ASYMPTOTIC CONVERGENCE ANALYSIS OF THE FORWARD-BACKWARD SPLITTING ALGORITHM

Ciyou Zhu *

Department of Mathematical Sciences
Johns Hopkins University, Baltimore, MD 21218

September 1991; revised April 1993.

Abstract. The asymptotic convergence of the forward-backward splitting algorithm for solving equations of type $0 \in T(z)$ is analyzed, where T is a multivalued maximal monotone operator in the n-dimensional Euclidean space. When the problem has a nonempty solution set, and T is split in the form $T = \mathcal{T} + h$ with $\mathcal{T}$ being maximal monotone and h being co-coercive with modulus greater than $\frac{1}{2}$, convergence rates are shown, under mild conditions, to be linear, superlinear, or sublinear depending on how rapidly $\mathcal{T}^{-1}$ and h^{-1} grow in the neighborhoods of certain specific points.

As a special case, when both $\mathcal{T}$ and h are polyhedral functions, we get R-linear convergence and 2-step Q-linear convergence without any further assumptions on the strict monotonicity on T or on the uniqueness of the solution. As another special case when $h = 0$, the splitting algorithm reduces to the proximal point algorithm, and we get new results, which complement R. T. Rockafellar's and F. J. Luque's earlier results on the proximal point algorithm.

Keywords. Monotone operator, multivalued equation, variational inequality, splitting algorithm, proximal point algorithm, asymptotic convergence.

---

* This work was initiated when the author was at University of Washington and were completed when the author was at the Johns Hopkins University, supported by Eliezer Naddor Postdoctoral Fellowship in Mathematical Sciences during the year 1991–92.

Currrent address: Mathematics and Computer Science Division, Argonne National Laboratory, Argonne, IL 60439.

1. Introduction.

Let $\mathbb{R}^n$ be the n-dimensional Euclidean space, and let the Euclidean inner product and norm be denoted by $\langle \cdot, \cdot \rangle$ and $|\cdot|$ respectively. A multifunction $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is said to be a *monotone operator* if

$$\langle z - z', w - w' \rangle \geq 0 \text{ whenever } w \in T(z),\ w' \in T(z'). \tag{1.1}$$

It is said to be *maximal monotone* if, in addition, the graph $\{(z, w) \in \mathbb{R}^n \times \mathbb{R}^n \mid w \in T(z)\}$ is not properly contained in the graph of any other monotone operator $T' : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$.

Such operators have been studied extensively because of their role in convex analysis and certain other fields. A fundamental problem is that of determining an element z such that $0 \in T(z)$. Some of the most important problems in the area of convex programming and related fields, such as variational inequality problems, can all be cast into this general framework. (See e.g. [19].)

We denote by $\bar{Z}$ the solution set of the equation $0 \in T(z)$, and let T^{-1} be the *inverse* of T, i.e., $T^{-1}(w) = \{z \in \mathbb{R}^n \mid w \in T(z)\}$. Obviously T^{-1} is maximal monotone if and only if T is maximal monotone. The *effective domain* of T is defined by the set $\{z \in \mathbb{R}^n \mid T(z) \neq \emptyset\}$. Suppose the operator T can be written in the split form

$$T = \mathcal{T} + h, \tag{1.2}$$

where $\mathcal{T} : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is a maximal monotone operator and $h : \mathbb{R}^n \to \mathbb{R}^n$ is a single-valued function. Then the original problem of finding z such that $0 \in T(z)$ is equivalent to the following problem:

$$\text{Find a } \bar{z} \in \mathbb{R}^n \text{ satisfying } 0 \in \mathcal{T}(\bar{z}) + h(\bar{z}). \tag{1.3}$$

Suppose $\bar{Z} \neq \emptyset$. Consider the following iteration

$$z^{k+1} = (I + \mathcal{T})^{-1}(I - h)(z^k), \quad k = 0, 1, 2, \cdots. \tag{1.4}$$

It has been shown by Minty [11] that the *proximal mapping* $(I + \mathcal{T})^{-1}$ is a single-valued mapping and its effective domain is all of $\mathbb{R}^n$. Therefore, the iteration (1.4) is well defined. It is easy to see that $\bar{z} \in \bar{Z}$ if and only if $\bar{z}$ is a fixed point of the algorithmic mapping defined by (1.4).

As pointed out by Tseng [22, 23] and Chen and Rockafellar [3–6], the iteration (1.4), despite its simplicity, is a powerful tool for the development of decomposition methods. Numerous existing algorithms for convex programming and variational inequality problems can be shown to be special cases of this iteration. Hence, any results on the convergence of (1.4) have numerous implications for all these algorithms. As a special case, it is easy to see that the proximal point algorithm is also included in the family (1.4) by taking $h = 0$.

Convergence of the splitting algorithm has been extensively studied [1b, 3–7, 9, 16, 23]. As to the rate of convergence, Chen and Rockafellar [6] proved, under some commonly used hypothesis, that the iteration scheme converges linearly from the very beginning if h is strongly monotone (which implies in turn that T itself is strongly monotone). In [3–6], Chen and Rockafellar also explored the possibility of introducing "step sizes" in the iteration and replacing the identity operator I with some specifically chosen mappings to enhance the convergence under the same strong monotonicity assumption. But the assumption that T is strongly monotone (or less stringently, strictly monotone) often excludes some important applications. For instance, the assumption certainly does not hold when the solution $\bar{z}$ is not unique.

In this paper, we establish asymptotic rate-of-convergence results without such a strong monotonicity assumption on T. Instead, we relate the rate of convergence to some "growth conditions" on $\mathcal{T}^{-1}$ and h^{-1} at some specific points, as Luque [10] did for the proximal point algorithm. With a careful study of the geometrical aspects of the convergence of the sequence $\{z^k\}$ generated by (1.4), we are able to draw conclusions about the rate of convergence of $\{z^k\}$ itself, while Luque's conclusions on the proximal point algorithm are mostly on the convergence of $\{\mathrm{dist}(\bar{Z}, z^k)\}$ to 0 [10]. Hence we even get new results on the proximal point algorithm as a special case of the iteration scheme (1.4) when $h = 0$.

We focus on the fundamental iteration scheme (1.4). The result that is most useful to us is the following Proposition 1.1 given by Gabay, where the convergence is established on the assumption of some *co-coercive* property of h. A function $h : Z \to \mathbb{R}^n$ is said to be co-coercive with modulus $\lambda > 0$ if

$$\langle h(z) - h(z'), z - z' \rangle \geq \lambda |h(z) - h(z')|^2 \quad \forall z \in Z,\ \forall z' \in Z. \tag{1.5}$$

Notice that a co-coercive function is Lipschitz continuous and monotone. Moreover, a co-coercive function with modulus $\lambda \geq 1$ is firmly nonexpansive.

Proposition 1.1 [7, Section 6]**.** *If h is a co-coercive function with modulus greater than $\frac{1}{2}$, then the sequence $\{z^k\}$ generated by the iteration (1.4) converges to a solution of (1.4) from any starting point z^0 in Z.*

Since we are going to investigate the asymptotic rate of convergence of the sequences generated by the splitting iteration (1.4) when h satisfies the condition in Proposition 1.1, we make the following blanket assumptions regarding $\mathcal{T}$ and h.

