Contemporary Mathematics
Volume **00**, 19xx

The method of cyclic projections for closed convex sets in Hilbert space

Heinz H. Bauschke, Jonathan M. Borwein, and Adrian S. Lewis

Abstract. The method of cyclic projections is a powerful tool for solving convex feasibility problems in Hilbert space. Although in many applications, in particular in the field of image reconstruction (electron microscopy, computed tomography), the convex constraint sets do not necessarily intersect, the method of cyclic projections is still employed. Results on the behaviour of the algorithm for this general case are improved, unified, and reviewed. The analysis relies on key concepts from convex analysis and the theory of nonexpansive mappings. The notion of the angle of a tuple of subspaces is introduced. New linear convergence results follow for the case when the constraint sets are closed subspaces whose orthogonal complements have a closed sum; this holds, in particular, for hyperplanes or in Euclidean space.

1. The problem

1.1. Introduction and formulation. A frequent problem in various areas of mathematics and physical sciences consists of determining a "solution" in form of a point satisfying some convex "constraints", i.e. lying in the intersection of certain convex sets. This problem is usually referred to as the *Convex Feasibility Problem.*

Often these constraint sets are closed subspaces in some Hilbert space; an immensely successful algorithm for solving this problem is the method of *cyclic projections*, where a sequence is generated by projecting cyclically onto the constraint subspaces. The fundamental result is due to von Neumann and Halperin:

Fact 1.1.1. Suppose $C_1, \ldots, C_N$ are intersecting closed affine subspaces with corresponding projections $P_1, \ldots, P_N$ in some Hilbert space H. If $x_0 \in H$, then the sequence

$$(x_0, x_1 = P_1 x_0, x_2 = P_2 x_1, \ldots, x_N = P x_{N-1}, x_{N+1} = P_1 x_N, \ldots)$$

converges in norm to the projection of x_0 onto $\bigcap_{i=1}^N C_i$.

1991 *Mathematics Subject Classification.* Primary 47H09, 49M45, 90C25; Secondary 46N10, 47N10, 52A05, 65F10, 92C55.

Research partially supported by NSERC.

In 1933, von Neumann [**81**, Theorem 13.7] proved this result for two closed subspaces; it was later extended to finitely many closed subspaces by Halperin [**52**]. (The generalization to intersecting closed affine subspaces is easy.)

An impressive survey on applications of the method of cyclic projections for intersecting closed affine subspaces can be found in Deutsch's [**28**]. In particular, if each C_i is a hyperplane, then one obtains the well-known method of Kaczmarz [**57**] for solving systems of linear equations.

In some important applications, however, Fact 1.1.1 cannot be used since either the constraints are not affine (i.e. the C_i are not all closed affine subspaces) or the constraints are inconsistent (i.e. the intersection $\bigcap_{i=1}^N C_i$ is empty) or even both.

Non-affine constraints occur, for instance, when the desired solution is supposed to be nonnegative, or bounded by some prescribed value, or a solution of a linear inequality; we refer the interested reader to the surveys [**8, 24, 25**].
In theory the mathematical models employed almost always possess consistent constraints; in practice, however, due (unavoidably) to inaccurate measurements, one is often confronted with *inconsistent constraints*.

Inconsistent (non-affine) constraints constitute a typical example of an *ill-posed problem*. They occur frequently in the field of *image reconstruction*: for instance in *signal reconstruction* (Goldburg and Marks [**48**]), in *electron microscopy* (Sezan [**75**]), in *limited-angle tomography* (Boyd and Little [**14**]), and in *medical imaging* (Viergever [**80**]); in particular in *computed tomography* (Censor [**21**]).

Our objective in this paper is to improve, unify, and review results on the method of cyclic projections for possibly non-intersecting closed convex nonempty sets in Hilbert space. We will also speculate a little on some open questions and hope very much to transmit some of our interest and enthusiasm to the reader.

