An. Şt. Univ. Ovidius Constanţa Vol. 12(2), 2004, 83–92

ON THE SPLITTING METHODS AND THE PROXIMAL POINT ALGORITHM FOR MAXIMAL MONOTONE OPERATORS

Corina L. Chiriac

To Professor Dan Pascali, at his 70's anniversary

Abstract

The theory of maximal set-valued monotone operators provides a powerful general framework for the study of convex programming and variational inequalities. A fundamental algorithm for finding a root of a monotone operator is the proximal point algorithm.

A lot of papers have been dedicated to this subject. Two principal classes of splitting methods are Peaceman-Rachford, and Douglas-Rachford algorithms. Eckstein has presented a generalized form of the proximal point algorithm – created by synthesizing the work of Rockafellar with that of Golshtein and Tretyakov – and has shown how it gives rise to a new method, generalized Douglas-Rachford splitting. Some results, about a connection between the proximal algorithm and Douglas-Rachford splitting will be given.

We give a proof that Douglas-Rachford splitting is an application of the proximal point algorithm. Using this fact we prove that Peaceman-Rachford splitting is equivalent to applying the generalized proximal point algorithm.

2000 Mathematics Subject Classification: 46T99, 47J25, 49J40.

Introduction.

For many maximal monotone operators T, the evaluation of inverses for operators of the form $I + \lambda T$, where $\lambda > 0$, may be difficult. Now suppose that we can choose two maximal monotone operators W and V such that $W + V = T$, but J_W^λ and J_V^λ are easier to evaluate than J_T^λ. A *splitting*

Key Words: proximal point algorithm, generalized Douglas-Rachford splitting method.

algorithm is a method that employs the resolvents J_W^λ, J_V^λ of W and V, but does not use the resolvent J_T^λ of the original operator T. Here we consider theDouglas -Rachford scheme of Lions and Mercier [9].

We shall present a result, which establishes a relation between two well-known algorithms: proximal point algorithm and Douglas-Rachford splitting algorithm.

Preliminary results.

We enumerate some concepts and main results, which will be used to get our results.

Let H be a real Hilbert space with inner product $(\cdot,\cdot)$ and associated norm $||\cdot||$. We consider a multi-valued operator $T : H \to 2^H$. First we recall some properties of the monotone and maximal monotone operators.

Theorem 1 (Minty [10]). *A monotone operator $T : H \to 2^H$is maximal if and only if $R(I+T) = H$.*

For alternative proofs of Theorem 1,or stronger related theorems, see [12],[2] or [7].

Given any operator A, let J_A denote the operator $(I + A)^{-1}$. Given any positive scalar λ and an operator$T$, $J_T^\lambda = (I + \lambda T)^{-1}$ is called the *resolvent* of T. An operator $B : H \to 2^H$ is said to be *nonexpansive* if

$$||y\prime - y|| \leq ||x\prime - x|| \text{ for all } [x, y], [x\prime, y\prime] \in G(B).$$

Note that nonexpansive operators are necessarily single-valued and Lipschitz continuous (see [11]).

An operator $C : H \to 2^H$ is said to be *firmly nonexpansive* if

$$||y\prime - y|| \leq (x\prime - x, y\prime - y) \text{ for all } [x, y], [x\prime, y\prime] \in G(C).$$

The following lemma summarizes some well-known properties of firmly nonexpansive operators.

Lemma 2 (Rockafellar [13]). *Let $T : H \to 2^H$ be an operator. The following statements are hold:*

(i) *All firmly nonexpansive operators are nonexpansive.*

(ii) *T is firmly nonexpansive if and only if $2T - I$ is nonexpansive.*

(iii) *T is firmly nonexpansive if and only if it is of the form $\frac{1}{2}(U + I)$, where U is nonexpansive.*

(iv) *T is firmly nonexpansive if and only* if *$I - T$ is firmly nonexpansive.*

We now give a critical theorem. The "only if" part of the following theorem has been well-known for some time (see [13]), but the "if" part has appeared in

[4]. The purpose here is to stress the complete symmetry that exists between (maximal) monotone operators and (full-domained) firmly nonexpansive operators over any Hilbert space.