Assumption 1.2 (blanket assumptions)**.**

(a) *The multivalued mapping $\mathcal{T} : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is a maximal monotone operator.*

(b) *The function $h : \mathbb{R}^n \to \mathbb{R}^n$ is co-coercive with modulus $\lambda > \frac{1}{2}$.*

(c) *The solution set $\bar{Z}$ of problem (1.3) is nonempty.*

The rest of the paper is organized as follows. In Section 2, we present some fundamental facts about the splitting iteration (1.4) and give the structure of the solution set $\bar{Z}$ of the problem, together with the definitions of some terminology that will be used repeatedly throughout the paper. In Section 3, we establish the Q- and R-rates of convergence of the sequence generated by the splitting algorithm. At the end of the section, we also point out some significant consequences of these new results. Finally, in Section 4, we specialize to the case of $h = 0$ and derive new results regarding to the proximal point algorithm.

2. Fundamental Properties.

Let

$$y^k := (I - h)(z^k), \quad k = 0, 1, 2, \cdots. \tag{2.1}$$

Then $z^{k+1} = (I + \mathcal{T})^{-1}(y^k)$, and we have

$$y^{k+1} = (I - h)(I + \mathcal{T})^{-1} y^k, \quad k = 0, 1, 2, \cdots. \tag{2.2}$$

The sequence $\{y^k\}$ as a by-product of the iteration will play an important role in our analysis. In addition, we shall make use of the notation

$$P := (I + \mathcal{T})^{-1} \text{ and } Q := I - P = (I + \mathcal{T}^{-1})^{-1}. \tag{2.3}$$

Chen and Rockafellar [4, Proposition 3.4] observed that the algorithmic mapping defined by (1.4) is nonexpansive. The following proposition gives more detailed estimates on the sequences $\{z^k\}$ and $\{y^k\}$. Define

$$w^{k+1} := Q(y^k) \in \mathcal{T}(z^{k+1}), \quad k = 0, 1, 2, \cdots. \tag{2.4}$$

Proposition 2.1 (some useful inequalities)**.** *Under the blanket assumption 1.2, for any $\bar{z} \in \bar{Z}$ there hold*

$$|z^{k+1} - \bar{z}|^2 + |w^{k+1} - \bar{w}|^2 \le |y^k - \bar{y}|^2 \le |z^k - \bar{z}|^2 - \rho|h(z^k) - h(\bar{z})|^2 \tag{2.5}$$

and

$$|y^{k+1} - \bar{y}|^2 + |w^{k+1} - \bar{w}|^2 + \rho|h(z^{k+1}) - h(\bar{z})|^2 \le |y^k - \bar{y}|^2, \tag{2.6}$$

for $k = 0, 1, \cdots$, where $\bar{w} := Q(\bar{y}) \in \mathcal{T}(\bar{z})$, $\bar{y} := (I - h)(\bar{z})$ and $\rho = 2\lambda - 1$.

Proof. Let $y := (I - h)z$ and $y' := (I - h)z'$ for arbitrary z and z' in $\mathbb{R}^n$. Then $y - y' = z - z' - \big(h(z) - h(z')\big)$, and we have

$$\begin{aligned} |y - y'|^2 &= |z - z'|^2 - 2\langle z - z', h(z) - h(z')\rangle + |h(z) - h(z')|^2 \\ &\le |z - z'|^2 - (2\lambda - 1)|h(z) - h(z')|^2. \end{aligned}$$

Let $z = z^k$ and $z' = \bar{z}$, we get the second half of (2.5). The first half of (2.5) follows directly from [19, Proposition 1]. Applying (2.5) to two consecutive k's, we get (2.6). □

Proposition 2.2 (structure of $\bar{Z}$)**.** *The set $h(\bar{Z})$ is a singleton. Denoting the unique element in $h(\bar{Z})$ by $\bar{v}$, we have $-\bar{v} \in \mathcal{T}(\bar{z})$ for all $\bar{z} \in \bar{Z}$ and*

$$\bar{Z} = \mathcal{T}^{-1}(-\bar{v}) \cap h^{-1}(\bar{v}). \tag{2.7}$$

Proof. Let $\bar{z}$ and $\bar{z}'$ be two arbitrary elements in $\bar{Z}$. Then both $\bar{z}$ and $\bar{z}'$ are fixed points of the algorithmic mapping defined by (1.4). Let $z^k = \bar{z}'$ in (2.5). Then $z^{k+1} = \bar{z}'$, and it follows from (2.5) that $h(\bar{z}') = h(\bar{z})$. Now for every $\bar{z} \in \bar{Z}$ we have $0 \in \mathcal{T}(\bar{z}) + h(\bar{z})$, which is equivalent to $-\bar{v} \in \mathcal{T}(\bar{z})$, and (2.7) follows directly. □

Introducing the notation

$$Z_{\mathcal{T}} := \mathcal{T}^{-1}(-\bar{v}) \text{ and } Z_h := h^{-1}(\bar{v}), \tag{2.8}$$

we can write $\bar{Z} = Z_{\mathcal{T}} \cap Z_h$. Note that both $Z_{\mathcal{T}}$ and Z_h are closed convex sets in $\mathbb{R}^n$ because $\mathcal{T}$ and h are maximal monotone. (See e.g. [24].) Let $\mathbb{B}(v, \delta)$ denote the open ball in $\mathbb{R}^n$ with center at v and radius δ.

Definition 2.3 (growth conditions). *The multifunction $\mathcal{T}^{-1}$ is said to satisfy the growth condition (in some neighborhoods of $-\bar{v}$) with the pair of positive constants (r, α) if there exists $\delta > 0$ such that*

$$\forall w \in \mathbb{B}(-\bar{v}, \delta), \quad \forall z \in \mathcal{T}^{-1}(w) \qquad \operatorname{dist}(Z_{\mathcal{T}}, z) \leq \alpha |w - (-\bar{v})|^r. \tag{2.9}$$

Similarly, the multifunction h^{-1} is said to satisfy the growth condition (in some neighborhoods of $\bar{v}$) with the pair of positive constants (s, β) if there exists $\delta > 0$ such that

$$\forall w \in \mathbb{B}(\bar{v}, \delta), \quad \forall z \in h^{-1}(w) \qquad \operatorname{dist}(Z_h, z) \leq \beta |w - \bar{v}|^s. \tag{2.10}$$

For polyhedral $\mathcal{T}$ and h, there exists positive α and β such that these conditions are satisfied with $(1, \alpha)$ and $(1, \beta)$ respectively by Robinson [17]. Actually, when $r = s = 1$, (2.9) and (2.10) reduce to the *locally upper Lipschitz conditions* introduced in [17] with modulus α and β on $\mathcal{T}^{-1}$ and h^{-1} respectively. Hence the growth conditions here may be viewed as a slight generalization of that concept. For instance, (2.9) may be called as a locally upper Lipschitz condition with modulus α and order r on $\mathcal{T}^{-1}$.

