Directional asymptotics of Fejér monotone sequences

Heinz H. Bauschke^{*}, Manish Krishan Lal[†], and Xianfu Wang[‡]
June 29, 2021

Abstract

The notion of Fejér monotonicity is instrumental in unifying the convergence proofs of many iterative methods, such as the Krasnoselskii–Mann iteration, the proximal point method, the Douglas-Rachford splitting algorithm, and many others. In this paper, we present directionally asymptotical results of strongly convergent subsequences of Fejér monotone sequences. We also provide examples to show that the sets of directionally asymptotic cluster points can be large and that weak convergence is needed in infinite-dimensional spaces.

2020 Mathematics Subject Classification: Primary 47H09, 47J26, 90C25; Secondary 47H05, 65K10. **Keywords:** Fejér monotone sequence, firmly nonexpansive mapping, normal cone.

1 Introduction

One of the most important tools in studying convergence of iterative methods in optimization and convex analysis is Fejér monotonicity; see, e.g., [1, Chapters 5, 26, 28], [5], [2], [7], [8], [6]. Recently, among many important advances, Rockafellar showed in [11] that in a finite-dimensional Hilbert space the sequences generated by the proximal point algorithm enjoy directionally asymptotic properties. In this paper, we study directional asymptotics of Fejér monotone sequences in Hilbert spaces. Consequently, many iteration methods in [1] and [5], whose convergence analysis relies on the Fejér monotonicity, have these directionally asymptotic behaviour.

The paper is organized as follows. In Section 2, we provide some preliminary results on Fejér monotone sequences useful in subsequent proofs. Our main results on directional asymptotics of Fejér monotone sequences are presented in Section 3. In Section 4, we show that the sets of directional asymptotics of Fejér monotone sequences can be large. We conclude the paper with an infinite-dimensional example illustrating weak without strong convergence in Section 5.

^{*}Mathematics, University of British Columbia, Kelowna, B.C. V1V 1V7, Canada. E-mail: heinz.bauschke@ubc.ca.

[†]Mathematics, University of British Columbia, Kelowna, B.C. V1V 1V7, Canada. E-mail: manish.krishanlal@ubc.ca.

[‡]Mathematics, University of British Columbia, Kelowna, B.C. V1V 1V7, Canada. E-mail: shawn.wang@ubc.ca.

The notation that we employ is for the most part standard and follows [1]; however, a partial list is provided for the reader's convenience. Throughout this paper, we assume that

X is a real Hilbert space,

with inner product $\langle \cdot, \cdot \rangle$ and induced norm $\| \cdot \|$. We use $\mathbb{N} := \{0, 1, 2, \ldots\}$ for set of natural numbers.

Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in X. We denote the set of (weak) cluster points of $(x_n)_{n\in\mathbb{N}}$ by

$$\mathcal{C}((x_n)_{n\in\mathbb{N}}) := \{x \in X \mid x \text{ is the weak limit of some subsequence of } (x_n)_{n\in\mathbb{N}} \}.$$

Of course, if X is finite-dimensional, then $C((x_n)_{n\in\mathbb{N}})$ is the same as the set of strong cluster points of $(x_n)_{n\in\mathbb{N}}$. We write $x_n \to x$ if $(x_n)_{n\in\mathbb{N}}$ converges strongly to x, and $x_n \to x$ if $(x_n)_{n\in\mathbb{N}}$ converges weakly to x. Let C be a subset of X and let $(x_n)_{n\in\mathbb{N}}$ be a sequence in X. Then $(x_n)_{n\in\mathbb{N}}$ is $Fej\acute{e}r$ monotone with respect to C if

$$(\forall c \in C)(\forall n \in \mathbb{N}) \|x_{n+1} - c\| \leqslant \|x_n - c\|,$$

and we also call C a Fejér monotone set of $(x_n)_{n\in\mathbb{N}}$. The corresponding support function of C is defined by $\sigma_C(x) := \sup \langle C, x \rangle$ while the corresponding distance function is $d_C(x) := \inf \|C - x\|$, for every $x \in X$. The polar cone of C is $C^{\ominus} := \{u \in X \mid \sigma_C(u) \leq 0\}$; note that if $z \in C$, then $N_C(z) = (C - z)^{\ominus}$ is the normal cone of C at x. For a set-valued monotone operator $A: X \rightrightarrows X$, the corresponding resolvent is $J_A := (\operatorname{Id} + A)^{-1}$. Finally, the unit sphere is abbreviated by

$$S := \{ x \in X \mid ||x|| = 1 \}.$$

2 Auxiliary results

We start with some preparatory results on Fejér monotone sequences.

Lemma 2.1. Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in X. Then the following hold:

(i) The largest Fejér monotone set of $(x_n)_{n\in\mathbb{N}}$ is the (possibly empty) closed convex set

$$\bigcap_{n\in\mathbb{N}} \{z \in X \mid 2\langle x_n - x_{n+1}, z \rangle \leq ||x_n||^2 - ||x_{n+1}||^2 \},\,$$

and is closed convex.

- (ii) If C is a Fejér monotone set of $(x_n)_{n\in\mathbb{N}}$, then $\overline{\operatorname{conv}}$ C, the closed convex hull of C, is a Fejér monotone set of $(x_n)_{n\in\mathbb{N}}$.
- (iii) If C_1 , C_2 are Fejér monotone sets of $(x_n)_{n\in\mathbb{N}}$, then $C_1 \cup C_2$ is a Fejér monotone set of $(x_n)_{n\in\mathbb{N}}$.
- (iv) If C is a Fejér monotone set of $(x_n)_{n\in\mathbb{N}}$, then C is a Fejér monotone set of every subsequence of $(x_n)_{n\in\mathbb{N}}$.