The mathematical model we analyze is described as follows:

SETTING. Throughout the paper, we assume that H is a real Hilbert space and $C_1, \dots, C_N$ are closed convex nonempty subsets of H with corresponding projections $P_1, \dots, P_N$. (We discuss projections in Section 2.)
For an arbitrary *starting point* $x^0 \in H$, the *method of cyclic projections* generates N sequences $(x_i^n)_n$ by

$$\boxed{\begin{array}{llll} x_1^1 := P_1 x^0, & x_2^1 := P_2 x_1^1, & \dots, & x_N^1 := P_N x_{N-1}^1, \\ x_1^2 := P_1 x_N^1, & x_2^2 := P_2 x_1^2, & \dots, & x_N^2 := P_N x_{N-1}^2, \\ x_1^3 := P_1 x_N^2, & \quad \dots . & & \end{array}}$$

We collect these sequences cyclically in one sequence $(x^0, x_1^1, x_2^1, \dots, x_N^1, x_1^2, x_2^2, \dots)$ to which we refer as the *orbit generated by* x^0 or the *orbit with starting point* x^0. We further define

$$\boxed{Q_1 := P_1 P_N P_{N-1} \cdots P_2,\ Q_2 := P_2 P_1 P_N \cdots P_3,\ \dots,\ Q_N := P_N P_{N-1} \cdots P_1,}$$

which allows us to write more concisely

$$x_i^n := Q_i^{n-1} x_i^1, \text{ for all } n \geq 1 \text{ and every } i;$$

for convenience, we set $P_0 := P_N$, $P_{N+1} := P_1$, $x_0^n := x_N^{n-1}$, and $x_{N+1}^n := x_1^{n+1}$, so that

$$x_{i+1}^n = P_{i+1}x_i^n, \text{ for all } n \geq 1 \text{ and every } i.$$

When appropriate, we will similarly identify $i = 0$ with $i = N$ and $i = N + 1$ with $i = 1$.

OVERVIEW.

Basic facts on *projections* and some examples are contained in Section 2. We touch on *quasi-projections*, a useful concept related to projections and recently introduced by Baillon and Bruck [**5**].

In Section 3, we provide a "toolbox" which makes the subsequent analysis clearer: some of our results rely indirectly on the notion of a *polar cone*. *Recession cones* are intimately connected with convex sets and their projections. One important example of a convex set is a *convex polyhedron*; here the study of the mappings Q_i naturally involves *polyhedral multifunctions*. We give some basic facts on convex *Kadec/Klee sets*, a notion which generalize the better-known concept of (locally) uniformly convex sets. A discussion on *(bounded) (linear) regularity* of N-tuples of closed convex sets follows; this key condition is sufficient for norm convergence of any orbit in the consistent case. We then introduce the notion of the *angle of an N-tuple* $(L_1, \ldots, L_N)$ *of closed subspaces* and show that the angle is positive if and only if $L_1^\perp + \cdots + L_N^\perp$ is closed. This strengthens results on the well-known two set case (Deutsch [**27**], Bauschke and Borwein [**9**]) and is used to yield new linear convergence results. The section concludes with a short review on *Fejér monotone sequences.*

Each projection P_i is *nonexpansive* and so are the compositions Q_i. Moreover, they actually satisfy various stronger conditions introduced by Bruck and Reich [**19**], by De Pierro and Iusem [**71**], and by Bauschke and Borwein [**8**]. These conditions, which yield significant information on the behaviour of orbits, are discussed in Section 4.

Section 5 contains our main results. After considering the *geometry of the problem*, we give a *dichotomy result on orbits* which roughly says that if each Q_i is fixed point free, then the orbit has no bounded subsequence; otherwise, each subsequence (x_i^n) converges weakly to some fixed point of Q_i. (Throughout this section, we specialize our general results to the *two set case* ($N = 2$) and to the *consistent case* ($\bigcap_{i=1}^N C_i \neq \emptyset$) to allow comparison with known results for these important special cases.) Two central questions arise:

1. When does each Q_i have a fixed point?
2. If each Q_i has a fixed point, when do the subsequences (x_i^n) converge in norm (or even linearly)?

Concerning Question 1, we provide sufficient conditions for the existence of fixed points or of *approximate fixed points* (that is, $\inf_{x\in H} \|x - Q_i x\| = 0$, for each i). It follows that while fixed points of Q_i need not exist for non-intersecting closed affine subspaces, approximate fixed points must.

