Chapter 4

FORWARD-BACKWARD SPLITTING METHODS AND THEIR CONVERGENCE

From the structure of the extended linear-quadratic programming problems in Chapter 3, we know that the most suitable splitting methods for those problems are forward-backward splitting types. In this chapter we analyze a general forward-backward method and its convergence structure.

4.1. Forward-Backward Splitting Methods

Returning to the general case of $T : \mathcal{H} \rightrightarrows \mathcal{H}$ and a splitting $T = T_1 + T_2$, where T_1 and T_2 are maximal monotone, we define that for any $\lambda_k > 0$ the *forward-backward operator* G_k as

$$G_k = (I + \lambda_k T_2)^{-1}(I - \lambda_k T_1). \tag{4.1.1}$$

Proposition 2.10.1 says that z is a zero point of T if and only if z is a fixed point of G_k for $k \geq 0$. That is, $z = G_k(z)$.

From Section 2.10, we know that the scheme (2.10.1) is just $z^{k+1} \in G_k(z^k)$. In this chapter, we shall consider slightly different general case from the scheme in Section 2.10. Let us define a general forward-backward splitting algorithm for this chapter first.

Definition 4.1.1. *Let $\mathcal{H}$ be a real Hilbert space and T_1, T_2 be maximal monotone operators on $\mathcal{H}$ such that $T = T_1 + T_2$ is maximal. For a sequence $\{\lambda_k\}_{k=0}^{\infty}$ of positive scalars, a sequence $\{z^k\}_{k=0}^{\infty} \subset \mathcal{H}$ is said to be generated by the forward-backward splitting (FBS) algorithm if for any initial z^0 there are $\varepsilon_k \geq 0$ such that*

$$z^{k+1} \in G_k(z^k) + \varepsilon_k \mathbb{B}, \qquad \forall k \geq 0, \tag{4.1.2}$$

and

$$\sum_{k=0}^{\infty} \varepsilon_k < \infty. \tag{4.1.3}$$

Where G_k is given by (4.1.1) and $\mathbb{B}$ is the closed unit ball (constant operator) in $\mathcal{H}$:

$$\mathbb{B} = \{z \in \mathcal{H} \,|\, \langle z, z\rangle \leq 1\}. \tag{4.1.4}$$

When $\varepsilon_k \equiv 0$ in (4.1.2), the above FBS algorithm is just the BFBS algorithm in Section 2.10. From the computational opinion, we can only get an approximation for the element of $G_k(z^k)$ in practice. So we need a looser assumption than that in (2.10.1). This is the motivation for (4.1.2).

From (4.1.2), it is obvious that we can obtain a z^{k+1} if and only if $G_k(z^k) \neq \emptyset$ for $k = 0, 1, \ldots$ which means that (4.1.2) is logical. Let us see what conditions can guarantee $G_k(z^k) \neq \emptyset$.

We have already seen that the proximal point operator $(I+\lambda_k T_2)^{-1}$ is single-valued and has full domain with T_2 maximal monotone. If $G_k(z^k) \neq \emptyset$, then from (4.1.2)

$$z^{k+1} \in \mathrm{rge}((I + \lambda_k T_2)^{-1}) + \varepsilon_k \mathbb{B} \quad \text{for} \quad k = 0, 1, \ldots. \tag{4.1.5}$$

Proposition 4.1.1. *$G_k(z) \neq \emptyset$ if and only if $z \in \mathrm{dom}(T_1)$ with T_2 maximal monotone. That is,*

$$\mathrm{dom}(G_k) = \mathrm{dom}(T_1).$$

Hence, a necessary condition for (4.1.2) being logical is that $\{z^k\} \subset \mathrm{dom}(T_1)$.

Proof. Since $(I+\lambda_k T_2)^{-1}$ is single-valued and has full domain, so $G_k(z) \neq \emptyset$ is equivalent to $(I - \lambda_k T_1)(z) \neq \emptyset$ which is again the same as $T_1(z) \neq \emptyset$. That is, $z \in \mathrm{dom}(T_1)$. □

Proposition 4.1.2. *The range of $(I + \lambda_k T_2)^{-1}$ is just the domain of T_2. That is,*

$$\mathrm{rge}((I + \lambda_k T_2)^{-1}) = \mathrm{dom}(T_2). \tag{4.1.6}$$

Proof.

$$\begin{aligned}
w \in \mathrm{dom}(T_2) &\Longleftrightarrow T_2(w) \neq \emptyset, \\
&\Longleftrightarrow (I + \lambda_k T_2)(w) \neq \emptyset, \\
&\Longleftrightarrow \exists z \quad \text{such that} \quad z \in (I + \lambda_k T_2)(w), \\
&\Longleftrightarrow \exists z \quad \text{such that} \quad w = (I + \lambda_k T_2)^{-1}(z), \\
&\Longleftrightarrow w \in \mathrm{rge}((I + \lambda_k T_2)^{-1}).
\end{aligned}$$

□

Thus, if for any $z^0 \in \mathcal{H}$ (4.1.2) makes sense, then T_1 should have full domain, i.e., $\mathrm{dom}(T_1) = \mathcal{H}$. In general, we have the following.

Corollary 4.1.3. *If (4.1.2) is logical for given $z^0 \in \mathrm{dom}(T_1)$, then from (4.1.5) and the previous propositions, we see*

$$z^k \in \left[\mathrm{dom}(T_2) + \varepsilon_{k-1}\,\mathbb{B}\right] \bigcap \mathrm{dom}(T_1) \quad \text{for} \quad k \geq 1, \tag{4.1.7}$$

$$\{z^k\}_{k=1}^{\infty} \subset \left[\mathrm{dom}(T_2) + \bigcup_{k=0}^{\infty} \varepsilon_k\,\mathbb{B}\right] \bigcap \mathrm{dom}(T_1). \tag{4.1.8}$$

Proposition 4.1.4. *A sufficient condition for (4.1.2) being logical with $z^0 \in \mathrm{dom}(T_1)$ is*

$$\left[cl[\mathrm{dom}(T_2)] + \bigcup_{k=0}^{\infty} \varepsilon_k\,\mathbb{B}\right] \subset \mathrm{dom}(T_1). \tag{4.1.9}$$

When $\mathrm{dom}(T_1) = \mathcal{H}$, (4.1.9) is automatically true. Then, (4.1.2) is always logical at this case.

Proof. If we can guarantee $G_k(z^k) \neq \emptyset$ for $k \geq 1$, then (4.1.2) will be logical, but this is true from (4.1.8), (4.1.9), the propositions 4.1.1 and 4.1.2. □

Fortunately, T_1 always has full domain of the problems under Consideration in this dissertation. In the following discussion, we suppose that (4.1.9) is always true.

Now, let us consider a special case with T_1 single-valued in (4.1.2).

Proposition 4.1.5. *If T_1 is single-valued on $\mathrm{dom}(T_1)$, then G_k is also single-valued on $\mathrm{dom}(T_1)$, for all $k \geq 0$, and (4.1.2) is equivalent to*

$$\|z^{k+1} - G_k(z^k)\| \leq \varepsilon_k, \forall k \geq 0. \tag{4.1.10}$$

Conversely, if all G_k are single-valued, T_1 is single-valued.

Proof. Since $(I + \lambda_k T_2)^{-1}$ is single-valued and has full domain, all G_k are single-valued if and only if, all $I - \lambda_k T_1$ are single-valued, which is equivalent that T_1 is single-valued. Thus, (4.1.2) becomes

$$z^{k+1} - G_k(z^k) \in \varepsilon_k\,\mathbb{B} \qquad \forall k \geq 0$$

which is the same as (4.1.10). □

(4.1.10) and (4.1.3) are similar to the approximate calculation (2.5.4) of Rockafellar (1976) to the proximal point step $P_k(z^k)$.

Before we proceed to the convergence of the FBS algorithm, please note a couple of useful lemmas for the proofs of some convergence theorems.

Lemma 4.1.6. *If T_1 is single-valued and continuous, then G_k is also single-valued and continuous for every $k \geq 0$.*

Proof. G_k is single-valued from Proposition 4.1.5. Now, since T_2 is maximal monotone, we have the following for any $z, z' \in \mathrm{dom}(T_1)$ such that

$$\begin{aligned} \|G_k(z) - G_k(z')\| &\leq \|z - z' - \lambda_k[T_1(z) - T_1(z')]\| \\ &\leq \|z - z'\| + \lambda_k \|T_1(z) - T_1(z')\|. \end{aligned} \tag{4.1.11}$$

This completes the proof. □

Lemma 4.1.7. *Let $\{z^k\}$ be a sequence generated by the FBS algorithm. If $\{z^k\}$ converges to a point $\bar{z}$, then $\bar{z} \in \mathrm{cl}(\mathrm{dom}(T_2))$.*

Proof. Since $\{z^k\}$ is generated by the FBS algorithm, we have

$$z^k \in \mathrm{dom}(T_2) + \varepsilon_k \mathbb{B}.$$

Therefore

$$\bar{z} = \lim_{k \to \infty} z^k \in \lim_{k \to \infty} \left[\mathrm{dom}(T_2) + \varepsilon_k \mathbb{B} \right] = \mathrm{cl}(\mathrm{dom}(T_2)).$$

□

Note: Even if T_2 is maximal monotone, $\mathrm{dom}(T_2)$ may still not be closed.

Example 4.1.1. Let $\mathcal{H} = \mathbb{R}$ and

$$T(x) = \begin{cases} \tan(x), & \text{if } -\frac{1}{2}\pi < x < \frac{1}{2}\pi \\ \infty, & \text{if } \frac{1}{2}\pi \leq x \\ -\infty, & \text{otherwise.} \end{cases}$$

Then T is maximal monotone, but $\mathrm{dom}(T) = (-\frac{1}{2}\pi, \frac{1}{2}\pi)$ is open.

Lemma 4.1.8. *Consider two sequences $\{z^k\}$ and $\{y^k\}$ in $\mathcal{H}$.*

(i) $\{z^k\}$ weakly converges to $\bar{z}$ and $\{y^k\}$ strongly converges to $\bar{y}$.

(ii) $\{z^k\}$ strongly converges to $\bar{z}$ and $\{y^k\}$ weakly converges to $\bar{y}$.

(a) If either Condition (i) or Condition (ii) is satisfied, then

$$\lim_{k\to\infty} \langle z^k, y^k \rangle = \langle \bar{z}, \bar{y} \rangle. \tag{4.1.12}$$

(b) Furthermore, if T is maximal monotone and $y^k \in T(z^k)$ as well, then $\bar{y} \in T(\bar{z})$.

Proof. Consider Condition (i) being satisfied here. Since

$$\langle z^k, y^k \rangle = \langle z^k, \bar{y} \rangle + \langle z^k, y^k - \bar{y} \rangle, \tag{4.1.13}$$

then from Condition (i)

$$\lim_{k\to\infty} \langle z^k, \bar{y} \rangle = \langle \bar{z}, \bar{y} \rangle, \qquad \lim_{k\to\infty} \|y^k - \bar{y}\| = 0 \tag{4.1.14}$$

and $\{\|z^k\|\}$ is bounded. However,

$$|\langle z^k, y^k - \bar{y} \rangle| \leq \|z^k\| \|y^k - \bar{y}\|. \tag{4.1.15}$$

Therefore, (4.1.13), (4.1.14), and (4.1.15) imply (4.1.12).

Now, if T is maximal monotone and $y^k \in T(z^k)$, then for any $z \in \mathcal{H}$ and $y \in T(z)$ we have

$$0 \leq \langle z^k - z, y^k - y \rangle \quad \text{for any} \quad k \geq 0. \tag{4.1.16}$$

Condition (i) implies that $\{z^k - z\}$ weakly converges to $\bar{z} - z$ and $\{y^k - y\}$ strongly converges to $\bar{y} - y$. Hence, (4.1.12) and (4.1.16) give that

$$0 \leq \lim_{k\to\infty} \langle z^k - z, y^k - y \rangle = \langle \bar{z} - z, \bar{y} - y \rangle.$$

But T is maximal, so $\bar{y} \in T(\bar{z})$.

Similar proof can be given for Condition (ii). $\square$

To conclude this section, we may propose the following definition:

Definition 4.1.2. *The convergence of a sequence* $\{z^k\}$ *to a point* z^* *is said to be Q-linear, with ratio,* θ *if there is an index* k_0 *with*

$$\|z^{k+1} - z^*\| \leq \theta \|z^k - z^*\| \quad \textit{for all} \quad k \geq k_0, \quad \textit{with} \quad \theta \in [0, 1). \tag{4.1.17}$$

We may say the convergence is Q-linear, with global ratio θ*, if this holds for* $k_0 = 0$.

4.2. The Global Convergence and Rate of the Basic FBS Algorithm

We are familiar with some convergence theorems for the BFBS algorithm in Section 2.10. Those theorems are for global convergence. It is easy to see that all convergence conclusions for the BFBS algorithm in Section 2.10 are decided by the conditions on T_1. In the dynamic models of Chapter 3 these conclusions may not be applicable.

For example, the splitting strategy I and the splitting strategy II in Section 3.4 and Section 3.5 are attractive due to their decomposable properties and simple structure. However, no known theorem in Section 2.10 can guarantee the convergence of $\{z^k\}$ from (4.1.2) with T_1 and T_2, if T_1 is not a zero operator, even if $\varepsilon_k \equiv 0$ for all $k \geq 0$. The reason is neither T_1 nor T_1^{-1} is strongly monotone in those cases due to the following (ref. Chen and Rockafellar 1992):

$$\langle z' - z'', T_1(z') - T_1(z'') \rangle \equiv 0 \qquad \forall z', z'' \in \mathcal{H}. \tag{4.2.1}$$

Therefore, we need to prove some new theorems which specially fit our situation. In addition, we need to extend new and known theorems to the form in Section 4.1. We shall deal with the general case in the next section. This section will concentrate mainly on the special case of $\varepsilon_k \equiv 0$. That is, the BFBS algorithm will be considered here.

Now we shall give several convergence theorems that depend strongly on T_2. Before considering the theorems, we introduce a property from functional analysis.

Lemma 4.2.1. *If an operator $T : \mathcal{H} \rightrightarrows \mathcal{H}$ is strongly monotone, then $T^{-1}(0)$ is nonempty and consists of exact one point in $\mathcal{H}$.*

Proof. Let $\alpha > 0$ be the strongly monotone modulus of T. Consider

$$z^{k+1} = (I+T)^{-1}(z^k) \quad \text{for any} \quad z^0 \in \mathcal{H}.$$

So, there exists a $w^{k+1} \in T(z^{k+1})$ such that

$$z^{k+1} + w^{k+1} = z^k \quad \text{for all} \quad k \geq 0. \tag{4.2.1}$$

In (4.2.1), the equality for k subtracting the equality for $k-1$ gives

$$z^{k+1} - z^k + w^{k+1} - w^k = z^k - z^{k-1} \quad \text{for} \quad k \geq 1.$$

Taking the norm of both sides implies

$$\begin{aligned}
\|z^k - z^{k-1}\|^2 &= \|z^{k+1} - z^k + w^{k+1} - w^k\|^2 \\
&= \|z^{k+1} - z^k\|^2 + 2\langle z^{k+1} - z^k, w^{k+1} - w^k\rangle + \|w^{k+1} - w^k\|^2 \\
&\geq \|z^{k+1} - z^k\|^2 + 2\alpha\|z^{k+1} - z^k\|^2 + \alpha^2\|z^{k+1} - z^k\|^2 \\
&= (1+\alpha)^2\|z^{k+1} - z^k\|^2,
\end{aligned}$$

where the inequality comes from the assumption and Proposition 2.1.3. Thus,

$$\|z^{k+1} - z^k\| \leq \frac{\|z^k - z^{k-1}\|}{1+\alpha} \leq \frac{\|z^1 - z^0\|}{(1+\alpha)^k}.$$

Hence, for any $k > 0$ and $m \geq 0$, one has the following:

$$\begin{aligned}
\|z^{k+m} - z^m\| &\leq \sum_{i=m}^{m+k-1} \|z^{i+1} - z^i\| \\
&\leq \sum_{i=m}^{m+k-1} (1+\alpha)^{-i}\|z^1 - z^0\| \\
&\leq (1+\alpha)^{-m}\frac{1+\alpha}{\alpha}\|z^1 - z^0\|.
\end{aligned} \tag{4.2.2}$$

Therefore, for any $k > 0$, one obtains from (4.2.2)

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

That is, the sequence $\{z^k\}$ strongly converges to a point z^∞ in $\mathcal{H}$ since $\mathcal{H}$ is a Hilbert space. Therefore,

$$z^\infty = \lim_{k\to\infty} z^{k+1} = (I+T)^{-1}(\lim_{k\to\infty} z^k) = (I+T)^{-1}(z^\infty).$$

This again implies that $z^\infty \in T^{-1}(0)$. Thus, $T^{-1}(0) \neq \emptyset$.