Proof. (i): Let $z \in X$ and $n \in \mathbb{N}$. Then $||x_{n+1} - z|| \leq ||x_n - z|| \Leftrightarrow ||x_{n+1} - z||^2 \leq ||x_n - z||^2 \Leftrightarrow ||x_{n+1}||^2 + ||z||^2 - 2\langle x_{n+1}, z \rangle \leq ||x_n||^2 + ||z||^2 - 2\langle x_n, z \rangle \Leftrightarrow 2\langle x_n - x_{n+1}, z \rangle \leq ||x_n||^2 - ||x_{n+1}||^2$. (ii)&(iii): These follow from (i). (iv): Obvious from the definition of Fejér monotonicity.

Lemma 2.2. Let C be a nonempty closed convex subset of X, let $\overline{z} \in X$, and let $(x_n)_{n \in \mathbb{N}}$ be a sequence in X. Suppose that $(x_n)_{n \in \mathbb{N}}$ is Fejér monotone with respect to C. Then the following hold for all n, m in \mathbb{N} such that $m \ge n + 1$:

$$(\forall z \in C) \quad \langle x_n - x_{n+1}, z - \overline{z} \rangle \leqslant \frac{1}{2} (\|x_n - \overline{z}\|^2 - \|x_{n+1} - \overline{z}\|^2)$$
 (2.1a)

$$= \langle x_{n+1} - \overline{z}, x_n - x_{n+1} \rangle + \frac{1}{2} ||x_n - x_{n+1}||^2$$
 (2.1b)

$$\leq \|x_{n+1} - \bar{z}\| \|x_n - x_{n+1}\| + \frac{1}{2} \|x_n - x_{n+1}\|^2$$
 (2.1c)

and

$$(\forall z \in C) \quad \langle x_n - x_m, z - \overline{z} \rangle \leqslant \frac{1}{2} (\|x_n - \overline{z}\|^2 - \|x_m - \overline{z}\|^2). \tag{2.2}$$

Proof. Let $z \in C$ and let $k \in \{n, n+1, ..., m\}$. Because $(x_n)_{n \in \mathbb{N}}$ is Fejér monotone with respect to C, we have

$$0 \le \|x_k - z\|^2 - \|x_{k+1} - z\|^2$$

$$= \|x_k - x_{k+1}\|^2 + 2\langle x_k - x_{k+1}, x_{k+1} - z\rangle$$

$$= \|x_k - x_{k+1}\|^2 + 2\langle x_k - x_{k+1}, (x_{k+1} - \overline{z}) - (z - \overline{z})\rangle.$$

Therefore,

$$\langle x_k - x_{k+1}, z - \overline{z} \rangle \le \frac{1}{2} ||x_k - x_{k+1}||^2 + \langle x_k - x_{k+1}, x_{k+1} - \overline{z} \rangle$$
 (2.3a)

$$= \frac{1}{2} (\|x_k - \bar{z}\|^2 - \|x_{k+1} - \bar{z}\|^2)$$
 (2.3b)

which yields (2.1a) and (2.1b). Next, (2.1c) is just Cauchy-Schwarz. Finally, (2.2) follows by summing (2.3) from k = n to k = m and telescoping.

We now localize the set of weak cluster points of a sequence.

Lemma 2.3. Let C be a nonempty subset of X and let $(x_n)_{n\in\mathbb{N}}$ be a sequence in X. Suppose that

$$\overline{\lim}_{n\to\infty}\sigma_C(x_n)\leqslant 0.$$

Then

$$C((x_n)_{n\in\mathbb{N}})\subseteq C^{\Theta}.$$

Proof. Suppose that $x \in \mathcal{C}((x_n)_{n \in \mathbb{N}})$ and to the contrary that $x \notin \mathcal{C}^{\ominus}$. Then $\sigma_{\mathcal{C}}(x) > 0$ and there exists a weakly convergent subsequence $(x_{k_n})_{n \in \mathbb{N}}$ of $(x_n)_{n \in \mathbb{N}}$ such that $x_{k_n} \to x$. The weak lower semicontinuity of $\sigma_{\mathcal{C}}$ now implies

$$0 < \sigma_C(x) \leqslant \underline{\lim}_{n \to \infty} \sigma_C(x_{k_n}) \leqslant \overline{\lim}_{n \to \infty} \sigma_C(x_{k_n}) \leqslant \overline{\lim}_{n \to \infty} \sigma_C(x_n) \leqslant 0,$$

which is absurd!

Lemma 2.4. Suppose that X is finite-dimensional, let C be a nonempty closed subset of X, and let $(x_n)_{n\in\mathbb{N}}$ be a bounded sequence in X. Then

$$d_C(x_n) \to 0 \iff C((x_n)_{n \in \mathbb{N}}) \subseteq C.$$

Proof. " \Rightarrow ": Let x be a cluster point in $\mathcal{C}((x_n)_{n\in\mathbb{N}})$, say $x_{k_n} \to x$. The continuity of d_C and the assumption yield

$$d_C(x) = \lim_{n \to \infty} d_C(x_{k_n}) = \lim_{n \to \infty} d_C(x_n) = 0.$$

Hence $x \in C$ because C is closed.

" \Leftarrow ": Suppose to the contrary that $\overline{\lim}_{n\to\infty} d_C(x_n) > 0$. Then there exists a subsequence $(x_{k_n})_{n\in\mathbb{N}}$ of $(x_n)_{n\in\mathbb{N}}$ such that

$$\lim_{n\to\infty} d_C(x_{k_n}) = \overline{\lim}_{n\to\infty} d_C(x_n) > 0.$$
 (2.4)

Recall that $(x_n)_{n\in\mathbb{N}}$ is bounded and X is finite-dimensional. Using Bolzano-Weierstrass and after passing to another subsequence and relabeling, we may and do assume that $x_{k_n} \to x$. By assumption, $x \in C$. But then $d_C(x_{k_n}) \to d_C(x) = 0$ which contradicts (2.4).

We end this section with results on linear isometries.

