Heinz H. Bauschke
Patrick L. Combettes

Convex Analysis and Monotone Operator Theory in Hilbert Spaces

Springer

Foreword

This self-contained book offers a modern unifying presentation of three basic areas of nonlinear analysis, namely convex analysis, monotone operator theory, and the fixed point theory of nonexpansive mappings.

This turns out to be a judicious choice. Showing the rich connections and interplay between these topics gives a strong coherence to the book. Moreover, these particular topics are at the core of modern optimization and its applications.

Choosing to work in Hilbert spaces offers a wide range of applications, while keeping the mathematics accessible to a large audience. Each topic is developed in a self-contained fashion, and the presentation often draws on recent advances.

The organization of the book makes it accessible to a large audience. Each chapter is illustrated by several exercises, which makes the monograph an excellent textbook. In addition, it offers deep insights into algorithmic aspects of optimization, especially splitting algorithms, which are important in theory and applications.

Let us point out the high quality of the writing and presentation. The authors combine an uncompromising demand for rigorous mathematical statements and a deep concern for applications, which makes this book remarkably accomplished.

Montpellier (France), October 2010 *Hédy Attouch*

Preface

Three important areas of nonlinear analysis emerged in the early 1960s: convex analysis, monotone operator theory, and the theory of nonexpansive mappings. Over the past four decades, these areas have reached a high level of maturity, and an increasing number of connections have been identified between them. At the same time, they have found applications in a wide array of disciplines, including mechanics, economics, partial differential equations, information theory, approximation theory, signal and image processing, game theory, optimal transport theory, probability and statistics, and machine learning.

The purpose of this book is to present a largely self-contained account of the main results of convex analysis, monotone operator theory, and the theory of nonexpansive operators in the context of Hilbert spaces. Authoritative monographs are already available on each of these topics individually. A novelty of this book, and indeed, its central theme, is the tight interplay among the key notions of convexity, monotonicity, and nonexpansiveness. We aim at making the presentation accessible to a broad audience, and to reach out in particular to the applied sciences and engineering communities, where these tools have become indispensable. We chose to cast our exposition in the Hilbert space setting. This allows us to cover many applications of interest to practitioners in infinite-dimensional spaces and yet to avoid the technical difficulties pertaining to general Banach space theory that would exclude a large portion of our intended audience. We have also made an attempt to draw on recent developments and modern tools to simplify the proofs of key results, exploiting for instance heavily the concept of a Fitzpatrick function in our exposition of monotone operators, the notion of Fejér monotonicity to unify the convergence proofs of several algorithms, and that of a proximity operator throughout the second half of the book.

The book in organized in 29 chapters. Chapters 1 and 2 provide background material. Chapters 3 to 7 cover set convexity and nonexpansive operators. Various aspects of the theory of convex functions are discussed in Chapters 8 to 19. Chapters 20 to 25 are dedicated to monotone operator theory. In addition to these basic building blocks, we also address certain themes from different angles in several places. Thus, optimization theory is discussed in Chapters 11, 19, 26, and 27. Best approximation problems are discussed in Chapters 3, 19, 27, 28, and 29. Algorithms are also present in various parts of the book: fixed point and convex feasibility algorithms in Chapter 5, proximal-point algorithms in Chapter 23, monotone operator splitting algorithms in Chapter 25, optimization algorithms in Chapter 27, and best approximation algorithms in Chapters 27 and 29. More than 400 exercises are distributed throughout the book, at the end of each chapter.

Preliminary drafts of this book have been used in courses in our institutions and we have benefited from the input of postdoctoral fellows and many students. To all of them, many thanks. In particular, HHB thanks Liangjin Yao for his helpful comments. We are grateful to Hédy Attouch, Jon Borwein, Stephen Simons, Jon Vanderwerff, Shawn Wang, and Isao Yamada for helpful discussions and pertinent comments. PLC also thanks Oscar Wesler. Finally, we thank the Natural Sciences and Engineering Research Council of Canada, the Canada Research Chair Program, and France's Agence Nationale de la Recherche for their support.

Kelowna (Canada) — *Heinz H. Bauschke*
Paris (France) — *Patrick L. Combettes*
October 2010

Contents

Chapter 5
Fejér Monotonicity and Fixed Point Iterations

A sequence is Fejér monotone with respect to a set C if each point in the sequence is not strictly farther from any point in C than its predecessor. Such sequences possess very attractive properties that greatly simplify the analysis of their asymptotic behavior. In this chapter, we provide the basic theory for Fejér monotone sequences and apply it to obtain in a systematic fashion convergence results for various classical iterations involving nonexpansive operators.

5.1 Fejér Monotone Sequences

The following notion is central in the study of various iterative methods, in particular in connection with the construction of fixed points of nonexpansive operators.

Definition 5.1 Let C be a nonempty subset of $\mathcal{H}$ and let $(x_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{H}$. Then $(x_n)_{n\in\mathbb{N}}$ is *Fejér monotone* with respect to C if

$$(\forall x \in C)(\forall n \in \mathbb{N}) \quad \|x_{n+1} - x\| \leq \|x_n - x\|. \tag{5.1}$$

Example 5.2 Let $(x_n)_{n\in\mathbb{N}}$ be a bounded sequence in $\mathbb{R}$ that is increasing (respectively decreasing). Then $(x_n)_{n\in\mathbb{N}}$ is Fejér monotone with respect to $[\sup\{x_n\}_{n\in\mathbb{N}}, +\infty[$ (respectively $]-\infty, \inf\{x_n\}_{n\in\mathbb{N}}]$).

Example 5.3 Let D be a nonempty subset of $\mathcal{H}$, let $T\colon D \to D$ be a quasinonexpansive—in particular, nonexpansive—operator such that $\operatorname{Fix} T \neq \varnothing$, and let $x_0 \in D$. Set $(\forall n \in \mathbb{N})\ x_{n+1} = Tx_n$. Then $(x_n)_{n\in\mathbb{N}}$ is Fejér monotone with respect to $\operatorname{Fix} T$.