Now let us prove the uniqueness of $T^{-1}(0)$.

Suppose $w_1 \in T^{-1}(0)$ and $w_2 \in T^{-1}(0)$. Then

$$0 = \langle w_1 - w_2, 0 - 0\rangle \geq \alpha\|w_1 - w_2\|^2.$$

Hence, we must have $w_1 = w_2$. □

Before proving the following theorems, we shall review two concepts regarding convergence rate.

Definition 4.2.1. *The convergence of a sequence $\{z^k\}_{k=0}^\infty$ to a point $\bar{z}$ is said to be linear with ratio θ, if there is an index $\bar{k}$ and a constant $C_{\bar{k}}$ with*

$$\|z^k - \bar{z}\| \leq \theta^{k-\bar{k}} C_{\bar{k}} \quad \textit{for all} \quad k \geq \bar{k}, \qquad \textit{with} \quad \theta \in [0,1).$$

Moreover, the above convergence is said to be Q-linear with ratio θ, if there is an index $\bar{k}$ with

$$\|z^{k+1} - \bar{z}\| \leq \theta\|z^k - \bar{z}\| \quad \textit{for all} \quad k \geq \bar{k}, \qquad \textit{with} \quad \theta \in [0,1).$$

We may say that the convergence is linear or Q-linear with global ratio θ, if the related inequality holds for $\bar{k} = 0$.

It is easily seen that Q-linear convergence is stronger than linear convergence. In other words, Q-linear convergence sequences must be linear convergence sequences. The other direction is not necessarily true.

Theorem 4.2.2. *If T_1 is Lipschitz continuous, with constant M, and T_2 is strongly monotone, with modulus α, then $\{z^k\}$, from (2.10.1), strongly converges* *to the unique zero point of* $T = T_1 + T_2$ *for any starting point* $z^0 \in \mathrm{dom}(T_1)$*, when* $\{\lambda_k\} \subset \mathbb{R}$ *satisfies for any* λ^- *and* λ^+ *that*

$$0 < \lambda^- \leq \lambda_k \leq \lambda^+ < \lambda^* := \begin{cases} \frac{2\alpha}{M^2-\alpha^2}, & \textit{for } M > \alpha; \\ \infty, & \textit{for } M \leq \alpha \end{cases} \tag{4.2.3}$$

and the convergence rate will be Q-linear with the global ratio θ*. Where*

$$\theta^2 = \max_{\lambda^- \leq \lambda \leq \lambda^+} \Theta(\lambda) = \max\{\Theta(\lambda^-), \Theta(\lambda^+)\} \tag{4.2.4}$$

and

$$\Theta(\lambda) = \frac{1 + M^2\lambda^2}{(1+\alpha\lambda)^2} \in (0,1) \quad \textit{when} \quad 0 < \lambda < \lambda^*. \tag{4.2.5}$$

Proof. Since T_2 is strongly monotone, $T = T_1 + T_2$ is also strongly monotone. Hence, there exists a z^*, such that $T^{-1}(0) = \{z^*\}$, from Lemma 4.2.1. Now, T_1 is single-valued from its Lipschitz continuity. Let $y^* = -T_1(z^*)$. Then, $y^* \in T_2(z^*)$. So,

$$z^* + \lambda_k y^* = z^* - \lambda_k T_1(z^*) \quad \text{for all} \quad k \geq 0. \tag{4.2.6}$$

Let $y^{k+1} = \frac{1}{\lambda_k}(z^k - z^{k+1}) - T_1(z^k)$. Then, (2.10.1) implies $y^{k+1} \in T_2(z^{k+1})$, and

$$z^{k+1} + \lambda_k y^{k+1} = z^k - \lambda_k T_1(z^k). \tag{4.2.7}$$

For fixed λ_k in (4.2.6), subtracting (4.2.6) from (4.2.7) gives

$$z^{k+1} - z^* + \lambda_k(y^{k+1} - y^*) = z^k - z^* - \lambda_k(T_1(z^k) - T_1(z^*)) \tag{4.2.8}$$

Taking square norm to both sides of (4.2.8) implies

$$\begin{aligned} &\|z^{k+1} - z^*\|^2 + \lambda_k^2\|y^{k+1} - y^*\|^2 + 2\lambda_k\langle z^{k+1} - z^*, y^{k+1} - y^*\rangle = \\ &\|z^k - z^*\|^2 + \lambda_k^2\|T_1(z^k) - T_1(z^*)\|^2 - 2\lambda_k\langle z^k - z^*, T_1(z^k) - T_1(z^*)\rangle. \end{aligned} \tag{4.2.9}$$

Now, since T_1 is Lipschitz continuous and monotone, and T_2 is strongly monotone, we can deduce from (4.2.9) and Proposition 2.1.3 that

$$(1 + \alpha\lambda_k)^2\|z^{k+1} - z^*\|^2 \leq (1 + M^2\lambda_k^2)\|z^k - z^*\|^2$$

That is

$$\|z^{k+1} - z^*\|^2 \leq \frac{1 + M^2\lambda_k^2}{(1 + \alpha\lambda_k)^2}\|z^k - z^*\|^2 = \Theta(\lambda_k)\|z^k - z^*\|^2 \tag{4.2.10}$$

When (4.2.3) is satisfied, we have

$$\sup_{k\geq 0}\{\frac{1 + M^2\lambda_k^2}{(1 + \alpha\lambda_k)^2}\} \leq \theta^2 < 1$$

Therefore, $z^k \to z^*$ strongly as $k \to \infty$, from (4.2.10), and the rate is Q-linear with the global ratio θ. □

Significance of Theorem 4.2.2 is: it does not require T_1 or the inverse of T_1 to be strongly monotone and hence is applicable to the saddle-point problems of extended linear-quadratic programming in Chapter 3, Chapter 7, Chapter 8, and Chapter 9, with keeping the good parallel computing structure.

Theorem 4.2.3. *If T_1 is Lipschitz continuous, with constant M, and T_1, T_2 are strongly monotone, with moduli α_1 and α_2, respectively, then $\{z^k\}$, from (2.10.1), strongly converges to the unique zero point of $T_1 + T_2$ for any starting point $z^0 \in \mathrm{dom}(T_1)$, if for any given λ^- and λ^+,*

$$0 < \lambda^- \leq \lambda_k \leq \lambda^+ < \lambda^* := \begin{cases} \frac{2(\alpha_1+\alpha_2)}{M^2-\alpha_2^2}, & \textit{for } M > \alpha_2; \\ \infty, & \textit{for } M \leq \alpha_2 \end{cases} \tag{4.2.11}$$

and the convergence rate is Q-linear with global ratio θ. Where

$$\theta^2 = \max_{\lambda^-\leq\lambda\leq\lambda^+} \Theta(\lambda) = \max\{\Theta(\lambda^-), \Theta(\lambda^+)\} \tag{4.2.12}$$

and

$$\Theta(\lambda) = \frac{1 - 2\alpha_1\lambda + M^2\lambda^2}{(1 + \alpha_2\lambda)^2} \in (0, 1) \quad \textit{when} \quad 0 < \lambda < \lambda^*. \tag{4.2.13}$$

Proof. Similar to the proofs of Theorem 4.2.2, we can deduce from (4.2.9) that

$$\|z^{k+1} - z^*\|^2 \leq \frac{1 + M^2\lambda_k^2 - 2\lambda_k\alpha_1}{(1 + \alpha_2\lambda_k)^2}\|z^k - z^*\|^2 = \Theta(\lambda_k)\|z^k - z^*\|^2. \tag{4.2.14}$$

When (4.2.11) is satisfied we obtain

$$\sup_{k\geq 0}\{\Theta(\lambda_k)\} \leq \theta^2 < 1.$$

Therefore, $z^k \to z^*$ strongly and Q-linearly as $k \to \infty$, from (4.2.14). □

In general Hilbert spaces, Part (b) of Gabay's Theorem can not be changed to converge *strongly*. This is due to O.Guler's work (1989). He found a counter example in which the sequence from the FBS method converges weakly, but does not converge strongly, for the special case of $T_1 = 0$ and $T_2 = T$, where $\beta = \infty$.

Before proceeding to the next theorem, consider a lemma related to a condition such that $T^{-1}(0) \neq \emptyset$.

Lemma 4.2.4. *Suppose T, T_1, and T_2 are maximal monotone such that $T = T_1 + T_2$. If $\mathrm{cl}(\mathrm{dom}(T_2))$ is a compact set in $\mathcal{H}$, and T_1 is single-valued and continuous, then $T^{-1}(0) \neq \emptyset$ over $\mathrm{dom}(T_2)$.*

Proof. Given any $\lambda_k > 0$, consider the G_k defined by (4.1.1). Lemma 4.1.6 tells us that G_k is continuous. Proposition 4.1.1 shows that $\mathrm{dom}(G_k) = \mathrm{dom}(T_1) \supset \mathrm{cl}(\mathrm{dom}(T_2))$. However, $\mathrm{rge}(G_k) = \mathrm{dom}(T_2)$. Therefore, there is a point $\bar{z} \in \mathrm{dom}(T_2)$, such that $\bar{z} = G_k(\bar{z})$ for Schauder Theorem of general functional analysis theory, through the compactness of $\mathrm{cl}(\mathrm{dom}(T_2))$. That is, $0 \in T(\bar{z})$. □

We improve Gabay's Theorem (b) by adding the following condition:

Theorem 4.2.5. *Let T, T_1, and T_2 be maximal monotone operators on a Hilbert space $\mathcal{H}$, such that $T = T_1 + T_2$, and T_1^{-1} is strongly monotone with modulus β. If $\{\lambda_k\}_{k=0}^{\infty} \subset (0,\infty)$ satisfies (2.10.9), and $\mathrm{cl}(\mathrm{dom}(T_2))$ is compact, then the sequence $\{z^k\}$ from (2.10.1) for any starting point $z^0 \in \mathrm{dom}(T_1)$ converges **strongly** to a zero of $T = T_1 + T_2$.*

Proof. Proposition 2.1.4 and Proposition 2.1.5 imply that T_1 is single-valued and Lipschitz continuous. Thus, Lemma 4.2.4 gives $T^{-1}(0) \neq \emptyset$. Let z^* be any member of $T^{-1}(0)$, and let $y^* = -T_1(z^*)$. Then $y^* \in T_2(z^*)$. So

$$z^* + \lambda_k y^* = z^* - \lambda_k T_1(z^*) \quad \text{for all} \quad k \geq 0. \tag{4.2.15}$$

Let $y^{k+1} = \frac{1}{\lambda_k}(z^k - z^{k+1}) - T_1(z^k)$. Then, (2.10.1) implies $y^{k+1} \in T_2(z^{k+1})$ and

$$z^{k+1} + \lambda_k y^{k+1} = z^k - \lambda_k T_1(z^k). \tag{4.2.16}$$

For fixed λ_k in (4.2.15), (4.2.16) subtracting (4.2.15) gives

$$z^{k+1} - z^* + \lambda_k(y^{k+1} - y^*) = z^k - z^* - \lambda_k(T_1(z^k) - T_1(z^*)) \tag{4.2.17}$$

Taking square norm to both sides of (4.2.17) implies

$$\begin{aligned}&\|z^{k+1} - z^*\|^2 + \lambda_k^2\|y^{k+1} - y^*\|^2 + 2\lambda_k\langle z^{k+1} - z^*, y^{k+1} - y^*\rangle = \\ &\|z^k - z^*\|^2 + \lambda_k^2\|T_1(z^k) - T_1(z^*)\|^2 - 2\lambda_k\langle z^k - z^*, T_1(z^k) - T_1(z^*)\rangle.\end{aligned} \tag{4.2.18}$$

Let

$$\eta = \min\{\inf_{k\geq 0}\{\lambda_k\}, 2\beta - \sup_{k\geq 0}\{\lambda_k\}\}$$

Then, (2.10.9) implies $\eta > 0$. Since T_1^{-1} is strongly monotone and T_2 is maximal monotone, we deduce from (4.2.18) that

$$\begin{aligned}\|z^{k+1} - z^*\|^2 + \eta^2\|y^{k+1} - y^*\|^2 &\leq \|z^k - z^*\|^2 \\ &+ \lambda_k^2\|T_1(z^k) - T_1(z^*)\|^2 - 2\lambda_k\langle z^k - z^*, T_1(z^k) - T_1(z^*)\rangle \\ &\leq \|z^k - z^*\|^2 - \lambda_k(2\beta - \lambda_k)\|T_1(z^k) - T_1(z^*)\|^2 \\ &\leq \|z^k - z^*\|^2 - \eta^2\|T_1(z^k) - T_1(z^*)\|^2\end{aligned} \tag{4.2.19}$$

Adding (4.2.19) over $k = 0, 1, \ldots, m-1$, we obtain

$$\begin{aligned}&\|z^m - z^*\|^2 + \eta^2\sum_{k=0}^{m-1}\|y^{k+1} - y^*\|^2 \\ \leq&\|z^0 - z^*\|^2 - \eta^2\sum_{k=0}^{m-1}\|T_1(z^k) - T_1(z^*)\|^2\end{aligned} \tag{4.2.20}$$

(4.2.20) implies that $\|z^k - z^*\| \leq \|z^0 - z^*\|$ for any $k \geq 0$ and the two positive term series are convergent as $m \to \infty$. Hence

$$\sum_{k=0}^{\infty}\|T_1(z^k) - T_1(z^*)\|^2 < \infty, \qquad \sum_{k=0}^{\infty}\|y^k - y^*\|^2 < \infty, \tag{4.2.21}$$

$$\lim_{k\to\infty} T_1(z^k) = T_1(z^*), \qquad \lim_{k\to\infty} y^k = y^*. \tag{4.2.22}$$

Since $\{z^k\} \subset \mathrm{dom}(T_2)$ and $\mathrm{cl}(\mathrm{dom}(T_2))$ is compact, there exists a strongly convergent subsequence $\{z^{k(j)}\}$ of $\{z^k\}$. Let z^∞ be the limit of $\{z^{k(j)}\}$. Since T_1 and T_2 are maximal monotone, from (4.2.22) and Lemma 4.1.8 we obtain

$$0 = y^* + T_1(z^*) = \lim_{j\to\infty} y^{k(j)} + T_1(z^{k(j)}) \in (T_1 + T_2)(\lim_{j\to\infty} z^{k(j)}) = T(z^\infty).$$

Therefore, $z^\infty \in T^{-1}(0)$. Now, replacing z^* by z^∞ in (4.2.19) makes no difference and implies that $\{z^k\}$ actually strongly converges to z^∞. □

Let us consider the convergence rates of Gabay's Theorem and Theorem 4.2.5.