Lemma 2.5. Let $A: X \to X$ be a linear isometry. Then the following hold:

- (i) A is injective.
- (ii) If $T := \frac{1}{2} \operatorname{Id} + \frac{1}{2} A$, then $(\forall x \in X) Tx \perp (x Tx)$.
- (iii) If $A^* = A^{-1} = -A$, then $J_A = \frac{1}{2} \operatorname{Id} \frac{1}{2} A$ and $(\forall x \in X) J_A x \perp (x J_A x)$.

Proof. Let *x* ∈ *X*. (i): If Ax = 0, then 0 = ||Ax|| = ||x|| and so x = 0.

(ii): Note that $\operatorname{Id} -T = \frac{1}{2}\operatorname{Id} - \frac{1}{2}$. Hence

$$4\langle Tx, x - Tx \rangle = \langle x + Ax, x - Ax \rangle = ||x||^2 - ||Ax||^2 = 0.$$

(iii): Clearly, $\pm A$ is monotone and $A^2 = -\operatorname{Id}$. Hence, [3, Proposition 2.10] yields $J_A = \frac{1}{2}\operatorname{Id} - \frac{1}{2}A = \operatorname{Id} - T$, where $T = \frac{1}{2}\operatorname{Id} + \frac{1}{2}A$. Now apply (ii).

Corollary 2.6. Let $A: X \to X$ be a linear operator such that $A^* = A^{-1} = -A$, let $x_0 \in X \setminus \{0\}$, and set

$$(\forall n \in \mathbb{N}) \quad x_{n+1} := J_A x_n = \frac{1}{2} x_n - \frac{1}{2} A x_n.$$

Then $x_n \to 0$ and $(\forall n \in \mathbb{N})$ $x_{n+1} \neq x_n$, and

$$(\forall n \in \mathbb{N}) \left\langle \frac{x_{n+1}}{\|x_{n+1}\|}, \frac{x_n - x_{n+1}}{\|x_n - x_{n+1}\|} \right\rangle = 0.$$
 (2.5)

Proof. Clearly, A is a maximally monotone isometry. The formula for J_Ax_n is a consequence of Lemma 2.5(iii) which also yields (2.5) after we prove that the quotients are well defined which we do next. Let $x \in X$. Then $(\operatorname{Id} + A)^{-1}x = J_Ax = 0 \Leftrightarrow x = (\operatorname{Id} + A)(0) = 0$ and $(\operatorname{Id} + A)^{-1}x = J_Ax = x \Leftrightarrow x = (\operatorname{Id} + A)x \Leftrightarrow Ax = 0 \Leftrightarrow x = 0$. We have shown that if $x \neq 0$, then $J_Ax \neq 0$ and $J_Ax \neq x$. A straightforward induction yields $(\forall n \in \mathbb{N}) \ x_n \neq 0$ and $x_{n+1} \neq x_n$, as claimed. Finally, [4, Corollary 1.2] implies that $x_n \to 0$.

Remark 2.7. When $X = \mathbb{R}^2$ and

$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},$$

then Corollary 2.6 recovers [11, the example on page 11]. Note that (see [10, page 206]) a linear isometry need not be surjective.

3 Directional asymptotics of Fejér monotone sequences

We are now ready for our main results on the directionally asymptotic behaviour of Fejér monotone sequences. The proofs significantly extend the reach of those brought to light by Rockafellar in [11, Theorem 2.3].

Theorem 3.1. Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in X that is Fejér monotone with respect to some nonempty closed convex subset Z of X. Suppose that $x_n \to \overline{z} \in X$ and that $(\forall n \in \mathbb{N})$ $x_{n+1} \neq x_n \neq \overline{z}$. Then

$$C\left(\left(\frac{x_n - x_{n+1}}{\|x_n - x_{n+1}\|}\right)_{n \in \mathbb{N}}\right) \cup C\left(\left(\frac{x_n - \overline{z}}{\|x_n - \overline{z}\|}\right)_{n \in \mathbb{N}}\right) \subseteq (Z - \overline{z})^{\ominus}; \tag{3.1}$$

in particular, if $\bar{z} \in Z$, then we may replace $(Z - \bar{z})^{\Theta}$ by $N_Z(\bar{z})$ in (3.1).

Proof. Let $n \in \mathbb{N}$. Taking the supremum over $z \in Z$ in (2.1) yields

$$\sigma_{Z-\bar{z}}(x_n - x_{n+1}) \le \|x_{n+1} - \bar{z}\| \|x_n - x_{n+1}\| + \frac{1}{2} \|x_n - x_{n+1}\|^2.$$
(3.2)

Dividing (3.2) by $||x_n - x_{n+1}||$ and using the positive homogeneity of $\sigma_{Z-\bar{z}}$, we have

$$\sigma_{Z-\bar{z}}\left(\frac{x_n - x_{n+1}}{\|x_n - x_{n+1}\|}\right) \leqslant \|x_{n+1} - \bar{z}\| + \frac{1}{2}\|x_n - x_{n+1}\|. \tag{3.3}$$

Because $x_n \to \overline{z}$ and so $x_n - x_{n+1} \to 0$, we let $n \to \infty$ in (3.3) to learn that

$$\overline{\lim}_{n \to \infty} \sigma_{Z - \overline{z}} \left(\frac{x_n - x_{n+1}}{\|x_n - x_{n+1}\|} \right) \le 0. \tag{3.4}$$

Combining (3.4) with Lemma 2.3, we obtain

$$C\left(\left(\frac{x_n - x_{n+1}}{\|x_n - x_{n+1}\|}\right)_{n \in \mathbb{N}}\right) \subseteq (Z - \bar{z})^{\Theta}.$$
(3.5)

Next, let $m \ge n + 1$. Taking the supermum over $z \in Z$ in (2.2) yields

$$\sigma_{Z-\bar{z}}(x_n - x_m) \leqslant \frac{1}{2} (\|x_n - \bar{z}\|^2 - \|x_m - \bar{z}\|^2). \tag{3.6}$$