We start with some basic properties.

Proposition 5.4 *Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{H}$ and let C be a nonempty subset of $\mathcal{H}$. Suppose that $(x_n)_{n\in\mathbb{N}}$ is Fejér monotone with respect to C. Then the following hold:*

(i) *$(x_n)_{n\in\mathbb{N}}$ is bounded.*
(ii) *For every $x \in C$, $(\|x_n - x\|)_{n\in\mathbb{N}}$ converges.*
(iii) *$(d_C(x_n))_{n\in\mathbb{N}}$ is decreasing and converges.*

Proof. (i): Let $x \in C$. Then (5.1) implies that $(x_n)_{n\in\mathbb{N}}$ lies in $B(x; \|x_0 - x\|)$.
(ii): Clear from (5.1).
(iii): Taking the infimum in (5.1) over $x \in C$ yields $(\forall n \in \mathbb{N})\ d_C(x_{n+1}) \leq d_C(x_n)$. ☐

The next result concerns weak convergence.

Theorem 5.5 *Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{H}$ and let C be a nonempty subset of $\mathcal{H}$. Suppose that $(x_n)_{n\in\mathbb{N}}$ is Fejér monotone with respect to C and that every weak sequential cluster point of $(x_n)_{n\in\mathbb{N}}$ belongs to C. Then $(x_n)_{n\in\mathbb{N}}$ converges weakly to a point in C.*

Proof. The result follows from Proposition 5.4(ii) and Lemma 2.39. ☐

Example 5.6 Suppose that $\mathcal{H}$ is infinite-dimensional and let $(x_n)_{n\in\mathbb{N}}$ be an orthonormal sequence in $\mathcal{H}$. Then $(x_n)_{n\in\mathbb{N}}$ is Fejér monotone with respect to $\{0\}$. As seen in Example 2.25, $x_n \rightharpoonup 0$ but $x_n \not\to 0$.

While a Fejér monotone sequence with respect to a closed convex set C may not converge strongly, its "shadow" on C always does.

Proposition 5.7 *Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{H}$ and let C be a nonempty closed convex subset of $\mathcal{H}$. Suppose that $(x_n)_{n\in\mathbb{N}}$ is Fejér monotone with respect to C. Then the* shadow sequence *$(P_C x_n)_{n\in\mathbb{N}}$ converges strongly to a point in C.*

Proof. It follows from (5.1) and (3.6) that, for every m and n in $\mathbb{N}$,

$$\begin{aligned}\|P_C x_n - P_C x_{n+m}\|^2 &= \|P_C x_n - x_{n+m}\|^2 + \|x_{n+m} - P_C x_{n+m}\|^2 \\ &\quad + 2\langle P_C x_n - x_{n+m} \mid x_{n+m} - P_C x_{n+m}\rangle \\ &\leq \|P_C x_n - x_n\|^2 + d_C^2(x_{n+m}) \\ &\quad + 2\langle P_C x_n - P_C x_{n+m} \mid x_{n+m} - P_C x_{n+m}\rangle \\ &\quad + 2\langle P_C x_{n+m} - x_{n+m} \mid x_{n+m} - P_C x_{n+m}\rangle \\ &\leq d_C^2(x_n) - d_C^2(x_{n+m}).\end{aligned} \tag{5.2}$$

Consequently, since $(d_C(x_n))_{n\in\mathbb{N}}$ was seen in Proposition 5.4(iii) to converge, $(P_C x_n)_{n\in\mathbb{N}}$ is a Cauchy sequence in the complete set C. ☐

Corollary 5.8 *Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{H}$, let C be a nonempty closed convex subset of $\mathcal{H}$, and let $x \in C$. Suppose that $(x_n)_{n\in\mathbb{N}}$ is Fejér monotone with respect to C and that $x_n \rightharpoonup x$. Then $P_C x_n \to x$.*

Proof. By Proposition 5.7, $(P_C x_n)_{n\in\mathbb{N}}$ converges strongly to some point $y \in C$. Hence, since $x - P_C x_n \to x - y$ and $x_n - P_C x_n \rightharpoonup x - y$, it follows from Theorem 3.14 and Lemma 2.41(iii) that $0 \geq \langle x - P_C x_n \mid x_n - P_C x_n\rangle \to \|x-y\|^2$. Thus, $x = y$. ☐

For sequences that are Fejér monotone with respect to closed affine subspaces, Proposition 5.7 can be strengthened.

Proposition 5.9 *Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{H}$ and let C be a closed affine subspace of $\mathcal{H}$. Suppose that $(x_n)_{n\in\mathbb{N}}$ is Fejér monotone with respect to C. Then the following hold:*

(i) $(\forall n \in \mathbb{N})\ P_C x_n = P_C x_0$.

(ii) *Suppose that every weak sequential cluster point of $(x_n)_{n\in\mathbb{N}}$ belongs to C. Then $x_n \rightharpoonup P_C x_0$.*

Proof. (i): Fix $n \in \mathbb{N}$, $\alpha \in \mathbb{R}$, and set $y_\alpha = \alpha P_C x_0 + (1-\alpha) P_C x_n$. Since C is an affine subspace, $y_\alpha \in C$, and it therefore follows from Corollary 3.20(i) and (5.1) that

$$\begin{aligned}\alpha^2\|P_C x_n - P_C x_0\|^2 &= \|P_C x_n - y_\alpha\|^2\\ &\leq \|x_n - P_C x_n\|^2 + \|P_C x_n - y_\alpha\|^2\\ &= \|x_n - y_\alpha\|^2\\ &\leq \|x_0 - y_\alpha\|^2\\ &= \|x_0 - P_C x_0\|^2 + \|P_C x_0 - y_\alpha\|^2\\ &= d_C^2(x_0) + (1-\alpha)^2\|P_C x_n - P_C x_0\|^2. \end{aligned} \tag{5.3}$$

Consequently, $(2\alpha - 1)\|P_C x_n - P_C x_0\|^2 \leq d_C^2(x_0)$ and, letting $\alpha \to +\infty$, we conclude that $P_C x_n = P_C x_0$.