Theorem 4.2.6. *Suppose T, T_1, and T_2 are general maximal monotone operators with $T = T_1 + T_2$. Then*

(a) under the conditions of Part (a) in Gabay's Theorem, the convergence rate is Q-linear with ratio $\theta = \lambda^2 M^2 + 1 - 2\alpha\lambda < 1$.

(b) under the conditions of Part (b) in Gabay's Theorem, or Theorem 4.2.5, the estimation of the convergence rate is as follows: there are $y^k \in T_2(z^k)$ with $z^* \in T^{-1}(0)$ such that

$$\liminf_{k\to\infty} \sqrt{k}[T_1(z^k) - T_1(z^*)] = 0, \quad \liminf_{k\to\infty} k\langle z^k - z^*, y^k + T_1(z^*)\rangle = 0, \tag{4.2.23}$$

$$\liminf_{k\to\infty} \sqrt{k}[y^k + T_1(z^*)] = 0, \quad \liminf_{k\to\infty} k\langle z^k - z^*, T_1(z^k) - T_1(z^*)\rangle = 0. \tag{4.2.24}$$

Proof of (a). Let $y^k \in T_2(z^k)$ and $x^k = T_1(z^k)$ such that

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

If $z^* \in T^{-1}(0)$, then $y^* = -x^* = -T_1(z^*) \in T_2(z^*)$ and

$$z^{k+1} - z^* + \lambda(y^{k+1} - y^*) = z^k - z^* - \lambda(x^k - x^*).$$

Since T_1 is strongly monotone with α and Lipschitz continuous with M, we have from the above that

$$\begin{aligned}
\|z^{k+1} - z^*\|^2 &\le \|z^{k+1} - z^* + \lambda(y^{k+1} - y^*)\|^2 \\
&= \|z^k - z^* - \lambda(x^k - x^*)\|^2 \\
&\le \|z^k - z^*\|^2 + \lambda^2\|x^k - x^*\|^2 - 2\lambda\langle z^k - z^*, x^k - x^*\rangle \\
&\le (\lambda^2 M^2 + 1 - 2\alpha\lambda)\|z^k - z^*\|^2
\end{aligned}$$

The condition on λ implies that

$$\theta = \lambda^2 M^2 + 1 - 2\alpha\lambda < 1.$$

Therefore, the convergence rate is Q-linear.

Proof of (b). Follow the proof of Theorem 4.2.5 beginning from (4.2.15) through (4.2.21). Note that $\lambda_k \equiv \lambda$ under the conditions of the part (b) in Gabay's Theorem.

Now, let i be any non-negative integer. Adding (4.2.18) over $k = m, m+1, \ldots, m+i$ gives

$$\begin{aligned} 0 \leq & \|z^{m+i+1} - z^*\|^2 + \sum_{k=m}^{m+i} [\lambda_k^2 \|y^{k+1} - y^*\|^2 \\ & \qquad + 2\lambda_k \langle z^{k+1} - z^*, y^{k+1} - y^* \rangle] \\ = & \|z^m - z^*\|^2 + \sum_{k=m}^{m+i} [\lambda_k^2 \|T_1(z^k) - T_1(z^*)\|^2 \\ & \qquad - 2\lambda_k \langle z^k - z^*, T_1(z^k) - T_1(z^*) \rangle] \\ \leq & \|z^0 - z^*\|^2. \end{aligned} \tag{4.2.25}$$

For $m = 0$ and $i \to \infty$, (4.2.21) and the condition on $\{\lambda_k\}$ imply that

$$\sum_{k=0}^{\infty} \lambda_k^2 \|T_1(z^k) - T_1(z^*)\|^2 < \infty, \qquad \sum_{k=0}^{\infty} \lambda_k^2 \|y^{k+1} - y^*\|^2 < \infty,$$

hence

$$\sum_{k=0}^{\infty} \langle z^k - z^*, T_1(z^k) - T_1(z^*) \rangle < \infty, \quad \sum_{k=1}^{\infty} \langle z^k - z^*, y^k - y^* \rangle < \infty. \tag{4.2.26}$$

Since each of the four series in (4.2.21) and (4.2.26) are of positive terms, (4.2.21) and (4.2.26) imply (4.2.23) and (4.2.24) because

$$\sum_{k=1}^{\infty} \frac{1}{k} = \infty.$$

$\square$

4.3. Nonexpansivity of G_k and the Convergence of the BFBS Method

Note from Section 4.1 that we can restrict the conditions to only dom(T_2) and obtain the same convergence theorems of the FBS algorithm in Section 2.10 and Section 4.2. Let us summarize the number of convergence conditions we have now for the BFBS algorithm. Since the conditions in G-L-P Theorem of Section 2.10 and the conditions in Theorem 4.2.5 are special cases of the conditions in the part (b) of Gabay's Theorem, with λ_k satisfying (2.10.9), we shall not describe them again.

Let T_1 and T_2 be maximal monotone operators on a Hilbert space $\mathcal{H}$ and $\{\lambda_k\}_{k=0}^{\infty} \subset (0, \infty)$

Condition A.

(a) T_1 is Lipschitz continuous with constant M,

(b) T_2 is strongly monotone with modulus α,

(c) $\{\lambda_k\} \subset \mathbb{R}$ satisfies (4.2.3).

Condition B.

(1) T_1 is Lipschitz continuous with constant M,

(2) T_1, T_2 are strongly monotone with moduli α_1, α_2 respectively,

(3) $\{\lambda_k\}$ satisfies (4.2.11).

Condition C.

(a) If T_1 is Lipschitz-continuous, with constant M, and strongly monotone, with modulus α,

(b)

$$0 < \inf\{\lambda_k\} \leq \sup\{\lambda_k\} < 2\alpha M^{-2}. \tag{4.3.1}$$

Condition D.

(i) T_1^{-1} is strongly monotone with modulus β,

(ii) $\{\lambda_k\}$ satisfies (2.10.9),

(iii) $T^{-1}(0) = (T_1 + T_2)^{-1}(0) \neq \emptyset$.

In the above, Condition A and Condition B are newly developed in Section 4.2, while Condition C and Condition D are from Gabay's theorem and Tseng's theorem.

From Section 2.10 and Section 4.2, we know that a sequence $\{z^k\}$ generated by the BFBS method *strongly* converges to a zero point of T under Condition A, Condition B, or Condition C while it *weakly* converges under Condition D in general Hilbert spaces.

Now we define several concepts intended for use. To begin, G_k is the forward-backward operator given by (4.1.1).

Definition 4.3.1. *If an operator T satisfies for a constant c*

$$\|w' - w''\| \leq c\|z' - z''\|, \quad \text{whenever} \quad w' \in T(z'), w'' \in T(z'') \tag{4.3.2}$$

for any $z', z'' \in Z \subset \mathcal{H}$, then the operator T is called nonexpansive over Z when $c = 1$ and contractive over Z with c when $c < 1$. T is maximal nonexpansive over $\mathcal{H}$ if there does not exist a nonexpansive operator T' over $\mathcal{H}$ such that $\mathrm{gph}(T) \subset \mathrm{gph}(T')$, $\mathrm{gph}(T) \neq \mathrm{gph}(T')$.

Lemma 4.3.1. *If T is nonexpansive over Z, then T is single-valued and Lipschitz continuous over Z.*

Proof. Lipschitz continuity is obvious. If $w' \in T(z)$ and $w'' \in T(z)$, for $z \in Z$, then

$$\|w' - w''\| \leq \|z - z\| = 0$$

Therefore, $w' = w''$. □

Lemma 4.3.2. *If $T : \mathcal{H} \rightrightarrows \mathcal{H}$ is contractive over a closed set Z with $T(Z) \subset Z$, then T has a unique fixed point $z^* \in Z$, i.e., $z^* = T(z^*)$. Furthermore, any sequence $\{z^k\}$ generated by $z^{k+1} = T(z^k)$ for any starting point $z^0 \in Z$ converges strongly to z^*.*

Proof. Let $\{z^k\}$ be any sequence from $z^{k+1} = T(z^k)$, with $z^0 \in Z$, and c be the contractive constant of T over Z. Then

$$\|z^{k+1} - z^k\| \leq c\|z^k - z^{k-1}\| \leq c^k\|z^1 - z^0\|.$$

$c < 1$ implies that $c^k \to 0$, when $k \to \infty$, and $\{z^k\}$ strongly converges to a point $z^* \in Z$, similar to the proof of Lemma 4.2.1. Thus

$$z^* = \lim_{k\to\infty} z^{k+1} = T(\lim_{k\to\infty} z^k) = T(z^*).$$

That is, $z^* \in Z$ is a fixed point of T. Hence, T has a fixed point. Now, if there is another fixed point $\bar{z} \in Z$, then

$$\|z^* - \bar{z}\| = \|T(z^*) - T(\bar{z})\| \leq c\|z^* - \bar{z}\|.$$

But $c < 1$. Thus, $\bar{z} = z^*$, and the fixed point of T is unique. □

If we replace the *contractive* by *nonexpansive* in Lemma 4.3.2, then we may not even have a fixed point in Z and $\{z^k\}$ may diverge even if there is a fixed point in Z.

Example 4.3.1. Let $Z = \mathbb{R}^2$ and T be

$$T(z) = z + \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 + z_1 \\ z_2 \end{pmatrix} \quad \text{when} \quad z \in Z.$$

Then Z is closed and T is nonexpansive over Z, however, T does not have a fixed point in Z.

Example 4.3.2. Let $Z = \mathbb{R}^2$ and T be

$$T = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

Then Z is closed and T is nonexpansive with unique fixed point $z^* = 0 \in \mathbb{R}^2$, however, $z^{k+1} = T(z^k)$ will diverge for any $z^0 \neq 0$.

Definition 4.3.2. *For a set of operators $\{T_i \,|\, i \in S\}$ with subscript set S, they are uniformly contractive with c over a set $Z \subset \mathcal{H}$, if every T_i is contractive over Z with the same constant $c < 1$ for all $i \in S$.*

Theorem 4.3.3. *If $\{T_k\}_{k=0}^{\infty}$ is a uniformly contractive operator sequence with c over a closed set $Z \subset \mathcal{H}$, with $\mathrm{rge}(T_k) \subset Z$ and has the same fixed point, then* *any sequence $\{z^k\}$ generated by $z^{k+1} = T_k(z^k)$ converges strongly to their unique fixed point and the convergence rate is Q-linear with global ratio c.*

Proof. From Lemma 4.3.2, every T_k has a unique fixed point in Z. Let z^* be their common fixed point and c be the uniformly contractive constant. Then

$$\|z^k - z^*\| = \|T_k(z^{k-1}) - T_k(z^*)\| \le c\|z^{k-1} - z^*\| \le c^k\|z^0 - z^*\|.$$

$c < 1$ implies that $\{z^k\}$ Q-linearly converges to z^*. □

Now let us consider the relationship between the above convergence conditions and the nonexpansivity of the forward-backward operator G_k.

Theorem 4.3.4. *If Condition C is satisfied over a set Z, then the forward-backward operators $G_k = (I + \lambda_k T_2)^{-1}(I - \lambda_k T_1)$, for all $k \ge 0$, are uniformly contractive over $\mathrm{dom}(T_1) \bigcap Z$ and satisfy*

$$\|w_1 - w_2\|^2 \le (\lambda_k^2 M^2 + 1 - 2\alpha\lambda_k)\|z_1 - z_2\|^2 \tag{4.3.3}$$

where $w_i \in G_k(z_i)$ and $z_i \in Z$ for $i = 1, 2$.

Proof. Let $y_i \in T_2(w_i)$ and $x_i \in T_1(z_i)$ such that

$$w_i + \lambda_k y_i = z_i - \lambda_k x_i, i = 1, 2.$$

Thus

$$w_1 - w_2 + \lambda_k(y_1 - y_2) = z_1 - z_2 - \lambda_k(x_1 - x_2).$$

Since T_1 is strongly monotone with α and Lipschitz continuous with M, we have from the above that

$$\begin{aligned}\|w_1 - w_2\|^2 &\le \|w_1 - w_2 + \lambda_k(y_1 - y_2)\|^2 \\ &= \|z_1 - z_2 - \lambda_k(x_1 - x_2)\|^2 \\ &\le \|z_1 - z_2\|^2 + \lambda_k^2\|x_1 - x_2\|^2 - 2\lambda_k\langle z_1 - z_2, x_1 - x_2\rangle \\ &\le (\lambda_k^2 M^2 + 1 - 2\alpha\lambda_k)\|z_1 - z_2\|^2\end{aligned}$$

The uniform contractivity constant c is from (4.3.1) such that

$$c^2 = \sup_{k\ge 0}(\lambda_k^2 M^2 + 1 - 2\alpha\lambda_k) = 1 - \inf_{k\ge 0} M^2\lambda_k(2M^{-2}\alpha - \lambda_k) < 1.$$

□

Theorem 4.3.5. *When Condition A is satisfied over a set Z, the forward-backward operators G_k, for all $k \geq 0$, are uniformly contractive over Z and satisfy $w_i \in G_k(z_i)$ and $z_i \in Z$ with $i = 1, 2$ such that*

$$\|w_1 - w_2\|^2 \leq \frac{1 + \lambda_k^2 M^2}{(1 + \lambda_k \alpha)^2} \|z_1 - z_2\|^2. \tag{4.3.4}$$

Proof. Similar to the proof of Theorem 4.2.3, it is easy to prove (4.3.4) from the known conditions. The uniform contractivity exists because

$$c^2 = \sup_{k \geq 0} \{ \frac{1 + \lambda_k^2 M^2}{(1 + \lambda_k \alpha)^2} \} < 1$$

when $\{\lambda_k\}$ satisfies (4.2.3). □

Theorem 4.3.6. *Suppose Condition B is true over Z. Then the forward-backward operators G_k are uniformly contractive over Z for all $k \geq 0$ and satisfy $w_i \in G_k(z_i)$ and $z_i \in Z$, $i = 1, 2$ such that*

$$\|w_1 - w_2\|^2 \leq \frac{1 + \lambda_k^2 M^2 - 2\alpha_1 \lambda_k}{(1 + \lambda_k \alpha_2)^2} \|z_1 - z_2\|^2. \tag{4.3.5}$$

Proof. This proof is similar to argument of Theorem 4.2.4 and the proof above. □

Therefore, from Theorem 4.3.3, it is simple to show that the BFBS method strongly converges to the unique zero of T under Condition A, Condition B, or Condition C.

Since Theorem 4.2.1 covers the results of Tseng's Theorem and Part (b) of Gabay's Theorem, we shall only analyze the nonexpansivity of the forward-backward operators G_k under Condition D.