Theorem 3 (Eckstein [5]). *Let λ be any positive scalar. An operator $T : H \to 2^H$ is monotone if and only if its resolvent $J_T^\lambda = (I + \lambda T)^{-1}$ is firmly nonexpansive. Furthermore, T is maximal monotone if and only if J_T^λ is firmly nonexpansive and $D(J_T^\lambda) = H$.*

Corollary 4. *An operator T is firmly nonexpansive if and only if $T^{-1} - I$ is monotone. T is firmly nonexpansive with full domain if and only if $T^{-1} - I$ is maximal monotone.*

Corollary 5. *For any $\lambda > 0$, the resolvent J_T^λ of a monotone operator T is single-valued. If T is also maximal, then J_T^λ is defined on all of H.*

Corollary 6 (The Representation Lemma). *Let λ >0 and let $T : H \to 2^H$ be monotone. Then every element $z \in H$ can be written in at most one way as $x + \lambda y$, where $y \in Tx$. If T is maximal, then every element $z \in H$ can be written in exactly one way as $x + \lambda y$, where $y \in Tx$.*

Corollary 7. *The correspondence from an operator T into $(I + T)^{-1}$ is a bijection between the collection of maximal monotone operators on H and the collection of firmly nonexpansive operators on H.*

Remark 8. Corollary 7 reminds us a result of Minty [10], but it is not identical (Minty did not use the concept of firm nonexpansiveness; see also [6]).

A *root* or *zero* of an operator T is a point x such that

$$0 \in Tx.$$

Let $zer(T) = T^{-1}(0)$ denote the set of all such points. The zeroes of a monotone operator are precisely the fixed points of its resolvents. In other words the following result is true:

Lemma 9. *Given any maximal monotone operator* T*, real number λ >0, and $x \in H$, we have $0 \in Tx$ if and only if $J_T^\lambda(x) = x$.*

Decomposition: Douglas-Rachford splitting methods

We shall consider the Douglas-Rachford scheme of Lions and Mercier [9].

Let us fix some$\lambda > 0$ and two maximal monotone operators W and V. The sequence $\{z^k\}$ is said to *obey the Douglas-Rachford recursion for λ,W and V* if

$$z^{k+1} = J_W^\lambda(2J_V^\lambda - I)z^k + (I - J_V^\lambda)z^k.$$

Let $[x^k, v^k] \in G(V)$ be, for all $k \geq 0$, the unique element such that $x^k + \lambda v^k = z^k$(by Corollary 6). Then, for all k, one has

$$(I - J_V^\lambda)z^k = x^k + \lambda v^k - x^k = \lambda v^k,$$

$$(2J_V^\lambda - I)z^k = 2x^k - (x^k + \lambda v^k) = x^k - \lambda v^k.$$

Similarly, if $[y^k, u^k] \in G(W)$, then $J_W^\lambda(y^k + \lambda u^k) = y^k$.

In view of these identities, one may give the following alternative prescription for finding z^{k+1} from z^k:

(i) Find the unique $[y^{k+1}, u^{k+1}] \in G(W)$ such that $y^{k+1} + \lambda u^{k+1} = x^k - \lambda v^k$.

(ii) Find the unique $[x^{k+1}, v^{k+1}] \in G(V)$ such that $x^{k+1} + \lambda v^{k+1} = y^{k+1} + \lambda v^k$.

The analysis is centered on the operator

$$S_{W,V}^\lambda = J_W^\lambda \circ (2J_V^\lambda - I) + (I - J_V^\lambda),$$

where "$\circ$" denotes mapping composition.

Thus the Douglas-Rachford recursion can be written as

$$z^{k+1} = S_{W,V}^\lambda(z^k).$$

Lions and Mercier [9] showed that $S_{W,V}^\lambda$ is firmly nonexpansive, from which they obtained the convergence of $\{z^k\}$. Their analysis can be extended by exploiting the connection between firm nonexpansiveness and maximal monotonicity.

