)x_\lambda) = \frac{1}{\lambda}(J_B^\lambda - J_A^\lambda(2J_B^\lambda - I))x_\lambda$$

$$= \frac{1}{\lambda}(x - J_A^\lambda(x - \lambda b))$$

$$= b + (u_\lambda - J_A^\lambda u_\lambda),$$

and we conclude, in the same way as for $F(\lambda)$, that $G(\lambda)$ is consistent with C, hence (32a). $\square$

Comments. In the linear case (A and B linear), it is well known that Algorithm I is of the second order, namely:

$$|u^n - u(t)| = O(\Delta t)^2 \quad \text{where } \Delta t = \frac{t}{n};$$

on the contrary, Algorithm II is only $O(\Delta t)$ (see e.g. [15], [16], [18]).

However, the first part devoted to the stationary problem showed that II is somewhat more stable than I in view of the counter-example given at Remark 6, even though both are unconditionally stable.

3. Examples and numerical results.

3.1. The obstacle problem. Let Ω be a regular bounded open set of $\mathbb{R}^d$ and $f \in L^2(\Omega)$ be given. We let $H = L^2(\Omega)$ and

$$Bv = -\Delta v + f,$$

where we specify $D(B) = H^2(\Omega) \cap H_0^1(\Omega)$. From Poincaré–Friedrichs' inequality, we have the existence of $\alpha > 0$ such that

$$(Bu - Bv, u - v) \geq \alpha |u - v|^2;$$

hence $B - \alpha I$ is monotone. We choose

$$K = \{v \in L^2(\Omega) : v(x) \geq 0 \text{ a.e. } x \in \Omega\},$$

and $A = \partial I_K$ so that $J_A^\lambda = P_K$ is the projection operator onto K and consists then of a single truncation:

$$(P_K v)(x) = \max(0, v(x)), \quad \text{a.e. } x \in \Omega.$$

At each step of Algorithms I and II, one has then only to make a truncation and then to compute the solution of a linear problem involving the Laplace operator. Note that, in this case, each step of the standard backward scheme

$$u^{n+1} = (I + \lambda C)^{-1} u^n$$

is as difficult to solve as the stationary problem itself. As for the algorithm (5), it is unstable for any λ in this continuous framework.

3.2. Problem solved. We have chosen $d = 1$, $\Omega =]0, 1[$ and $f(x) = -100$ for $x \leq \frac{1}{2}$, $f(x) = +100$ for $x \geq \frac{1}{2}$. The exact solution of the stationary problem is known analytically and consists of two pieces of parabola.

We have solved a finite difference approximation of the previous problem which corresponds to replace H by $\mathbb{R}^N$, and the Laplace operator by the matrix of the 3 point

formula

$$\frac{u_{i-1} - 2u_i + u_{i+1}}{h^2},$$

where $h = 1/(N+1)$. This problem may seem very simple, but the condition number of this matrix is $0(h^{-2})$. The stability condition for the forward scheme (5) is then $\lambda < Ch^{-2}$. In the case where $h = \frac{1}{40}$, we have been obliged to choose $\lambda \leq \lambda_0$ for the forward scheme with $\lambda_0 = 3 \times 10^{-4}$.

In Table 1, we give a comparison between the present algorithms for various λ and the forward scheme (5) with $\lambda = \lambda_0$. For values of λ about 30 times bigger than λ_0, the present algorithms give approximately the same results with a relative difference less than 1%.

TABLE 1

Values of the solution at $x = 0.7$ for various values of t and λ.

| t | $\lambda = 3 \cdot 10^{-4}$ | | | $\lambda = 3 \cdot 10^{-3}$ | | $\lambda = 6 \cdot 10^{-3}$ | | $\lambda = 3 \cdot 10^{-2}$ | |
	FWD*	I†	II‡	I	II	I	II	I	II
0.06	3.253	3.248	3.242	3.265	3.209	3.300	3.180	3.951	3.094
0.12	3.991	3.989	3.986	3.994	3.972	4.018	3.960	4.562	3.976
0.18	4.198	4.197	4.196	4.202	4.190	4.210	4.186	4.337	4.220
0.24	4.259	4.259	4.258	4.261	4.257	4.266	4.257	4.340	4.287
0.30	4.278	4.278	4.277	4.278	4.277	4.281	4.278	4.282	4.301
0.36	4.283	4.283	4.283	4.283	4.283	4.285	4.284	4.307	4.303
0.42	4.285	4.285	4.285	4.285	4.285	4.286	4.285	4.290	4.301
0.48	4.285	4.285	4.285	4.285	4.285	4.286	4.286	4.297	4.298
0.54	4.285	4.285	4.285	4.285	4.285	4.285	4.286	4.289	4.294
0.60	4.285	4.285	4.285	4.285	4.285	4.285	4.285	4.290	4.291
0.66	4.285	4.285	4.285	4.285	4.285	4.285	4.285	4.286	4.289

* FWD: forward scheme (5)
† I: Algorithm I
‡ II: Algorithm II

In Table 2, we give an idea of the efficiency of the present algorithms as a tool for solving the stationary problem. To get 4 exact digits, Algorithm II needs about 15 times less iterations than the forward scheme (5), and Algorithm I, 30 times less. A comparison with the case $h = \frac{1}{20}$ showed us that this difference increases as N increases.

