

CMS Books in Mathematics



Canadian
Mathematical Society
Société mathématique
du Canada

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 is 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 the-

ory. 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 Hédÿ 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)
Paris (France)
October 2010

Heinz H. Bauschke
Patrick L. Combettes

Contents

1	Background	1
1.1	Sets in Vector Spaces	1
1.2	Operators	2
1.3	Order	3
1.4	Nets	4
1.5	The Extended Real Line	4
1.6	Functions	5
1.7	Topological Spaces	7
1.8	Two-Point Compactification of the Real Line	9
1.9	Continuity	9
1.10	Lower Semicontinuity	10
1.11	Sequential Topological Notions	15
1.12	Metric Spaces	16
	Exercises	22
2	Hilbert Spaces	27
2.1	Notation and Examples	27
2.2	Basic Identities and Inequalities	29
2.3	Linear Operators and Functionals	31
2.4	Strong and Weak Topologies	33
2.5	Weak Convergence of Sequences	36
2.6	Differentiability	37
	Exercises	40
3	Convex Sets	43
3.1	Definition and Examples	43
3.2	Best Approximation Properties	44
3.3	Topological Properties	52
3.4	Separation	55
	Exercises	57

4	Convexity and Nonexpansiveness	59
4.1	Nonexpansive Operators	59
4.2	Projectors onto Convex Sets	61
4.3	Fixed Points of Nonexpansive Operators	62
4.4	Averaged Nonexpansive Operators	67
4.5	Common Fixed Points	71
	Exercises	72
5	Fejér Monotonicity and Fixed Point Iterations	75
5.1	Fejér Monotone Sequences	75
5.2	Krasnosel'skiĭ–Mann Iteration	78
5.3	Iterating Compositions of Averaged Operators	82
	Exercises	85
6	Convex Cones and Generalized Interiors	87
6.1	Convex Cones	87
6.2	Generalized Interiors	90
6.3	Polar and Dual Cones	96
6.4	Tangent and Normal Cones	100
6.5	Recession and Barrier Cones	103
	Exercises	104
7	Support Functions and Polar Sets	107
7.1	Support Points	107
7.2	Support Functions	109
7.3	Polar Sets	110
	Exercises	111
8	Convex Functions	113
8.1	Definition and Examples	113
8.2	Convexity–Preserving Operations	116
8.3	Topological Properties	120
	Exercises	125
9	Lower Semicontinuous Convex Functions	129
9.1	Lower Semicontinuous Convex Functions	129
9.2	Proper Lower Semicontinuous Convex Functions	132
9.3	Affine Minorization	133
9.4	Construction of Functions in $\Gamma_0(\mathcal{H})$	136
	Exercises	141
10	Convex Functions: Variants	143
10.1	Between Linearity and Convexity	143
10.2	Uniform and Strong Convexity	144
10.3	Quasiconvexity	148
	Exercises	151

11 Convex Variational Problems	155
11.1 Infima and Suprema	155
11.2 Minimizers	156
11.3 Uniqueness of Minimizers	157
11.4 Existence of Minimizers	157
11.5 Minimizing Sequences	160
Exercises	164
12 Infimal Convolution	167
12.1 Definition and Basic Facts	167
12.2 Infimal Convolution of Convex Functions	170
12.3 Pasch–Hausdorff Envelope	172
12.4 Moreau Envelope	173
12.5 Infimal Postcomposition	178
Exercises	178
13 Conjugation	181
13.1 Definition and Examples	181
13.2 Basic Properties	184
13.3 The Fenchel–Moreau Theorem	190
Exercises	194
14 Further Conjugation Results	197
14.1 Moreau’s Decomposition	197
14.2 Proximal Average	199
14.3 Positively Homogeneous Functions	201
14.4 Coercivity	202
14.5 The Conjugate of the Difference	204
Exercises	205
15 Fenchel–Rockafellar Duality	207
15.1 The Attouch–Brézis Theorem	207
15.2 Fenchel Duality	211
15.3 Fenchel–Rockafellar Duality	213
15.4 A Conjugation Result	217
15.5 Applications	218
Exercises	220
16 Subdifferentiability	223
16.1 Basic Properties	223
16.2 Convex Functions	227
16.3 Lower Semicontinuous Convex Functions	229
16.4 Subdifferential Calculus	233
Exercises	240