(ii): Combine Theorem 5.5, Corollary 5.8, and (i). ☐

We now turn our attention to strong convergence properties.

Proposition 5.10 *Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{H}$ and let C be a subset of $\mathcal{H}$ such that $\operatorname{int} C \neq \varnothing$. Suppose that $(x_n)_{n\in\mathbb{N}}$ is Fejér monotone with respect to C. Then $(x_n)_{n\in\mathbb{N}}$ converges strongly to a point in $\mathcal{H}$.*

Proof. Take $x \in \operatorname{int} C$ and $\rho \in \mathbb{R}_{++}$ such that $B(x;\rho) \subset C$. Define a sequence $(z_n)_{n\in\mathbb{N}}$ in $B(x;\rho)$ by

$$(\forall n \in \mathbb{N}) \quad z_n = \begin{cases} x, & \text{if } x_{n+1} = x_n;\\ x - \rho\dfrac{x_{n+1} - x_n}{\|x_{n+1} - x_n\|}, & \text{otherwise.}\end{cases} \tag{5.4}$$

Then (5.1) yields $(\forall n \in \mathbb{N})\ \|x_{n+1} - z_n\|^2 \leq \|x_n - z_n\|^2$ and, after expanding, we obtain

$$(\forall n \in \mathbb{N}) \quad \|x_{n+1} - x\|^2 \leq \|x_n - x\|^2 - 2\rho\|x_{n+1} - x_n\|. \tag{5.5}$$

Thus, $\sum_{n\in\mathbb{N}} \|x_{n+1} - x_n\| \leq \|x_0 - x\|^2/(2\rho)$ and $(x_n)_{n\in\mathbb{N}}$ is therefore a Cauchy sequence. □

Theorem 5.11 *Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{H}$ and let C be a nonempty closed convex subset of $\mathcal{H}$. Suppose that $(x_n)_{n\in\mathbb{N}}$ is Fejér monotone with respect to C. Then the following are equivalent:*

(i) *$(x_n)_{n\in\mathbb{N}}$ converges strongly to a point in C.*
(ii) *$(x_n)_{n\in\mathbb{N}}$ possesses a strong sequential cluster point in C.*
(iii) $\underline{\lim}\, d_C(x_n) = 0$.

Proof. (i)⇒(ii): Clear.

(ii)⇒(iii): Suppose that $x_{k_n} \to x \in C$. Then $d_C(x_{k_n}) \leq \|x_{k_n} - x\| \to 0$.

(iii)⇒(i): Proposition 5.4(iii) implies that $d_C(x_n) \to 0$. Hence, $x_n - P_C x_n \to 0$ and (i) follows from Proposition 5.7. □

We conclude this section with a linear convergence result.

Theorem 5.12 *Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{H}$ and let C be a nonempty closed convex subset of $\mathcal{H}$. Suppose that $(x_n)_{n\in\mathbb{N}}$ is Fejér monotone with respect to C and that for some $\kappa \in [0, 1[$,*

$$(\forall n \in \mathbb{N}) \quad d_C(x_{n+1}) \leq \kappa d_C(x_n). \tag{5.6}$$

Then $(x_n)_{n\in\mathbb{N}}$ converges linearly to a point $x \in C$; more precisely,

$$(\forall n \in \mathbb{N}) \quad \|x_n - x\| \leq 2\kappa^n d_C(x_0). \tag{5.7}$$

Proof. Theorem 5.11 and (5.6) imply that $(x_n)_{n\in\mathbb{N}}$ converges strongly to some point $x \in C$. On the other hand, (5.1) yields

$$\begin{aligned}(\forall n \in \mathbb{N})(\forall m \in \mathbb{N}) \quad \|x_n - x_{n+m}\| &\leq \|x_n - P_C x_n\| + \|x_{n+m} - P_C x_n\| \\ &\leq 2d_C(x_n).\end{aligned} \tag{5.8}$$

Letting $m \to +\infty$ in (5.8), we conclude that $\|x_n - x\| \leq 2d_C(x_n)$. □

5.2 Krasnosel'skiĭ–Mann Iteration

Given a nonexpansive operator T, the sequence generated by the Banach–Picard iteration $x_{n+1} = Tx_n$ of (1.66) may fail to produce a fixed point of T. A simple illustration of this situation is $T = -\mathrm{Id}$ and $x_0 \neq 0$. In this case, however, it is clear that the *asymptotic regularity* property $x_n - Tx_n \to 0$ does not hold. As we shall now see, this property is critical.

Theorem 5.13 *Let D be a nonempty closed convex subset of $\mathcal{H}$, let $T\colon D \to D$ be a nonexpansive operator such that* $\operatorname{Fix} T \neq \varnothing$*, and let $x_0 \in D$. Set*

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = Tx_n \tag{5.9}$$

and suppose that $x_n - Tx_n \to 0$. Then the following hold:

(i) *$(x_n)_{n\in\mathbb{N}}$ converges weakly to a point in* $\operatorname{Fix} T$.
(ii) *Suppose that $D = -D$ and that T is odd: $(\forall x \in D)$ $T(-x) = -Tx$. Then $(x_n)_{n\in\mathbb{N}}$ converges strongly to a point in* $\operatorname{Fix} T$.

Proof. From Example 5.3, $(x_n)_{n\in\mathbb{N}}$ is Fejér monotone with respect to $\operatorname{Fix} T$.

(i): Let x be a weak sequential cluster point of $(x_n)_{n\in\mathbb{N}}$, say $x_{k_n} \rightharpoonup x$. Since $Tx_{k_n} - x_{k_n} \to 0$, Corollary 4.18 asserts that $x \in \operatorname{Fix} T$. Appealing to Theorem 5.5, the assertion is proved.