In respect to Question 2, a variety of conditions guaranteeing norm convergence (in the presence of fixed points for each Q_i) is offered: one of the sets C_i has to be *(boundedly) compact* or all sets are *Kadec/Klee*, or *convex polyhedra*, or *closed*

affine subspaces. In the affine subspace case, each sequence (x_i^n) converges to the fixed point of Q_i nearest to x^0. Moreover, the convergence will be linear, whenever the angle of the N-tuple of the associated closed subspaces is positive. In contrast, if $N = 2$ and the angle is zero, then the convergence will be "arbitrarily slow". We draw connections to the theory of generalized inverses by specializing to hyperplanes with linearly independent normal vectors.
The analysis in Section 5 extends results by Bauschke and Borwein [**9**], by Cheney and Goldstein [**23**], by Gubin et al. [**51**], by De Pierro and Iusem [**71**], by Kosmol [**60**], and by Kosmol and Zhou [**61, 62**]. We end the paper by referring the reader to related algorithms.

1.2. Interpretations and notation. Because the orbit (x_i^n) is generated by the rule $x_i^n := P_i x_{i-1}^n$, the problem described in the previous subsection allows interpretations in the following (related) mathematical areas:

- *Difference Equations:* the orbit (x_i^n) is given by a *nonautonomous difference equation* ([**63**]).
- *Discrete Dynamical Systems:* the orbit (x_i^n) is the *orbit of a discrete dynamical system* ([**66**]).
- *Discrete Processes:* the orbit (x_i^n) is the *positive orbit of a difference equation* ([**64**]).

Less seriously, one can also think about our problem as a *deterministic magnetic billiard*: the starting point x^0 is a *metal billiard-ball* and the sets C_i are *electric magnets*. Now switch the magnets on and off cyclically: the trajectory of the billard ball corresponds exactly to the orbit (x_i^n).

NOTATION. The *norm* of a point $x \in H$ is given by $\|x\| := \sqrt{\langle x, x\rangle}$, where $\langle \cdot, \cdot \rangle$ is the *inner product* of H. We write I for the *identity mapping* on H, B_H for the *unit ball* $\{x \in H : \|x\| \leq 1\}$, and $B(x_0, r)$ for $x_0 + rB_H$, where $x_0 \in H$ and $r \geq 0$. Suppose C is a nonempty subset of H. The *interior* (resp. *intrinsic core, boundary, closure, span*) *of* C is denoted $\mathrm{int}(C)$ (resp. $\mathrm{icr}(C)$, $\mathrm{bd}(C)$, $\overline{C}$, $\mathrm{span}(C)$). The *indicator function of a* C, ι_C, is defined by $\iota_C(x) = 0$, if $x \in C$; $+\infty$, otherwise; its Fenchel conjugate, the *support function of* C, is $\iota_C^* = \sup\langle C, \cdot\rangle$. The *normal cone* $N_C(x)$ *at* $x \in C$ is given by $N_C(x) = \{x^* \in H : \langle x^*, C - x\rangle \leq 0\}$. If C is a subspace, then we write $C^\perp$ for its *orthogonal complement*. The subdifferential of a convex function f from H to $\mathbb{R}$ at $x \in H$ is abbreviated $\partial f(x) = \{x^* \in H : \langle x^*, h\rangle \leq f(x+h) - f(x)\ \forall h \in H\}$. If D is another nonempty subset of H, then the *gap between* C *and* D is defined by $d(C, D) := \inf_{c \in C,\, d \in D} \|c - d\|$. For sequences, we use the symbol $\to$ (resp. $\rightharpoonup$) to indicate *norm* (resp. *weak*) *convergence*. Finally, we write $[\cdot]$ for the *mod* function with values in $\{1, \ldots, N\}$.

2. Projections

2.1. Basic properties.

DEFINITION 2.1.1. Suppose C is a closed convex nonempty subset of H. Then for every $x \in H$, there exists a unique nearest point, say $P_C x$, to x in C:

$$\|x - P_C x\| \leq \|x - c\|, \text{ for all } c \in C.$$

The associated mapping P_C from H to C which sends every point to its nearest point in C is called the *projection onto* C. The function

$$d(\cdot, C) : H \to \mathbb{R} : x \mapsto \|x - P_C x\|$$

is called the *distance function (to C)*.