Algorithm I seems twice as fast as Algorithm II, and this can be explained by the fact that the first gives, at the nth step, an approximation of the solution of the evolution problem at the time $2n\lambda$ versus $n\lambda$ for the second algorithm.

TABLE 2

Number of iterations necessary for Algorithms I and II to give a solution with n exact digits.

| n | $\lambda = 3 \cdot 10^{-4}$ | | $\lambda = 3 \cdot 10^{-3}$ | | $\lambda = 6 \cdot 10^{-3}$ | | $\lambda = 3 \cdot 10^{-2}$ | |
	I	II/FWD	I	II	I	II	I	II
1	300	500	30	50	10	25	6	18
2	400	700	40	70	15	35	9	24
3	600	1100	60	110	25	55	14	36
4	700	1400	70	140	45	95	60	102
5	900	1900	90	190	70	140	100	186

3.3. An hyperbolic obstacle problem. We have also considered the same obstacle problem with $Bu = du/dx - f$ instead of $-d^2u/dx^2 + f$, and the specification

$$D(B) = \{u \in H^1(0, 1), u(0) = 0\}.$$

We notice that B is monotone but not coercive:

$$(Bu, u) = \tfrac{1}{2}u^2(1) \geq 0.$$

We chose $A = \partial I_K$, where $K = \{v \in L^2(0, 1) : v \leq 0\}$, $f(x) = +1$ for $x \leq \tfrac{1}{2}$ and $f(x) = -1$ for $x \geq \tfrac{1}{2}$.

The exact solution of the stationary problem is

$$u_\infty(x) = \begin{cases} 0 & \text{for } x \leq \tfrac{1}{2}, \\ \tfrac{1}{2} - x & \text{for } x \geq \tfrac{1}{2}. \end{cases}$$

The exact solution of the evolution problem can be evaluated by an integration along the characteristics.

For the discrete problem, H is replaced by $\mathbb{R}^N$, and du/dx by an $N \times N$ matrix such that

$$(Bu)_i = \frac{u_i - u_{i-1}}{h} - f_i,$$

where $h = 1/N$.

It happens, in this case, that for $\lambda = h$, the forward scheme (5) corresponds to an integration along the characteristics and gives then the exact solution of the evolution problem.

This is not the case however for $\lambda < h$, and the stability condition for (5) is exactly $\lambda \leq h$.

All the computations have been performed with $h = \tfrac{1}{100}$.

We give a comparison of Algorithms I and II on this academic problem (see Tables 3, 4, 5 and 6). The results show that Algorithm I is more accurate than II, but II is more stable and not subject to the same oscillations as I for λ five or ten times greater than h.

TABLE 3

Values of the solution for $x = 1$, with respect to t, λ, and the type of the algorithm (hyperbolic case where $u^0 = 0$ for $t = 0$).

t	FWD	FWD (exact)	$\lambda = 0.01$ I	II	$\lambda = 0.05$ I	II	$\lambda = 0.1$ I	II
	$\lambda = 0.005$							
0.2	−0.200	−0.200	−0.200	−0.200	−0.199	−0.199	−0.198	−0.194
0.4	−0.399	−0.400	−0.397	−0.393	−0.390	−0.373	−0.381	−0.353
0.6	−0.499	−0.500	−0.497	−0.490	−0.509	−0.468	−0.509	−0.446
0.8	−0.500	−0.500	−0.500	−0.499	−0.503	−0.495	−0.540	−0.485
1.0	−0.500	−0.500	−0.500	−0.500	−0.498	−0.499	−0.505	−0.496
1.2	−0.500	−0.500	−0.500	−0.500	−0.500	−0.500	−0.487	−0.499
1.4	−0.500	−0.500	−0.500	−0.500	−0.500	−0.500	−0.504	−0.500
1.6	−0.500	−0.500	−0.500	−0.500	−0.500	−0.500	−0.504	−0.500

TABLE 4

Number of iterations necessary for Algorithms
I and II to give a solution with n exact digits.

n	$\lambda = 0.01$ I	$\lambda = 0.01$ II	$\lambda = 0.05$ I	$\lambda = 0.05$ II	$\lambda = 0.1$ I	$\lambda = 0.1$ II
1	50	50	6	12	3	6
2	60	60	6	16	7	10
3	70	70	12	20	25	13
4	70	80	23	24	38	16

TABLE 5

Same as above but starting from $u^0 = -x$ (for $t = 0$).

t	FWD (exact)	$\lambda = 0.01$ I	$\lambda = 0.01$ II	$\lambda = 0.05$ I	$\lambda = 0.05$ II
0.1	−1.000	−1.000	−1.000	−1.000	−0.999
0.2	−1.000	−1.000	−1.000	−0.999	−0.998
0.3	−1.000	−0.999	−0.999	−0.996	−0.985
0.4	−1.000	−0.995	−0.986	−0.981	−0.946
0.5	−1.000	−0.944	−0.920	−0.933	−0.872
0.6	−0.800	−0.796	−0.785	−0.828	−0.772
0.7	−0.600	−0.620	−0.635	−0.672	−0.672
0.8	−0.500	−0.520	−0.541	−0.526	−0.593
1.0	−0.500	−0.500	−0.501	−0.485	−0.518
1.2	−0.500	−0.500	−0.500	−0.513	−0.502