Theorem 4.3.7. *If Condition D is satisfied over Z, then the forward-backward operators G_k from (4.1.1) are nonexpansive over Z for all $k \geq 0$ and for any $z_i \in Z$ with $w_i = G_k(z_i)$, $i = 1, 2$*

$$\|w_1 - w_2\|^2 \leq \|z_1 - z_2\|^2 - \lambda_k(2\beta - \lambda_k)\|T_1(z_1) - T_1(z_2)\|^2. \tag{4.3.6}$$

Proof. For any $k \geq 0$, let $y_i \in T_2(w_i)$ such that

$$w_i + \lambda_k y_i = z_i - \lambda_k T_1(z_i), i = 1, 2.$$

Thus

$$w_1 - w_2 + \lambda_k(y_1 - y_2) = z_1 - z_2 - \lambda_k[T_1(z_1) - T_1(z_2)].$$

Since T_1^{-1} is strongly monotone with modulus β and (2.10.9), so

$$\begin{aligned}
\|w_1 - w_2\|^2 &\leq \|w_1 - w_2 + \lambda_k(y_1 - y_2)\|^2 \\
&= \|z_1 - z_2 - \lambda_k[T_1(z_1) - T_1(z_2)]\|^2 \\
&\leq \|z_1 - z_2\|^2 - \lambda_k(2\beta - \lambda_k)\|T_1(z_1) - T_1(z_2)\|^2 \\
&\leq \|z_1 - z_2\|^2.
\end{aligned}$$

□

In general we are not able to guarantee the contractivity of G_k under Condition D.

Example 4.3.3. Let $T_2 = N_Z$ with $Z \neq \emptyset$ closed, convex and $T_1(z) \equiv d$ with d a fixed point in $\mathcal{H}$. Then T_1 and T_2 are maximal monotone and T_1^{-1} is strongly monotone with any positive scalar. But $T_1(z_1) - T_1(z_2) = 0$ for any z_1 and z_2 in $\mathcal{H}$. Hence

$$\|G_k(z_1) - G_k(z_2)\| \equiv \|z_1 - z_2\| \quad \text{for any} \quad z_1, z_2 \in Z; k \geq 0.$$

Thus G_k is not contractive for each $k \geq 0$.

Therefore, Condition D is weaker than the other three conditions from the perspective of nonexpansivity and contractivity.

4.4. Relationships among Convergence Conditions

At the end of Section 4.3 we concluded that Condition D is weak from one perspective of nonexpansivity. Now we shall prove that it is also true from another perspective.

Theorem 4.4.1. *If Condition C is true, then Condition D is also true with* $\beta = \alpha M^{-2}$.

Proof. Condition C implies that

$$\|T_1(y) - T_1(x)\| \le M\|y - x\|,$$
$$\alpha\|y - x\|^2 \le \langle y - x, T_1(y) - T_1(x)\rangle.$$

for any x and y in $\mathcal{H}$. Hence, we have that

$$\begin{aligned}\langle y - x, T_1(y) - T_1(x)\rangle &\ge \alpha\|y - x\|^2 \\ &\ge \alpha M^{-2}\|T_1(y) - T_1(x)\|^2.\end{aligned}$$

That is, T_1^{-1} is strongly monotone with modulus $\beta = \alpha M^{-2}$.

Now the conditions for $\{\lambda_k\}$ and T become trivial from the above and Lemma 4.2.2. □

Theorem 4.4.2. *If Condition A is true for* T_1, T_2, *and* $\{\lambda_k\}$, *then the following maximal monotone operators* $\bar{T}_1$, $\bar{T}_2$, *and step sizes* $\{\bar{\lambda}_k\}$ *will satisfy Condition D and these two groups deduce the exact same sequence* $\{z^k\}$ *from the FBS scheme (2.10.1) if* z^0 *is the same when*

$$\bar{T}_1 = T_1 + \alpha I, \qquad \bar{T}_2 = T_2 - \alpha I \quad \text{and} \quad \bar{\lambda}_k = \frac{\lambda_k}{1 + \alpha\lambda_k} \tag{4.4.1}$$

with the modulus

$$\beta = \frac{\alpha}{\max\{M^2, \alpha^2\} + \alpha^2}. \tag{4.4.2}$$

Proof. Since T_1 is maximal monotone and Lipschitz continuous with constant M and T_2 is strongly monotone with modulus α, we know that $\bar{T}_1$ is strongly monotone with α and $\bar{T}_2$ is maximal monotone. So $T = T_1 + T_2$ has a unique zero point from Lemma 4.2.2. Next we are going to prove that $\bar{T}_1^{-1}$ is strongly monotone with the modulus β in (4.4.2) and $\{\bar{\lambda}_k\}$ in (4.4.1) satisfies Condition D from (4.2.3).

Since

$$M\|y - x\| \ge \|T_1(y) - T_1(x)\|,$$

for any y and x in the space $\mathcal{H}$, we have the following:

$$\begin{aligned} &\|y-x\|^2 \geq M^{-2}\|T_1(y)-T_1(x)\|^2 \\ =&M^{-2}\left[\|\bar{T}_1(y)-\bar{T}_1(x)\|^2+\alpha^2\|y-x\|^2\right. \\ &\left.-2\alpha\langle y-x,\bar{T}_1(y)-\bar{T}_1(x)\rangle\right]. \end{aligned} \tag{4.4.3}$$

Therefore,

$$\|\bar{T}_1(y)-\bar{T}_1(x)\|^2-2\alpha\langle y-x,\bar{T}_1(y)-\bar{T}_1(x)\rangle \leq (M^2-\alpha^2)\|y-x\|^2 \tag{4.4.4}$$

for any y and x in $\mathcal{H}$.

Case (a). $M \leq \alpha$. In this case, (4.4.4) implies

$$\langle y-x,\bar{T}_1(y)-\bar{T}_1(x)\rangle \geq \frac{1}{2\alpha}\|\bar{T}_1(y)-\bar{T}_1(x)\|^2.$$

Thus, $\bar{T}_1^{-1}$ is strongly monotone with modulus $(2\alpha)^{-1}$. However this modulus is the same as the β in (4.4.2) in this case.

Now (4.2.3) implies

$$0 < \inf_{k\geq 0}\{\lambda_k\} \leq \sup_{k\geq 0}\{\lambda_k\} < \infty$$

and (4.4.1) implies

$$\lambda_k = \frac{\bar{\lambda}_k}{1-\alpha\bar{\lambda}_k}. \tag{4.4.5}$$

These two inequalities again imply that $0 < \inf\{\bar{\lambda}_k\}$ and $0 < \inf\{1-\alpha\bar{\lambda}_k\}$. That is,

$$0 < \inf_{k\geq 0}\{\bar{\lambda}_k\} \leq \sup_{k\geq 0}\{\bar{\lambda}_k\} < \frac{1}{\alpha} = 2\beta.$$

Hence, Condition D is true under (4.4.1) and (4.4.2).

Case (b). $M > \alpha$. $\bar{T}_1$ is strongly monotone with α. i.e.

$$\langle y-x,\bar{T}_1(y)-\bar{T}_1(x)\rangle \geq \alpha\|y-x\|^2.$$

Combining it with (4.4.3) gives

$$(M^2+2\alpha^2)\langle y-x,\bar{T}_1(y)-\bar{T}_1(x)\rangle \geq \alpha\|\bar{T}_1(y)-\bar{T}_1(x)\|^2+\alpha^3\|y-x\|^2.$$

Putting (4.4.4) into the above again gives

$$\frac{M^2(M^2+\alpha^2)}{M^2-\alpha^2}\langle y-x, \bar{T}_1(y)-\bar{T}_1(x)\rangle \geq \frac{\alpha M^2}{M^2-\alpha^2}\|\bar{T}_1(y)-\bar{T}_1(x)\|^2.$$

That is,

$$\langle y-x, \bar{T}_1(y)-\bar{T}_1(x)\rangle \geq \frac{\alpha}{M^2+\alpha^2}\|\bar{T}_1(y)-\bar{T}_1(x)\|^2$$

for any y and x in $\mathcal{H}$. Hence, $\bar{T}_1^{-1}$ is strongly monotone with modulus $\alpha(M^2+\alpha^2)^{-1}$. However, this modulus is the same as the β in (4.4.2) at the case here.

Now (4.2.3) and (4.4.5) imply

$$0 < \inf_{k\geq 0}\{\frac{\bar{\lambda}_k}{1-\alpha\bar{\lambda}_k}\} \leq \sup_{k\geq 0}\{\frac{\bar{\lambda}_k}{1-\alpha\bar{\lambda}_k}\} < \frac{2\alpha}{M^2-\alpha^2}.$$

Thus, $0<\inf\{\bar{\lambda}_k\}$ and

$$\frac{\bar{\lambda}_k}{1-\alpha\bar{\lambda}_k} \leq \eta\frac{2\alpha}{M^2-\alpha^2}$$

for any $k\geq 0$ and some η with $0<\eta<1$. That is,

$$(M^2-\alpha^2+2\eta\alpha^2)\bar{\lambda}_k \leq 2\eta\alpha.$$

Since $M^2+\alpha^2>2\alpha^2$ and

$$M^2+\alpha^2-2\alpha^2(1-\eta) > (M^2+\alpha^2)(1-1+\eta) = (M^2+\alpha^2)\eta,$$

we see that

$$\sup_{k\geq 0}\{\bar{\lambda}_k\} \leq \frac{2\eta\alpha}{M^2+\alpha^2-2\alpha^2(1-\eta)} < \frac{2\eta\alpha}{\eta(M^2+\alpha^2)} = 2\beta.$$

Therefore, Condition D is true under (4.4.1) and (4.4.2).

The two groups deduce the same sequence from (2.10.1) because of the following:

$$\begin{aligned} z^{k+1} &= (I+\lambda_k T_2)^{-1}(I-\lambda_k T_1)(z^k) \\ &= (I+\frac{\bar{\lambda}_k}{1-\alpha\bar{\lambda}_k}T_2)^{-1}(I-\frac{\bar{\lambda}_k}{1-\alpha\bar{\lambda}_k}T_1)(z^k) \\ &= (I-\alpha\bar{\lambda}_k I+\bar{\lambda}_k T_2)^{-1}(I-\alpha\bar{\lambda}_k I-\bar{\lambda}_k T_1)(z^k) \\ &= (I+\bar{\lambda}_k\bar{T}_2)^{-1}(I-\bar{\lambda}_k\bar{T}_1)(z^k). \end{aligned}$$

Thus, the proof is complete. □

Theorem 4.4.3. *If Condition B is true for T_1, T_2, and $\{\lambda_k\}$, then the following maximal monotone operators $\bar{T}_1$, $\bar{T}_2$, and step sizes $\{\bar{\lambda}_k\}$ will satisfy Condition D and these two groups deduce the exact same sequence $\{z^k\}$ from the FBS scheme (2.10.1) if z^0 is same when*

$$\bar{T}_1 = T_1 + \alpha_2 I, \qquad \bar{T}_2 = T_2 - \alpha_2 I \quad \text{and} \quad \bar{\lambda}_k = \frac{\lambda_k}{1+\alpha_2\lambda_k} \tag{4.4.6}$$

with the modulus

$$\beta = \frac{\alpha_1 + \alpha_2}{\max\{M^2, \alpha_2^2\} + \alpha_2^2 + 2\alpha_1\alpha_2}. \tag{4.4.7}$$

Proof. Similar to the proof of Theorem 4.4.2, since T_1 is strongly monotone with modulus α_1, and Lipschitz continuous with constant M, and T_2 is strongly monotone with modulus α_2, we know that $\bar{T}_1$ is strongly monotone with $\alpha_1 + \alpha_2$, and $\bar{T}_2$ is maximal monotone. So $T = T_1 + T_2$ has a unique zero point from Lemma 4.2.2. Next we are going to prove that $\bar{T}_1^{-1}$ is strongly monotone with the modulus β in (4.4.7) and $\{\bar{\lambda}_k\}$ in (4.4.6) satisfies Condition D from (4.2.11).

Since

$$M\|y-x\| \geq \|T_1(y) - T_1(x)\|,$$

for any y and x in the space $\mathcal{H}$, we have the following,

$$\begin{aligned} \|y-x\|^2 &\geq M^{-2}\|T_1(y) - T_1(x)\|^2 \\ &= M^{-2}\Big[\|\bar{T}_1(y) - \bar{T}_1(x)\|^2 + \alpha_2^2\|y-x\|^2 \\ &\qquad - 2\alpha_2\langle y-x, \bar{T}_1(y) - \bar{T}_1(x)\rangle\Big]. \end{aligned} \tag{4.4.8}$$

Therefore,

$$\|\bar{T}_1(y) - \bar{T}_1(x)\|^2 - 2\alpha_2\langle y-x, \bar{T}_1(y) - \bar{T}_1(x)\rangle \leq (M^2 - \alpha_2^2)\|y-x\|^2 \tag{4.4.9}$$

for any y and x in $\mathcal{H}$.

Case (a). $M \leq \alpha_2$**.** In this case, (4.4.9) implies

$$\langle y-x, \bar{T}_1(y) - \bar{T}_1(x)\rangle \geq \frac{1}{2\alpha_2}\|\bar{T}_1(y) - \bar{T}_1(x)\|^2.$$

Thus, $\bar{T}_1^{-1}$ is strongly monotone with modulus $(2\alpha_2)^{-1}$. However this modulus is the same as the β in (4.4.7) in this case.

Now (4.2.11) implies

$$0 < \inf_{k \geq 0} \{\lambda_k\} \leq \sup_{k \geq 0} \{\lambda_k\} < \infty$$

and (4.4.6) implies

$$\lambda_k = \frac{\bar{\lambda}_k}{1 - \alpha_2 \bar{\lambda}_k}. \tag{4.4.10}$$

These two inequalities again imply that $0 < \inf\{\bar{\lambda}_k\}$ and $0 < \inf\{1 - \alpha_2 \bar{\lambda}_k\}$. That is,

$$0 < \inf_{k \geq 0} \{\bar{\lambda}_k\} \leq \sup_{k \geq 0} \{\bar{\lambda}_k\} < \frac{1}{\alpha_2} = 2\beta.$$

Hence, Condition D is true for $\bar{T}_1$ and $\bar{T}_2$ with β under (4.4.6) and (4.4.7).

Case (b). $M > \alpha_2$**.** $\bar{T}_1$ is strongly monotone with $\alpha_1 + \alpha_2$. i.e.

$$\langle y - x, \bar{T}_1(y) - \bar{T}_1(x) \rangle \geq (\alpha_1 + \alpha_2)\|y - x\|^2.$$

Combining it with (4.4.8) gives

$$\begin{aligned}
&[M^2 + 2\alpha_2(\alpha_1 + \alpha_2)]\langle y - x, \bar{T}_1(y) - \bar{T}_1(x)\rangle \geq \\
&(\alpha_1 + \alpha_2)\|\bar{T}_1(y) - \bar{T}_1(x)\|^2 + \alpha_2^2(\alpha_1 + \alpha_2)\|y - x\|^2.
\end{aligned}$$

Putting (4.4.9) into the above again gives

$$\begin{aligned}
&[M^2 + 2\alpha_2(\alpha_1 + \alpha_2) + \frac{2\alpha_2^3(\alpha_1 + \alpha_2)}{M^2 - \alpha_2^2}]\langle y - x, \bar{T}_1(y) - \bar{T}_1(x)\rangle \\
&= \frac{M^2(M^2 + 2\alpha_1\alpha_2 + \alpha_2^2)}{M^2 - \alpha_2^2}\langle y - x, \bar{T}_1(y) - \bar{T}_1(x)\rangle \\
&\geq (\alpha_1 + \alpha_2)(1 + \frac{\alpha_2}{M^2 - \alpha_2^2})\|\bar{T}_1(y) - \bar{T}_1(x)\|^2 \\
&= \frac{M^2(\alpha_1 + \alpha_2)}{M^2 - \alpha_2^2}\|\bar{T}_1(y) - \bar{T}_1(x)\|^2.
\end{aligned}$$

That is,

$$\langle y - x, \bar{T}_1(y) - \bar{T}_1(x)\rangle \geq \frac{\alpha_1 + \alpha_2}{M^2 + 2\alpha_1\alpha_2 + \alpha_2^2}\|\bar{T}_1(y) - \bar{T}_1(x)\|^2$$

for any y and x in $\mathcal{H}$. Hence, $\bar{T}_1^{-1}$ is strongly monotone with modulus $(\alpha_1 + \alpha_2)(M^2 + 2\alpha_1\alpha_2 + \alpha_2^2)^{-1}$. However, this modulus is the same as the β in (4.4.7) in the case here.