17	Differentiability of Convex Functions	241
	17.1 Directional Derivatives	241
	17.2 Characterizations of Convexity	244
	17.3 Characterizations of Strict Convexity	246
	17.4 Directional Derivatives and Subgradients	247
	17.5 Gâteaux and Fréchet Differentiability	251
	17.6 Differentiability and Continuity	257
	Exercises	258
18	Further Differentiability Results	261
	18.1 The Ekeland–Lebourg Theorem	261
	18.2 The Subdifferential of a Maximum	264
	18.3 Differentiability of Infimal Convolutions	266
	18.4 Differentiability and Strict Convexity	267
	18.5 Stronger Notions of Differentiability	268
	18.6 Differentiability of the Distance to a Set	271
	Exercises	273
19	Duality in Convex Optimization	275
	19.1 Primal Solutions via Dual Solutions	275
	19.2 Parametric Duality	279
	19.3 Minimization under Equality Constraints	283
	19.4 Minimization under Inequality Constraints	285
	Exercises	291
20	Monotone Operators	293
	20.1 Monotone Operators	293
	20.2 Maximally Monotone Operators	297
	20.3 Bivariate Functions and Maximal Monotonicity	302
	20.4 The Fitzpatrick Function	304
	Exercises	308
21	Finer Properties of Monotone Operators	311
	21.1 Minty’s Theorem	311
	21.2 The Debrunner–Flor Theorem	315
	21.3 Domain and Range	316
	21.4 Local Boundedness and Surjectivity	318
	21.5 Kenderov’s Theorem and Fréchet Differentiability	320
	Exercises	321
22	Stronger Notions of Monotonicity	323
	22.1 Para, Strict, Uniform, and Strong Monotonicity	323
	22.2 Cyclic Monotonicity	326
	22.3 Rockafellar’s Cyclic Monotonicity Theorem	327
	22.4 Monotone Operators on \mathbb{R}	329
	Exercises	330

23 Resolvents of Monotone Operators 333

23.1 Definition and Basic Identities 333

23.2 Monotonicity and Firm Nonexpansiveness 335

23.3 Resolvent Calculus 337

23.4 Zeros of Monotone Operators 344

23.5 Asymptotic Behavior 346

Exercises 349

24 Sums of Monotone Operators 351

24.1 Maximal Monotonicity of a Sum 351

24.2 3^* Monotone Operators 354

24.3 The Brézis–Haraux Theorem 357

24.4 Parallel Sum 359

Exercises 361

25 Zeros of Sums of Monotone Operators 363

25.1 Characterizations 363

25.2 Douglas–Rachford Splitting 366

25.3 Forward–Backward Splitting 370

25.4 Tseng’s Splitting Algorithm 372

25.5 Variational Inequalities 375

Exercises 378

26 Fermat’s Rule in Convex Optimization 381

26.1 General Characterizations of Minimizers 381

26.2 Abstract Constrained Minimization Problems 383

26.3 Affine Constraints 386

26.4 Cone Constraints 387

26.5 Convex Inequality Constraints 389

26.6 Regularization of Minimization Problems 393

Exercises 395

27 Proximal Minimization 399

27.1 The Proximal-Point Algorithm 399

27.2 Douglas–Rachford Algorithm 400

27.3 Forward–Backward Algorithm 405

27.4 Tseng’s Splitting Algorithm 407

27.5 A Primal–Dual Algorithm 408

Exercises 411

28 Projection Operators 415

28.1 Basic Properties 415

28.2 Projections onto Affine Subspaces 417

28.3 Projections onto Special Polyhedra 419

28.4 Projections Involving Convex Cones 425

28.5 Projections onto Epigraphs and Lower Level Sets 427

- Exercises 429
- 29 Best Approximation Algorithms** 431
 - 29.1 Dykstra’s Algorithm 431
 - 29.2 Haugazeau’s Algorithm..... 436
 - Exercises 440
- Bibliographical Pointers** 441
- Symbols and Notation** 443
- References** 449

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

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

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 $T: D \rightarrow D$ be a quasicontractive—in particular, nonexpansive—operator such that $\text{Fix } T \neq \emptyset$, 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 $\text{Fix } T$.

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 $(\forall n \in \mathbb{N}) d_C(x_{n+1}) \leq d_C(x_n)$. \square

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. \square

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 $x_n \not\rightarrow 0$.

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} ,

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

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 . \square

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 $P_C x_n \rightarrow x$.*

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

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 $x_n \rightarrow P_C x_0$.*

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

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

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

- (ii): Combine Theorem 5.5, Corollary 5.8, and (i). \square

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 $\text{int } C \neq \emptyset$. 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 $x \in \text{int } C$ 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

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

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

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

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. \square

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) \Rightarrow (ii): Clear.

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

(iii) \Rightarrow (i): Proposition 5.4(iii) implies that $d_C(x_n) \rightarrow 0$. Hence, $x_n - P_C x_n \rightarrow 0$ and (i) follows from Proposition 5.7. \square

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 $\kappa \in [0, 1[$,*

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

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

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

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

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

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

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 $T = -\text{Id}$ and $x_0 \neq 0$. In this case, however, it is clear that the *asymptotic regularity* property $x_n - Tx_n \rightarrow 0$ 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 $T: D \rightarrow D$ be a nonexpansive operator such that $\text{Fix } T \neq \emptyset$, and let $x_0 \in D$. Set*

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

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

- (i) $(x_n)_{n \in \mathbb{N}}$ converges weakly to a point in $\text{Fix } T$.
- (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 $\text{Fix } T$.