We list basic facts on projections. (i) and (ii) are folklore (see for instance [**84**, Section 1]) whereas (iii) is a simple but useful expansion.

PROPOSITION 2.1.2. Suppose C is a closed convex nonempty subset of H. Then:

(i) The projection onto C is well-defined and the distance function to C is continuous and convex, hence weakly lower semicontinuous.

(ii) For any two points $x, c^* \in H$, we have $c^* = P_C x$ if and only if

$$c^* \in C \text{ and } \langle C - c^*, x - c^* \rangle \leq 0.$$

(iii) For any $x, y \in H$, the formula below contains exclusively nonnegative terms:

$$\begin{aligned} \|x-y\|^2 &= \|(I-P_C)x - (I-P_C)y\|^2 + \|P_C x - P_C y\|^2 \\ &\quad + 2\langle (I-P_C)x - (I-P_C)y, P_C x - P_C y\rangle \\ &= \|(x-y) - (P_C x - P_C y)\|^2 + \|P_C x - P_C y\|^2 \\ &\quad + 2\langle x - P_C x, P_C x - P_C y\rangle + 2\langle y - P_C y, P_C y - P_C x\rangle; \end{aligned}$$

in particular, $\langle x - y, P_C x - P_C y\rangle \geq \|P_C x - P_C y\|^2$.

COROLLARY 2.1.3. Suppose (x_i^n) (resp. (y_i^n)) is the orbit generated by some x^0 (resp. y^0) in H. Then the sequences

$$(\|(x_i^n - y_i^n) - (x_{i+1}^n - y_{i+1}^n)\|^2),\ (\langle x_i^n - x_{i+1}^n, x_{i+1}^n - y_{i+1}^n\rangle),\ (\langle y_i^n - y_{i+1}^n, y_{i+1}^n - x_{i+1}^n\rangle)$$

are nonnegative and tend to zero for each i. In fact, the series generated by the terms of these sequences is convergent.

2.2. Examples.

EXAMPLES 2.2.1. Suppose C is a closed convex nonempty subset of H and x is an arbitrary point in H.

(i) If $C = B_H$, then $P_C x = \begin{cases} x, & \text{if } \|x\| \leq 1; \\ x/\|x\|, & \text{otherwise.} \end{cases}$

(ii) If C is a hyperplane, say $C = \{x \in H : \langle a, x\rangle = b\}$ with $a \in H \setminus \{0\}$, then $P_C x = x - ((\langle a, x\rangle - b)/\|a\|^2)a$.

(iii) If C is a halfspace, say $C = \{x \in H : \langle a, x\rangle \leq b\}$ with $a \in H \setminus \{0\}$, then $P_C x = x - ((\langle a, x\rangle - b)^+/\|a\|^2)a$.

(iv) If H is a Hilbert lattice with lattice cone H^+, then $P_C x = x^+$.

(v) Suppose C is a closed convex cone. If the negative polar cone is denoted $C^\ominus$ (see Subsection 3.1), then $P_{C^\ominus}(x) = x - P_C x$. In particular, if C is a closed subspace, then $P_{C^\perp}(x) = x - P_C x$.

(vi) Suppose Y is another Hilbert space and $A : H \to Y$ is a continuous linear operator with closed range. Let $C = \text{range}(A^*)$. If the generalized inverse of A is denoted $A^\dagger$ (see Definiton 5.7.10), then $P_C x = A^\dagger A x$.

(vii) Suppose $Y = \mathbb{R}^N$ and a continuous linear operator A is given by $Ah = (\langle a_1, h\rangle, \ldots, \langle a_N, h\rangle)$ for some vectors $a_1, \ldots, a_N \in H$. Suppose C is equal to $\text{span}\{a_1, \ldots, a_N\}$. Then $P_C x = A^\dagger A x$. Furthermore, if the vectors $a_1, \ldots, a_N$ are linearly independent, then $P_C x = A^*(AA^*)^{-1}Ax$.