Now (4.2.11) and (4.4.10) imply

$$0 < \inf_{k\geq 0}\{\frac{\bar{\lambda}_k}{1-\alpha_2\bar{\lambda}_k}\} \leq \sup_{k\geq 0}\{\frac{\bar{\lambda}_k}{1-\alpha_2\bar{\lambda}_k}\} < \frac{2(\alpha_1+\alpha_2)}{M^2-\alpha_2^2}.$$

Thus, $0 < \inf\{\bar{\lambda}_k\}$ and

$$\frac{\bar{\lambda}_k}{1-\alpha_2\bar{\lambda}_k} \leq \eta\frac{2(\alpha_1+\alpha_2)}{M^2-\alpha_2^2}$$

for any $k \geq 0$ and some η with $0 < \eta < 1$. This can again imply for any $k \geq 0$ that

$$[M^2 - \alpha_2^2 + 2\eta\alpha_2(\alpha_1+\alpha_2)]\bar{\lambda}_k \leq 2\eta(\alpha_1+\alpha_2). \tag{4.4.11}$$

Since $M > \alpha_2$, so

$$M^2 + \alpha_2^2 + 2\alpha_1\alpha_2 \geq 2\alpha_2(\alpha_1+\alpha_2).$$

Hence

$$\begin{aligned}
&M^2 - \alpha_2^2 + 2\eta\alpha_2(\alpha_1+\alpha_2)\\
=&M^2 + \alpha_2^2 + 2\alpha_1\alpha_2 - 2\alpha_2(\alpha_1+\alpha_2)(1-\eta)\\
>&(M^2 + \alpha_2^2 + 2\alpha_1\alpha_2)[1-(1+\eta)]\\
=&\eta(M^2 + \alpha_2^2 + 2\alpha_1\alpha_2).
\end{aligned}$$

we see from the above and (4.4.11) that

$$\begin{aligned}
\sup_{k\geq 0}\{\bar{\lambda}_k\} \leq& \frac{2\eta(\alpha_1+\alpha_2)}{M^2-\alpha_2^2+2\eta\alpha_2(\alpha_1+\alpha_2)}\\
<& \frac{2(\alpha_1+\alpha_2)}{M^2+\alpha_2^2+2\alpha_1\alpha_2} = 2\beta.
\end{aligned}$$

Therefore, Condition D is true under (4.4.6) and (4.4.7).

The two groups deduce the same sequence from (2.10.1) because of (4.4.6) and the following:

$$\begin{aligned} z^{k+1} &= (I + \lambda_k T_2)^{-1}(I - \lambda_k T_1)(z^k) \\ &= (I + \frac{\bar{\lambda}_k}{1-\alpha_2\bar{\lambda}_k} T_2)^{-1}(I - \frac{\bar{\lambda}_k}{1-\alpha_2\bar{\lambda}_k} T_1)(z^k) \\ &= (I - \alpha_2\bar{\lambda}_k I + \bar{\lambda}_k T_2)^{-1}(I - \alpha_2\bar{\lambda}_k I - \bar{\lambda}_k T_1)(z^k) \\ &= (I + \bar{\lambda}_k \bar{T}_2)^{-1}(I - \bar{\lambda}_k \bar{T}_1)(z^k). \end{aligned}$$

Hence, the proof is complete. □

4.5. The Global Convergence of the FBS Algorithm

In Section 2.10 and Section 4.2, all the convergence results are related to the BFBS case with $\varepsilon_k \equiv 0$ in the FBS algorithm. They are inadequate for the general case in Section 4.1. We shall extend to the general FBS case in this section. Note that we always assume that (4.1.9) is true.

Proposition 4.5.1. *Suppose T_1 is single-valued and continuous on* $\mathrm{dom}(T_1)$. *Let $\{z^k\}$ be a sequence generated by the FBS algorithm (4.1.2) for a starting point $z^0 \in \mathrm{dom}(T_1)$. If $\{z^k\}$ strongly converges to a point z^*, and $\inf\{\lambda_k\} > 0$, then $z^* \in T^{-1}(0)$.*

Proof. Let $w^k = G_k(z^k)$ since T_1 is single-valued. Then

$$\|w^k - z^{k+1}\| \le \varepsilon_k.$$

$\{z^k\}$ converges to z^* implies, from (4.1.3), that

$$\lim_{k\to\infty} w^k - z^* = 0, \qquad \lim_{k\to\infty} w^k - z^k = 0.$$

But, $w^k = G_k(z^k)$, and the continuity of T_1, imply that

$$\lambda_k^{-1}(z^k - w^k) - T_1(z^k) \in T_2(w^k),$$

hence

$$0 - \lim_{k\to\infty} T_1(z^k) = -T_1(z^*) \in T_2(\lim_{k\to\infty} w^k) = T_2(z^*).$$

That is, $0 \in T_1(z^*) + T_2(z^*) = T(z^*)$. □

We shall see if the conditions in Section 4.3 can directly extend to the FBS algorithm. Before entering our main proof, we need to prove a lemma first.

Lemma 4.5.2. *If $\{\varepsilon_k\}_{k=0}^{\infty} \subset (0,\infty)$ satisfies (4.1.3) and $0 \le c < 1$, then*

$$\gamma_k = \sum_{i=0}^{k} c^{k-i}\varepsilon_i \to 0, \quad \text{as} \quad k \to \infty. \tag{4.5.1}$$

Proof. $\{\varepsilon_k\}$ satisfying (4.1.3) implies that $\varepsilon_k \to 0$ as $k \to \infty$ and there is a $\bar{\varepsilon}_k$ for every $k \ge 0$ such that

$$\bar{\varepsilon}_k = \max_{i \ge k} \varepsilon_i.$$

$\varepsilon_k \to 0$ implies that $\bar{\varepsilon}_k \to 0$ when $k \to \infty$. Let $\bar{k}$ be the integer satisfying that $(k-1)/2 \le \bar{k} \le k/2$. Then

$$\begin{aligned}\gamma_k &= \sum_{i=0}^{\bar{k}} c^{k-i}\varepsilon_i + \sum_{i=\bar{k}+1}^{k} c^{k-i}\varepsilon_i \\ &\le c^{\bar{k}} \sum_{i=0}^{\infty} \varepsilon_i + (1-c)^{-1}\bar{\varepsilon}_{\bar{k}}.\end{aligned}$$

Since $\bar{k} \to \infty$ when $k \to \infty$, $c^{\bar{k}} \to 0$ and $\bar{\varepsilon}_{\bar{k}} \to 0$ also when $k \to \infty$. Therefore,

$$\lim_{k\to\infty} \gamma_k = 0.$$

□

Theorem 4.5.3. *If Condition A, or Condition B, or Condition C is satisfied on* $\mathrm{dom}(T_1)$*, the* $\{z^k\}$ *generated by the FBS algorithm will strongly converge to the unique zero point of* T*. Moreover, if there are a* $\eta < 1$ *and a constant* C_0*, such that* $\varepsilon_k \le \eta^k C_0$ *for every* $k \ge 0$*, then the convergence is globally linear.*

Proof. T_1 is Lipschitz continuous with constant M. Thus on $\mathrm{dom}(T_2)$ G_k is single-valued and

$$\|z^{k+1} - G_k(z^k)\| \le \varepsilon_k. \tag{4.5.2}$$

Let z^* be the zero point of T. Then $z^* = G_k(z^*)$ for all $k \ge 0$ and

$$\|z^{k+1} - z^*\| \le \|z^{k+1} - G_k(z^k)\| + \|G_k(z^k) - z^*\|. \tag{4.5.3}$$

Theorem 4.3.4, or Theorem 4.3.5, or Theorem 4.3.6 shows that there is a constant $0 \le c < 1$ such that $\{G_k\}$ is uniformly contractive with c. That is,

$$\|G_k(z^k) - z^*\| = \|G_k(z^k) - G_k(z^*)\| \le c\|z^k - z^*\|.$$

Putting this and (4.5.2) into (4.5.3) gives

$$\begin{aligned}\|z^{k+1} - z^*\| &\le \varepsilon_k + c\|z^k - z^*\| \\ &\le \sum_{i=0}^{k} c^{k-i}\varepsilon_i + c^{k+1}\|z^0 - z^*\|.\end{aligned}$$

Lemma 4.5.2 and $c < 1$ imply that

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

Therefore, $\{z^k\}$ strongly converges to z^*.

Now, if $\varepsilon_k \le \eta^k C_0$ for every $k \ge 0$, then we have

$$\begin{aligned}\|z^{k+1} - z^*\| &\le \sum_{i=0}^{k} c^{k-i}\eta^i C_0 + c^{k+1}\|z^0 - z^*\| \\ &\le \bar{c}^{k+1}\bar{C}_0.\end{aligned}$$

Where

$$\bar{c} = \max\left\{\sqrt{c}, \sqrt{\eta}\right\} < 1, \quad \bar{C}_0 = \sum_{k=0}^{\infty}(\bar{c}^{k-1}C_0 + \bar{c}^{k+1}\|z^0 - z^*\|) < \infty.$$

□

Under Condition D the situation is more complicated than under the other 3 conditions. However, we can still deduce the weak convergence.

Theorem 4.5.4. *If Condition D is satisfied on* dom(T_1), *then*

a) *the* $\{z^k\}$ *generated by the FBS algorithm for any starting point* $z^0 \in \mathrm{dom}(T_1)$ *will weakly converge to a zero point of* T,

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

c) Furthermore, the convergence will be strong if cl[dom(T_2)] *is compact.*

Proof. Let z^* be an arbitrary zero point of T. Under Condition D, if we follow the same method of proof as in Theorem 4.5.3, and take $c = 1$, we see that

$$\|z^{k+1} - z^*\| \leq \sum_{i=0}^{k} \varepsilon_i + \|z^0 - z^*\|.$$

(4.1.3) and the above guarantee that $\{z^k\}$ is a bounded sequence. Thus there exists a subsequence $\{z^{k(j)}\}$ which weakly converges to a point z^∞. Let

$$w^k = G_k(z^k) = (I + \lambda_k T_2)^{-1}(I - \lambda_k T_1)(z^k).$$

Then

$$\|z^{k+1} - w^k\| \leq \varepsilon_k, \|w^k - z^*\| \leq \|z^k - z^*\|. \tag{4.5.4}$$

Hence $\{w^k\}$ is also a bounded sequence. From (4.5.4), the subsequence $\{w^{k(j)-1}\}$ weakly converges to z^∞. Let $y^k \in T_2(w^k)$ and $y^* = -T_1(z^*)$ such that

$$w^k + \lambda_k y^k = z^k - \lambda_k T_1(z^k),$$

$$z^* + \lambda_k y^* = z^* - \lambda_k T_1(z^*).$$

Then

$$\begin{aligned}
&\|w^k - z^* + \lambda_k(y^k - y^*)\|^2 \\
=&\|z^k - z^* - \lambda_k(T_1(z^k) - T_1(z^*))\|^2 \\
\leq&\|z^k - z^*\|^2 - \lambda_k(2\beta - \lambda_k)\|T_1(z^k) - T_1(z^*)\|^2
\end{aligned}$$

However

$$\begin{aligned}
&\|w^k - z^* + \lambda_k(y^k - y^*)\|^2 \\
\geq&\|w^k - z^*\|^2 + \lambda_k^2\|y^k - y^*\|^2 \\
\geq&\|z^{k+1} - z^*\|^2 - 2\|z^{k+1} - z^*\|\|w^k - z^{k+1}\| + \lambda_k^2\|y^k - y^*\|^2 \\
\geq&\|z^{k+1} - z^*\|^2 - 2\varepsilon_k\|z^{k+1} - z^*\| + \lambda_k^2\|y^k - y^*\|^2.
\end{aligned}$$

Therefore, if $\eta = \min\{\inf_{k\geq 0} \lambda_k, 2\beta - \sup_{k\geq 0} \lambda_k\}$, then $\eta > 0$ and

$$\begin{aligned}
&\|z^{k+1} - z^*\|^2 + \eta^2\|y^k - y^*\|^2 - 2\varepsilon_k\|z^{k+1} - z^*\| \\
\leq&\|z^k - z^*\|^2 - \eta^2\|T_1(z^k) - T_1(z^*)\|^2.
\end{aligned} \tag{4.5.5}$$

Adding both sides of (4.5.5) for $k = 0, 1, \ldots, m$ gives

$$\begin{aligned} &\|z^{m+1} - z^*\|^2 + \eta^2 \sum_{k=0}^{m} \|y^k - y^*\|^2 - 2\sum_{k=0}^{m} \varepsilon_k \|z^{k+1} - z^*\| \\ \le &\|z^0 - z^*\|^2 - \eta^2 \sum_{k=0}^{m} \|T_1(z^k) - T_1(z^*)\|^2. \end{aligned} \tag{4.5.6}$$

Thus, from (4.1.3) and $\{\|z^{k+1} - z^*\|\}$ bounded, (4.5.6) implies

$$\sum_{k=0}^{\infty} \|y^k - y^*\|^2 < \infty,$$

$$\sum_{k=0}^{\infty} \|T_1(z^k) - T_1(z^*)\|^2 < \infty.$$

So

$$\lim_{k\to\infty} y^k = y^*, \qquad \lim_{k\to\infty} T_1(z^k) = T_1(z^*).$$

Therefore, Lemma 4.1.8 implies that $y^* \in T_2(z^\infty)$ and $T_1(z^*) = T_1(z^\infty)$. Hence $0 \in T_1(z^\infty) + T_2(z^\infty) = T(z^\infty)$. Now, replacing z^* by z^∞ gives

$$z^{k+1} - z^\infty = z^k - z^\infty - \lambda_k[y^k + T_1(z^k)] + z^{k+1} - w^k. \tag{4.5.7}$$

But

$$y^k + T_1(z^k) \to 0 \quad \text{and} \quad z^{k+1} - w^k \to 0 \quad \text{strongly when} \quad k \to \infty.$$

Hence (4.5.7) implies that $\{z^{k(j)}\}$ weakly (or strongly) converges to z^∞, if and only if, $\{z^{k(j)+1}\}$ weakly (or strongly) converges to z^∞. This fact implies that $\{z^k\}$ weakly converges to z^∞, i.e., a) is true.

Since (4.5.7) also implies that

$$\|z^{k+1} - z^k\| \le \lambda_k \|y^k + T_1(z^k)\| + \|z^{k+1} - w^k\|.$$

Therefore, conclusion b) is also true.