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

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

(ii): Since $D = -D$ is convex, $0 \in D$ and, since T is odd, $0 \in \text{Fix } T$. 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 $n \in \mathbb{N}$,

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

and, by the parallelogram identity,

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

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

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 $T: D \rightarrow D$ be a nonexpansive operator such that $\text{Fix } T \neq \emptyset$, 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*

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

Then the following hold:

- (i) $(x_n)_{n \in \mathbb{N}}$ is Fejér monotone with respect to $\text{Fix } T$.
- (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 $\text{Fix } T$.

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 $y \in \text{Fix } T$ and $n \in \mathbb{N}$,

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

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

(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 $n \in \mathbb{N}$,

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

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

(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 $x \in \text{Fix } T$. In view of Theorem 5.5, the proof is complete. \square

Proposition 5.15 *Let $\alpha \in]0, 1[$, let $T: \mathcal{H} \rightarrow \mathcal{H}$ be an α -averaged operator such that $\text{Fix } T \neq \emptyset$, 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*

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

Then the following hold:

- (i) $(x_n)_{n \in \mathbb{N}}$ is Fejér monotone with respect to $\text{Fix } T$.
- (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 $\text{Fix } T$.

Proof. Set $R = (1 - 1/\alpha)\text{Id} + (1/\alpha)T$ and $(\forall n \in \mathbb{N}) \mu_n = \alpha\lambda_n$. Then $\text{Fix } R = \text{Fix } T$ and R is nonexpansive by Proposition 4.25. In addition, we rewrite (5.15) as $(\forall n \in \mathbb{N}) x_{n+1} = x_n + \mu_n(Rx_n - x_n)$. 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. \square

Corollary 5.16 *Let $T: \mathcal{H} \rightarrow \mathcal{H}$ be a firmly nonexpansive operator such that $\text{Fix } T \neq \emptyset$, 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 $(\forall n \in \mathbb{N}) x_{n+1} = x_n + \lambda_n(Tx_n - x_n)$. Then the following hold:*

- (i) $(x_n)_{n \in \mathbb{N}}$ is Fejér monotone with respect to $\text{Fix } T$.
- (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 $\text{Fix} T$.

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

Example 5.17 Let $T: \mathcal{H} \rightarrow \mathcal{H}$ be a firmly nonexpansive operator such that $\text{Fix} T \neq \emptyset$, 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 $\text{Fix} T$.

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 $\bigcap_{i \in I} \text{Fix} T_i \neq \emptyset$, 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 $i: \{(k, l) \mid k \in \{1, \dots, p\}, l \in \{1, \dots, m_k\}\} \rightarrow I$ is surjective and that $\sum_{k=1}^p \omega_k = 1$. For every $k \in \{1, \dots, p\}$, set $I_k = \{i(k, 1), \dots, i(k, m_k)\}$, and set

$$\alpha = \max_{1 \leq k \leq p} \rho_k, \quad \text{where } (\forall k \in \{1, \dots, p\}) \quad \rho_k = \frac{m_k}{m_k - 1 + \frac{1}{\max_{i \in I_k} \alpha_i}}, \quad (5.16)$$

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

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

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

Proof. Set $T = \sum_{k=1}^p \omega_k R_k$, where $(\forall k \in \{1, \dots, p\}) R_k = T_{i(k,1)} \cdots T_{i(k,m_k)}$. 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 $\text{Fix} T = \bigcap_{i \in I} \text{Fix} T_i$. 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 $\text{Fix} R_k = \bigcap_{i \in I_k} \text{Fix} T_i$. In turn, we derive from Proposition 4.30 and (5.16) that T is α -averaged and, from Proposition 4.34, that $\text{Fix} T = \bigcap_{k=1}^p \text{Fix} R_k = \bigcap_{k=1}^p \bigcap_{i \in I_k} \text{Fix} T_i = \bigcap_{i \in I} \text{Fix} T_i$. \square

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 $C = \bigcap_{i \in I} C_i \neq \emptyset$. For every $i \in I$, let $\beta_i \in]0, 2[$ and set $T_i = (1 - \beta_i)\text{Id} + \beta_i P_{C_i}$. 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 $i: \{(k, l) \mid k \in \{1, \dots, p\}, l \in \{1, \dots, m_k\}\} \rightarrow I$ is surjective and that $\sum_{k=1}^p \omega_k = 1$. Furthermore, let $x_0 \in \mathcal{H}$ and set

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

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

Proof. For every $i \in I$, set $\alpha_i = \beta_i/2 \in]0, 1[$. 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 $k \in \{1, \dots, p\}$, $\max_{i \in I_k} \alpha_i \in]0, 1[$, which implies that $\rho_k \in]0, 1[$ and thus that $\alpha \in]0, 1[$. Altogether, the result follows from Corollary 5.18 with $\lambda_n \equiv 1$. \square

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 $C = \bigcap_{i \in I} C_i \neq \emptyset$, 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

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

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 $\text{Fix } T = \bigcap_{i \in I} \text{Fix } P_i = \bigcap_{i \in I} C_i = C$. Alternatively, apply Corollary 5.18. \square