PROOF. (i), (ii), (iii), and (v) are easy to check. (iv) is proved by Borwein and Yost's [**13**, Theorem 8]. (vi) is due to Desoer and Whalen; see [**26**]. (vii) is a special case of (vi) and was used in [**9**, Section 5]. □

The following proposition is useful for building more examples of projections. We omit the simple proof.

PROPOSITION 2.2.2. Suppose C is a closed convex nonempty subset of H and x is an arbitrary point in H. Then:

(i) (translation formula) If $z \in H$, then $P_{z+C}(x) = z + P_C(x - z)$.
(ii) (dilation formula) If $\rho \neq 0$, then $P_{\rho C}(x) = \rho P_C(x/\rho)$.

2.3. Baillon and Bruck's quasi-projection.

DEFINITION 2.3.1. (Baillon and Bruck's quasi-projection [**5**]) Suppose C is a closed convex nonempty subset of H. For every $x \in H$, define the *quasi-projection of x onto C*, $\mathcal{Q}_C x$, by

$$\mathcal{Q}_C x = \{c^* \in C : \|c^* - c\| \leq \|x - c\|, \text{ for all } c \in C\}.$$

It is easy to verify the following proposition (see also [**7**]):

PROPOSITION 2.3.2. Suppose C is a closed convex nonempty subset of H and x is a point in H. Then:

(i) $\mathcal{Q}_C x$ is a bounded closed convex nonempty subset of C; moreover:

$$P_C x \in \mathcal{Q}_C x \subseteq C \cap B(P_C x, d(x, C)).$$

Consequently, if $x \in C$, then $\mathcal{Q}_C x = \{P_C x\}$.
(ii) (translation formula) $\mathcal{Q}_{z+C} x = z + \mathcal{Q}_C(x - z)$, for all $z \in H$.
(iii) If C is a closed affine subspace, then $\mathcal{Q}_C x = \{P_C x\}$.

3. A toolbox

3.1. Polar cones.

DEFINITION 3.1.1. The *positive* (resp. *negative*) *polar cone* $S^{\oplus}$ (resp. $S^{\ominus}$) of a nonempty subset S of H is given by

$$S^{\oplus} := \{x \in H : \langle x, S\rangle \geq 0\} \text{ (resp. } S^{\ominus} := \{x \in H : \langle x, S\rangle \leq 0\} = -S^{\oplus}).$$

The next proposition is readily verified:

PROPOSITION 3.1.2.

(i) Every positive (resp. negative) polar cone is a closed convex nonempty cone.
(ii) If $S_1, S_2, \ldots, S_M$ are nonempty subsets of H, then $\overline{S_1^{\oplus} + S_2^{\oplus} + \cdots + S_M^{\oplus}} \subseteq (S_1 \cap S_2 \cap \cdots \cap S_M)^{\oplus}$.

The class of closed subspaces can be characterized in the class of closed convex nonempty cones:

PROPOSITION 3.1.3. Suppose S is a closed convex nonempty cone in H. Then S is a subspace if and only if $S \cap S^{\oplus} = \{0\}$.

PROOF. "$\Longrightarrow$": If S is a subspace, then $S^{\oplus} = S^{\perp}$ and hence $S \cap S^{\oplus} = \{0\}$.
"$\Longleftarrow$": By Examples 2.2.1.(v) and Proposition 2.2.2, $x = P_S x - P_{S^{\oplus}}(-x)$, for all $x \in H$. Let us assume that S is not a subspace. Then there exists some vector $x \in S$ such that $-x \notin S$, that is $P_S(-x) \neq -x$. Thus $-x = P_S(-x) - P_{S^{\oplus}} x$, where $P_{S^{\oplus}} x \neq 0$. Since $P_{S^{\oplus}} x = P_S(-x) + x \in S$, we have found $P_{S^{\oplus}} x \in (S \cap S^{\oplus}) \setminus \{0\}$. The proof is complete. $\square$