Consider the operator

$$Q_{W,V}^\lambda = (S_{W,V}^\lambda)^{-1} - I.$$

Using the above algorithmic description (i)-(ii), we obtain the following expression for the graph of $S_{W,V}^\lambda$

$$G(S^{\lambda}_{W,V}) = \{[x + \lambda v, y + \lambda v] | [x, v] \in G(V), [y, u] \in G(W), y + \lambda u = x - \lambda v\}.$$

A simple computation provides an expression for $Q^{\lambda}_{W,V} = (S^{\lambda}_{W,V})^{-1} - I$,wiyh its graph:

$$G(Q^{\lambda}_{W,V}) = \{[y + \lambda v, x - y] | [x, v] \in G(V), [y, u] \in G(W), y + \lambda u = x - \lambda v\}.$$

Given any Hilbert space H, a scalar $\lambda > 0$, and the operators W and V on H, we define $Q^{\lambda}_{W,V}$ to be the *splitting operator* of W and V with respect to λ. The following theorem establishes the maximal monotonicity of $Q^{\lambda}_{W,V}$:

Theorem 10. *If W and V are monotone then $Q^{\lambda}_{W,V}$ is monotone. If W and V are maximal monotone then $Q^{\lambda}_{W,V}$ is maximal monotone.*

Combining Theorems 10 and 3 , we have the key Lions-Mercier result.

Corollary 11. *If W and V are maximal monotone, then $S^{\lambda}_{W,V} = (I + Q^{\lambda}_{W,V})^{-1}$ is firmly nonexpansive and is defined on all of H.*

There is also a relationship between the zeroes of $Q^{\lambda}_{W,V}$ and those of $W+V$.

Theorem 12. *Given $\lambda > 0$ and the operators W and V on H,*we have:

$$zer(Q^{\lambda}_{W,V}) = Z_{\lambda} = \{x+\lambda v | v \in Vx, -v \in Wx\} \subset \{x+\lambda v | x \in zer(W+V), v \in Vx\}.$$

In conclusion, given any zero z of $Q^{\lambda}_{W,V}$, $J^{\lambda}_{V}(z)$ is a zero of $W+V$. Thus one may imagine finding a zero of$W + V$ by using the proximal point algorithm on $Q^{\lambda}_{W,V}$ and then applying the operator J^{λ}_{V} to the result. In fact, this is precisely what the Douglas-Rachford splitting method does.

Theorem 13. *The Douglas-Rachford iteration*

$$z^{k+1} = J^{\lambda}_{W}(2J^{\lambda}_{V} - I)z^k + (I - J^{\lambda}_{V})z^k$$

is equivalent to applying proximal point algorithm to the maximal monotone operator $Q^{\lambda}_{W,V}$ with the proximal point stepsizes λ_k fixed at 1, and exact evaluation of the resolvents.

In conclusion the Douglas-Rachford splitting method is a special case of the proximal point algorithm as applied to the splitting operator $Q^{\lambda}_{W,V}$.

Generalized Proximal Point Algorithm

We present a scheme due to Golshtein and Tretyakov [6], which generalizes proximal point algorithm. They consider iterations of the form

$$z^{k+1} = (I - \rho_k) z^k + \rho_k J_T^{\lambda}(z^k),$$

(1)
where $\{\rho_k\}_{k=0}^{\infty} \subset (0, 2)$ is a sequence of *over*-or *under-relaxation* factors.

Golshtein and Tretyakov also allow resolvents to be evaluated approximatively, but, unlike Rockafellar, do not allow the stepsize λ to vary with k, restrict H to be finite-dimensional, and do not consider the case in which $zer(T) = \emptyset$. The following theorem combines the results of Rockafellar and Golshtein-Tretyakov.

Theorem 14 (Eckstein [5]). *LetT be a maximal monotone operator on* H, *and let* $\{z^k\}$ *be such that*

$$z^{k+1} = (I - \rho_k) z^k + \rho_k w^k \textit{ for all } k \geq 0,$$

where

$$||w^k - (I + \lambda_k T)^{-1}(z^k)|| \leq \varepsilon_k \textit{ for all } k \geq 0,$$

and $\{\varepsilon_k\}$, $\{\rho_k\}$, $\{\lambda_k\} \subset [0, +\infty)$ *are sequences such that*

$$\mathrm{E}_1 = \sum_{k=0}^{\infty} \varepsilon_k < \infty,\ \Delta_1 = \inf_{k \geq 0} \rho_k > 0,\ \Delta_2 = \sup_{k \geq 0} \rho_k < 2,$$
$$\overline{\lambda} = \inf_{k \geq 0} \lambda_k > 0.$$

Such a sequence $\{z^k\}$ *is said to be conform to the generalized proximal point algorithm. If* T *possesses a zero, then* $\{z^k\}$ *converges weakly to a zero of* T. *If* T *has no zeroes, then* $\{z^k\}$ *is an unbounded sequence.*

We make some **remarks**:

- Theorem 14 states also that, in a general Hilbert space, the proximal point algorithm produces an unbounded sequence when applied to a maximal monotone operator that has no zeroes.