Passing to the limit as $m \to \infty$ and using the lower semicontinuity of $\sigma_{Z-\bar{z}}$ in (3.6), we obtain

$$\sigma_{Z-\bar{z}}(x_n - \bar{z}) \leqslant \underline{\lim}_{m \to \infty} \sigma_{Z-\bar{z}}(x_n - x_m) \leqslant \frac{1}{2} \|x_n - \bar{z}\|^2. \tag{3.7}$$

Dividing (3.7) by $||x_n - \overline{z}||$ and using the positive homogeneity of $\sigma_{Z-\overline{z}}$, we have

$$\sigma_{Z-\bar{z}}\left(\frac{x_n-\bar{z}}{\|x_n-\bar{z}\|}\right) \leqslant \frac{1}{2}\|x_n-\bar{z}\|. \tag{3.8}$$

Because $x_n \to \overline{z}$, we let $n \to \infty$ in (3.8) and get

$$\overline{\lim}_{n\to\infty} \sigma_{Z-\bar{z}} \left(\frac{x_n - \bar{z}}{\|x_n - \bar{z}\|} \right) \le 0. \tag{3.9}$$

Combining (3.9) with our trusted Lemma 2.3, we obtain

$$C\left(\left(\frac{x_n - \overline{z}}{\|x_n - \overline{z}\|}\right)_{n \in \mathbb{N}}\right) \subseteq (Z - \overline{z})^{\Theta}.$$
(3.10)

Altogether, (3.5) and (3.10) imply (3.1).

Although ostensibly more general, the following result is actually an easy consequence of Theorem 3.1:

Corollary 3.2. Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in X that is Fejér monotone with respect to some nonempty closed convex subset Z of X. Suppose that $(x_{k_n})_{n\in\mathbb{N}}$ is a subsequence of $(x_n)_{n\in\mathbb{N}}$ such that $x_{k_n} \to \overline{z} \in X$ and that $(\forall n \in \mathbb{N})$ $x_{k_{n+1}} \neq x_{k_n} \neq \overline{z}$. Then

$$\mathcal{C}\left(\left(\frac{x_{k_n} - x_{k_{n+1}}}{\|x_{k_n} - x_{k_{n+1}}\|}\right)_{n \in \mathbb{N}}\right) \cup \mathcal{C}\left(\left(\frac{x_{k_n} - \overline{z}}{\|x_{k_n} - \overline{z}\|}\right)_{n \in \mathbb{N}}\right) \subseteq (Z - \overline{z})^{\Theta}; \tag{3.11}$$

in particular, if $\bar{z} \in Z$, then we may replace $(Z - \bar{z})^{\Theta}$ by $N_Z(\bar{z})$ in (3.11).

Proof. Recalling Lemma 2.1(iv), we simply apply Theorem 3.1 to
$$(x_{k_n})_{n \in \mathbb{N}}$$
.

When *X* is finite-dimensional, we have the following two nice results:

Corollary 3.3. Suppose that X is finite-dimensional. Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in X that is Fejér monotone with respect to some nonempty closed convex subset Z of X. Suppose that $(x_{k_n})_{n\in\mathbb{N}}$ is a subsequence of $(x_n)_{n\in\mathbb{N}}$ such that $x_{k_n} \to \overline{z} \in Z$ and that $(\forall n \in \mathbb{N})$ $x_{k_{n+1}} \neq x_{k_n} \neq \overline{z}$. Then

$$\mathcal{C}\left(\left(\frac{x_{k_n}-x_{k_{n+1}}}{\|x_{k_n}-x_{k_{n+1}}\|}\right)_{n\in\mathbb{N}}\right)\cup\mathcal{C}\left(\left(\frac{x_{k_n}-\bar{z}}{\|x_{k_n}-\bar{z}\|}\right)_{n\in\mathbb{N}}\right)\subseteq\mathbb{S}\cap N_Z(\bar{z});$$

equivalently,

$$\lim_{n\to\infty}d_{S\cap N_Z(\overline{z})}\left(\frac{x_{k_n}-x_{k_{n+1}}}{\|x_{k_n}-x_{k_{n+1}}\|}\right)=0\quad and\quad \lim_{n\to\infty}d_{S\cap N_Z(\overline{z})}\left(\frac{x_{k_n}-\overline{z}}{\|x_{k_n}-\overline{z}\|}\right)=0.$$

Proof. Combine Corollary 3.2 with Lemma 2.4.

Corollary 3.4 (no zigzagging). Suppose that X is finite-dimensional and let $(x_n)_{n\in\mathbb{N}}$ be a sequence that is Fejér monotone with respect to some closed convex subset Z of X. Suppose that $x_n \to \overline{z} \in Z$, that $(\forall n \in \mathbb{N}) \ x_{n+1} \neq x_n \neq \overline{z}$, that int $Z \neq \emptyset$, and that $N_Z(\overline{z})$ is a ray. Then

$$\lim_{n \to \infty} \frac{x_n - x_{n+1}}{\|x_n - x_{n+1}\|} = \lim_{n \to \infty} \frac{x_n - \overline{z}}{\|x_n - \overline{z}\|} \in \mathbb{S} \cap N_Z(\overline{z}). \tag{3.12}$$

Proof. Clear from Corollary 3.3 because $S \cap N_Z(\bar{z})$ is a singleton when $N_Z(\bar{z})$ is a ray.

4 Large sets of directionally asymptotic cluster points

In this section, we give an example illustrating that the sets of directionally asymptotic cluster points can be large. It also shows that without the interiority assumption in Corollary 3.4, (3.12) can go quite wrong.

We start with a fact from real analysis:

Fact 4.1 (Dirichlet). (See, e.g., [12, page 88]) Let $\alpha \in \mathbb{R} \setminus \mathbb{Q}$. Then the set $\{n\alpha - \lfloor n\alpha \rfloor \mid n \in \mathbb{N}\}$ is dense in [0,1].