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 $I =$*

$\{1, \dots, m\}$, let $(T_i)_{i \in I}$ be a family of nonexpansive operators from D to D such that $\text{Fix}(T_1 \cdots T_m) \neq \emptyset$, 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

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

Then $x_n - T_1 \cdots T_m x_n \rightarrow 0$, and there exist points $y_1 \in \text{Fix} T_1 \cdots T_m$, $y_2 \in \text{Fix} T_2 \cdots T_m T_1$, \dots , $y_m \in \text{Fix} T_m T_1 \cdots T_{m-1}$ such that

$$x_n \rightharpoonup y_1 = T_1 y_2, \quad (5.21)$$

$$T_m x_n \rightharpoonup y_m = T_m y_1, \quad (5.22)$$

$$T_{m-1} T_m x_n \rightharpoonup y_{m-1} = T_{m-1} y_m, \quad (5.23)$$

$$\vdots$$

$$T_3 \cdots T_m x_n \rightharpoonup y_3 = T_3 y_4, \quad (5.24)$$

$$T_2 \cdots T_m x_n \rightharpoonup y_2 = T_2 y_3. \quad (5.25)$$

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

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

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

$$(\text{Id} - T_m) x_n - (\text{Id} - T_m) y \rightarrow 0, \quad (5.27)$$

$$(\text{Id} - T_{m-1}) T_m x_n - (\text{Id} - T_{m-1}) T_m y \rightarrow 0, \quad (5.28)$$

$$\vdots$$

$$(\text{Id} - T_2) T_3 \cdots T_m x_n - (\text{Id} - T_2) T_3 \cdots T_m y \rightarrow 0, \quad (5.29)$$

$$(\text{Id} - T_1) T_2 \cdots T_m x_n - (T_2 \cdots T_m y - y) \rightarrow 0. \quad (5.30)$$

Upon adding (5.27)–(5.30), we obtain $x_n - T x_n \rightarrow 0$. 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 $y_1 \in \text{Fix} T$, which provides (5.21). On the other hand, (5.27) yields $T_m x_n - x_n \rightarrow T_m y_1 - y_1$. So altogether $T_m x_n \rightharpoonup T_m y_1 = y_m$, and we obtain (5.22). In turn, since (5.28) asserts that $T_{m-1} T_m x_n - T_m x_n \rightarrow T_{m-1} y_m - y_m$, we ob-

tain $T_{m-1}T_mx_n \rightarrow T_{m-1}y_m = y_{m-1}$, hence (5.23). Continuing this process, we arrive at (5.25). \square

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 $\text{Fix}(P_1 \cdots P_m) \neq \emptyset$ and set*

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

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

Proof. This follows from Proposition 4.8 and Theorem 5.22. \square

Remark 5.24 In Corollary 5.23, suppose that, for some $j \in I$, C_j is bounded. Then $\text{Fix}(P_1 \cdots P_m) \neq \emptyset$. 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 $C = \bigcap_{i \in I} C_i \neq \emptyset$, 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 $x_n \rightarrow y_1 \in \text{Fix}(P_1 \cdots P_m) = \bigcap_{i \in I} \text{Fix} P_i = C$. Alternatively, this is a special case of Example 5.20. \square

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 $x_n \rightarrow_{PC} x_0$. 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 $V = \text{Fix } T$ and $(\forall n \in \mathbb{N}) x_{n+1} = Tx_n$. Then $x_n \rightarrow P_V x_0 \Leftrightarrow x_n - x_{n+1} \rightarrow 0$.*

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

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

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

Then $x_n \rightarrow P_C x_0$.

Proof. Set $T = P_1 \cdots P_m$. Then T is nonexpansive, and $\text{Fix } T = C$ 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 $x_n - Tx_n \rightarrow 0$. Thus, by Proposition 5.27, $x_n \rightarrow P_C x_0$.

We now turn our attention to the general affine case. Since $C \neq \emptyset$, 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

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

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

Exercises

Exercise 5.1 Find a nonexpansive operator $T: \mathcal{H} \rightarrow \mathcal{H}$ 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., $x_n - x_{n+1} \rightarrow 0$.

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 $\overline{\text{conv}} C$.