3.2. Recession cones.

DEFINITION 3.2.1. The *recession cone* $\operatorname{rec}(C)$ of a closed convex nonempty subset C of H is given by

$$\operatorname{rec}(C) := \{x \in H : x + C \subseteq C\}.$$

PROPOSITION 3.2.2. Suppose C is a closed convex nonempty subset of H. Then:

(i) (Zarantonello's [**84**, Section 3]) $\operatorname{rec}(C)$ is a closed convex nonempty cone.
(ii) A point $x \in H$ belongs to $\operatorname{rec}(C)$ if and only if there exists a sequence (c_n) in C and a sequence of reals (r_n) tending to $+\infty$ such that (c_n/r_n) converges weakly to x.
(iii) If D is a closed convex superset of C, then $\operatorname{rec}(C) \subseteq \operatorname{rec}(D)$.

PROOF. (ii): "$\Longrightarrow$": Since $\operatorname{rec}(C)$ is cone, points of the form $c + nx$ belong to C, for all $c \in C$ and $n \geq 1$. Let us choose $c_n := c + nx$ and $r_n := n$. Then the sequence $(c_n/r_n) = (c/n + x)$ converges to x, as desired.
"$\Longleftarrow$": Fix any $c \in C$; then $(1 - 1/r_n)c + c_n/r_n \in C$, for all large n. Since C is weakly closed, the weak limit $c + x$ lies in C. Finally, (iii) is obvious from (ii). $\square$

There is an intimate connection between the projection, the recession cone, and the indicator function of a given closed convex nonempty set:

FACT 3.2.3. Suppose C is a closed convex nonempty subset of H. Then:

(i) (Castaing and Valadier's [**20**, Proposition I-7]) $\operatorname{rec}(C) = (\operatorname{domain}(\iota_C^*))^{\ominus}$.
(ii) (Zarantonello's [**84**, Theorem 3.1]) $\overline{\operatorname{range}}(P_C - I) = (\operatorname{rec}(C))^{\oplus}$.

REMARK 3.2.4. The set $\operatorname{domain}(\iota_C^*) = \{x \in H : \sup\langle x, C\rangle < \infty\}$ is also known as the *barrier cone of* C; see for instance [**4**].

3.3. Convex polyhedra.

DEFINITION 3.3.1. The intersection of finitely many halfspaces is called a *convex polyhedron.*

PROPOSITION 3.3.2. Suppose each C_i is a convex polyhedron given as $\{x \in H : \langle a_{i,j}, x\rangle \leq b_{i,j}\rangle\}$ for finitely many vectors $a_{i,j}$ and real numbers $b_{i,j}$. Suppose further K is a closed subspace of $\bigcap_{i,j} \operatorname{kernel}(a_{i,j})$. Define $D_i := C_i \cap K^{\perp}$. Then:

(i) Each D_i is a convex polyhedron in $K^{\perp}$.
(ii) $P_i = P_K + P_{D_i} P_{K^{\perp}}$, for each i.
(iii) $I - Q_N = P_{K^{\perp}} - P_{D_N} \cdots P_{D_1} P_{K^{\perp}}$.
(iv) $\inf_{x \in H} \|x - Q_N x\|$ is attained if and only if $\inf_{y \in K^{\perp}} \|y - P_{D_N} \cdots P_{D_1} y\|$ is.

PROOF. (i) is obvious. (ii): Fix $x \in H$ and an index i. We check Proposition 2.1.2.(ii): since $C_i = K \oplus D_i$, the point $P_K x + P_{D_i} P_{K^{\perp}} x$ belongs to C_i. Pick any $k \in K$ and $d_i \in D_i$. Then

$$\begin{aligned} & \langle k + d_i - (P_K x + P_{D_i} P_{K^\perp} x), x - (P_K x + P_{D_i} P_{K^\perp} x)\rangle \\ = \ & \langle k - P_K x, (x - P_K x) - P_{D_i} P_{K^\perp} x\rangle + \langle d_i - P_{D_i} P_{K^\perp} x, P_{K^\perp} x - P_{D_i} P_{K^\perp} x\rangle \\ \leq \ & 0, \end{aligned}$$

as desired. Hence (ii) holds. (iii) follows from (ii) inductively whereas (iv) is obvious from (iii). $\square$

REMARK 3.3.3. It follows that range$(I - P_i)$ is a subset of $K^\perp$. Ky Fan ([**40**, Theorem 14]) showed that range$(I - P_i)$ is actually contained in the cone generated by the $a_{i,j}$; this will, however, not be needed in the sequel.