For the remainder of this section, R_{α} denotes the counterclockwise rotator in the Euclidean plane by α .

Example 4.2. Suppose that $X = \mathbb{R}^2$, let $0 < \theta \notin \frac{1}{2}\pi\mathbb{N}$, Then $T := \frac{1}{2}\operatorname{Id} + \frac{1}{2}R_{2\theta} = \cos(\theta)R_{\theta}$ is firmly nonexpansive, with $Z := \operatorname{Fix} T = \{0\}$. Let $x_0 \in X \setminus \{0\}$, and set

$$(n \in \mathbb{N}) \quad x_{n+1} := Tx_n. \tag{4.1}$$

Then $x_n \to \bar{z} := 0$, and $(\forall n \in \mathbb{N})$ $x_{n+1} \neq x_n \neq \bar{z}$ and $\langle x_n - x_{n+1}, x_{n+1} \rangle = 0$. Moreover, we have the following dichotomoy:

(i) $\theta \in 2\pi \mathbb{Q}$ and

$$\mathcal{C}\left(\left(\frac{x_n-x_{n+1}}{\|x_n-x_{n+1}\|}\right)_{n\in\mathbb{N}}\right)\cup\mathcal{C}\left(\left(\frac{x_n}{\|x_n\|}\right)_{n\in\mathbb{N}}\right)\quad\text{is a finite subset of }\mathbb{S}.$$

(ii) $\theta \notin 2\pi \mathbb{Q}$ and

$$C\left(\left(\frac{x_n-x_{n+1}}{\|x_n-x_{n+1}\|}\right)_{n\in\mathbb{N}}\right)=C\left(\left(\frac{x_n}{\|x_n\|}\right)_{n\in\mathbb{N}}\right)=S.$$

Proof. Because $\theta \notin 2\pi \mathbb{Z}$, we have Fix $T = \{0\}$. Note that

$$T^n = (\cos \theta)^n R_{n\theta} = (\cos \theta)^n \begin{pmatrix} \cos n\theta & -\sin n\theta \\ \sin n\theta & \cos n\theta \end{pmatrix}$$
 and $x_n = T^n x_0$.

Since $0 < |\cos \theta| < 1$ and $R_{n\theta}$ is an isometry, we have $||T^n|| \to 0$. Thus $x_n \to 0$. By Lemma 2.5(ii), we have $\langle x_n - x_{n+1}, x_{n+1} \rangle = 0$. Moreover, $||x_{n+1}|| = ||\cos \theta R_{\theta} x_n|| = |\cos \theta|||x_n|| < ||x_n|||$ because $x_0 \neq 0$, so $(\forall n \in \mathbb{N})$ $x_{n+1} \neq x_n$ and $x_n \neq 0$.

To study the set of cluster points of $(T^n x_0/\|T^n x_0\|)_{n\in\mathbb{N}}$, we consider two cases.

Case 1: $\cos \theta > 0$. We have

$$\frac{T^n x_0}{\|T^n x_0\|} = R_{n\theta} \frac{x_0}{\|x_0\|} = \begin{pmatrix} \cos n\theta & -\sin n\theta \\ \sin n\theta & \cos n\theta \end{pmatrix} \frac{x_0}{\|x_0\|}.$$
 (4.2)

We proceed with two subcases.

Subcase 1: $\frac{\theta}{2\pi} \in \mathbb{Q}$. Then $\theta = 2\pi \frac{k}{l}$ with $k, l \in \mathbb{N}$ and $l \neq 0$, and $\cos n\theta = \cos \frac{n}{l}(2k\pi)$ and $\sin n\theta = \sin \frac{n}{l}(2k\pi)$. By using $n = ml, ml + 1, \ldots, ml + l - 1$ with $m \in \mathbb{N}$, the set

$${R_{n\theta} \mid n \in \mathbb{N}} = {R_{t2k\pi/l} \mid t = 0, \ldots, l-1}.$$

The sequence $(R_{n\theta})_{n\in\mathbb{N}}$ has at most l cluster points. From (4.2) we see that $(T^nx_0/\|T^nx_0\|)_{n\in\mathbb{N}}$ has at most l cluster points. In fact, if k=2, then there are precisely l cluster points.

Subcase 2: $\frac{\theta}{2\pi} \notin \mathbb{Q}$. Then $\theta = 2\pi\alpha$ with $\alpha \in \mathbb{R}_{++} \setminus \mathbb{Q}$, and

$$\cos n\theta = \cos n(2\pi\alpha) = \cos(n\alpha)(2\pi) = \cos(n\alpha - \lfloor n\alpha \rfloor)(2\pi)$$
, and

$$\sin n\theta = \sin n(2\pi\alpha) = \sin(n\alpha)(2\pi) = \sin(n\alpha - \lfloor n\alpha \rfloor)(2\pi),$$

By Fact 4.1, $\{n\alpha - \lfloor n\alpha \rfloor \mid n \in \mathbb{N}\}$ is dense in [0,1]. Hence the set of cluster points of $\{R_{n\theta} \mid n \in \mathbb{N}\}$ is $\{R_{\beta} \mid \beta \in [0,2\pi]\}$. By (4.2), the set of cluster points of $(T^n x_0/\|T^n x_0\|)_{n \in \mathbb{N}}$ is S. Case 2: $\cos \theta < 0$. We have

$$\frac{T^n x_0}{\|T^n x_0\|} = (-1)^n R_{n\theta} \frac{x_0}{\|x_0\|} = (-1)^n \begin{pmatrix} \cos n\theta & -\sin n\theta \\ \sin n\theta & \cos n\theta \end{pmatrix} \frac{x_0}{\|x_0\|}.$$
 (4.3)

We proceed with two subcases.