Exercise 5.5 Let $T: \mathcal{H} \rightarrow \mathcal{H}$ be a nonexpansive operator such that $\text{Fix } T \neq \emptyset$, and let $(x_n)_{n \in \mathbb{N}}$ be a sequence in \mathcal{H} such that

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

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

Exercise 5.6 Find a nonexpansive operator $T: \mathcal{H} \rightarrow \mathcal{H}$ 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 $I = \{1, \dots, m\}$, let $(C_i)_{i \in I}$ be a family of closed convex subsets of \mathcal{H} such that $C = \bigcap_{i \in I} C_i \neq \emptyset$, 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

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

- (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

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

- (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

- 3^* monotone, 354–356
- accretive operator, 293
- acute cone, 105
- addition of functions, 6
- addition of set-valued operators, 2
- adjoint of a linear operator, 31
- affine constraint, 386
- affine hull, 1, 93
- affine minorant, 133, 134, 168, 184, 191, 192, 207, 223, 224
- affine operator, 3, 35, 44, 418
- affine projector, 48, 62
- affine subspace, 1, 43, 108
- almost surely, 29
- alternating minimizations, 162
- alternating projection method, 406, 432
- Anderson–Duffin theorem, 361
- angle bounded operator, 361, 362
- Apollonius’s identity, 30, 46
- approximating curve, 64
- asymptotic center, 117, 159, 195
- asymptotic regularity, 78
- at most single-valued, 2
- Attouch–Brézis condition, 210, 216
- Attouch–Brézis theorem, 207, 209
- autoconjugate function, 190, 239, 303
- averaged nonexpansive operator, 67, 72, 80–82, 294, 298, 440
- Baillon–Haddad theorem, 270
- ball, 43
- Banach–Alaoglu theorem, 34
- Banach–Picard theorem, 20
- Banach–Steinhaus theorem, 31
- barrier cone, 103, 156
- base of a topology, 7, 9, 16, 33
- base of neighborhoods, 23
- best approximation, 44, 278, 410
- best approximation algorithm, 410, 411, 431
- biconjugate function, 181, 185, 190, 192
- biconjugation theorem, 190
- bilinear form, 39, 384, 385
- Bishop–Phelps theorem, 107
- Boltzmann–Shannon entropy, 137, 139
- boundary of a set, 7
- bounded below, 134, 184
- Brézis–Haraux theorem, 358
- Bregman distance, 258
- Browder–Göhde–Kirk theorem, 64
- Brøndsted–Rockafellar theorem, 236
- Bunt–Motzkin theorem, 318
- Burg’s entropy, 138, 151, 244
- Cantor’s theorem, 17
- Cauchy sequence, 17, 24, 34, 46
- Cauchy–Schwarz inequality, 29
- Cayley transform, 349
- chain, 3
- chain rule, 40
- characterization of minimizers, 381
- Chebyshev center, 249, 250
- Chebyshev problem, 47
- Chebyshev set, 44–47, 318
- closed ball, 16
- closed convex hull, 43
- closed range, 220
- closed set, 7, 8
- closure of a set, 7, 22
- cluster point, 7, 8

- cocoercive operator, 60, 61, 68, 70, 270, 294, 298, 325, 336, 339, 355, 370, 372, 377, 379, 438
- coercive function, 158–161, 165, 202–204, 210
- cohyppomonotone operator, 337
- common fixed points, 71
- compact set, 7, 8, 13
- complementarity problem, 376
- complementary slackness, 291, 422
- complete metric space, 17
- composition of set-valued operators, 2
- concave function, 113
- cone, 1, 87, 285, 287, 376, 387, 389, 425
- conical hull, 87
- conjugate function, 181
- conjugation, 181, 197, 226, 230
- constrained minimization problem, 283, 285, 383
- continuity, 9
- continuous affine minorant, 168
- continuous convex function, 123, 136
- continuous function, 11
- continuous linear functional, 31
- continuous operator, 9, 63
- convergence of a net, 7
- convex combination, 44
- convex cone, 87, 89, 179, 183, 285, 425
- convex feasibility problem, 81, 84
- convex function, 113, 155
- convex hull, 43, 44
- convex integrand, 118, 138, 193, 238
- convex on a set, 114, 125
- convex programming problem, 290
- convex set, 43
- convexity with respect to a cone, 285
- core, 90, 95, 123, 207, 210, 214, 216, 241, 271
- counting measure, 140
- cyclically monotone operator, 326
- Debrunner–Flor theorem, 315
- decreasing function, 5
- decreasing sequence of convex sets, 48, 417
- demiclosedness principle, 63
- dense hyperplane, 123
- dense set, 7, 33, 123, 232
- descent direction, 248, 249
- descent lemma, 270
- diameter of a set, 16
- directed set, 3, 4, 22, 27
- directional derivative, 241, 247
- discontinuous linear functional, 32, 123, 169
- discrete entropy, 140
- distance, 27
- distance to a set, 16, 20, 24, 32, 34, 44, 49, 98, 167, 170, 173, 177, 183, 185, 188, 238, 271, 272
- domain of a function, 5, 6, 113
- domain of a set-valued operator, 2
- domain of continuity, 11
- Douglas–Rachford algorithm, 366, 376, 401, 404
- dual cone, 96
- dual optimal value, 214
- dual problem, 212, 214, 275, 279, 408
- dual solution, 275, 279
- duality, 211, 213, 275
- duality gap, 212, 214–216, 221
- Dykstra’s algorithm, 431, 432
- Eberlein–Šmulian theorem, 35
- effective domain of a function, 6
- Ekeland variational principle, 19
- Ekeland–Lebourg theorem, 263
- enlargement of a monotone operator, 309
- entropy of a random variable, 139
- epi-sum, 167
- epigraph, 5, 6, 12, 15, 113, 119, 133, 168
- equality constraint, 283