(ii): Since $D = -D$ is convex, $0 \in D$ and, since T is odd, $0 \in \operatorname{Fix} T$. Therefore, by Fejér monotonicity, $(\forall n \in \mathbb{N})$ $\|x_{n+1}\| \leq \|x_n\|$. Thus, there exists $\ell \in \mathbb{R}_+$ such that $\|x_n\| \downarrow \ell$. Now let $m \in \mathbb{N}$. Then, for every $n \in \mathbb{N}$,

$$\|x_{n+1+m} + x_{n+1}\| = \|Tx_{n+m} - T(-x_n)\| \leq \|x_{n+m} + x_n\|, \tag{5.10}$$

and, by the parallelogram identity,

$$\|x_{n+m} + x_n\|^2 = 2\big(\|x_{n+m}\|^2 + \|x_m\|^2\big) - \|x_{n+m} - x_n\|^2. \tag{5.11}$$

However, since $Tx_n - x_n \to 0$, we have $\lim_n \|x_{n+m} - x_n\| = 0$. Therefore, since $\|x_n\| \downarrow \ell$, (5.10) and (5.11) yield $\|x_{n+m} + x_n\| \downarrow 2\ell$ as $n \to +\infty$. In turn, we derive from (5.11) that $\|x_{n+m} - x_n\|^2 \leq 2(\|x_{n+m}\|^2 + \|x_m\|^2) - 4\ell^2 \to 0$ as $m, n \to +\infty$. Thus, $(x_n)_{n\in\mathbb{N}}$ is a Cauchy sequence and $x_n \to x$ for some $x \in D$. Since $x_{n+1} \to x$ and $x_{n+1} = Tx_n \to Tx$, we have $x \in \operatorname{Fix} T$. ☐

We now turn our attention to an alternative iterative method, known as the *Krasnosel'skiĭ–Mann algorithm*.

Theorem 5.14 (Krasnosel'skiĭ–Mann algorithm) *Let D be a nonempty closed convex subset of $\mathcal{H}$, let $T\colon D \to D$ be a nonexpansive operator such that* $\operatorname{Fix} T \neq \varnothing$*, let $(\lambda_n)_{n\in\mathbb{N}}$ be a sequence in $[0,1]$ such that $\sum_{n\in\mathbb{N}} \lambda_n(1-\lambda_n) = +\infty$, and let $x_0 \in D$. Set*

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = x_n + \lambda_n\big(Tx_n - x_n\big). \tag{5.12}$$

Then the following hold:

(i) *$(x_n)_{n\in\mathbb{N}}$ is Fejér monotone with respect to* $\operatorname{Fix} T$.
(ii) *$(Tx_n - x_n)_{n\in\mathbb{N}}$ converges strongly to 0.*
(iii) *$(x_n)_{n\in\mathbb{N}}$ converges weakly to a point in* $\operatorname{Fix} T$.

Proof. Since $x_0 \in D$ and D is convex, (5.12) produces a well-defined sequence in D.

(i): It follows from Corollary 2.14 and the nonexpansiveness of T that, for every $y \in \operatorname{Fix} T$ and $n \in \mathbb{N}$,

$$\begin{aligned}\|x_{n+1} - y\|^2 &= \|(1-\lambda_n)(x_n - y) + \lambda_n(Tx_n - y)\|^2 \\ &= (1-\lambda_n)\|x_n - y\|^2 + \lambda_n\|Tx_n - Ty\|^2 \\ &\quad - \lambda_n(1-\lambda_n)\|Tx_n - x_n\|^2 \\ &\leq \|x_n - y\|^2 - \lambda_n(1-\lambda_n)\|Tx_n - x_n\|^2. \end{aligned} \tag{5.13}$$

Hence, $(x_n)_{n\in\mathbb{N}}$ is Fejér monotone with respect to $\operatorname{Fix} T$.

(ii): We derive from (5.13) that $\sum_{n\in\mathbb{N}} \lambda_n(1-\lambda_n)\|Tx_n - x_n\|^2 \leq \|x_0 - y\|^2$. Since $\sum_{n\in\mathbb{N}} \lambda_n(1-\lambda_n) = +\infty$, we have $\underline{\lim}\, \|Tx_n - x_n\| = 0$. However, for every $n \in \mathbb{N}$,

$$\begin{aligned}\|Tx_{n+1} - x_{n+1}\| &= \|Tx_{n+1} - Tx_n + (1-\lambda_n)(Tx_n - x_n)\| \\ &\leq \|x_{n+1} - x_n\| + (1-\lambda_n)\|Tx_n - x_n\| \\ &= \|Tx_n - x_n\|. \end{aligned} \tag{5.14}$$

Consequently, $(\|Tx_n - x_n\|)_{n\in\mathbb{N}}$ converges and we must have $Tx_n - x_n \to 0$.

(iii): Let x be a weak sequential cluster point of $(x_n)_{n\in\mathbb{N}}$, say $x_{k_n} \rightharpoonup x$. Then it follows from Corollary 4.18 that $x \in \operatorname{Fix} T$. In view of Theorem 5.5, the proof is complete. ☐

Proposition 5.15 *Let* $\alpha \in \,]0,1[$, *let* $T\colon \mathcal{H} \to \mathcal{H}$ *be an* α*-averaged operator such that* $\operatorname{Fix} T \neq \varnothing$, *let* $(\lambda_n)_{n\in\mathbb{N}}$ *be a sequence in* $[0, 1/\alpha]$ *such that* $\sum_{n\in\mathbb{N}} \lambda_n(1-\alpha\lambda_n) = +\infty$, *and let* $x_0 \in \mathcal{H}$. *Set*

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = x_n + \lambda_n\big(Tx_n - x_n\big). \tag{5.15}$$

Then the following hold:

(i) $(x_n)_{n\in\mathbb{N}}$ *is Fejér monotone with respect to* $\operatorname{Fix} T$.
(ii) $(Tx_n - x_n)_{n\in\mathbb{N}}$ *converges strongly to* 0.
(iii) $(x_n)_{n\in\mathbb{N}}$ *converges weakly to a point in* $\operatorname{Fix} T$.