The growth condition is certainly much weaker than the strong monotonicity condition. On one hand we have the following proposition.

Proposition 2.3. *If a multifunction $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is strongly monotone with modulus $\eta > 0$, i.e.,*

$$\langle z - z', w - w' \rangle \geq \eta |z - z'|^2 \quad \forall w \in F(z),\ \forall w' \in F(z'), \tag{2.11}$$

then

$$|z - z'| \leq \eta^{-1} |w - w'| \quad \forall z \in F^{-1}(w),\ \forall z' \in F^{-1}(w'), \tag{2.12}$$

or in other words, the inequality in the growth condition for F holds with $(1, \eta^{-1})$ globally everywhere (not only within certain neighborhood of some specific point).

Proof. Note that $w \in F(z)$ and $w' \in F(z')$ are equivalent to $z \in F^{-1}(w)$ and $z' \in F^{-1}(w')$ respectively. Hence (2.12) follows directly from (2.11) and the inequality $|z - z'||w - w'| \geq \langle z - z', w - w' \rangle$. $\square$

On the other hand, there is no lack of examples where the growth condition is satisfied while the strong monotonicity condition is not.

Example 2.1. Consider the function $F : \mathbb{R} \to \mathbb{R}$ defined as

$$F(z) = \begin{cases} \log z, & \text{if } z > 1, \\ 0, & \text{if } -1 \le z \le 1, \\ -\log(-z), & \text{if } z < -1. \end{cases}$$

Obviously F is maximal monotone, but not strongly monotone. With some elementary calculus, it is easy to verify that F is co-coercive with modulus 1 and that within a certan neighborhood of $w = 0$, the growth condition for F^{-1} is satisfied with the pair $(1, 2)$.

Next, we give a multidimensional example of optimization. Recall that the normal cone $N_C(z)$ of a closed convex set C at $z \in C$ is defined by

$$N_C(z) = \{u \in \mathbb{R}^n \mid \langle u, u' - z\rangle \le 0, u' \in C\}. \tag{2.13}$$

Moreover, $N_C(z)$ is the subgradient of the corresponding indicator function δ_C of C at z [18].

Example 2.2. Let $f : \mathbb{R}^2 \to \mathbb{R}^2$ be the function defined as

$$f(z_1, z_2) = \begin{cases} (z_1 \log z_1 - z_1), & \text{if } z_1 > 1, \\ -1, & \text{if } -1 \le z_1 \le 1, \\ \big(-z_1 \log(-z_1) + z_1\big), & \text{if } z_1 < -1. \end{cases}$$

Let S be the closed convex set $S = \{z \in \mathbb{R}^2 \,|\, |z| \le 2\}$. Consider the convex programming problem

$$\min f(z) \text{ subject to } z \in S.$$

Let $\mathcal{T} = N_S$, and let $h = \nabla f$, i.e.,

$$h(z) = \begin{cases} (\log z_1, 0), & \text{if } z_1 > 1, \\ (0, 0), & \text{if } -1 \le z_1 \le 1, \\ \big(-\log(-z_1), 0\big), & \text{if } z_1 < -1. \end{cases}$$

Then the problem of minimization is equivalent to the equation $0 \in \mathcal{T}(z) + h(z)$. Obviously, h in this problem is not strongly monotone. We show, in the following, that all the blanket conditions and the growth conditions for the problem are satisfied. Note that $\bar{Z} = S \cap \{ z \in \mathbb{R}^2 \mid -1 \le z_1 \le 1 \}$ and $h(\bar{Z}) = 0$ in this problem. Hence according to (2.8),

$$Z_{\mathcal{T}} = \mathcal{T}^{-1}(0) = S \text{ and } Z_h = h^{-1}(0) = \{ z \in \mathbb{R}^2 \mid -1 \le z_1 \le 1 \}.$$

It follows from Example 2.1 that h is co-coercive with modulus 1 and that the growth condition (2.10) on h^{-1} is satisfied with the pair $(1, 2)$. Now for the multifunction $\mathcal{T} = N_S$, the set $Z_{\mathcal{T}}$ coincides with the effective domain of $\mathcal{T}$. Hence $\operatorname{dist}(Z_{\mathcal{T}}, z) = 0$ for all $z \in \mathcal{T}^{-1}(w)$, whatever w is. Therefore, the growth condition (2.9) on $\mathcal{T}^{-1}$ is satisfied with any positive pair (r, α).

Recall that the tangent cone $T_C(z)$ of C at $z \in C$ is the polar of the normal cone

$$T_C(z) = N_C^{\circ}(z) = \{u \in \mathbb{R}^n \mid \langle u, u'' \rangle \le 0, u'' \in N_C(z)\}. \tag{2.14}$$

Obviously

$$C \subset T_C(z) + z \quad \forall z \in C. \tag{2.15}$$

Hence for any $\bar{z} \in \bar{Z}$, we have $\bar{Z} \cap (N_{\bar{Z}}(\bar{z}) + \bar{z}) = \{\bar{z}\}$, which is equivalent to $Z_{\mathcal{T}} \cap Z_h \cap (N_{\bar{Z}}(\bar{z}) + \bar{z}) = \{\bar{z}\}$, or

$$(Z_{\mathcal{T}} - \bar{z}) \cap (Z_h - \bar{z}) \cap N_{\bar{Z}}(\bar{z}) = \{0\}. \tag{2.16}$$

A slightly more stringent condition on the sets $Z_{\mathcal{T}}$, Z_h and $N_{\bar{Z}}(\bar{z})$, which will be used in our analysis, is as follows.

Definition 2.4 (regularity condition). *The sets $Z_{\mathcal{T}}$ and Z_h are said to satisfy the regularity condition if*

$$T_{Z_{\mathcal{T}}}(\bar{z}) \cap T_{Z_h}(\bar{z}) \cap N_{\bar{Z}}(\bar{z}) = \{0\} \quad \forall \bar{z} \in \bar{Z}. \tag{2.17}$$

Observe that $T_{\bar{Z}}(\bar{z}) \cap N_{\bar{Z}}(\bar{z}) = \{0\}$. Hence (2.17) will be true if the inclusion $T_{Z_{\mathcal{T}} \cap Z_h}(\bar{z}) \subset T_{Z_{\mathcal{T}}}(\bar{z}) \cap T_{Z_h}(\bar{z})$ holds actually as an equality

$$T_{Z_{\mathcal{T}} \cap Z_h}(\bar{z}) = T_{Z_{\mathcal{T}}}(\bar{z}) \cap T_{Z_h}(\bar{z}). \tag{2.18}$$

For polyhedral $\mathcal{T}$ and h, both $Z_{\mathcal{T}}$ and Z_h are polyhedral convex sets. Therefore $T_{Z_{\mathcal{T}}}(\bar{z})$ and $T_{Z_h}(\bar{z})$ coincide with $Z_{\mathcal{T}} - \bar{z}$ and $Z_h - \bar{z}$ respectively in some neighborhood of $\bar{z}$, and the regularity condition is automatically satisfied. In the general case, a sufficient condition for (2.18) to hold for all $\bar{z} \in \bar{Z}$ is [1a, Table 4.3]

$$0 \in \operatorname{int}(Z_{\mathcal{T}} - Z_h). \tag{2.19}$$

Specifically, the problem in Example 2.2 satisfies (2.19), as could be easily verified.