Now, if cl[dom(T_2)] is compact, then $\{z^{k(j)}\}$ strongly converges to z^∞. Therefore, (4.5.7) will give the strong convergence of $\{z^k\}$. □

From the results in Section 4.3, we know that all four conditions make G_k contractive or nonexpansive, and T_1 is Lipschitz continuous (so it is single-valued) in all four conditions. Now we shall give a general theorem which includes the above convergent results as special cases. First a lemma proof:

Lemma 4.5.5. *Suppose T_1 and T_2 are monotone and T_1 is single-valued. If $\lambda_k \to \bar{\lambda} > 0$, then for any $z \in \mathcal{H}$*

$$\lim_{k\to\infty}(I - \lambda_k T_1)(z) = (I - \bar{\lambda} T_1)(z), \tag{4.5.8}$$

$$\lim_{k\to\infty}(I + \lambda_k T_2)^{-1}(z) = (I + \bar{\lambda} T_2)^{-1}(z), \tag{4.5.9}$$

$$\lim_{k\to\infty} G_k(z) = \bar{G}(z) = (I + \bar{\lambda} T_2)^{-1}(I - \bar{\lambda} T_1)(z). \tag{4.5.10}$$

Proof. Let $z \in \mathcal{H}$. Then

$$\lim_{k\to\infty} \|(I - \lambda_k T_1)(z) - (I - \bar{\lambda} T_1)(z)\| = \lim_{k\to\infty} |\lambda_k - \bar{\lambda}| \|T_1(z)\| = 0.$$

Let $w_k = (I + \lambda_k T_2)^{-1}(z)$ and $w = (I + \bar{\lambda} T_2)^{-1}(z)$. Then, there exists $y_k \in T_2(w_k)$ and $y \in T_2(w)$, such that

$$w_k + \lambda_k y_k = z = w + \bar{\lambda} y.$$

Hence

$$w_k - w + \lambda_k (y_k - y) = (\bar{\lambda} - \lambda_k) y.$$

From the monotonicity of T_2, we have

$$\begin{aligned}\lim_{k\to\infty} \|w_k - w\|^2 &\leq \lim_{k\to\infty} \|w_k - w + \lambda_k(y_k - y)\|^2 \\ &= \lim_{k\to\infty} (\bar{\lambda} - \lambda_k)^2 \|y\|^2 = 0.\end{aligned}$$

Thus we have proved (4.5.8) and (4.5.9). For proving (4.5.10), let $u_k = G_k(z)$ and $u = \bar{G}(z)$. Then, there exists $v_k \in T_2(u_k)$ and $v \in T_2(u)$, such that

$$u_k + \lambda_k (v_k + T_1(z)) = z = u + \bar{\lambda}(v + T_1(z)).$$

Hence

$$u_k - u + \lambda_k (v_k - v) = (\bar{\lambda} - \lambda_k)(v + T_1(z)).$$

Therefore

$$\lim_{k\to\infty} \|u_k - u\| \leq \lim_{k\to\infty} |\bar{\lambda} - \lambda_k| \|v + T_1(z)\| = 0.$$

This completes the proof of (4.5.10). $\square$

Definition 4.5.1. *We say an operator T is (globally) strongly monotone at $\bar{z}$ (with modulus α) if*

$$\langle z-\bar{z}, y-\bar{y}\rangle \geq \alpha\|z-\bar{z}\|^2 \quad \text{for any} \quad z \quad \text{and} \quad y \in T(z),\ \bar{y} \in T(\bar{z}). \tag{4.5.11}$$

Similarly, we say T^{-1} is (globally) strongly monotone at $\bar{z}$ (with modulus β), if

$$\langle z-\bar{z}, y-\bar{y}\rangle \geq \alpha\|z-\bar{z}\|^2 \quad \text{for any} \quad z \quad \text{and} \quad y \in T(z),\ \bar{y} \in T(\bar{z}). \tag{4.5.12}$$

Theorem 4.5.6. *Suppose $\{z^k\}$ is a sequence generated by the FBS algorithm (4.1.2) with $\inf\{\lambda_k\} > 0$. Then*

(a) $\{z^k\}$ strongly converges to the unique point $\bar{z} \in T^{-1}(0)$ if $\{G_k\}$ in (4.1.1) is uniformly contractive.

(b) $z^* \in T^{-1}(0)$ if $\{\lambda_k\}$ has a upper bound and $\{z^k\}$ strongly converges to z^*.

(c) $\{z^k\}$ strongly converges to a $\bar{z} \in T^{-1}(0)$ if

  (1) $T^{-1}(0) \neq \emptyset$,

  (2) all G_k are nonexpansive,

  (3) $\mathrm{cl}[\mathrm{dom}(T_2)]$ is compact.

  (4) T_1^{-1} is globally strongly monotone at a point $z^* \in T^{-1}(0)$.

(d) $\{z^k\}$ weakly converges to a $\bar{z} \in T^{-1}(0)$, if $T^{-1}(0) \neq \emptyset$, all G_k are nonexpansive and T_1^{-1} is globally strongly monotone at a point $z^* \in T^{-1}(0)$.

Proof. Part (a): Let $w^k = G_k(z^k)$ and $\bar{z} \in T^{-1}(0)$. Then, there is a $c < 1$ such that

$$\|w^k - z^{k+1}\| \leq \varepsilon_k, \|w^k - \bar{z}\| \leq c\|z^k - \bar{z}\|, \tag{4.5.13}$$

$$\|z^{k+1} - \bar{z}\| \leq \varepsilon_k + c\|z^k - \bar{z}\| \leq c^{k+1}\|z^0 - \bar{z}\| + \sum_{i=0}^{k} c^{k-i}\varepsilon_i. \tag{4.5.14}$$

Lemma 4.5.2 and $c < 1$ imply that $\{z^k\}$ strongly converges to $\bar{z}$.

Part (b): If $\{z^k\}$ strongly converges to a point z^*, and $\{\lambda_k\}$ is a bounded sequence in *reals*, there exists a cluster point $\bar{\lambda}$, and a subsequence $\{\lambda_{k(i)}\}$ of $\{\lambda_k\}$, such that

$$\lim_{i\to\infty} \lambda_{k(i)} = \bar{\lambda} > 0. \tag{4.5.15}$$

Let $\bar{G} = (I + \bar{\lambda}T_2)^{-1}(I - \bar{\lambda}T_1)$. Then Lemma 4.5.5 and (4.5.15) imply that

$$\lim_{i\to\infty} \|G_{k(i)}(z^{k(i)}) - \bar{G}(z^*)\| \le \lim_{i\to\infty} [\|z^{k(i)} - z^*\| + \|G_{k(i)}(z^*) - \bar{G}(z^*)\|] = 0.$$

However, $w^k = G_k(z^k)$ and $\lim_{k\to\infty} w^k = \lim_{k\to\infty} z^{k+1} = z^*$. Therefore, $z^* = \bar{G}(z^*)$. That is, $z^* \in T^{-1}(0)$.

Part (c): Let $w^k = G_k(z^k)$ and $z^* \in T^{-1}(0)$. Then

$$\|w^k - z^{k+1}\| \le \varepsilon_k, \|w^k - z^*\| \le \|z^k - z^*\|, \tag{4.5.16}$$

$$\|z^{k+1} - z^*\| \le \varepsilon_k + \|z^k - z^*\| \le \|z^0 - z^*\| + \sum_{i=0}^{k} \varepsilon_i, \tag{4.5.17}$$

and, there is a $y^k \in T_2(w^k)$ for each $k \ge 0$, with $y^* = -T_1(z^*) \in T_2(z^*)$, such that

$$w^k - z^* + \lambda_k(y^k - y^*) = z^k - z^* - \lambda_k[T_1(z^k) - T_1(z^*)]. \tag{4.5.18}$$

Let

$$\begin{aligned} f_k = {} & \lambda_k\|y^k - y^*\|^2 + 2\langle w^k - z^*, y^k - y^*\rangle \\ & + 2\langle z^k - z^*, T_1(z^k) - T_1(z^*)\rangle - \lambda_k\|T_1(z^k) - T_1(z^*)\|^2. \end{aligned} \tag{4.5.19}$$

Then, from (4.5.18), one sees that

$$\|w^k - z^*\|^2 = \|z^k - z^*\|^2 - \lambda_k f_k. \tag{4.5.20}$$

Since G_k is nonexpansive, we obtain $f_k \ge 0$ for all $k \ge 0$. Now, let

$$h_k = \|w^k - z^{k+1}\|^2 + 2\langle w^k - z^{k+1}, z^{k+1} - z^*\rangle.$$

Then, from (4.5.16), (4.5.17), and (4.1.3)

$$\sum_{k=0}^{\infty} h_k < \infty \tag{4.5.21}$$

and we have

$$\|w^k - z^*\|^2 = \|z^{k+1} - z^*\|^2 + h_k.$$

Hence, from (4.5.20),

$$\|z^{k+1} - z^*\|^2 = \|z^k - z^*\|^2 - h_k - \lambda_k f_k.$$

Adding the above from $k = m$ to $k = m + i$, for $m \geq 0$ and $i \geq 0$, gives

$$\|z^{m+i+1} - z^*\|^2 = \|z^m - z^*\|^2 - \sum_{k=m}^{m+i} (h_k + \lambda_k f_k). \tag{4.5.22}$$

If $m = 0$ and $i \to \infty$, (4.5.21) and (4.5.22) imply that

$$\sum_{k=0}^{\infty} \lambda_k f_k < \infty, \qquad \sum_{k=0}^{\infty} f_k < \infty \tag{4.5.23}$$

Now, T_1^{-1} is strongly monotone at z^*, hence, we obtain from (4.5.19) and (4.5.23) that

$$\lim_{k\to\infty} \|y^k - y^*\| = 0, \qquad \lim_{k\to\infty} \|T_1(z^k) - T_1(z^*)\| = 0. \tag{4.5.24}$$

(4.1.3) and (4.5.17) imply that $\{z^k - z^*\}$ is bounded. Since cl[dom(T_2)] is compact, there is a subsequence $\{z^{k(j)}\}$ of $\{z^k\}$, such that, $\{z^{k(j)}\}$ strongly converges to a point $\bar{z} \in$ cl[dom(T_2)]. We know from (4.5.16) that $\{w^{k(j)-1}\}$ also strongly converges $\bar{z}$. Lemma 4.1.8 and (4.5.24) imply that $y^* \in T_2(\bar{z})$ and $T_1(z^*) = T_1(\bar{z})$. Therefore, $0 \in T(\bar{z})$. In (4.5.16), we can replace z^* by $\bar{z}$ now and obtain the same inequalities. This implies that $\{z^k\}$ actually strongly converges to $\bar{z}$.

Part (d): The same proof as Part (c) until we arrive to (4.5.24). Now (4.1.3) and (4.5.17) imply that $\{z^k - z^*\}$ is bounded. Hence, there is a subsequence $\{z^{k(j)}\}$ of $\{z^k\}$, such that, $\{z^{k(j)}\}$ weakly converges to a point $\bar{z}$. We know from (4.5.16) that $\{w^{k(j)-1}\}$ also weakly converges $\bar{z}$.Lemma 4.1.8 and (4.5.24) imply that $y^* \in T_2(\bar{z})$ and $T_1(z^*) = T_1(\bar{z})$. Therefore, $0 \in T(\bar{z})$. In (4.5.16), we can replace z^* by $\bar{z}$ now and obtain the same inequalities. Combining with (4.5.18) and (4.5.24), implies that $\{z^k\}$ actually weakly converges to $\bar{z}$. $\square$

It is obvious that Part (a) of Theorem 4.5.6 covers the results of Theorem 4.5.3 and Part (c) and Part (d), of Theorem 4.5.6, cover the results of Theorem 4.5.4.

A sufficient condition for the nonexpansivity of G_k is that $I - \lambda_k T_1$ is nonexpansive. It is not necessary. In Section 4.6 we shall prove that $I - \lambda_k T_1$ is nonexpansive, if and only if, T_1^{-1} is strongly monotone with modulus $\lambda_k/2$.

4.6. Nonexpansivity and Conditions for Strong Monotonicity

In Section 4.5 we have known that the conditions A to D, for the convergence of the FBS method, mainly depend on the strong monotonicity of the operators (may include Lipschitz continuity) or inverse operators. We shall give some equivalent conditions for (strong) monotonicity of operators by nonexpansivity or contractivity of operators.

We first introduce a theorem recently offered by Rockafellar on the relationship between nonexpansive (NE) and monotone (MT) operators.

NE-MT THEOREM. *The following relations between two operators S and T on $\mathcal{H}$ are equivalent and such that S is nonexpansive, if and only if, T is monotone:*

(a) $S = 2(I+T)^{-1} - I$,

(b) $S = I - 2(I+T^{-1})^{-1}$,

(c) $T = (I+S)^{-1} \cdot 2I - I$,

(d) $\mathrm{gph}(S) = J(\mathrm{gph}(T))$, *where* $J(x,v) = (x+v, x-v)$. *Moreover,*

$$\bar{x} \in S(\bar{x}) \Longleftrightarrow 0 \in T(\bar{x}).$$

□

Now let us consider conditions for monotonicity of T and T^{-1}.

Theorem 4.6.1. *T^{-1} is strongly monotone with modulus $\lambda > 0$, if and only if, $T^{-1} - \lambda I$ is monotone iff $I - 2\lambda T$ is nonexpansive.*

Proof. From Part (b) of NE-MT Theorem, $I - 2\lambda T$ is nonexpansive

$$\Longleftrightarrow I - 2\lambda T = I - 2(I+R)^{-1} \quad \text{for some} \quad R \quad \text{monotone},$$

$$\Longleftrightarrow \lambda T = (I+R)^{-1} \quad \text{for some} \quad R \quad \text{monotone},$$

$$\Longleftrightarrow T^{-1} = (I+R)\cdot\lambda I \quad \text{for some} \quad R \quad \text{monotone},$$

$$\Longleftrightarrow T^{-1} - \lambda I = R\cdot\lambda I \quad \text{for some} \quad R \quad \text{monotone},$$

$$\Longleftrightarrow T^{-1} - \lambda I \quad \text{monotone},$$

$$\Longleftrightarrow T^{-1} \quad \text{strongly monotone with modulus} \quad \lambda.$$

□

Theorem 4.6.2. *An operator T on $\mathcal{H}$ is Lipschitz continuous and strongly monotone, if and only if, there exists a $\lambda > 0$ such that $I - \lambda T$ is contractive.*

Proof. (Necessity:)
Let M be the Lipschitz constant of T and α be the modulus for the strong monotonicity of T. Then, similar to the proof of Theorem 4.3.4 with $T_1 = T$ and $T_2 = 0$, for any $y_i \in (I - \lambda T)(x_i)$ with $i = 1, 2$ we see that

$$\|y_1 - y_2\|^2 \leq (\lambda^2 M^2 + 1 - 2\alpha\lambda)\|x_1 - x_2\|^2$$

where

$$\lambda^2 M^2 + 1 - 2\alpha\lambda < 1$$

when $0 < \lambda < \frac{2\alpha}{M^2}$.
(Sufficiency:)
If there exist constants λ and c, such that $0 \leq c < 1$, and

$$\|y_1 - y_2\|^2 \leq c\|x_1 - x_2\|^2 \tag{4.6.1}$$

where $y_i \in (I - \lambda T)(x_i)$ with $i = 1, 2$, then there exists $z_i \in T(x_i)$, such that $y_i = x_i - \lambda z_i$ for $i = 1, 2$. Hence (4.6.1) is

$$\begin{aligned} c\|x_1 - x_2\|^2 &\geq \|x_1 - x_2 - \lambda(z_1 - z_2)\|^2 \\ &= \|x_1 - x_2\|^2 + \lambda^2\|z_1 - z_2\|^2 - 2\lambda\langle x_1 - x_2, z_1 - z_2\rangle \end{aligned}$$