Proof. Set $R = (1 - 1/\alpha)\mathrm{Id} + (1/\alpha)T$ and $(\forall n \in \mathbb{N})$ $\mu_n = \alpha\lambda_n$. Then $\operatorname{Fix} R = \operatorname{Fix} T$ and R is nonexpansive by Proposition 4.25. In addition, we rewrite (5.15) as $(\forall n \in \mathbb{N})$ $x_{n+1} = x_n + \mu_n\big(Rx_n - x_n\big)$. Since $(\mu_n)_{n\in\mathbb{N}}$ lies in $[0,1]$ and $\sum_{n\in\mathbb{N}} \mu_n(1-\mu_n) = +\infty$, the results follow from Theorem 5.14. ☐

Corollary 5.16 *Let* $T\colon \mathcal{H} \to \mathcal{H}$ *be a firmly nonexpansive operator such that* $\operatorname{Fix} T \neq \varnothing$, *let* $(\lambda_n)_{n\in\mathbb{N}}$ *be a sequence in* $[0,2]$ *such that* $\sum_{n\in\mathbb{N}} \lambda_n(2-\lambda_n) = +\infty$, *and let* $x_0 \in \mathcal{H}$. *Set* $(\forall n \in \mathbb{N})$ $x_{n+1} = x_n + \lambda_n\big(Tx_n - x_n\big)$. *Then the following hold:*

(i) $(x_n)_{n\in\mathbb{N}}$ *is Fejér monotone with respect to* $\operatorname{Fix} T$.
(ii) $(Tx_n - x_n)_{n\in\mathbb{N}}$ *converges strongly to* 0.

(iii) $(x_n)_{n\in\mathbb{N}}$ *converges weakly to a point in* $\operatorname{Fix} T$.

Proof. In view of Remark 4.24(iii), apply Proposition 5.15 with $\alpha = 1/2$. □

Example 5.17 Let $T\colon \mathcal{H} \to \mathcal{H}$ be a firmly nonexpansive operator such that $\operatorname{Fix} T \neq \varnothing$, let $x_0 \in \mathcal{H}$, and set $(\forall n \in \mathbb{N})\ x_{n+1} = Tx_n$. Then $(x_n)_{n\in\mathbb{N}}$ converges weakly to a point in $\operatorname{Fix} T$.

The following type of iterative method involves a mix of compositions and convex combinations of nonexpansive operators.

Corollary 5.18 *Let* $(T_i)_{i\in I}$ *be a finite family of nonexpansive operators from* $\mathcal{H}$ *to* $\mathcal{H}$ *such that* $\bigcap_{i\in I} \operatorname{Fix} T_i \neq \varnothing$, *and let* $(\alpha_i)_{i\in I}$ *be real numbers in* $]0,1[$ *such that, for every* $i \in I$, T_i *is* α_i*-averaged. Let* p *be a strictly positive integer, for every* $k \in \{1,\dots,p\}$, *let* m_k *be a strictly positive integer and* ω_k *be a strictly positive real number, and suppose that* $\mathrm{i}\colon \big\{(k,l) \bigm| k \in \{1,\dots,p\}, l \in \{1,\dots,m_k\}\big\} \to I$ *is surjective and that* $\sum_{k=1}^{p} \omega_k = 1$. *For every* $k \in \{1,\dots,p\}$, *set* $I_k = \{\mathrm{i}(k,1),\dots,\mathrm{i}(k,m_k)\}$, *and set*

$$\alpha = \max_{1\le k\le p} \rho_k, \quad \textit{where} \quad (\forall k \in \{1,\dots,p\}) \quad \rho_k = \frac{m_k}{m_k - 1 + \dfrac{1}{\max\limits_{i\in I_k} \alpha_i}}, \tag{5.16}$$

and let $(\lambda_n)_{n\in\mathbb{N}}$ *be a sequence in* $[0,1/\alpha]$ *such that* $\sum_{n\in\mathbb{N}} \lambda_n(1-\alpha\lambda_n) = +\infty$. *Furthermore, let* $x_0 \in \mathcal{H}$ *and set*

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = x_n + \lambda_n \left(\sum_{k=1}^{p} \omega_k T_{\mathrm{i}(k,1)} \cdots T_{\mathrm{i}(k,m_k)} x_n - x_n \right). \tag{5.17}$$

Then $(x_n)_{n\in\mathbb{N}}$ *converges weakly to a point in* $\bigcap_{i\in I} \operatorname{Fix} T_i$.

Proof. Set $T = \sum_{k=1}^{p} \omega_k R_k$, where $(\forall k \in \{1,\dots,p\})\ R_k = T_{\mathrm{i}(k,1)} \cdots T_{\mathrm{i}(k,m_k)}$. Then (5.17) reduces to (5.15) and, in view of Proposition 5.15, it suffices to show that T is α-averaged and that $\operatorname{Fix} T = \bigcap_{i\in I} \operatorname{Fix} T_i$. For every $k \in \{1,\dots,p\}$, it follows from Proposition 4.32 and (5.16) that R_k is ρ_k-averaged and, from Corollary 4.37 that $\operatorname{Fix} R_k = \bigcap_{i\in I_k} \operatorname{Fix} T_i$. In turn, we derive from Proposition 4.30 and (5.16) that T is α-averaged and, from Proposition 4.34, that $\operatorname{Fix} T = \bigcap_{k=1}^{p} \operatorname{Fix} R_k = \bigcap_{k=1}^{p} \bigcap_{i\in I_k} \operatorname{Fix} T_i = \bigcap_{i\in I} \operatorname{Fix} T_i$. □

Remark 5.19 It follows from Remark 4.24(iii) that Corollary 5.18 is applicable to firmly nonexpansive operators and, a fortiori, to projection operators by Proposition 4.8.

Corollary 5.18 provides an algorithm to solve a *convex feasibility problem*, i.e., to find a point in the intersection of a family of closed convex sets. Here are two more examples.