3. Rates of Convergence for the Splitting Algorithm.

Lemma 3.1. *Let $\{z^k\}$ be an infinite sequence generated by the splitting iteration (1.4) such that $z^k \to \bar{z}$ and $z^k \neq \bar{z}$ for all k. Suppose the regularity condition on $Z_{\mathcal{T}}$ and Z_h is satisfied. Let α^k and β^k be defined by the following equations for all k*

$$\begin{aligned} \operatorname{dist}(T_{Z_{\mathcal{T}}}(\bar{z}), z^k - \bar{z}) &= \alpha^k |z^k - \bar{z}|, \\ \operatorname{dist}(T_{Z_h}(\bar{z}), z^k - \bar{z}) &= \beta^k |z^k - \bar{z}|. \end{aligned} \tag{3.1}$$

Then there exists $\sigma > 0$ such that

$$\liminf_{k\to\infty} \big(\max\{\alpha^k, \beta^k\}\big) \geq \sigma > 0. \tag{3.2}$$

Proof. We prove the lemma by contradiction. Suppose that such a σ does not exist. Then there is a subsequence $\{z^k\}_{\mathcal{K}}$, where $\mathcal{K}$ is an infinite subset of the set of all nonnegative integers, such that

$$\max\{\alpha^k, \beta^k\} \to 0 \ \text{ as } \ k \to \infty, \ k \in \mathcal{K}. \tag{3.3}$$

Define

$$u^k := \frac{z^k - \bar{z}}{|z^k - \bar{z}|}.$$

Then all u^k's are in the compact set $\{u \in \mathbb{R}^n \mid |u| = 1\}$. Therefore there exists some $\bar{u}$ with $|\bar{u}| = 1$ such that

$$u^k \to \bar{u} \ \text{ as } \ k \to \infty, \ k \in \mathcal{K}',$$

where $\mathcal{K}'$ is an infinite subset of $\mathcal{K}$. Observe that the distance function $\mathrm{dist}(C, \cdot)$ is positive homogeneous when C is a cone. Hence (3.3) implies

$$\mathrm{dist}(T_{Z_\mathcal{T}}(\bar{z}), u^k) \to 0 \text{ as } k \to \infty,\ k \in \mathcal{K}',$$
$$\mathrm{dist}(T_{Z_h}(\bar{z}), u^k) \to 0 \text{ as } k \to \infty,\ k \in \mathcal{K}',$$

which in turn implies

$$\mathrm{dist}(T_{Z_\mathcal{T}}(\bar{z}), \bar{u}) = \mathrm{dist}(T_{Z_h}(\bar{z}), \bar{u}) = 0, \tag{3.4}$$

because the distance function $\mathrm{dist}(C, \cdot)$ is continuous. Notice that the tangent cones $T_{Z_\mathcal{T}}(\bar{z})$ and $T_{Z_h}(\bar{z})$ are closed sets in $\mathbb{R}^n$. Hence it follows from (3.4) that

$$\bar{u} \in T_{Z_\mathcal{T}}(\bar{z}) \text{ and } \bar{u} \in T_{Z_h}(\bar{z}). \tag{3.5}$$

Now we claim that $\bar{u} \in N_{\bar{Z}}(\bar{z})$. Indeed, if $\bar{u} \notin N_{\bar{Z}}(\bar{z})$, then there exists some $z' \in \bar{Z}$ such that

$$\langle z' - \bar{z}, \bar{u} \rangle = \varepsilon > 0.$$

Hence for sufficiently large $k \in \mathcal{K}'$, there holds

$$\langle z' - \bar{z}, \frac{z^k - \bar{z}}{|z^k - \bar{z}|} \rangle \geq \tfrac{1}{2}\varepsilon.$$

But $z^k - z' = z^k - \bar{z} + \bar{z} - z'$. Then

$$\begin{aligned} |z^k - z'|^2 &= |\bar{z} - z'|^2 + |z^k - \bar{z}|^2 + 2\langle \bar{z} - z', z^k - \bar{z} \rangle \\ &< |\bar{z} - z'|^2 + |z^k - \bar{z}|^2 - \varepsilon |z^k - \bar{z}| \\ &< |\bar{z} - z'|^2 \text{ for sufficiently large } k \in \mathcal{K}', \end{aligned}$$

which makes $z^k \to \bar{z}$ impossible because the algorithm mapping defined by (1.4) is nonexpansive (by [4, Proposition 3.4] or by (2.5)), and z' is a fixed point of that mapping. Therefore we get

$$\bar{u} \in T_{Z_\mathcal{T}}(\bar{z}) \cap T_{Z_h}(\bar{z}) \cap N_{\bar{Z}}(\bar{z}) \text{ with } \bar{u} \neq 0,$$

which contradicts the regularity condition 2.4. $\square$

Lemma 3.2. *If in addition to the conditions of Lemma 3.1, the growth conditions on $\mathcal{T}^{-1}$ and h^{-1} are satisfied with (r,α) and (s,β) respectively, then for any $t>0$, there hold, for $\bar{y}=(I-h)(\bar{z})$ and sufficiently large k,*

$$\frac{|y^{k+1}-\bar{y}|}{|y^k-\bar{y}|^t} \le \Big(|y^{k+1}-\bar{y}|^{2(1-1/t)} + (\frac{\alpha^{k+1}}{\alpha})^{2/r}|y^{k+1}-\bar{y}|^{2(1/r-1/t)} + \rho(\frac{\beta^{k+1}}{\beta})^{2/s}|y^{k+1}-\bar{y}|^{2(1/s-1/t)}\Big)^{-t/2}, \tag{3.6}$$

and

$$\frac{|z^{k+2}-\bar{z}|}{|z^k-\bar{z}|^t} \le \Big(|z^{k+2}-\bar{z}|^{2(1-1/t)} + (\frac{\alpha^{k+2}}{\alpha})^{2/r}|z^{k+2}-\bar{z}|^{2(1/r-1/t)} + \rho(\frac{\beta^{k+2}}{\beta})^{2/s}|z^{k+2}-\bar{z}|^{2(1/s-1/t)}\Big)^{-t/2}. \tag{3.7}$$