- Euclidean space, 28
- even function, 186
- eventually in a set, 4
- evolution equation, 313
- exact infimal convolution, 167, 170, 171, 207, 209, 210
- exact infimal postcomposition, 178
- exact modulus of convexity, 144–146
- existence of minimizers, 157, 159
- expected value, 29
- extended real line, 4
- extension, 297
- F_σ set, 24
- Farkas's lemma, 99, 106
- farthest-point operator, 249, 296
- Fejér monotone, 75, 83, 86, 160, 400
- Fenchel conjugate, 181
- Fenchel duality, 211
- Fenchel–Moreau theorem, 190
- Fenchel–Rockafellar duality, 213, 275, 282, 408
- Fenchel–Young inequality, 185, 226
- Fermat's rule, 223, 235, 381
- firmly nonexpansive operator, 59, 61–63, 68, 69, 73, 80, 81, 176, 270, 294, 298, 335, 337, 436
- first countable space, 23
- Fitzpatrick function, 304, 311, 351
- Fitzpatrick function of order n , 330
- fixed point, 62, 79–81
- fixed point iterations, 75
- fixed point set, 20, 62–64, 436
- forward–backward algorithm, 370, 377, 405, 438, 439
- forward–backward–forward algorithm, 375
- Fréchet derivative, 38, 39, 257
- Fréchet differentiability, 38, 176, 177, 243, 253, 254, 268–270, 320
- Fréchet gradient, 38
- Fréchet topological space, 23
- frequently in a set, 4
- function, 5
- G_δ set, 19, 263, 320
- Gâteaux derivative, 37
- Gâteaux differentiability, 37–39, 243, 244, 246, 251, 252, 254, 257, 267
- gauge, 120, 124, 202
- generalized inverse, 50, 251, 360, 361, 395, 418
- generalized sequence, 4
- global minimizer, 223
- gradient, 38, 176, 243, 244, 266, 267, 382
- gradient operator, 38
- graph, 5
- graph of a set-valued operator, 2
- Hölder continuous gradient, 269
- half-space, 32, 33, 43, 419, 420
- Hamel basis, 32
- hard thresholder, 61
- Haugazeau's algorithm, 436, 439
- Hausdorff distance, 25
- Hausdorff space, 7, 16, 33
- hemicontinuous operator, 298, 325
- Hessian, 38, 243, 245, 246
- Hilbert direct sum, 28, 226
- Hilbert space, 27
- Huber's function, 124
- hyperplane, 32, 34, 48, 123
- increasing function, 5
- increasing sequence of convex sets, 416
- indicator function, 12, 113, 173, 227
- inequality constraint, 285, 389
- infimal convolution, 167, 187, 207, 210, 237, 266, 359
- infimal postcomposition, 178, 187, 199, 218, 237
- infimum, 5, 157, 159, 184, 188

- infimum of a function, 6
- infinite sum, 27
- initial condition, 295
- integral function, 118, 138, 193, 238
- interior of a set, 7, 22, 90, 123
- inverse of a monotone operator, 295
- inverse of a set-valued operator, 2, 231
- inverse strongly monotone operator, 60
- Jensen's inequality, 135
- Karush–Kuhn–Tucker conditions, 393
- Kenderov theorem, 320
- kernel of a linear operator, 32
- Kirszbraun–Valentine theorem, 337
- Krasnosel'skiĭ–Mann algorithm, 78, 79
- Lagrange multiplier, 284, 287, 291, 386–388, 391
- Lagrangian, 280, 282
- Lax–Milgram theorem, 385
- least element, 3
- least-squares solution, 50
- Lebesgue measure, 29
- Legendre function, 273
- Legendre transform, 181
- Legendre–Fenchel transform, 181
- level set, 5, 6, 12, 15, 132, 158, 203, 383
- limit inferior, 5
- limit superior, 5
- line segment, 1, 43, 54, 132
- linear convergence, 20, 78, 372, 377, 406, 407
- linear equations, 50
- linear functional, 32
- linear monotone operator, 296–298, 355
- Lipschitz continuous, 20, 31, 59, 123, 176, 229, 339
- Lipschitz continuous gradient, 269–271, 405–407, 439
- local minimizer, 156
- locally bounded operator, 316, 319, 344
- locally Lipschitz continuous, 20, 122
- lower bound, 3
- lower level set, 5, 6, 148, 427
- lower semicontinuity, 10, 129
- lower semicontinuous, 10, 12
- lower semicontinuous convex envelope, 130, 185, 192, 193, 207
- lower semicontinuous convex function, 122, 129, 132, 185
- lower semicontinuous envelope, 14, 23
- lower semicontinuous function, 129
- lower semicontinuous infimal convolution, 170, 210
- marginal function, 13, 120, 152
- max formula, 248
- maximal element, 3
- maximal monotone operator, 297
- maximal monotonicity and continuity, 298
- maximal monotonicity of a sum, 351
- maximally cyclically monotone operator, 326
- maximally monotone extension, 316, 337
- maximally monotone operator, 297, 298, 311, 335, 336, 338, 339, 438
- maximum of a function, 6
- measure space, 28, 295
- metric space, 16
- metric topology, 16, 33, 34
- metrizable topology, 16, 23, 34
- midpoint convex function, 141