Example 5.20 (string-averaged relaxed projections) Let $(C_i)_{i\in I}$ be a finite family of closed convex sets such that $C = \bigcap_{i\in I} C_i \neq \varnothing$. For every $i \in I$, let $\beta_i \in \left]0,2\right[$ and set $T_i = (1-\beta_i)\mathrm{Id} + \beta_i P_{C_i}$. Let p be a strictly positive integer; for every $k \in \{1,\ldots,p\}$, let m_k be a strictly positive integer and ω_k be a strictly positive real number, and suppose that i: $\big\{(k,l) \bigm| k \in \{1,\ldots,p\},\, l \in \{1,\ldots,m_k\}\big\} \to I$ is surjective and that $\sum_{k=1}^{p} \omega_k = 1$. Furthermore, let $x_0 \in \mathcal{H}$ and set

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = \sum_{k=1}^{p} \omega_k T_{\mathrm{i}(k,1)} \cdots T_{\mathrm{i}(k,m_k)} x_n. \tag{5.18}$$

Then $(x_n)_{n\in\mathbb{N}}$ converges weakly to a point in C.

Proof. For every $i \in I$, set $\alpha_i = \beta_i/2 \in \left]0,1\right[$. Since, for every $i \in I$, Proposition 4.8 asserts that P_{C_i} is firmly nonexpansive, Corollary 4.29 implies that T_i is α_i-averaged. Borrowing notation from Corollary 5.18, we note that for every $k \in \{1,\ldots,p\}$, $\max_{i\in I_k} \alpha_i \in \left]0,1\right[$, which implies that $\rho_k \in \left]0,1\right[$ and thus that $\alpha \in \left]0,1\right[$. Altogether, the result follows from Corollary 5.18 with $\lambda_n \equiv 1$. ☐

Example 5.21 (parallel projection algorithm) Let $(C_i)_{i\in I}$ be a finite family of closed convex subsets of $\mathcal{H}$ such that $C = \bigcap_{i\in I} C_i \neq \varnothing$, let $(\lambda_n)_{n\in\mathbb{N}}$ be a sequence in $[0,2]$ such that $\sum_{n\in\mathbb{N}} \lambda_n(2-\lambda_n) = +\infty$, let $(\omega_i)_{i\in I}$ be strictly positive real numbers such that $\sum_{i\in I} \omega_i = 1$, and let $x_0 \in \mathcal{H}$. Set

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = x_n + \lambda_n \Bigg(\sum_{i\in I} \omega_i P_i x_n - x_n \Bigg), \tag{5.19}$$

where, for every $i \in I$, P_i denotes the projector onto C_i. Then $(x_n)_{n\in\mathbb{N}}$ converges weakly to a point in C.

Proof. This is an application of Corollary 5.16(iii) with $T = \sum_{i\in I} \omega_i P_i$. Indeed, since the operators $(P_i)_{i\in I}$ are firmly nonexpansive by Proposition 4.8, their convex combination T is also firmly nonexpansive by Example 4.31. Moreover, Proposition 4.34 asserts that $\operatorname{Fix} T = \bigcap_{i\in I} \operatorname{Fix} P_i = \bigcap_{i\in I} C_i = C$. Alternatively, apply Corollary 5.18. ☐

5.3 Iterating Compositions of Averaged Operators

Our first result concerns the asymptotic behavior of iterates of a composition of averaged nonexpansive operators with possibly no common fixed point.

Theorem 5.22 *Let D be a nonempty weakly sequentially closed (e.g., closed and convex) subset of $\mathcal{H}$, let m be a strictly positive integer, set $I =$*

$\{1,\ldots,m\}$, let $(T_i)_{i\in I}$ be a family of nonexpansive operators from D to D such that $\operatorname{Fix}(T_1\cdots T_m)\neq\varnothing$*, and let $(\alpha_i)_{i\in I}$ be real numbers in* $]0,1[$ *such that, for every $i\in I$, T_i is α_i-averaged. Let $x_0\in D$ and set*

$$(\forall n\in\mathbb{N})\quad x_{n+1}=T_1\cdots T_m x_n. \tag{5.20}$$

Then $x_n - T_1\cdots T_m x_n\to 0$, and there exist points $y_1\in\operatorname{Fix}T_1\cdots T_m$, $y_2\in\operatorname{Fix}T_2\cdots T_mT_1$, \ldots, $y_m\in\operatorname{Fix}T_mT_1\cdots T_{m-1}$ such that

$$x_n\rightharpoonup y_1=T_1y_2, \tag{5.21}$$
$$T_mx_n\rightharpoonup y_m=T_my_1, \tag{5.22}$$
$$T_{m-1}T_mx_n\rightharpoonup y_{m-1}=T_{m-1}y_m, \tag{5.23}$$
$$\vdots$$
$$T_3\cdots T_mx_n\rightharpoonup y_3=T_3y_4, \tag{5.24}$$
$$T_2\cdots T_mx_n\rightharpoonup y_2=T_2y_3. \tag{5.25}$$

Proof. Set $T=T_1\cdots T_m$ and $(\forall i\in I)$ $\beta_i=(1-\alpha_i)/\alpha_i$. Now take $y\in\operatorname{Fix}T$. The equivalence (i)⇔(iii) in Proposition 4.25 yields

$$\begin{aligned}\|x_{n+1}-y\|^2&=\|Tx_n-Ty\|^2\\&\leq\|T_2\cdots T_mx_n-T_2\cdots T_my\|^2\\&\quad-\beta_1\|(\mathrm{Id}-T_1)T_2\cdots T_mx_n-(\mathrm{Id}-T_1)T_2\cdots T_my\|^2\\&\leq\|x_n-y\|^2-\beta_m\|(\mathrm{Id}-T_m)x_n-(\mathrm{Id}-T_m)y\|^2\\&\quad-\beta_{m-1}\|(\mathrm{Id}-T_{m-1})T_mx_n-(\mathrm{Id}-T_{m-1})T_my\|^2-\cdots\\&\quad-\beta_2\|(\mathrm{Id}-T_2)T_3\cdots T_mx_n-(\mathrm{Id}-T_2)T_3\cdots T_my\|^2\\&\quad-\beta_1\|(\mathrm{Id}-T_1)T_2\cdots T_mx_n-(T_2\cdots T_my-y)\|^2.\end{aligned} \tag{5.26}$$