- In view of Theorems 14 and 12, we immediately obtain the following Lions-Mercier convergence result:

If $W + V$ *has a zero, then the Douglas-Rachford splitting method produces a sequence* $\{z^k\}$ *weakly convergent to a limit* z *of the form* $x + \lambda v$, *where* $x \in zer(W + V)$, $v \in Vx$,*and* $-v \in Wx$.

- Using Remark 15, we deduce the following result:
Suppose W and V are maximal monotone operators and $zer(W+V)=\emptyset$. Then the sequence $\{z^k\}$ produced by Douglas-Rachford splitting is unbounded.

We intend **to establish a relation between the Peaceman-Rachford algorithm and the generalized proximal point algorithm presented above.**

The following result will be used in the next presentation. We adapt a theorem, which was stated and proved in [1], in view of our goal.

Theorem 18. *Assume that T is a maximal monotone operator on H and* $zer(T)$ be a nonempty set. *We consider that the following statements hold:*
(i) $0<\underline{\lambda}\leq\lambda_k$ *for all* $k\in\mathbf{N}^*$,
(ii) $0<\overline{\rho}\leq\rho_k\leq 2$ *for all* $k\in\mathbf{N}^*$.
Then the sequence $\{z^k\}$ generated by the rule (1) weakly converges to an element of $zer(T)$ and it is such that

$$\lim_{k\to\infty}||z^k-z^{k-1}||=0.$$

In the following analysis, we use the *Peaceman-Rachford scheme* of Lions and Mercier [9]. Let us consider some $\lambda>0$ and two maximal monotone operators W and V. The sequence $\{z^k\}$ is obtained by Peaceman-Rachford algorithm if

$$z^{k+1}=(2J_W^\lambda-I)(2J_V^\lambda-I)z^k.$$

(2)
Given any sequence satisfying (2), let $[z^k,v^k]$ be, for all $k\geq 0$, the unique element of $G(V)$ such that

$$x^k+\lambda v^k=z^k.$$

The existence and uniqueness of this element follow from Corollaries 5, 6. Then for all k, one has

$$(2J_V^\lambda-I)z^k=2x^k-(x^k+\lambda v^k)=x^k-\lambda v^k$$

.

Similarly, if $[y^k,u^k]\in G(W)$, then

$$J_W^\lambda(y^k + \lambda u^k) = y^k.$$

Using these relations, we can give the following alternative scheme for finding z^{k+1} from z^k:

(i) Find the unique element $[y^{k+1}, u^{k+1}] \in G(W)$such that

$$y^{k+1} + \lambda u^{k+1} = x^k - \lambda v^k,$$

(ii) Find the unique element $[x^{k+1}, v^{k+1}] \in G(V)$ such that

$$x^{k+1} + \lambda v^{k+1} = y^{k+1} - \lambda v^{k+1}$$

From (2) we obtain

$$z^{k+1} = 2J_W^\lambda(2J_V^\lambda - I)z^k + 2(I - J_V^\lambda)z^k - z^k.$$

This relation suggests us to use the operator

$$S_{W,V}^\lambda = J_W^\lambda \circ (2J_V^\lambda - I) + (I - J_V^\lambda).$$

The Peaceman-Rachford recursion (2) can be written as follows:

$$z^{k+1} = 2S_{W,V}^\lambda(z^k) - z^k = (2S_{W,V}^\lambda - I)z^k$$

(3)