Subcase 1: $\frac{\theta}{2\pi} \in \mathbb{Q}$. Due to $(-1)^n$, we have to consider n being even and odd. When n is even, write n = 2k with $k \in \mathbb{N}$,

$$R_{n\theta} = \begin{pmatrix} \cos k(2\theta) & -\sin k(2\theta) \\ \sin k(2\theta) & \cos k(2\theta) \end{pmatrix}.$$

Since $\frac{2\theta}{2\pi} \in \mathbb{Q}$, similar arguments as in Case 1 subcase 1 show that $(R_{k(2\theta)})_{k \in \mathbb{N}}$ has a finite number of cluster points. When n is odd, let n = 2k + 1 and $\theta = m2\pi/l$ with $k, l, m \in \mathbb{N}$ and $l \neq 0$. If we set k = tl + s for $t, s \in \mathbb{N}$ and $0 \leq s \leq l - 1$, then

$$(2k+1)\frac{m2\pi}{l} = \frac{2(tl+s)+1}{l}m2\pi = \left(2t+\frac{2s+1}{l}\right)m2\pi$$

so that

$$\cos(2k+1)\theta = \cos\left(\frac{2s+1}{l}m2\pi\right), \quad \sin(2k+1)\theta = \sin\left(\frac{2s+1}{l}m2\pi\right)$$

where $0 \le s \le l-1$. This shows that when n is odd, we have most l cluster points. Combining the even and odd cases, the set $\{R_{n\theta} \mid n \in \mathbb{N}\}$ has at most a finite number of cluster points, so is $(T^n x_0/\|T^n x_0\|)_{n\in\mathbb{N}}$ by (4.3).

Subcase 2: $\frac{\theta}{2\pi} \notin \mathbb{Q}$. Put $\theta = \alpha(2\pi)$ with $\alpha \notin \mathbb{Q}$. When n is even, write n = 2k with $k \in \mathbb{N}$. Then $R_{n\theta} = R_{k(2\theta)}^{-1}$, similar arguments as in Case 1 subcase 2 show that the set of cluster points of $\{R_{k(2\theta)} \mid k \in \mathbb{N}\}$ is $\{R_{\beta} \mid 0 \leq \beta \leq 2\pi\}$, because $\frac{2\theta}{2\pi} \in \mathbb{R}_{++} \setminus \mathbb{Q}$. When n is odd, write n = 2k + 1 with $k \in \mathbb{N}$. Since that

$$\cos[(2k+1)\alpha 2\pi] = \cos[k(2\alpha)2\pi + 2\alpha\pi] = \cos[(k(2\alpha) - \lfloor k(2\alpha) \rfloor)2\pi + 2\alpha\pi], \text{ and}$$
$$\sin[(2k+1)\alpha 2\pi] = \sin[k(2\alpha)2\pi + 2\alpha\pi] = \sin[(k(2\alpha) - \lfloor k(2\alpha) \rfloor)2\pi + 2\alpha\pi],$$

and that

$$\{k(2\alpha) - \lfloor k(2\alpha) \rfloor \mid k \in \mathbb{N} \}$$
 is dense in $[0, 1]$

we see that the set of cluster points of $\{R_{(2k+1)\theta} \mid k \in \mathbb{N}\}$ is $\{R_{\beta} \mid 2\alpha\pi \leqslant \beta \leqslant 2\alpha\pi + 2\pi\}$. Hence, in both cases the set of cluster points of $(T^n x_0/||T^n x_0||)_{n \in \mathbb{N}}$ is \mathbb{S} .

Finally, we consider the set of cluster points of

$$\left(\frac{x_n-x_{n+1}}{\|x_n-x_{n+1}\|}\right)_{n\in\mathbb{N}}.$$

Now

$$Id - T = \frac{Id - R_{2\theta}}{2} = \sin\theta \begin{pmatrix} \sin\theta & \cos\theta \\ -\cos\theta & \sin\theta \end{pmatrix}$$
$$= \sin\theta \begin{pmatrix} \cos(\theta + 3\pi/2) & -\sin(\theta + 3\pi/2) \\ \sin(\theta + 3\pi/2) & \cos(\theta + 3\pi/2) \end{pmatrix} = \sin\theta R_{(\theta + 3\pi/2)},$$

so that

$$\frac{x_{n} - x_{n+1}}{\|x_{n} - x_{n+1}\|} = \frac{(\text{Id} - T)x_{n}}{\|(\text{Id} - T)x_{n}\|}$$

$$= \begin{cases} R_{(\theta + 3\pi/2)} \frac{x_{n}}{\|x_{n}\|}, & \text{if } \sin \theta > 0; \\ -R_{(\theta + 3\pi/2)} \frac{x_{n}}{\|x_{n}\|}, & \text{if } \sin \theta < 0. \end{cases}$$

Then the set of cluster points of

$$\left(\frac{x_n - x_{n+1}}{\|x_n - x_{n+1}\|}\right)_{n \in \mathbb{N}}$$

is just $\pm R_{(\theta+3\pi/2)}$ rotations of the set of cluster points of $(x_n/\|x_n\|)_{n\in\mathbb{N}}$. Consequently, the set of cluster points of $((x_n-x_{n+1})/\|x_n-x_{n+1}\|)_{n\in\mathbb{N}}$ is a finite set if $\frac{\theta}{2\pi}\in\mathbb{Q}$; and is \mathbb{S} if $\frac{\theta}{2\pi}\notin\mathbb{Q}$.

Remark 4.3. We do not consider the case when $\theta \in \frac{1}{2}\pi\mathbb{N}$ because then $T = \operatorname{Id}$ or 0 in which case one has finite convergence of $(x_n)_{n \in \mathbb{N}}$.