- midpoint convex set, 57
- minimax, 218
- minimization in a product space, 403
- minimization problem, 13, 156, 381, 393, 401, 402, 404–406
- minimizer, 156, 157, 159, 163, 243, 384
- minimizing sequence, 6, 13, 160, 399
- minimum of a function, 6, 243
- Minkowski gauge, 120, 124, 202
- Minty's parametrization, 340
- Minty's theorem, 311
- modulus of convexity, 144
- monotone extension, 297
- monotone linear operator, 296
- monotone operator, 244, 293, 311, 351, 363
- monotone set, 293
- Moore–Penrose inverse, 50
- Moreau envelope, 173, 175, 176, 183, 185, 187, 197, 198, 270, 271, 276, 277, 334, 339, 342
- Moreau's conical decomposition, 98
- Moreau's decomposition, 198
- Moreau–Rockafellar theorem, 204
- negative orthant, 5
- negative real number, 4
- neighborhood, 7
- net, 4, 5, 22, 27, 53, 314
- nonexpansive operator, 59, 60, 62, 63, 79, 159, 270, 294, 298, 336, 348, 439
- nonlinear equation, 325
- norm, 27, 35, 40, 115, 118, 144, 147, 150, 151, 183, 199, 231, 252
- norm topology, 33
- normal cone, 101, 227, 230, 238, 272, 304, 334, 354, 383, 389
- normal equation, 50
- normal vector, 107
- obtuse cone, 105
- odd operator, 79, 379
- open ball, 16
- open set, 7
- operator splitting algorithm, 366
- Opial's condition, 41
- optimal value, 214
- order, 3
- orthogonal complement, 27
- orthonormal basis, 27, 37, 161, 301, 313, 344
- orthonormal sequence, 34
- outer normal, 32
- parallel linear subspace, 1
- parallel projection algorithm, 82
- parallel splitting algorithm, 369, 404
- parallel sum of monotone operators, 359
- parallelogram identity, 29
- parametric duality, 279
- paramonotone operator, 323, 385
- partial derivative, 259
- partially ordered set, 3
- Pasch–Hausdorff envelope, 172, 179
- periodicity condition, 295
- perspective function, 119, 184
- POCS algorithm, 84
- pointed cone, 88, 105
- pointwise bounded operator family, 31
- polar cone, 96, 110
- polar set, 110, 202, 206, 428
- polarization identity, 29
- polyhedral cone, 388, 389
- polyhedral function, 216–218, 381, 383, 388, 389
- polyhedral set, 216, 383
- polyhedron, 419
- positive operator, 60
- positive orthant, 5, 426
- positive real number, 4
- positive semidefinite matrix, 426

- positively homogeneous function, 143, 201, 229, 278
- positively homogeneous operator, 3
- power set, 2
- primal optimal value, 214
- primal problem, 212, 214, 275, 279, 408
- primal solution, 275, 279
- primal–dual algorithm, 408
- probability simplex, 426
- probability space, 29, 139
- product topology, 7
- projection algorithm, 431, 439
- projection onto a ball, 47
- projection onto a convex cone, 97, 98, 425
- projection onto a half-space, 419
- projection onto a hyperplane, 48, 419
- projection onto a hyperslab, 419
- projection onto a linear subspace, 49
- projection onto a lower level set, 427
- projection onto a polar set, 428
- projection onto a ray, 426
- projection onto a set, 44
- projection onto an affine subspace, 48, 77, 417
- projection onto an epigraph, 133, 427
- projection operator, 44, 61, 62, 175, 177, 334, 360, 361, 415
- projection theorem, 46, 238
- projection-gradient algorithm, 406
- projector, 44, 61
- proper function, 6, 132
- proximal average, 199, 205, 271, 307
- proximal mapping, 175
- proximal minimization, 399
- proximal-gradient algorithm, 405, 439
- proximal-point algorithm, 345, 399, 438
- proximal set, 44–46
- proximity operator, 175, 198, 199, 233, 243, 244, 271, 334, 339, 342–344, 375, 381, 382, 401, 402, 404, 405, 415, 428
- pseudocontractive operator, 294
- pseudononexpansive operator, 294
- quadratic function, 251
- quasiconvex function, 148, 157, 160, 165
- quasinonexpansive operator, 59, 62, 71, 75
- quasirelative interior, 91
- random variable, 29, 135, 139, 194
- range of a set-valued operator, 2
- range of a sum of operators, 357, 358
- recession cone, 103
- recession function, 152
- recovery of primal solutions, 275, 408
- reflected resolvent, 336, 363, 366
- regularization, 393
- regularized minimization problem, 393
- relative interior, 90, 96, 123, 210, 216, 234
- resolvent, 333, 335, 336, 366, 370, 373
- reversal of a function, 186, 236, 342
- reversal of an operator, 3, 340
- Riesz–Fréchet representation, 31
- right-shift operator, 330, 356
- Rådström’s cancellation, 58
- saddle point, 280–282
- scalar product, 27
- second Fréchet derivative, 38
- second Gâteaux derivative, 38
- second-order derivative, 245, 246