That is

$$(1-c)\|x_1 - x_2\|^2 + \lambda^2\|z_1 - z_2\|^2 \leq 2\lambda\langle x_1 - x_2, z_1 - z_2\rangle$$

The above implies that T is strongly monotone with modulus $\frac{(1-c)}{2\lambda}$ and that T is Lipschitz continuous with constant $\frac{2}{\lambda}$ since

$$\langle x_1 - x_2, z_1 - z_2\rangle \leq \|x_1 - x_2\|\|z_1 - z_2\|$$

□

Lemma 4.6.3. *The operator $I - 2T$ on $\mathcal{H}$ is nonexpansive if and only if there exists a monotone operator S such that*

$$I - T = (I + S)^{-1} \quad \text{and} \quad T = (I + S^{-1})^{-1}. \tag{4.6.2}$$

Proof. NE-MT Theorem implies that $I - 2T$ is nonexpansive if and only if there exists a monotone S such that

$$I - 2T = 2(I + S)^{-1} - I = I - 2(I + S^{-1})^{-1}.$$

However, this is equivalent to (4.6.2). □

Lemma 4.6.4. *If operator T is monotone, then*

$$(I + T)^{-1} = I - (I + T^{-1})^{-1}. \tag{4.6.3}$$

Proof. (4.6.3) comes from the NE-MT Theorem since

$$2(I + T)^{-1} - I = I - 2(I + T^{-1})^{-1}.$$

□

Lemma 4.6.5. *If an operator T on $\mathcal{H}$ is (maximal or strongly) monotone, then for each $\lambda > 0$ $(I + \lambda T)^{-1}$ is (maximal nonexpansive or contractive) nonexpansive.*

Proof. Let $y_i = (I + \lambda T)^{-1}(x_i)$ for $i = 1, 2$. Then

$$x_i - y_i \in \lambda T(y_i) \quad \text{for} \quad i = 1, 2.$$

Hence

$$\begin{aligned}\|x_1 - x_2\|^2 &= \|y_1 - y_2 + x_1 - y_1 - (x_2 - y_2)\|^2 \\ &\geq \|y_1 - y_2\|^2 + 2\langle y_1 - y_2, x_1 - y_1 - (x_2 - y_2)\rangle \\ &\geq \|y_1 - y_2\|^2.\end{aligned}$$

This implies the (maximal) nonexpansivity.

If there exists a constant $\alpha > 0$ such that for any $z_i \in T(x_i)$

$$\langle x_1 - x_2, z_1 - z_2\rangle \geq \alpha\|x_1 - x_2\|^2,$$

then for any $\lambda > 0$ let $y_i = x_i + \lambda z_i$ with $i = 1, 2$. It is easy to see that $x_i = (I + \lambda T)^{-1}(y_i)$ and

$$\begin{aligned}\|y_1 - y_2\|^2 &= \|x_1 - x_2 + \lambda(z_1 - z_2)\|^2 \\ &= \|x_1 - x_2\|^2 + 2\lambda\langle x_1 - x_2, z_1 - z_2\rangle + \lambda^2\|z_1 - z_2\|^2 \\ &\geq (1 + 2\lambda\alpha + \lambda^2\alpha^2)\|x_1 - x_2\|^2 \\ &= (1 + \lambda\alpha)^2\|x_1 - x_2\|^2\end{aligned}$$

Therefore

$$\|x_1 - x_2\|^2 \leq \frac{1}{(1 + \lambda\alpha)^2}\|y_1 - y_2\|^2$$

That is, $(I + \lambda T)^{-1}$ is contractive. □

Proposition 4.6.6. *If T is Lipschitz Continuous with M and strongly monotone with α, then T^{-1} is also strongly monotone with modulus $\beta = \alpha M^{-2}$.*

Proof. For any $z_i \in T(x_i)$ with $i = 1, 2$, by assumptions we have

$$\langle x_1 - x_2, z_1 - z_2\rangle \geq \alpha\|x_1 - x_2\|^2,$$

$$\|z_1 - z_2\| \leq M\|x_1 - x_2\|.$$

Hence

$$\begin{aligned}\langle x_1 - x_2, z_1 - z_2\rangle &\geq \frac{\alpha}{M^2}M^2\|x_1 - x_2\|^2 \\ &\geq \frac{\alpha}{M^2}\|z_1 - z_2\|^2\end{aligned}$$

That is, T^{-1} is strongly monotone with modulus αM^{-2}. $\square$

4.7. The Local Convergence and Rate for the FBS Method

Definition 4.7.1. *A monotone multivalue operator T is said to be (locally) strongly monotone at z^* with modulus α, if $\alpha > 0$ and there is a $\delta > 0$, such that*

$$\langle z - z^*, w - \bar{w} \rangle \geq \alpha \|z - z^*\|^2 \quad \forall w \in T(z),\ \bar{w} \in T(z^*),\quad \|z - z^*\| \leq \delta. \qquad (4.7.1)$$

Similarly, we have the following:

Definition 4.7.2. *We declare an operator T to be (locally) Lipschitz continuous at z^* with constant M, if there is a $\eta > 0$ such that*

$$\|w - \bar{w}\| \leq M\|z - z^*\| \quad \forall w \in T(z), \quad \bar{w} \in T(z^*), \quad \|z - z^*\| \leq \eta. \qquad (4.7.2)$$

Actually we may replace $\bar{w}$ by $T(z^*)$ since it is evident that T is single-valued at z^* from (4.7.2).

Proposition 4.7.1. *Let $T = T_1 + T_2$ with T_1 and T_2 monotone. If T_1 or T_2 is strongly monotone at z^* with modulus α, then T is also strongly monotone at z^* with α and there exists a $\delta > 0$ such that $T(z) \cap T(z^*) = \emptyset$ for any $z \neq z^*$ and $\|z - z^*\| \leq \delta$. Furthermore,*

$$\|w - \bar{w}\| \geq \alpha\|z - z^*\| \quad \forall w \in T(z), \quad \bar{w} \in T(z^*), \quad \|z - z^*\| \leq \delta. \qquad (4.7.3)$$

Proof. The strong monotonicity of T at z^* is obvious. Hence there is a δ such that (4.7.1) is true. Now assume that there is a $z \neq z^*$ and $\|z - z^*\| \leq \delta$ such that $T(z) \cap T(z^*) \neq \emptyset$. Then let $w \in T(z) \cap T(z^*)$ and we see from (4.7.1) that

$$0 = \langle z - z^*, w - w \rangle \geq \alpha\|z - z^*\|^2.$$

But this is impossible since $z \neq z^*$. Thus, our assumption is wrong.

(4.7.3) is a direct implication of (4.7.1). $\square$

Definition 4.7.3. *If an operator T satisfies at a point z^* for a constant c*

$$\|w - \bar{w}\| \leq c\|z - z^*\|, \quad \forall w \in T(z), \quad \bar{w} \in T(z^*), \quad \|z - z^*\| \leq \kappa \tag{4.7.4}$$

for some $\kappa > 0$, then the operator T is declared (locally) nonexpansive at z^*, if $c = 1$ and (locally) contractive at z^* with constant c if $c < 1$.

Lemma 4.7.2. *If T is nonexpansive at z^*, or contractive at z^* with constant c, then T is single-valued at z^* and Lipschitz continuous at z^* with constant $M = 1$, or $M = c$.*

Proof. Lipschitz continuity at z^* is obvious with $\eta = \kappa$ in (4.7.2). If $w \in T(z^*)$ and $w' \in T(z^*)$, then

$$\|w - w'\| \leq \|z^* - z^*\| = 0$$

Therefore, $w = w'$. □

Lemma 4.7.3. *Suppose $T : \mathcal{H} \rightrightarrows \mathcal{H}$ is contractive at z^* with constant c and z^* is a fixed point of T, i.e., $z^* = T(z^*)$. Then z^* is an isolated fixed point and any sequence $\{z^k\}$ generated by $z^{k+1} = T(z^k)$, for some $\kappa > 0$, and any starting point z^0 with $\|z^0 - z^*\| \leq \kappa$ converges to z^* Q-linearly with ratio c.*

Proof. Since T is contractive at z^* with c, there is a $\kappa > 0$ such that (4.7.4) is true. Let $\{z^k\}$ be any sequence from $z^{k+1} = T(z^k)$ with $\|z^0 - z^*\| \leq \kappa$. Then

$$\|z^k - z^*\| \leq c\|z^{k-1} - z^*\| \leq c^k\|z^0 - z^*\|.$$

$c < 1$ implies that $c^k \to 0$ when $k \to \infty$ and $\{z^k\}$ Q-linearly converges to z^* with ratio c. Now if there is another fixed point $\bar{z}$ of T with $\|\bar{z} - z^*\| \leq \kappa$, then

$$\|z^* - \bar{z}\| = \|T(z^*) - T(\bar{z})\| \leq c\|z^* - \bar{z}\|.$$

But $c < 1$. Thus $\bar{z} = z^*$. That is, the fixed point z^* of T is unique in the neighborhood $\{z : \|z - z^*\| \leq \kappa\}$ of z^*. □

Definition 4.7.4. *For a set of operators $\{T_i \,|\, i \in S\}$ with subscript set S, we declare they are (locally) uniformly contractive at a point z^*, if every T_i is contractive at z^* with the same constant $c < 1$ and the same κ in (4.7.4) for all $i \in S$.*

Theorem 4.7.4. *If $\{T_k\}_{k=0}^{\infty}$ is a uniformly contractive operator sequence at z^* with constant c, and z^* is a fixed point for every T_k, then any sequence $\{z^k\}$ generated by $z^{k+1} = T_k(z^k)$ Q-linearly converges to their isolated fixed point z^* with ratio c for some $\kappa > 0$ and any $\|z^0 - z^*\| \le \kappa$.*

Proof. From Lemma 4.7.3, z^* is an isolated fixed point of every T_k. Let κ be the constant such that (4.7.4) is true for every T_k. Then when $\|z^0 - z^*\| \le \kappa$,

$$\|z^k - z^*\| = \|T_{k-1}(z^{k-1}) - T_{k-1}(z^*)\| \le c\|z^{k-1} - z^*\| \le c^k\|z^0 - z^*\|.$$

$c < 1$ implies that $\{z^k\}$ Q-linearly converges to z^* with ratio c. □

Now let us consider the local convergence of the FBS method by applying the above theorem from local contractivity of the forward-backward operator G_k.

Theorem 4.7.5. *Let $z^* \in T^{-1}(0)$ and $T = T_1 + T_2$ with T_1 and T_2 monotone. If Condition C is locally satisfied at z^* for T_1 and $\{\lambda_k\}$, then*

(a) z^* is a fixed point of every forward-backward operator $G_k = (I + \lambda_k T_2)^{-1}(I - \lambda_k T_1)$ and $\{G_k\}$ is locally uniformly contractive at z^* with constant θ:

$$\theta^2 = \sup_{k \ge 0}(\lambda_k^2 M^2 + 1 - 2\alpha\lambda_k) < 1;$$

(b) there exists a $\kappa > 0$, such that the sequence $\{z^k\}$ generated through the BFBS method $z^{k+1} = G_k(z^k)$ will Q-linearly converge to z^* with ratio θ for any starting point z^0 with $\|z^0 - z^*\| \le \kappa$.

Proof. Proposition 2.10.1 proved that z^* is a fixed point of every G_k. Since T_1 is strongly monotone with α at z^* and Lipschitz continuous with M at z^*, there are a $\delta > 0$ and a $\eta > 0$ such that (4.7.1) and (4.7.2) hold. Let $\kappa = \min\{\delta, \eta\}$, $\bar{w} = T_1(z^*)$, $y^* = -T_1(z^*) \in T_2(z^*)$ and $w \in T_1(z)$ with $\|z - z^*\| \le \kappa$. If $x^k = G_k(z)$, there is a $y^k \in T_2(x^k)$ for any $k \ge 0$, such that

$$x^k + \lambda_k y^k = z - \lambda_k w.$$

Thus

$$x^k - z^* + \lambda_k(y^k - y^*) = z - z^* - \lambda_k(w - \bar{w}).$$

Hence

$$\begin{aligned}\|x^k - z^*\|^2 &\le \|x^k - z^* + \lambda_k(y^k - y^*)\|^2 \\ &= \|z - z^* - \lambda_k(w - \bar{w})\|^2 \\ &\le \|z - z^*\|^2 + \lambda_k^2\|w - \bar{w}\|^2 - 2\lambda_k\langle z - z^*, w - \bar{w}\rangle \\ &\le (\lambda_k^2 M^2 + 1 - 2\alpha\lambda_k)\|z - z^*\|^2\end{aligned}$$

The uniform contractivity constant θ from (4.3.1) so

$$\theta^2 = \sup_{k\ge 0}(\lambda_k^2 M^2 + 1 - 2\alpha\lambda_k) = 1 - \inf_{k\ge 0} M^2\lambda_k(2M^{-2}\alpha - \lambda_k) < 1.$$

Part (b) is just the direct corollary of the above part (a) and Theorem 4.7.4 for $\kappa = \min\{\delta, \eta\} > 0$. ☐

Theorem 4.7.6. *Let $z^* \in T^{-1}(0)$ and $T = T_1 + T_2$ with T_1 and T_2 monotone. If Condition A of Section 4.3 is locally satisfied at z^* for T_1, T_2 and $\{\lambda_k\}$, then*

(a) z^* is a fixed point of every forward-backward operator $G_k = (I+\lambda_k T_2)^{-1}(I-\lambda_k T_1)$, and $\{G_k\}$ is locally uniformly contractive at z^* with constant θ:

$$\theta^2 = \sup_{k\ge 0}\frac{1 + \lambda_k^2 M^2}{(1 + \alpha\lambda_k)^2} < 1;$$

(b) there exists a $\kappa > 0$ such that the sequence $\{z^k\}$ generated through the BFBS method $z^{k+1} = G_k(z^k)$ will Q-linearly converge to z^* with ratio θ for any starting point z^0 with $\|z^0 - z^*\| \le \kappa$.