5 Missing the sphere: an infinite-dimensional example

It is interesting to ask whether in infinite-dimensional Hilbert spaces the nonempty sets of weak cluster points in Theorem 3.1 lie in the sphere S. It turns out that the answer is negative, and the sequence provided is obtained by iterating a resolvent. To this end, we assume in this section that

$$X=\ell^2\big(\{1,2,\ldots\}\big),\,$$

with the standard Schauder basis $e_1 := (1,0,0,...), e_2 := (0,1,0,0,...)$, and so on. We define the *right-shift operator* by

$$R: X \to X: (\xi_1, \xi_2, \ldots) \mapsto (0, \xi_1, \xi_2, \ldots),$$

Then R is a linear isometry with Fix $R = \{0\}$. We shall also require the following classical identity

Fact 5.1 (Vandermonde's identity). (See [9, Section 5.1].) *Let* m, n, r *be in* \mathbb{N} . *Then*

$$\binom{m+n}{r} = \sum_{k=0}^{r} \binom{m}{r} \binom{n}{r-k}.$$

Example 5.2. Define the firmly nonexpansive operator $T: X \to X$ by

$$T:=\frac{1}{2}\operatorname{Id}+\frac{1}{2}R,$$

set $x_0 := e_1$, and $(x_n)_{n \in \mathbb{N}} := (T^n x_0)_{n \in \mathbb{N}}$ with $x_0 = e_1$. Then the following hold:

- (i) $(x_n)_{n\in\mathbb{N}}$ is Fejér monotone with respect to Fix $T=\{0\}$, and $x_n\to 0$. Moreover, $(\forall n\in\mathbb{N})$ $x_{n+1}\neq x_n\neq 0$.
- (ii) $(\forall n \in \mathbb{N}) \langle x_{n+1}, x_n x_{n+1} \rangle = 0.$
- (iii) Both

$$\left(\frac{x_n}{\|x_n\|}\right)_{n\in\mathbb{N}}$$
 and $\left(\frac{x_n-x_{n+1}}{\|x_n-x_{n+1}\|}\right)_{n\in\mathbb{N}}$

converge weakly — but not strongly — to $0 \notin S$.

Proof. (i): It is well known that $(x_n)_{n\in\mathbb{N}}$ is Fejér monotone with respect to Fix $T=\{0\}$, because T is nonexpansive. Since $R^ke_1=e_{k+1}$, we have

$$x_n = T^n x_0 = \frac{1}{2^n} (\operatorname{Id} + R)^n x_0 = \frac{1}{2^n} \sum_{k=0}^n \binom{n}{k} R^k x_0 = \frac{1}{2^n} \sum_{k=0}^n \binom{n}{k} e_{k+1}.$$
 (5.1)

Hence, by Fact 5.1,

$$\|x_n\|^2 = \frac{1}{4^n} \sum_{k=0}^n \binom{n}{k}^2 = \frac{1}{4^n} \binom{2n}{n} = \frac{1}{4^n} \frac{(2n)!}{(n!)^2};$$

in particular, $x_n \neq 0$. $x_{n+1} \neq x_n$ because x_{n+1} contains a nonzero term of e_{n+2} and $e_{n+2} \perp e_{k+1}$ for $1 \leq k \leq n$.

Now recall Stirling's formula (see, e.g., [13, Theorem 5.44]) which states that

$$n! \approx \sqrt{2\pi n} \frac{n^n}{e^n} \tag{5.2}$$

for large n, and which implies

$$||x_n||^2 = \frac{1}{4^n} \binom{2n}{n} = \frac{(2n)!}{4^n (n!)^2} \approx \frac{1}{4^n} \frac{\sqrt{2\pi (2n)} (2n/e)^{2n}}{(\sqrt{2\pi n})^2 (n/e)^{2n}} = \frac{1}{\sqrt{\pi}} \frac{1}{\sqrt{n}} \to 0.$$

(The qualitative fact that $x_n \to 0$ also follows from [1, Example 5.29] or [4, Corollary 1.2].)

(ii): Since *R* is an isometry, this follows from Lemma 2.5(ii).

(iii). For *fixed k*, we have from Stirling's formula (5.2) that

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} \approx \frac{1}{k!} \frac{\sqrt{2\pi n} (n/e)^n}{\sqrt{2\pi (n-k)} ((n-k)/e)^{n-k}} \approx \frac{n^k}{k!}$$
 (5.3)

for large *n*. Hence

$$\frac{x_n}{\|x_n\|} \approx \sqrt[4]{\pi} \sqrt[4]{n} \frac{1}{2^n} \sum_{k=0}^n \binom{n}{k} e_{k+1} \approx \sum_{k=0}^n \frac{\sqrt[4]{\pi} \sqrt[4]{n}}{2^n} \frac{n^k}{k!} e_{k+1} \to 0$$

because that $(e_k)_{k \in \mathbb{N}}$ is a total set in $\ell^2(\mathbb{N})$ and that for each fixed $k \in \mathbb{N}$ the coefficient of e_{k+1} in $x_n/\|x_n\|$ clearly converges to 0 as $n \to \infty$; see, e.g., [10, Example 4.8-6].

Next,

$$x_n - x_{n+1} = \frac{1}{2^n} \sum_{k=0}^n \binom{n}{k} e_{k+1} - \frac{1}{2^{n+1}} \sum_{k=0}^{n+1} \binom{n+1}{k} e_{k+1}.$$
 (5.4)

Since $e_{n+2} \perp e_{k+1}$ for $0 \le k \le n$, by Fact 5.1 and (5.1) we have

$$\langle x_n, x_{n+1} \rangle = \left\langle \frac{1}{2^n} \sum_{k=0}^n \binom{n}{k} e_{k+1}, \frac{1}{2^{n+1}} \sum_{k=0}^{n+1} \binom{n+1}{k} e_{k+1} \right\rangle$$
 (5.5a)

$$= \frac{1}{2} \frac{1}{4^n} \left\langle \sum_{k=0}^n \binom{n}{k} e_{k+1}, \sum_{k=0}^n \binom{n+1}{k} e_{k+1} \right\rangle$$
 (5.5b)

$$= \frac{1}{2} \frac{1}{4^n} \sum_{k=0}^n \binom{n}{k} \binom{n+1}{k} = \frac{1}{2} \frac{1}{4^n} \sum_{k=0}^n \binom{n}{n-k} \binom{n+1}{k} = \frac{1}{2} \frac{1}{4^n} \binom{2n+1}{n}.$$
 (5.5c)