- selection of a set-valued operator, 2
- self-conjugacy, 183, 185
- self-dual cone, 96, 186
- self-polar cone, 186
- separable Hilbert space, 27, 194
- separated sets, 55
- separation, 55
- sequential cluster point, 7, 15, 33
- sequential topological space, 16, 23
- sequentially closed, 15, 16, 53, 231, 300, 301
- sequentially compact, 15, 16, 36
- sequentially continuous, 15, 16
- sequentially lower semicontinuous, 15, 129
- set-valued operator, 2
- shadow sequence, 76
- sigma-finite measure space, 194
- Slater condition, 391
- slope, 168
- soft thresholder, 61, 199
- solid cone, 88, 105
- span of a set, 1
- splitting algorithm, 375, 401, 402, 404, 405
- Stampacchia's theorem, 384, 395
- standard unit vectors, 28, 89, 92
- steepest descent direction, 249
- strict contraction, 64
- strict epigraph, 180
- strictly convex function, 114, 144, 161, 267, 324
- strictly convex on a set, 114
- strictly convex set, 157
- strictly decreasing function, 5
- strictly increasing function, 5
- strictly monotone operator, 323, 344
- strictly negative real number, 4
- strictly nonexpansive operator, 325
- strictly positive operator, 246
- strictly positive orthant, 5
- strictly positive real number, 4
- strictly quasiconvex function, 149, 157
- strictly quasinonexpansive operator, 59, 71
- string-averaged relaxed projections, 82
- strong convergence, 33, 37
- strong relative interior, 90, 95, 96, 209, 210, 212, 215, 217, 234, 236, 381
- strong separation, 55
- strong topology, 33
- strongly convex function, 144, 159, 188, 197, 270, 276, 324, 406
- strongly monotone operator, 323, 325, 336, 344, 372
- subadditive function, 143
- subdifferentiable function, 223, 247
- subdifferential, 223, 294, 304, 312, 324, 326, 354, 359, 381, 383
- subdifferential of a maximum, 264
- subgradient, 223
- sublinear function, 143, 153, 156, 241
- subnet, 4, 8, 22
- sum of linear subspaces, 33
- sum of monotone operators, 351
- sum rule for subdifferentials, 234
- summable family, 27
- supercoercive function, 158, 159, 172, 203, 210, 229
- support function, 109, 156, 183, 195, 201, 229, 240
- support point, 107, 109, 164
- supporting hyperplane, 107, 109
- supremum, 5, 129, 188
- supremum of a function, 6
- surjective monotone operator, 318, 320, 325, 358
- tangent cone, 100
- time-derivative operator, 295, 312, 334
- Toland-Singer duality, 205

- topological space, 7
- topology, 7
- totally ordered set, 3
- trace of a matrix, 28
- translation of an operator, 3
- Tseng's splitting algorithm, 373, 378, 407
- Tykhonov regularization, 393
- unbounded net, 314
- uniform boundedness principle, 31
- uniformly convex function, 144, 147, 324, 394, 399
- uniformly convex on a set, 144, 147, 324, 407
- uniformly convex set, 164, 165
- uniformly monotone on a set, 324
- uniformly monotone on bounded sets, 346, 367, 373, 376, 378, 408
- uniformly monotone operator, 323, 325, 344, 354, 358, 367, 373, 376, 378, 408
- uniformly quasiconvex function, 149, 163
- upper bound, 3
- upper semicontinuous function, 11, 124, 281
- value function, 279, 289
- variational inequality, 375–378, 383
- Volterra integration operator, 308
- von Neumann's minimax theorem, 218
- von Neumann–Halperin theorem, 85
- weak closure, 53
- weak convergence, 33, 36, 79–81
- weak sequential closure, 53
- weak topology, 33
- weakly closed, 33–35, 45, 53
- weakly compact, 33–35
- weakly continuous operator, 33, 35, 62, 418
- weakly lower semicontinuous, 35, 129
- weakly lower semicontinuous function, 33
- weakly open, 33
- weakly sequentially closed, 33–35, 53
- weakly sequentially compact, 33, 35
- weakly sequentially continuous operator, 343, 426
- weakly sequentially lower semicontinuous, 129
- Weierstrass theorem, 13
- Yosida approximation, 333, 334, 336, 339, 345, 347, 348
- zero of a monotone operator, 344, 345, 347, 381, 438
- zero of a set-valued operator, 2
- zero of a sum of operators, 363, 366, 369, 375