Proof. By Proposition 2.2, the inequalities (2.6) and (2.5) can be written as

$$|y^{k+1}-\bar{y}|^2 + |w^{k+1}-(-\bar{v})|^2 + \rho|h(z^{k+1})-\bar{v}|^2 \le |y^k-\bar{y}|^2, \tag{3.8}$$

$$|z^{k+1}-\bar{z}|^2 + |w^{k+1}-(-\bar{v})|^2 \le |y^k-\bar{y}|^2 \le |z^k-\bar{z}|^2 - \rho|h(z^k)-\bar{v}|^2, \tag{3.9}$$

where $w^{k+1} := Q(y^k) \in \mathcal{T}(z^{k+1})$. The growth condition 2.3, together with (2.15) and (3.1), yields, for sufficiently large k,

$$\begin{aligned}\alpha^k|z^k-\bar{z}| &= \mathrm{dist}(T_{Z_{\mathcal{T}}}(\bar{z})+\bar{z}, z^k) \le \mathrm{dist}(Z_{\mathcal{T}}, z^k) \le \alpha|w^k-(-\bar{v})|^r,\\ \beta^k|z^k-\bar{z}| &= \mathrm{dist}(T_{Z_h}(\bar{z})+\bar{z}, z^k) \le \mathrm{dist}(Z_h, z^k) \le \beta|h(z^k)-\bar{v}|^s,\end{aligned}$$

or equivalently

$$(\alpha^k/\alpha)^{1/r}|z^k-\bar{z}|^{1/r} \le |w^k-(-\bar{v})|, \tag{3.10}$$

$$(\beta^k/\beta)^{1/s}|z^k-\bar{z}|^{1/s} \le |h(z^k)-\bar{v}|. \tag{3.11}$$

Substituting (3.10) and (3.11) in (3.8) and (3.9), and noticing that $|y^k-\bar{y}| \le |z^k-\bar{z}|$ by the second half of (3.9), we get

$$|y^{k+1}-\bar{y}|^2 + (\frac{\alpha^{k+1}}{\alpha})^{2/r}|y^{k+1}-\bar{y}|^{2/r} + \rho(\frac{\beta^{k+1}}{\beta})^{2/s}|y^{k+1}-\bar{y}|^{2/s} \le |y^k-\bar{y}|^2, \tag{3.12}$$

$$|z^{k+1}-\bar{z}|^2+\big(\frac{\alpha^{k+1}}{\alpha}\big)^{2/r}|z^{k+1}-\bar{z}|^{2/r}\le|z^k-\bar{z}|^2-\rho\big(\frac{\beta^k}{\beta}\big)^{2/s}|z^k-\bar{z}|^{2/s}, \quad (3.13)$$

for sufficiently large k. By (3.12), we have

$$\frac{|y^{k+1}-\bar{y}|}{|y^k-\bar{y}|}\le\frac{1}{\Big(1+\big(\frac{\alpha^{k+1}}{\alpha}\big)^{2/r}|y^{k+1}-\bar{y}|^{2/r-2}+\rho\big(\frac{\beta^{k+1}}{\beta}\big)^{2/s}|y^{k+1}-\bar{y}|^{2/s-2}\Big)^{1/2}},$$

from which (3.6) follows. Applying (3.13) to two consecutive k's and noticing that $|z^{k+2}-\bar{z}|\le|z^{k+1}-\bar{z}|$, we have

$$\frac{|z^{k+2}-\bar{z}|}{|z^k-\bar{z}|}\le\frac{1}{\Big(1+\big(\frac{\alpha^{k+2}}{\alpha}\big)^{2/r}|z^{k+2}-\bar{z}|^{2/r-2}+\rho\big(\frac{\beta^{k+2}}{\beta}\big)^{2/s}|z^{k+2}-\bar{z}|^{2/s-2}\Big)^{1/2}},$$

from which (3.7) follows. $\square$

An infinite sequence $\{u^k\}$ with $u^k\to\bar{u}$, $u^k\ne\bar{u}$, is said to converge at least Q-linearly, Q-superlinearly with order $\bar{t}$ or Q-sublinearly if

$$\limsup_{k\to\infty}\frac{|u^{k+1}-\bar{u}|}{|u^k-\bar{u}|}\le\gamma<1,$$

$$\limsup_{k\to\infty}\frac{|u^{k+1}-\bar{u}|}{|u^k-\bar{u}|^t}=0, \text{ for all } t \text{ satisfying } \bar{t}>t>1$$

or

$$\limsup_{k\to\infty}\frac{|u^{k+1}-\bar{u}|}{|u^k-\bar{u}|^t}=0, \text{ for some } t \text{ satisfying } 0<t<1$$

correspondingly. (cf. [12]. The definition of Q-sublinear convergence here is somehow different from the one given in [12].) Similarly, $\{u^k\}$ will be said to converge at least R-linearly, R-superlinearly with order $\bar{t}$ or R-sublinearly if

$$\limsup_{k\to\infty}|u^k-\bar{u}|^{1/k}\le\gamma<1,$$

$$\limsup_{k\to\infty}|u^k-\bar{u}|^{1/t^k}=0 \text{ for all } t \text{ satisfying } \bar{t}>t>1$$

or

$$\limsup_{k\to\infty}|u^k-\bar{u}|^{1/k}=1$$

correspondingly.

Theorem 3.3 (Q-rates on $\{y^k\}$). *Let $\{z^k\}$ be an infinite sequence generated by the splitting iteration (1.4) such that $z^k \to \bar{z}$ and $z^k \neq \bar{z}$ for all k. Suppose both the regularity condition on $Z_{\mathcal{T}}$ and Z_h and the growth conditions on $\mathcal{T}^{-1}$ and h^{-1} with (r, α) and (s, β) respectively are satisfied. Let $\bar{t} = \min\{r, s\}$, and let σ be the positive constant defined in (3.2). Then the sequence $\{y^k\}$ defined by (2.1) converges to $\bar{y} = (I - h)(\bar{z})$ at least Q-linearly with*

$$\limsup_{k\to\infty} \frac{|y^{k+1} - \bar{y}|}{|y^k - \bar{y}|} \leq \left(1 + \min\{(\sigma/\alpha)^{2/r}, \rho(\sigma/\beta)^{2/s}\}\right)^{-1/2}, \tag{3.14}$$

Q-superlinearly with order $\bar{t}$ or Q-sublinearly, according to whether $\bar{t} = 1$, $\bar{t} > 1$ or $0 < \bar{t} < 1$.

Proof. The convergence of $\{y^k\}$ follows directly from Proposition 2.1. If $\bar{t} = 1$, let $t = 1$ in (3.6) and take $\limsup$ as $k \to \infty$. We get (3.14) by Lemma 3.1. Hence $\{y^k\}$ converges at least Q-linearly.

If $\bar{t} > 1$, then for any t satisfying $\bar{t} > t > 1$, take $\limsup$ in (3.6) as $k \to \infty$. We have, by Lemma 3.1,

$$\limsup_{k\to\infty} \frac{|y^{k+1} - \bar{y}|}{|y^k - \bar{y}|^t} = 0. \tag{3.15}$$

Hence $\{y^k\}$ converges at least Q-superlinearly with order $\bar{t}$.