It follows from (5.1) and (5.5) that

$$||x_n - x_{n+1}||^2 = ||x_n||^2 + ||x_{n+1}||^2 - 2\langle x_n, x_{n+1} \rangle$$
(5.6a)

$$= \frac{1}{4^n} \binom{2n}{n} + \frac{1}{4^{n+1}} \binom{2(n+1)}{n+1} - 2\frac{1}{2} \frac{1}{4^n} \binom{2n+1}{n}$$
 (5.6b)

$$= \frac{1}{4^n} \left[\binom{2n}{n} + \frac{1}{4} \binom{2n+2}{n+1} - \binom{2n+1}{n} \right]$$
 (5.6c)

$$= \frac{1}{2(n+1)4^n} {2n \choose n} = \frac{1}{2(n+1)} ||x_n||^2$$
 (5.6d)

$$\approx \frac{1}{2n} \frac{1}{\sqrt{\pi n}} = \frac{1}{2\sqrt{\pi}} \frac{1}{n^{3/2}} \tag{5.6e}$$

for large n. Combining (5.4), (5.6), and (5.3), we obtain

$$\frac{x_n - x_{n+1}}{\|x_n - x_{n+1}\|} \approx \frac{\sqrt{2}\sqrt[4]{\pi}n^{3/4}}{2^n} \sum_{k=0}^n \binom{n}{k} e_{k+1} - \frac{\sqrt{2}\sqrt[4]{\pi}n^{3/4}}{2^{n+1}} \sum_{k=0}^{n+1} \binom{n+1}{k} e_{k+1}$$
 (5.7a)

$$\approx \sum_{k=0}^{n} \frac{\sqrt{2}\sqrt[4]{\pi}n^{3/4}}{2^{n}} \binom{n}{k} e_{k+1} - \sum_{k=0}^{n+1} \frac{\sqrt{2}\sqrt[4]{\pi}n^{3/4}}{2^{n+1}} \binom{n+1}{k} e_{k+1}$$
 (5.7b)

$$\approx \sum_{k=0}^{n} \frac{\sqrt{2}\sqrt[4]{\pi}n^{3/4}}{2^{n}} \frac{n^{k}}{k!} e_{k+1} - \sum_{k=0}^{n+1} \frac{\sqrt{2}\sqrt[4]{\pi}n^{3/4}}{2^{n+1}} \frac{(n+1)^{k}}{k!} e_{k+1}$$
 (5.7c)

$$\rightarrow 0,$$
 (5.7d)

because for every fixed $k \in \mathbb{N}$ the coefficients of e_{k+1} in $(x_n - x_{n+1})/\|x_n - x_{n+1}\|$ converge to 0 as $n \to \infty$.

In summary, both quotient limits converge weakly but not strongly to 0.

Acknowledgements

We thank Terry Rockafellar for his inspirational talk at the virtual West Coast Optimization Meeting in May 2021 which stimulated this research. We also thank Walaa Moursi for suggesting to investigate the operator *T* in Section 5. HHB and XW were partially supported by NSERC Discovery Grants. MKL was partially supported by HHB and XW's NSERC Discovery Grants.

References

- [1] H.H. BAUSCHKE AND P.L. COMBETTES: Convex Analysis and Monotone Operator Theory in Hilbert Spaces, second edition, Springer, 2017.
- [2] H.H. BAUSCHKE, M.N. DAO, AND W.M. MOURSI: On Fejér monotone sequences and nonexpansive mappings, *Linear and Nonlinear Analysis* 1 (2015), 287–295.
- [3] H.H. BAUSCHKE, W.L. HARE, AND W.M. MOURSI: Generalized solutions for the sum of two maximally monotone operators, *SIAM Journal on Control and Optimization* 52 (2014), 1034–1047.

- [4] R.E. BRUCK AND S. REICH: Nonexpansive projections and resolvents of accretive operators in Banach spaces, *Houston Journal of Mathematics* 3 (1977), 459–470.
- [5] A. CEGIELSKI: *Iterative Methods for Fixed Point Problems in Hilbert Spaces*, Lecture Notes in Mathematics, 2057, Springer, Heidelberg, 2012.
- [6] P.L. COMBETTES: Solving monotone inclusions via compositions of nonexpansive averaged operators, *Optimization* 53 (2004), 475–504.
- [7] P.L. COMBETTES: Fejér-monotonicity in convex optimization, C.A. Floudas, P.M. Pardalos (Eds.), *Encyclopedia of Optimization*, second ed., Springer, New York, 2009, pp. 1016–1024.
- [8] P.L. COMBETTES: Quasi-Fejérian analysis of some optimization algorithms, in: D. Butnariu, Y. Censor, S. Reich (Eds.), *Inherently Parallel Algorithms for Feasibility and Optimization*, Elsevier, New York, 2001, pp. 115–152.
- [9] R.L. GRAHAM, D.E. KNUTH, AND O. PATASHNIK: *Concrete Mathematics*, second edition, Addison-Wesley, 1994.
- [10] E. Kreyszig: Introductory Functional Analysis with Applications, Wiley, 1989.
- [11] R.T. ROCKAFELLAR: Advances in convergence and scope of the proximal point algorithm, *Journal of Nonlinear and Convex Analysis*, in press. https://sites.math.washington.edu/~rtr/papers/rtr257-ConvergencePPA.pdf
- [12] H.H. SOHRAB: *Basic Real Analysis*, second edition, Birkhäuser/Springer, New York, 2014.
- [13] K.R. Stromberg: Introduction to Classical Real Analysis, Wadsworth, 1981.