Therefore, $(x_n)_{n\in\mathbb{N}}$ is Fejér monotone with respect to $\operatorname{Fix}T$ and

$$(\mathrm{Id}-T_m)x_n-(\mathrm{Id}-T_m)y\to 0, \tag{5.27}$$
$$(\mathrm{Id}-T_{m-1})T_mx_n-(\mathrm{Id}-T_{m-1})T_my\to 0, \tag{5.28}$$
$$\vdots$$
$$(\mathrm{Id}-T_2)T_3\cdots T_mx_n-(\mathrm{Id}-T_2)T_3\cdots T_my\to 0, \tag{5.29}$$
$$(\mathrm{Id}-T_1)T_2\cdots T_mx_n-(T_2\cdots T_my-y)\to 0. \tag{5.30}$$

Upon adding (5.27)–(5.30), we obtain $x_n-Tx_n\to 0$. Hence, since T is nonexpansive as a composition of nonexpansive operators, it follows from Theorem 5.13(i) that $(x_n)_{n\in\mathbb{N}}$ converges weakly to some point $y_1\in\operatorname{Fix}T$, which provides (5.21). On the other hand, (5.27) yields $T_mx_n-x_n\to T_my_1-y_1$. So altogether $T_mx_n\rightharpoonup T_my_1=y_m$, and we obtain (5.22). In turn, since (5.28) asserts that $T_{m-1}T_mx_n-T_mx_n\to T_{m-1}y_m-y_m$, we obtain $T_{m-1}T_m x_n \rightharpoonup T_{m-1}y_m = y_{m-1}$, hence (5.23). Continuing this process, we arrive at (5.25). ☐

As noted in Remark 5.19, results on averaged nonexpansive operators apply in particular to firmly nonexpansive operators and projectors onto convex sets. Thus, by specializing Theorem 5.22 to convex projectors, we obtain the iterative method described in the next corollary, which is known as the POCS (Projections Onto Convex Sets) algorithm in the signal recovery literature.

Corollary 5.23 (POCS algorithm) *Let m be a strictly positive integer, set $I = \{1,\dots,m\}$, let $(C_i)_{i\in I}$ be a family of nonempty closed convex subsets of $\mathcal{H}$, let $(P_i)_{i\in I}$ denote their respective projectors, and let $x_0 \in \mathcal{H}$. Suppose that* $\operatorname{Fix}(P_1\cdots P_m) \neq \varnothing$ *and set*

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = P_1 \cdots P_m x_n. \tag{5.31}$$

Then there exists $(y_1,\dots,y_m) \in C_1 \times \cdots \times C_m$ such that $x_n \rightharpoonup y_1 = P_1 y_2$, $P_m x_n \rightharpoonup y_m = P_m y_1$, $P_{m-1}P_m x_n \rightharpoonup y_{m-1} = P_{m-1}y_m$, …, $P_3\cdots P_m x_n \rightharpoonup y_3 = P_3 y_4$, and $P_2\cdots P_m x_n \rightharpoonup y_2 = P_2 y_3$.

Proof. This follows from Proposition 4.8 and Theorem 5.22. ☐

Remark 5.24 In Corollary 5.23, suppose that, for some $j \in I$, C_j is bounded. Then $\operatorname{Fix}(P_1\cdots P_m) \neq \varnothing$. Indeed, consider the circular composition of the m projectors given by $T = P_j \cdots P_m P_1 \cdots P_{j-1}$. Then Proposition 4.8 asserts that T is a nonexpansive operator that maps the nonempty bounded closed convex set C_j to itself. Hence, it follows from Theorem 4.19 that there exists a point $x \in C_j$ such that $Tx = x$.

The next corollary describes a periodic projection method to solve a convex feasibility problem.

Corollary 5.25 *Let m be a strictly positive integer, set $I = \{1,\dots,m\}$, let $(C_i)_{i\in I}$ be a family of closed convex subsets of $\mathcal{H}$ such that $C = \bigcap_{i\in I} C_i \neq \varnothing$, let $(P_i)_{i\in I}$ denote their respective projectors, and let $x_0 \in \mathcal{H}$. Set* $(\forall n \in \mathbb{N})$ *$x_{n+1} = P_1\cdots P_m x_n$. Then $(x_n)_{n\in\mathbb{N}}$ converges weakly to a point in C.*

Proof. Using Corollary 5.23, Proposition 4.8, and Corollary 4.37, we obtain $x_n \rightharpoonup y_1 \in \operatorname{Fix}(P_1\cdots P_m) = \bigcap_{i\in I} \operatorname{Fix} P_i = C$. Alternatively, this is a special case of Example 5.20. ☐

Remark 5.26 If, in Corollary 5.25, all the sets are closed affine subspaces, so is C and we derive from Proposition 5.9(i) that $x_n \rightharpoonup P_C x_0$. Corollary 5.28 is classical, and it states that the convergence is actually strong in this case. In striking contrast, the example constructed in [146] provides a closed hyperplane and a closed convex cone in $\ell^2(\mathbb{N})$ for which alternating projections converge weakly but not strongly.

The next result will help us obtain a sharper form of Corollary 5.25 for closed affine subspaces.