Proof. Proposition 2.10.1 proved that z^* is a fixed point of every G_k. Since T_2 is strongly monotone with α at z^*, and T_1 is Lipschitz continuous with M at z^*, there exist a $\delta > 0$ and a $\eta > 0$, such that (4.7.1) for T_2 and (4.7.2) for T_1 hold respectively. Let $\kappa = \min\{\delta, \eta\}$, $\bar{w} = T_1(z^*)$, $y^* = -T_1(z^*) \in T_2(z^*)$ and $w \in T_1(z)$ with $\|z - z^*\| \le \kappa$. If $x^k = G_k(z)$, there exists a $y^k \in T_2(x^k)$ for any $k \ge 0$, such that

$$x^k + \lambda_k y^k = z - \lambda_k w.$$

Thus

$$x^k - z^* + \lambda_k(y^k - y^*) = z - z^* - \lambda_k(w - \bar{w}).$$

Hence

$$\begin{aligned}(1+\alpha\lambda_k)^2\|x^k - z^*\|^2 &\le \|x^k - z^* + \lambda_k(y^k - y^*)\|^2 \\ &= \|z - z^* - \lambda_k(w - \bar{w})\|^2 \\ &\le \|z - z^*\|^2 + \lambda_k^2\|w - \bar{w}\|^2 \\ &\le (1 + \lambda_k^2 M^2)\|z - z^*\|^2\end{aligned}$$

Therefore from the above and (4.2.3), the uniform contractivity constant θ is

$$\theta^2 = \sup_{k\ge 0} \frac{1 + \lambda_k^2 M^2}{(1+\alpha\lambda_k)^2} < 1.$$

Part (b) is just the direct corollary of the above part (a) and Theorem 4.7.4 for $\kappa = \min\{\delta, \eta\} > 0$. □

Theorem 4.7.7. *Let $z^* \in T^{-1}(0)$ and $T = T_1 + T_2$ with T_1 and T_2 monotone. If Condition C of Section 4.3 is locally satisfied at z^* for T_1, T_2, and $\{\lambda_k\}$, then*

(a) *z^* is a fixed point of every forward-backward operator $G_k = (I+\lambda_k T_2)^{-1}(I - \lambda_k T_1)$ and $\{G_k\}$ is locally uniformly contractive at z^* with constant θ:*

$$\theta^2 = \sup_{k\ge 0} \frac{1 + \lambda_k^2 M^2 - 2\alpha_1\lambda_k}{(1+\alpha_2\lambda_k)^2} < 1;$$

(b) *there exists a $\kappa > 0$ such that the sequence $\{z^k\}$ generated through the BFBS method $z^{k+1} = G_k(z^k)$ will Q-linearly converge to z^*, with ratio θ, for any starting point z^0 with $\|z^0 - z^*\| \le \kappa$.*

Proof. Proposition 2.10.1 proved that z^* is a fixed point of every G_k. Since T_2 is strongly monotone, with α_2 at z^*, and T_1 is Lipschitz continuous with M and strongly monotone with α_1 at z^*, there exist $\delta_1 > 0$, $\delta_2 > 0$, and a $\eta > 0$ such that (4.7.1) for T_1, T_2, and (4.7.2) for T_1 hold respectively. Let $\kappa = \min\{\delta_1, \delta_2, \eta\}$, $\bar{w} = T_1(z^*)$, $y^* = -T_1(z^*) \in T_2(z^*)$ and $w \in T_1(z)$ with $\|z - z^*\| \le \kappa$. If $x^k = G_k(z)$, there exists a $y^k \in T_2(x^k)$ for any $k \ge 0$ such that

$$x^k + \lambda_k y^k = z - \lambda_k w.$$

Thus

$$x^k - z^* + \lambda_k(y^k - y^*) = z - z^* - \lambda_k(w - \bar{w}).$$

Hence

$$\begin{aligned}(1+\alpha_2\lambda_k)^2\|x^k - z^*\|^2 &\le \|x^k - z^* + \lambda_k(y^k - y^*)\|^2 \\ &= \|z - z^* - \lambda_k(w - \bar{w})\|^2 \\ &\le (1 + \lambda_k^2 M^2 - 2\alpha_1\lambda_k)\|z - z^*\|^2\end{aligned}$$

Therefore, from the above and (4.2.11), the uniform contractivity constant θ is

$$\theta^2 = \sup_{k\ge 0} \frac{1 + \lambda_k^2 M^2 - 2\alpha_1\lambda_k}{(1+\alpha_2\lambda_k)^2} < 1.$$

Part (b) is just the direct corollary of the above part (a) and Theorem 4.7.4 for $\kappa = \min\{\delta_1, \delta_2, \eta\} > 0$. □

Therefore, from Theorem 4.7.4, it is simply seen that the BFBS method Q-linearly converges to the zero z^* of T, if Condition A, or Condition B, or Condition C is locally satisfied at z^*.

Now let us consider Condition D of Section 4.3.

Theorem 4.7.8. *Let $z^* \in T^{-1}(0)$ and $T = T_1 + T_2$ with T_1 and T_2 monotone. If z^* is an isolated zero point of T, and Condition D of Section 4.3 is locally satisfied at z^* for T_1 and $\{\lambda_k\}$, then the forward-backward operators G_k, from (4.1.1), are nonexpansive at z^* for all $k \ge 0$, and there exists a $\kappa > 0$, such that any sequence $\{z^k\}$ generated from the BFBS method, for any starting point z^0 with $\|z^0 - z^*\| \le \kappa$ weakly converges to z^*, with the same estimation rate as those in Part (b) of Theorem 4.2.6.*

Proof. Proposition 2.10.1 proved that z^* is a fixed point of every G_k. Since T_1^{-1} is strongly monotone, with β at z^*, and z^* is an isolated zero point of T, there exist $\delta > 0$ and $\eta > 0$, such that z^* is the only zero point of T, or the only fixed point of G_k, in $\{z : \|z - z^*\| \le \eta\}$ and for any z with $\|z - z^*\| \le \delta$

$$\langle z - z^*, w - w^*\rangle \ge \beta\|w - w^*\|^2 \quad \text{for all} \quad w \in T_1(z) \quad \text{and} \quad w^* = T_1(z^*).$$

Let $\kappa = \min\{\delta, \eta\}$, $y^* = -T_1(z^*) \in T_2(z^*)$ and $w \in T_1(z)$ with $\|z - z^*\| \leq \kappa$. If $x^k = G_k(z)$, there exists a $y^k \in T_2(x^k)$, for any $k \geq 0$, such that

$$x^k + \lambda_k y^k = z - \lambda_k w.$$

Thus

$$x^k - z^* + \lambda_k(y^k - y^*) = z - z^* - \lambda_k(w - w^*).$$

Hence

$$\begin{aligned} &\|x^k - z^*\|^2 + \lambda_k^2\|y^k - y^*\| \\ \leq &\|x^k - z^* + \lambda_k(y^k - y^*)\|^2 \\ = &\|z - z^* - \lambda_k(w - w^*)\|^2 \\ \leq &\|z - z^*\|^2 - \lambda_k(2\beta - \lambda_k)\|w - w^*\|^2. \end{aligned} \tag{4.7.5}$$

Therefore, from the above and (2.10.9), the nonexpansivity of G_k at z^* is obvious.

Now, if $\|z^0 - z^*\| \leq \kappa$ and $z^{k+1} = G_k(z^k)$, then replacing z, w and x^k in (4.7.5) by z^k, w^k and z^{k+1}, respectively, gives

$$\begin{aligned} &\|z^{k+1} - z^*\|^2 + \lambda_k^2\|y^k - y^*\|^2 \\ \leq &\|z^k - z^*\|^2 - \lambda_k(2\beta - \lambda_k)\|w^k - w^*\|^2. \end{aligned} \tag{4.7.6}$$

Thus, for any $k \geq 0$ $\|z^k - z^*\| \leq \|z^0 - z^*\| \leq \kappa$, hence, there exists a subsequence $\{z^{k(j)}\}$ of $\{z^k\}$ and a z^∞ such that

$$\|z^\infty - z^*\| \leq \kappa \quad \text{and} \quad z^{k(j)} \to z^\infty \quad \text{weakly as} \quad j \to \infty.$$

Let $\mu = \inf_{k \geq 0}\{\lambda_k, 2\beta - \lambda_k\}$. Then from (2.10.9) $\mu > 0$. Adding (4.7.6), from $k = 0$ to $k = m$, implies that

$$\begin{aligned} &\|z^{m+1} - z^*\|^2 + \mu^2 \sum_{k=0}^{m} \|y^k - y^*\|^2 \\ \leq &\|z^0 - z^*\|^2 - \mu^2 \sum_{k=0}^{m} \|w^k - w^*\|^2. \end{aligned} \tag{4.7.7}$$

Therefore,

$$\lim_{k \to \infty} y^k = y^*, \qquad \lim_{k \to \infty} w^k = w^*. \tag{4.7.8}$$

Lemma 4.1.8 implies that $y^* \in T_2(z^\infty)$ and $w^* \in T_1(z^\infty)$. That is, $0 \in T(z^\infty)$. Again z^* being the only zero point of T in $\{z : \|z - z^*\|\}$ gives $z^\infty = z^*$. So the only weak cluster point of $\{z^k\}$ is z^*, from (4.7.8) and Lemma 4.1.8. Therefore, $\{z^k\}$ weakly converges to z^* by the boundedness of $\{z^k\}$. The estimation rate of convergence is the same as those in Part (b) of Theorem 4.2.6 from (4.7.7). $\square$

4.8. On the Superlinear Convergence of the Basic FBS Method

Now let us consider the superlinear convergence. We declare that a convergent sequence $\{z^k\}$ converges *(Q-)superlinearly* to its limit $\bar{z}$, if

$$\lim_{k\to\infty} \|z^{k+1} - \bar{z}\|/\|z^k - \bar{z}\| = 0. \tag{4.8.1}$$

Theorem 4.8.1. *Suppose T_1 is single-valued and $\{z^k\}$ be a convergent sequence, with limit $z^* \in T^{-1}(0)$, generated through the BFBS algorithm for maximal monotone T_2. Let $y^* = -T_1(z^*)$ and $y^{k+1} = (1/\lambda_k)(z^k - z^{k+1}) - T_1(z^k)$. Then, $y^k \in T_2(z^k)$ for $k = 1, 2, \ldots,$ and for given $\{\lambda_k\}$*

(a) $\{z^k\}$ converges superlinearly, if and only if,

$$\begin{aligned}\lim_{k\to\infty} \Big[2\lambda_k \frac{\langle z^k - z^*, T_1(z^k) - T_1(z^*)\rangle + \langle z^{k+1} - z^*, y^{k+1} - y^*\rangle}{\|z^k - z^*\|^2} \\ + \lambda_k^2 \frac{\|y^{k+1} - y^*\|^2 - \|T_1(z^k) - T_1(z^*)\|^2}{\|z^k - z^*\|^2}\Big] = 1\end{aligned}. \tag{4.8.2}$$

(b) if

$$\lim_{k\to\infty} \frac{2\lambda_k\langle z^k - z^*, T_1(z^k) - T_1(z^*)\rangle - \lambda_k^2\|T_1(z^k) - T_1(z^*)\|^2}{\|z^k - z^*\|^2} = 1, \tag{4.8.3}$$

then $\{z^k\}$ converges superlinearly to z^*. Furthermore, if there exists a k_0 and a constant C such that

$$\|y^k - y^*\| \le C\|z^k - z^*\| \quad \textit{for} \quad k \ge k_0 \tag{4.8.4}$$

additionally, then (4.8.3) is also a necessary condition for the superlinear convergence of $\{z^k\}$.

Proof. From the proof of Theorem 4.2.2 we know that (4.2.9) is also true here. Hence, (4.2.9) and (4.8.1) imply Part (a) (here we do not even require that T_1 and T_2 be monotone for this conclusion).

For the proof of Part (b) we know that T_2 is maximal monotone. From this and (4.2.9) and (4.8.3), we obtain (4.8.1). That is, $\{z^k\}$ converges superlinearly. Conversely, (4.2.9) and the given additional condition (4.8.4) imply for $k \geq k_0$ that

$$\begin{aligned} &\|z^k - z^*\|^2 - 2\lambda_k \langle z^k - z^*, T_1(z^k) - T_1(z^*)\rangle \\ &+ \lambda_k^2 \|T_1(z^k) - T_1(z^*)\|^2 \leq (1 + C\lambda_k)^2 \|z^{k+1} - z^*\|^2. \end{aligned} \tag{4.8.5}$$

If (4.8.1) is true, then (4.8.1) and (4.8.5) imply (4.8.3). Hence, (4.8.3) is also necessary. □

Theorem 4.8.1 shows us the nature of the superlinear convergence of the BFBS algorithm. Thus, if we want to develop a superlinear convergence BFBS algorithm for a certain class of problems, we may consider development from (4.8.2) or (4.8.3). However, because we do not know what z^{k+1} is at the k-th iteration, (4.8.2) is harder to apply than (4.8.3). It seems that our choices are very limited. Our only choice is to determine the step sizes λ_k, since T_1 and T_2 are determined before the iteration begins. For the choice of λ_k, only the sufficient condition (4.8.3) for the superlinear convergence of the $\{z^k\}$ may be applicable.

Note, in (4.8.3), why did we use $\lim_{k\to\infty}$ rather than $\liminf_{k\to\infty}$? This was due to the following property.

Proposition 4.8.2. *Let* $\{z^k\}$ *be any sequence generated through the BFBS algorithm. Then, for* $z^* \in T^{-1}(0)$, *we have*

$$\limsup_{k\to\infty} \frac{2\lambda_k \langle z^k - z^*, T_1(z^k) - T_1(z^*)\rangle - \lambda_k^2 \|T_1(z^k) - T_1(z^*)\|^2}{\|z^k - z^*\|^2} \leq 1. \tag{4.8.6}$$

Hence

$$\liminf_{k\to\infty} \frac{2\lambda_k \langle z^k - z^*, T_1(z^k) - T_1(z^*)\rangle - \lambda_k^2 \|T_1(z^k) - T_1(z^*)\|^2}{\|z^k - z^*\|^2} = 1$$

if and only if (4.8.3) is true.

Proof. From (4.2.9) we have

$$1 - \frac{2\lambda_k\langle z^k - z^*, T_1(z^k) - T_1(z^*)\rangle - \lambda_k^2\|T_1(z^k) - T_1(z^*)\|^2}{\|z^k - z^*\|^2} \geq 0. \tag{4.8.7}$$

Therefore, (4.8.6) is true. □

Now let us consider a couple of general equivalent conditions for superlinear convergence.

Theorem 4.8.3. *A sequence $\{z^k\}$ superlinearly converges to a point $\bar{z}$, if and only if, one of the following conditions is satisfied:*

(a)

$$\lim_{k\to\infty} \frac{\|z^{k+1} - z^k\|^2 + 2\langle z^{k+1} - z^k, z^k - \bar{z}\rangle}{\|z^k - \bar{z}\|^2} = -1; \tag{4.8.8}$$

(b)

$$\lim_{k\to\infty} \frac{\|z^k - z^{k+1}\|}{\|z^k - \bar{z}\|} = 1, \qquad \lim_{k\to\infty} \frac{\langle z^k - z^{k+1}, z^k - \bar{z}\rangle}{\|z^k - \bar{z}\|^2} = 1; \tag{4.8.9}$$

(c)

$$\lim_{k\to\infty} \frac{\|z^{k+1} - \bar{z}\|}{\|z^k - z^{k+1}\|} = 0. \tag{4.8.10}$$

Proof. Since

$$\|z^{k+1} - \bar{z}\|^2 = \|z^{k+1} - z^k\|^2 + \|z^k - \bar{z}\|^2 + 2\langle z^{k+1} - z^k, z^k - \bar{z}\rangle, \tag{4.8.11}$$

(4.8.1) implies that $\{z^k\}$ superlinearly converges to $\bar{z}$, if and only if, Condition (a) is true.

On the other hand,

$$\|z^{k+1} - z^k\|^2 = \|z^{k+1} - \bar{z}\|^2 + \|z^k - \bar{z}\|^2 - 2\langle z^{k+1} - \bar{z}, z^k - \bar{z}\rangle. \tag{4.8.12}$$

Hence, (4.8.1) and (4.8.8) imply that condition (a) and Condition (b) are equivalent.

In the last case, (4.8.10) can easily be obtained from (4.8.9) and (4.8.1). Conversely, if (4.8.10) is true, (4.8.1) can be implied from

$$\frac{\|z^k - \bar{z}\|^2}{\|z^k - z^{k+1}\|^2} = 1 + \frac{\|z^{k+1} - \bar{z}\|^2}{\|z^k - z^{k+1}\|^2} + 2\frac{\langle z^k - z^{k+1}, z^{k+1} - \bar{z}\rangle}{\|z^k - z^{k+1}\|^2} \to 1 \tag{4.8.13}$$

as $k \to \infty$. This completes the proof. □