If $\bar{t} < 1$, then for any t satisfying $0 < t < \bar{t}$, take $\limsup$ in (3.6) as $k \to \infty$. We have, by Lemma 3.1,

$$\limsup_{k\to\infty} \frac{|y^{k+1} - \bar{y}|}{|y^k - \bar{y}|^t} = 0. \tag{3.16}$$

Hence $\{y^k\}$ converges at least Q-sublinearly. □

Theorem 3.4 (R-rates on $\{y^k\}$). *Suppose the conditions of Theorem 3.3 are satisfied. Then the sequence $\{y^k\}$ converges to $\bar{z}$ at least R-linearly or R-superlinearly with order $\bar{t}$, according to whether $\bar{t} = 1$ or $\bar{t} > 1$.*

Proof. The conclusions follows from the fact that the Q-linear or Q-superlinear convergence implies the corresponding R-convergence. (See [12].) □

Theorem 3.5 (two-step Q-rates on $\{z^k\}$). *Suppose the conditions of Theorem 3.3 are satisfied. Then the sequence $\{z^k\}$ converges to $\bar{z}$ at least two-step Q-linearly* *with*

$$\limsup_{k\to\infty} \frac{|z^{k+2}-\bar{z}|}{|z^k-\bar{z}|} \le \left(1+\min\{(\sigma/\alpha)^{2/r}, \rho(\sigma/\beta)^{2/s}\}\right)^{-1/2}, \tag{3.17}$$

two-step Q-superlinearly with order $\bar{t}$ or two-step Q-sublinearly, according to whether $\bar{t}=1$, $\bar{t}>1$ or $0<\bar{t}<1$.

Proof. Similar to the proof of Theorem 3.3. Use (3.7) instead of (3.6). $\square$

In the following, we single out an important special case when both $\mathcal{T}$ and h are polyhedral functions in the sense of Robinson [17].

Corollary 3.6 (special case for polyhedral $\mathcal{T}$ and h). *Suppose both $\mathcal{T}$ and h in (1.2) are polyhedral. Then there exist $\alpha>0$, $\beta>0$ and $\sigma>0$ such that for all $z^0\in\mathbb{R}^n$, the sequences $\{z^k\}$ and $\{y^k\}$ defined by (1.4) and (2.1) have the following properties*

(a) *$\{y^k\}$ converges to some $\bar{y}\in(\bar{Z}-\bar{v})$ at least Q-linearly with*

$$\limsup_{k\to\infty} \frac{|y^{k+1}-\bar{y}|}{|y^k-\bar{y}|} \le \left(1+\min\{(\sigma/\alpha)^2, \rho(\sigma/\beta)^2\}\right)^{-1/2} \tag{3.18}$$

(b) *$\{z^k\}$ converges to some $\bar{z}\in\bar{Z}$ at least R-linearly and two-step Q-linearly with*

$$\limsup_{k\to\infty} \frac{|z^{k+2}-\bar{z}|}{|z^k-\bar{z}|} \le \left(1+\min\{(\sigma/\alpha)^2, \rho(\sigma/\beta)^2\}\right)^{-1/2}, \tag{3.19}$$

provided that the iteration does not terminate finitely with some $z^k\in\bar{Z}$.

Proof. For polyhedral $\mathcal{T}$ and h, there exist $\alpha>0$ and $\beta>0$ such that the growth conditions 2.3 on $\mathcal{T}^{-1}$ and h^{-1} are satisfied with $(1,\alpha)$ and $(1,\beta)$ respectively by [17, Corollary]. Observe also that $Z_{\mathcal{T}}$, Z_h and $\bar{Z}$ are all polyhedral convex sets. Hence $T_{Z_{\mathcal{T}}}(\bar{z})$ and $T_{Z_h}(\bar{z})$ coincide with $Z_{\mathcal{T}}-\bar{z}$ and $Z_h-\bar{z}$ respectively in some neighborhood of $\bar{z}$. Therefore the regularity condition 2.4 is also satisfied.

Recall that any convex set C can be partitioned into the collection of relative interiors of all its *faces* [18] and that the tangent cone $T_C(u)$ for all $u\in C$ in the interior of the same face are same [2, Theorem 2.3]. Now there are only finitely many faces for each polyhedral convex set [18]. Hence there are only finitely many different $T_{Z_{\mathcal{T}}}(\bar{z})$'s and $T_{Z_h}(\bar{z})$'s for all $\bar{z}\in\bar{Z}$. Then it follows that there is a uniform

$\sigma > 0$ such that (3.2) in Lemma 3.1 holds for all $\bar{z} \in \bar{Z}$. Therefore we have the corollary by Theorems 3.3, 3.4 and 3.5. $\square$

Whether the blanket assumptions are satisfied, so that the iteration converges in the first place by Gabay's proposition, certainly depends on the particular splitting, as well as on the operator T in the original problem. However, the requirement that the modulus $\lambda > \frac{1}{2}$ in the co-coercivity blanket assumption 1.2(b) does not actually impose severe restrictions to the applications of the algorithm. If h is co-coercive, say, with (positive) modulus $\bar{\lambda} \leq \frac{1}{2}$, we can simply rescale the problem by multiplying the original equation $0 \in T(z)$ with $\bar{\lambda}$. This does not change the solution of the original problem. However, the function $(\bar{\lambda} h)$ in the scaled splitting

$$\bar{\lambda} T = \bar{\lambda} \mathcal{T} + \bar{\lambda} h$$

becomes co-coercive with modulus 1. Iteration (1.4) now takes the form

$$z^{k+1} = (I + \bar{\lambda}\mathcal{T})^{-1}(I - \bar{\lambda} h)(z^k), \quad k = 0, 1, 2, \cdots. \tag{3.20}$$

This amounts to introduce a "step length" in the iteration (1.4). We adopt the unscaled approach in the paper to keep the notation as simple as possible. The interested reader will have no difficulty extending all the results into the scaled form.

Now, a large number of numerical methods for convex programming and variational inequality problems (such as the projection method of Goldstein [8], certain asymmetric projection methods, and decomposition methods) can all be formulated as splitting iterations. With careful rescaling and appropriate splitting, most of them could be put in the framework of forward-backward iteration (1.4) with the co-coercivity blanket assumption being satisfied. (See e.g. [7, Theorem 6.1] and [22, 23].) Hence, when applied to affine variational problems, the corollary ascertains the linear rate of convergence for these algorithms without any further assumptions on the *strict* monotonicity on the function involved. This is something new compared with the previous results (cf. [6, 13, 14, 15]). The strict monotonicity assumption, if any (such as implied in the conditions of [15, Theorem 2.9]), would exclude some important applications of these algorithms, as pointed out by Tseng [22]. According to Corollary 3.6, we can claim, e.g., that the decomposition methods of Tseng [22] for extended linear-quadratic programming have the linear rate of convergence even when the problem is not *fully quadratic* [20, 25].

4. Rates of Convergence for the Proximal Point Algorithm.

In the special case when $h = 0$, the iteration (1.4) reduces to the *proximal point algorithm* [19], and we have $y^k = z^k$ for all k. Now $\bar{v} = h(\bar{z}) = 0$, $Z_h = h^{-1}(\bar{v}) = \mathbb{R}^n$, and $\mathcal{T} = T$, $Z_{\mathcal{T}} = \bar{Z}$. Hence (2.17) in the regularity condition holds automatically. Moreover, (2.10) in the growth condition holds for any positive pair (s, β), and particularly for $s = r$. Therefore we have $\bar{t} = r$ in Theorems 3.3, 3.4 and 3.5, with the right-hand sides of (3.14) and (3.17) being reduced to $\left(1 + (\sigma/\alpha)^{2/r}\right)^{-1/2}$. In the following, we are going to strengthen these results by first proving that actually $\alpha^k \to 1$ in Lemma 3.1 for this special case.