Consider the operator

$$Q_{W,V}^\lambda = (S_{W,V}^\lambda)^{-1} - I,$$

Since Theorem 10 implies that $Q_{W,V}^\lambda$ is maximal monotone, we can define the operator

$$P_{W,V}^\lambda = 2(I + Q_{W,V}^\lambda)^{-1} - I = 2(I + Q_{W,V}^\lambda)^{-1} + (1 - 2)I.$$

We rewrite (3) using $P^{\lambda}_{W,V}$, in the form

$$z^{k+1} = P^{\lambda}_{W,V}(z^k) \ = 2(I + Q^{\lambda}_{W,V}(z^k) + (1-2)z^k.$$

Theorem 19. *The Peaceman-Rachford iteration*

$$z^{k+1} = (2J^{\lambda}_{W} - I)(2J^{\lambda}_{V} - I)z^k$$

is equivalent to applying the generalized proximal point algorithm to the maximal monotone operator $Q^{\lambda}_{W,V}$ with the proximal point stepsizes λ_k fixed at 1 *and the relaxation factors $\rho_k = 2$ for all $k \geq 1$.*

Proof. The Peaceman-Rachford iteration is

$$z^{k+1} = P^{\lambda}_{W,V}(z^k),$$

which is just

$$z^{k+1} = (1-2)z^k + 2(I + Q^{\lambda}_{W,V})^{-1}(z^k),$$

that is the generalized proximal point scheme (1) with $\rho_k = 2$ for all $k \geq 1$.

In view of the Theorems 18 and 12, we immediately obtain the following result.

Corollary 20. *If $W+V$ has a zero, then the Peaceman-Rachford splitting method produces a sequence $\{z^k\}$ weakly convergent to a limit* z *of the form $x + \lambda v$, where $x \in zer(W+V)$, $v \in Vx$ and $-v \in Wx$.*

Proof. From the Theorem 18, we obtain that the sequence $\{z^k\}$ converges weakly to a limit $z \in zer(Q^{\lambda}_{W,V})$. Applying Theorem 12, we have

$$z = x + \lambda v,$$

where $x \in zer(W+V)$, $v \in Vx$ and $-v \in Wx$.

References

[1] P. Alexandre, V. H. Nguyen, P. Tossings, *The perturbed generalized proximal point algorithm*, Math. Modelling Numer. Anal.**32**(1998), 223-253.

[2] H. Brezis, *Operateurs Maximaux Monotones et Semi-groupes de Contractions dans les Espaces de Hilbert*, North-Holland, 1973.

[3] C. Chiriac,*A numerical analysis of variational inequalities,* PhD thesis, "Ovidius" University, 2002.

[4] J. Eckstein,*The Lions-Mercier algorithm and the alternating direction method are instances of the proximal point algorithm*, Report LIDS-P-17-69, Laboratory for Information and Decision Sciences, MIT Cambridge, MA, 1988.

[5] J. Eckstein, D. P. Bertsekas,*On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators*, Math. Programming **55** (1992), 293-318.

[6] E. G. Golshtein, N. V. Tretyakov,*Modified Lagrangians in convex programming and their generalizations,* Math. Programming Study **10** (1979), 86-97.

[7] M. C. Joshi, R. K. Bose,*Some Topics in Nonlinear Functional Analysis,* Halsted, Wiley, New Delhi, 1985.

[8] J. Lawrence, J. E. Spingarn,*On fixed points of nonexpansive piecewise isometric mappimgs,* Proc. London Math. Soc. **55** (1987), 605-624.

[9] P. L. Lions, B. Mercier, *Splitting algorithms for the sum of two nonlinear operators,* SIAM J. Numer. Anal. **16** (1979), 964-979.

[10] G.J. Minty,*Monotone (nonlinear) operators in Hilbert space,* Duke Math. Journal **29** (1962), 341-346.

[11] D. Pascali, S. Sburlan, *Nonlinear Mappings of Monotone Type,* Ed. Academiei, Bucuresti 1978.

[12] R. T. Rockafellar,*On the maximal monotonicity of subdifferential mappings,* Pacific J. Math. **33** (1970), 209-216.

[13] R. T. Rockafellar,*Monotone operators and the proximal point algorithm,* SIAM J. Control Optim. **14** (1976), 877-898.

"CORINA L. CHIRIAC
"George Baritiu" University, Brasov, Romania
corinalchiriac@yahoo.com