Proposition 5.27 *Let $T \in \mathcal{B}(\mathcal{H})$ be nonexpansive and let $x_0 \in \mathcal{H}$. Set $V = \operatorname{Fix} T$ and $(\forall n \in \mathbb{N})$ $x_{n+1} = Tx_n$. Then $x_n \to P_V x_0 \Leftrightarrow x_n - x_{n+1} \to 0$.*

Proof. If $x_n \to P_V x_0$, then $x_n - x_{n+1} \to P_V x_0 - P_V x_0 = 0$. Conversely, suppose that $x_n - x_{n+1} \to 0$. We derive from Theorem 5.13(ii) that there exists $v \in V$ such that $x_n \to v$. In turn, Proposition 5.9(i) yields $v = P_V x_0$. ☐

Corollary 5.28 (von Neumann–Halperin) *Let m be a strictly positive integer, set $I = \{1, \ldots, m\}$, let $(C_i)_{i\in I}$ be a family of closed affine subspaces of $\mathcal{H}$ such that $C = \bigcap_{i\in I} C_i \neq \varnothing$, let $(P_i)_{i\in I}$ denote their respective projectors, let $x_0 \in \mathcal{H}$, and set*

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = P_1 \cdots P_m x_n. \tag{5.32}$$

Then $x_n \to P_C x_0$.

Proof. Set $T = P_1 \cdots P_m$. Then T is nonexpansive, and $\operatorname{Fix} T = C$ by Corollary 4.37.

We first assume that each set C_i is a linear subspace. Then T is odd, and Theorem 5.22 implies that $x_n - Tx_n \to 0$. Thus, by Proposition 5.27, $x_n \to P_C x_0$.

We now turn our attention to the general affine case. Since $C \neq \varnothing$, there exists $y \in C$ such that for every $i \in I$, $C_i = y + V_i$, i.e., V_i is the closed linear subspace parallel to C_i, and $C = y + V$, where $V = \bigcap_{i\in I} V_i$. Proposition 3.17 implies that, for every $x \in \mathcal{H}$, $P_C x = P_{y+V} x = y + P_V(x - y)$ and $(\forall i \in I)$ $P_i x = P_{y+V_i} x = y + P_{V_i}(x - y)$. Using these identities repeatedly, we obtain

$$(\forall n \in \mathbb{N}) \quad x_{n+1} - y = (P_{V_1} \cdots P_{V_m})(x_n - y). \tag{5.33}$$

Invoking the already verified linear case, we get $x_n - y \to P_V(x_0 - y)$ and conclude that $x_n \to y + P_V(x_0 - y) = P_C x_0$. ☐

Exercises

Exercise 5.1 Find a nonexpansive operator $T\colon \mathcal{H} \to \mathcal{H}$ that is not firmly nonexpansive and such that, for every $x_0 \in \mathcal{H}$, the sequence $(T^n x_0)_{n\in\mathbb{N}}$ converges weakly but not strongly to a fixed point of T.

Exercise 5.2 Construct a non-Cauchy sequence $(x_n)_{n\in\mathbb{N}}$ in $\mathbb{R}$ that is asymptotically regular, i.e., $x_n - x_{n+1} \to 0$.

Exercise 5.3 Find an alternative proof of Theorem 5.5 based on Corollary 5.8 in the case when C is closed and convex.

Exercise 5.4 Let C be a nonempty subset of $\mathcal{H}$ and let $(x_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{H}$ that is Fejér monotone with respect to C. Show that $(x_n)_{n\in\mathbb{N}}$ is Fejér monotone with respect to $\overline{\operatorname{conv}}\, C$.

Exercise 5.5 Let $T\colon \mathcal{H} \to \mathcal{H}$ be a nonexpansive operator such that $\operatorname{Fix} T \neq \varnothing$, and let $(x_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{H}$ such that

(i) for every $x \in \operatorname{Fix} T$, $(\|x_n - x\|)_{n\in\mathbb{N}}$ converges;
(ii) $x_n - Tx_n \to 0$.

Show that $(x_n)_{n\in\mathbb{N}}$ converges weakly to a point in $\operatorname{Fix} T$.

Exercise 5.6 Find a nonexpansive operator $T\colon \mathcal{H} \to \mathcal{H}$ that is not firmly nonexpansive and such that, for every $x_0 \in \mathcal{H}$, the sequence $(T^n x_0)_{n\in\mathbb{N}}$ converges weakly but not strongly to a fixed point of T.

Exercise 5.7 Let m be a strictly positive integer, set $I = \{1, \ldots, m\}$, let $(C_i)_{i\in I}$ be a family of closed convex subsets of $\mathcal{H}$ such that $C = \bigcap_{i\in I} C_i \neq \varnothing$, and let $(P_i)_{i\in I}$ be their respective projectors. Derive parts (ii) and (iii) from (i) and Theorem 5.5, and also from Corollary 5.18.

(i) Let $i \in I$, let $x \in C_i$, and let $y \in \mathcal{H}$. Show that $\|P_i y - x\|^2 \leq \|y - x\|^2 - \|P_i y - y\|^2$.
(ii) Set $x_0 \in \mathcal{H}$ and

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = \frac{1}{m}\big(P_1 x_n + P_1 P_2 x_n + \cdots + P_1 \cdots P_m x_n\big). \tag{5.34}$$

(a) Let $x \in C$ and $n \in \mathbb{N}$. Show that $\|x_{n+1} - x\|^2 \leq \|x_n - x\|^2 - (1/m)\sum_{i\in I} \|P_i x_n - x\|^2$.
(b) Let x be a weak sequential cluster point of $(x_n)_{n\in\mathbb{N}}$. Show that $x \in C$.
(c) Show that $(x_n)_{n\in\mathbb{N}}$ converges weakly to a point in C.

(iii) Set $x_0 \in \mathcal{H}$ and

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = \frac{1}{m-1}\big(P_1 P_2 x_n + P_2 P_3 x_n + \cdots + P_{m-1} P_m x_n\big). \tag{5.35}$$

(a) Let $x \in C$ and $n \in \mathbb{N}$. Show that $\|x_{n+1} - x\|^2 \leq \|x_n - x\|^2 - \sum_{i=1}^{m-1} (\|P_{i+1} x_n - x_n\|^2 + \|P_i P_{i+1} x_n - P_{i+1} x_n\|^2)/(m-1)$.
(b) Let x be a weak sequential cluster point of $(x_n)_{n\in\mathbb{N}}$. Show that $x \in C$.
(c) Show that $(x_n)_{n\in\mathbb{N}}$ converges weakly to a point in C.

Index