To put the iteration in a more general scheme, we further allow the *proximal constant* $c_k > 0$ to vary with the iteration as Rockafellar did in [19]. Then we get some new results that complement Rockafellar's [19] and Luque's [10] earlier results on the proximal point algorithm.

With the notation

$$P_k := (I + c_k T)^{-1} \tag{4.1}$$

the proximal point iteration can be written as

$$z^{k+1} = P_k(z^k), \quad k = 0, 1, 2, \cdots. \tag{4.2}$$

Lemma 4.1. *Let $\{z^k\}$ be an infinite sequence generated by the proximal point iteration (4.1) such that $z^k \to \bar{z}$ and $z^k \neq \bar{z}$ for all k. Let α^k be defined by the following equation for all k*

$$\operatorname{dist}(T_{\bar{Z}}(\bar{z}), z^k - \bar{z}) = \alpha^k |z^k - \bar{z}|. \tag{4.3}$$

Then $\alpha^k \to 1$ as $k \to \infty$.

Proof. Obviously $\alpha^k \in [0, 1]$ for all k. Now we prove that 1 is the only cluster point of $\{\alpha^k\}$.

Suppose $\{\alpha^k\}$ has another cluster point $\tilde{\sigma} < 1$. Define

$$u^k := \frac{z^k - \bar{z}}{|z^k - \bar{z}|}.$$

Then (similar to the proof of Lemma 3.1,) there is a subsequence $\{\alpha^k\}_{\mathcal{K}}$, where $\mathcal{K}$ is an infinite subset of the set of all nonnegative integers, such that for some $\bar{u} \in \mathbb{R}^n$ with $|\bar{u}| = 1$, there hold

$$u^k \to \bar{u} \text{ and } \operatorname{dist}(T_{\bar{Z}}(\bar{z}), u^k) \to \tilde{\sigma} < 1 \text{ as } k \to \infty, \quad k \in \mathcal{K}. \tag{4.4}$$

Recall that in the proof of Lemma 3.1, we have already shown that such a $\bar{u}$ must be in $N_{\bar{Z}}(\bar{z})$. Therefore

$$\operatorname{dist}(N_{\bar{Z}}(\bar{z}), \bar{u}) = 0. \tag{4.5}$$

Now the tangent cone $T_{\bar{Z}}$ is the polar of the normal cone $N_{\bar{Z}}$. Hence it follows from Moreau decomposition [18, Theorem 31.5] that

$$\operatorname{dist}^2(T_{\bar{Z}}(\bar{z}), \bar{u}) = |\bar{u}|^2 - \operatorname{dist}^2(N_{\bar{Z}}(\bar{z}), \bar{u}) = 1, \tag{4.6}$$

which contradicts (4.4) in view of the continuity of the distance function. $\square$

Lemma 4.2. *If in addition to the conditions of Lemma 4.1, the growth condition on $T^{-1} = \mathcal{T}^{-1}$ is satisfied with (r, α), then for any $t > 0$, there holds, for sufficiently large k,*

$$\frac{|z^{k+1} - \bar{z}|}{|z^k - \bar{z}|^t} \le \Big(|z^{k+1} - \bar{z}|^{2(1-1/t)} + c_k^2\Big(\frac{\alpha^{k+1}}{\alpha}\Big)^{2/r}|z^{k+1} - \bar{z}|^{2(1/r-1/t)}\Big)^{-t/2}. \tag{4.7}$$

Proof. By [19, Proposition 1], we have

$$|z^{k+1} - \bar{z}|^2 + |w^{k+1}|^2 = |z^k - \bar{z}|^2, \tag{4.8}$$

where $w^{k+1} := Q_k(y^k) \in c_k T(z^{k+1})$ with $Q_k = I - P_k$, or equivalently $c_k^{-1} w^{k+1} \in T(z^{k+1})$. The growth condition, together with (2.15) and (4.3), yields, for sufficiently large k,

$$\alpha^k |z^k - \bar{z}| = \operatorname{dist}(T_{Z_{\mathcal{T}}}(\bar{z}) + \bar{z}, z^k) \le \operatorname{dist}(Z_{\mathcal{T}}, z^k) \le \alpha |c_k^{-1} w^k|^r. \tag{4.9}$$

Solving for $|w^k|$ and substituting in (4.8), we get

$$|z^{k+1} - \bar{z}|^2 + c_k^2\Big(\frac{\alpha^{k+1}}{\alpha}\Big)^{2/r}|z^{k+1} - \bar{z}|^{2/r} \le |z^k - \bar{z}|^2 \tag{4.10}$$

for sufficiently large k. Hence

$$\frac{|z^{k+1}-\bar{z}|}{|z^k-\bar{z}|} \le \frac{1}{\left(1+c_k^2\left(\frac{\alpha^{k+1}}{\alpha}\right)^{2/r}|z^{k+1}-\bar{z}|^{2/r-2}\right)^{1/2}}, \tag{4.11}$$

from which (4.7) follows. ◻

Theorem 4.3 (Q-rates on $\{z^k\}$). *Let $\{z^k\}$ be an infinite sequence generated by the proximal point iteration (4.2) such that $z^k \to \bar{z}$ and $z^k \neq \bar{z}$ for all k. Suppose the growth condition on $T^{-1} = \mathcal{T}^{-1}$ is satisfied with (r, α), and c_k is bounded away from 0, i.e.*

$$\liminf_{k\to\infty} c_k =: \bar{c} > 0. \tag{4.12}$$

(a) *If $\bar{c} < +\infty$, then the sequence $\{z^k\}$ converges to $\bar{z}$ at least Q-linearly with*

$$\limsup_{k\to\infty} \frac{|z^{k+1}-\bar{z}|}{|z^k-\bar{z}|} \le \frac{1}{(1+(\bar{c}/\alpha)^2)^{1/2}}, \tag{4.13}$$

Q-superlinearly with order r or Q-sublinearly, according to whether $r = 1$, $r > 1$ or $0 < r < 1$.

(b) *If $\bar{c} = +\infty$, then the sequence $\{z^k\}$ converges to $\bar{z}$ at least Q-superlinearly even when $r = 1$.*

Proof. If $r = 1$, let $t = 1$ in (4.7) and take $\limsup$ as $k \to \infty$. We get (4.13) by Lemma 4.1. Hence $\{z^k\}$ converges at least Q-linearly or Q-superlinearly, according to whether $\bar{c}$ is finite or $+\infty$.