TABLE 6

Number of iterations necessary
for Algorithms I and II to give a
solution with n exact digits.

n	$\lambda = 0.01$ I	$\lambda = 0.01$ II	$\lambda = 0.05$ I	$\lambda = 0.05$ II
1	40	80	8	16
2	45	90	13	22
3	45	110	25	26
4	50	120	35	32

For given t, the solution $u(x, t)$ given by Algorithms I and II is smoother than the exact solution: the same phenomenon occurs for the forward scheme (5) and $\lambda < h$.

The results show that it is possible, in this case, to take λ five times bigger than h with still fairly accurate results, which does not give a decisive superiority of Algorithms I and II over the standard forward scheme which requires clearly less computations in this special one dimensional case.

4. Conclusion. From the numerical results, it seems that the alternating direction Algorithms I and II are more interesting in parabolic problems where the stability

condition for the forward scheme is sharp; Algorithm I requires roughly half the computation of II.

For the hyperbolic problem considered in § 3.3, I seems to be more accurate than II, but II to be more stable than I. The comparison with a forward scheme is difficult, since the forward scheme gives an exact solution for $\lambda = h$ (integration along the characteristics).

The superiority of Algorithm II over the forward scheme (5) has been shown by [10], [11], for a wide class of problems of the type considered in § 1.4, where J_A^λ is linear and J_B^λ nonlinear but split in small nonlinear pieces. We conjecture that I should have some analogous properties in this case.

Acknowledgment. We would like to thank Jim Douglas and Todd Dupont for very helpful conversations, and the referee for his interesting remarks.

REFERENCES

[1] J. B. BAILLON, *Quelques propriétés de convergence asymptotique pour les contractions impaires*, C.R. Acad. Sci. Paris, Série A, 283 (1976), pp. 587–590.

[2] J. B. BAILLON AND B. MERCIER, Internal Report, Ecole Polytechnique, 1977.

[3] H. BREZIS, *Opérateurs maximaux monotones et semigroupes de contraction dans les espaces de Hilbert*, Lecture Notes no. 5, North-Holland, Amsterdam, 1973.

[4] ———, to appear. (Lecture Notes, Ecole Polytechnique, Unpublished.)

[5] F. E. BROWDER AND W. V. PETRYSHN, *Construction of fixed points of nonlinear mappings in Hilbert spaces*, J. Math. Anal. Appl., 20 (1967), pp. 197–228.

[6] R. E. BRUCK, *An iterative solution of a variational inequality for certain monotone operators in Hilbert spaces*, Bull. Amer. Math. Soc., 81 (1875), pp. 890–892.

[7] J. DOUGLAS AND J. E. GUNN, *A general formulation of alternating direction methods*, Numer. Math., 6 (1964), pp. 428–453.

[8] J. DOUGLAS AND H. H. RACHFORD, *On the numerical solution of the heat conduction problem in 2 and 3 space variables*, Trans. Amer. Math. Soc., 82 (1956), pp. 421–439.

[9] I. EKELAND AND R. TEMAM, *Convex Analysis and Variational Problems*, Springer-Verlag, New York, 1977.

[10] D. GABAY AND B. MERCIER, *A dual algorithm for the solution of nonlinear variational problems via finite element approximation*, Computer Math. Appl., 2 (1976), pp. 17–40.

[11] R. GLOWINSKI AND A. MAROCCO, *Approximation par éléments finis d'ordre un et résolution par pénalisation-dualité d'une classe de problèmes non linéaires*, R.A.I.R.O., R2 (1975), pp. 41–76.

[12] A. A. GOLDSTEIN, *Convex programming in Hilbert spaces*, Bull. Amer. Math. Soc., 70 (1964), pp. 709–710.

[13] J. LIEUTAUD, *Approximation par des méthodes de décomposition*, Doctoral Dissertation, Paris, 1971.

[14] P. L. LIONS, *Une méthode itérative de résolution d'inéquations variationnelles*, to appear.

[15] G. I. MARCHOUK, *Methods of Numerical Mathematics*, Springer-Verlag, New York, 1975.

[16] D. H. PEACEMAN AND H. H. RACHFORD, *The numerical solution of parabolic elliptic differential equations*, J. Soc. Indust. Appl. Math., 3 (1955), pp. 28–41.

[17] T. ROCKAFELLAR, *Convex Analysis*, Princeton University Press, Princeton, NJ, 1970.

[18] R. S. VARGA, *Matrix iterative analysis*, Prentice-Hall, Englewood Cliffs, NJ, 1966.

Note added in proof. Our references would be incomplete without the excellent paper by R. B. Kellogg, *Nonlinear alternating direction algorithm*, Math. Comp. 23 (1969), pp. 23–28 which deals with Algorithm I in the case of single valued monotone operators.