If $r > 1$, then for any t satisfying $r > t > 1$, take $\limsup$ in (4.7) as $k \to \infty$. We have, by Lemma 4.1,

$$\limsup_{k\to\infty} \frac{|z^{k+1}-\bar{z}|}{|z^k-\bar{z}|^t} = 0. \tag{4.14}$$

Hence $\{z^k\}$ converges at least with Q-order r.

If $r < 1$, then for any t satisfying $r < t < 1$, take $\limsup$ in (4.7) as $k \to \infty$. We have, by Lemma 3.1,

$$\limsup_{k\to\infty} \frac{|z^{k+1}-\bar{z}|}{|z^k-\bar{z}|^t} = 0. \tag{4.15}$$

Hence $\{z^k\}$ converges at least Q-sublinearly. ◻

When T is a polyhedral function, the growth condition is satisfied with some $(1, \alpha)$ by [17, Corollary]. Therefore we have the following corollary as a special case of Theorem 4.3.

Corollary 4.4 (special case for polyhedral T). *Let $\{z^k\}$ be an infinite sequence generated by the proximal point iteration (4.2) such that $z^k \to \bar{z}$ and $z^k \neq \bar{z}$ for all k. Suppose T is a polyhedral function and c_k is bounded away from 0 with* $\liminf_{k\to\infty} c_k =: \bar{c} > 0$.

(a) *If $\bar{c} < +\infty$, then the sequence $\{z^k\}$ converges to $\bar{z}$ at least Q-linearly and (4.13) holds.*

(b) *If $\bar{c} = +\infty$, then the sequence $\{z^k\}$ converges to $\bar{z}$ at least Q-superlinearly.*

Hence we conclude that when the proximal point algorithm is applied to the extended linear-quadratic programming problems [20, 21], the sequence generated by iteration (4.2) converges Q-linearly (or Q-superlinearly if $\bar{c} = +\infty$) even when the solution of the problem is not unique. This result is an improvement of earlier result of Rockafellar and Wets [21, Theorem 6] in the special case when the exact proximal point iterations is implemented.

The result in this section on the proximal point algorithm differs from that of Rockafellar [19] in the aspect that we here do not require the solution set $\bar{Z}$ to be a singleton. It also differs from that of Luque [10] in the aspect that conclusions here are on the rates of convergence of $\{z^k\}$ itself to the limit point $\bar{z}$, instead on the rate of convergence of $\{\text{dist}(\bar{Z}, z^k)\}$ to 0. But we here need the exact proximal point iteration being implemented, while [19] and [10] (in most cases) allow for some "inexact" proximal point iterations. Under certain additional assumptions, the results in this paper can also be extended to the inexact case. These topics will be treated elsewhere.

Acknowledgments. The author thanks Professor J.-S. Pang for his very helpful comments on an earlier version of the paper.

REFERENCES.

[1a] J.-P. Aubin and H. Frankowska, *Set-Valued Analysis*, Birkhäuser, Boston, 1990.

[1b] R. E. Bruck Jr., "An iterative solution of a variational inequality for certain monotone operators in Hilbert space," Bulletin of the American Mathematical Society 81 (1975), pp. 890–892.

[2] J. V. Burke and J. J. Moré, "On the identification of active constraints," SIAM J. Numer. Anal. 25 (1988), pp. 1197–1211.

[3] H.-G. Chen and R. T. Rockafellar, "Application of splitting algorithm to optimal control and extended linear-quadratic programming," technical report, Department of Applied Math, University of Washington, August 1990.

[4] H.-G. Chen and R. T. Rockafellar, "Convergence and structure of forward-backward splitting methods," technical report, Department of Applied Math, University of Washington, October 1990.

[5] H.-G. Chen and R. T. Rockafellar, "Extended forward-backward splitting methods and convergence," technical report, Department of Applied Math, University of Washington, November 1990.

[6] G. H.-G. Chen and R. T. Rockafellar, "Forward-backward splitting methods in Lagrangian optimization," technical report, Department of Applied Math, University of Washington, May 1992.

[7] D. Gabay, *Application of the Method of Multipliers to Variational Inequalities*, in: *Augmented Lagrangian Methods: Applications to the Numerical Solutions of Boundary-Value Problems*, M. Fortin and R. Glowinski, eds., North-Holland 1983.

[8] A. A. Goldstein, "Convex programming in Hilbert space," Bull. Amer. Math. Soc. 70 (1964), pp. 709–710.

[9] B. Lemaire, "Coupling optimization methods and variational convergence," in: *Trends in Mathematical Optimization*, K.-H. Hoffman, J.-B. Hiriart-Urruty, J. Zowe and C. Lemarechal, eds., Birkhäuser, Basel, 1988, pp. 163–179.

[10] F. J. Luque, "Asymptotic convergence analysis of the proximal point algorithm," SIAM J. Control Opt. 22 (1984), pp. 277–293.

[11] G. J. Minty, "Monotone (nonlinear) operators in Hilbert space", Duke Math. J. 29(1962), pp. 341–346.

[12] J. M. Ortega and W. C. Rheinboldt, *Iterative Solution of Nonlinear Equations in Several Variables*, Academic Press, New York 1970.

[13] J.-S. Pang, "Necessary and sufficient conditions for the convergence of iterative methods for the linear complementarity problem," J. of Optimization Theory and Applications 42 (1984), pp. 1–17.

[14] J.-S. Pang, "Asymmetric variational inequality problems over product sets: Applications and iterative methods," Math. Programming 31 (1985), pp. 206–219.

[15] J.-S. Pang and D. Chan, "Iterative methods for variational and complementarity problems," Math. Programming 24 (1982), pp. 284–313.

[16] G. B. Passty, "Ergodic convergence to a zero of the sum of monotone operators in Hilbert space," J. of Mathematical Analysis and Applications 72 (1979), pp. 383–390.

[17] S. M. Robinson, "Some continuity properties of polyhedral multifunctions," Math. Programming Studies 14 (1981), pp. 206–214.

[18] R. T. Rockafellar, *Convex Analysis*, Princeton Univ. Press, Princeton, N.J., 1970.

[19] R. T. Rockafellar, "Monotone operators and the proximal point algorithm," SIAM J. Control Opt. 14 (1976), pp. 877–898.

[20] R. T. Rockafellar, "Computational schemes for solving large-scale problems in extended linear-quadratic programming," Math. Programming 48 (1990), pp. 447–474.

[21] R. T. Rockafellar and R. J-B Wets, "A Lagrangian finite generation technique for solving linear-quadratic problems in stochastic programming," Math. Programming Studies 28 (1986), pp. 63–93.

[22] P. Tseng, "Further applications of a splitting algorithm to decomposition in variational inequalities and convex programming," Math. Programming 48 (1990), pp. 249–263.

[23] P. Tseng, "Applications of a splitting algorithm to decomposition in convex

programming and variational inequalities," SIAM J. Control Opt. 29 (1991), pp. 119–138.

[24] E. Zeidler, *Nonlinear Functional Analysis and its Applications*, Springer-Verlag, New York 1990.

[25] C. Zhu and R. T. Rockafellar, "Primal-dual projected gradient algorithms for extended linear-quadratic programming," SIAM J. Opt. 3 (1993), pp. 751–783.