3.4. Polyhedral multifunctions.

DEFINITION 3.4.1. Suppose X, Y are Euclidean spaces and $\Omega : X \rightrightarrows Y$ is a multifunction. If the graph of Ω can be written as the union of finitely many convex polyhedra, then Ω is called a *polyhedral multifunction*.

FACT 3.4.2. (Robinson [**73**, Section 1]) The class of polyhedral multifunctions is closed under addition, scalar multiplication, composition, and inversion.

FACT 3.4.3. (Robinson [**73**, Section 1]) If A is a linear mapping between the Euclidean spaces X and Y, then A is a polyhedral multifunction. If D is a convex polyhedron, then $\partial\iota_D(\cdot) = N_D(\cdot)$ is a polyhedral multifunction.

COROLLARY 3.4.4. If $D_1, \ldots, D_N$ are convex polyhedra in the Euclidean space X, then $I - P_{D_N} \cdots P_{D_1}$ is a single-valued polyhedral multifunction.

PROOF. $y = P_{D_i} x \Leftrightarrow x \in (I + \partial\iota_{D_i})(y) \Leftrightarrow y \in (I + \partial\iota_{D_i})^{-1}(x)$. Now I and $\partial\iota_{D_i}$ are polyhedral multifunctions by Fact 3.4.3; hence so is $I - P_{D_N} \cdots P_{D_1}$ by repeated application of Fact 3.4.2 . $\square$

FACT 3.4.5. (Robinson's [**73**, Proposition 2]) If Ω is a polyhedral multifunction between the Euclidean spaces X and Y and if W is a bounded subset of range(Ω), then there exists some bounded subset V of X such that $\Omega(V) \supseteq W$.

COROLLARY 3.4.6. If $D_1, \ldots, D_N$ are convex polyhedra in the Euclidean space X, then the infimum $\inf_{x\in X} \|x - P_{D_N} \cdots P_{D_1} x\|$ is attained.

PROOF. $\Omega := I - P_{D_N} \cdots P_{D_1}$ is polyhedral and single-valued (Corollary 3.4.4). Let $\beta > \inf_{x\in X} \|\Omega x\|$ and set $W := \text{range}(\Omega) \cap \beta B_X$. By the previous fact, we get $\alpha > 0$ such that $\Omega(\alpha B_X) \supseteq W$. Thus $\inf_{x\in X} \|\Omega x\| = \inf_{x\in \alpha B_X} \|\Omega x\|$; however, this latter infimum is attained, since Ω is continuous and αB_X is compact. $\square$

REMARK 3.4.7. In particular, if each D_i is a halfspace in the Euclidean space X, then the infimum $\inf_{x\in X} \|x - P_{D_N} \cdots P_{D_1} x\|$ is attained — this also follows from De Pierro and Iusem's [**71**, Proposition 13].

Combining Proposition 3.3.2.(iv) and Corollary 3.4.6 thus yields:

COROLLARY 3.4.8. If each C_i is a convex polyhedron, then the infimum $\inf_{x\in H} \|x - Q_N x\|$ is attained.

3.5. Kadec/Klee sets. The following definition generalizes well-known notions of rotundity of unit balls in Banach spaces; see Deville et al.'s recent monograph [**33**].

DEFINITION 3.5.1. Suppose C is a closed convex nonempty subset of H. We say that C is:

(i) *Kadec/Klee*, if $\left.\begin{array}{c}(x_n) \text{ in } C,\\ x_n \rightharpoonup x \in \operatorname{bd}(C)\end{array}\right\}$ implies $x_n \to x$.

(ii) *uniformly convex*, if $\left.\begin{array}{c}(x_n),(y_n) \text{ in } C,\\ d((x_n+y_n)/2, \operatorname{bd}(C)) \to 0\end{array}\right\}$ implies $x_n - y_n \to 0$.

(iii) *locally uniformly convex*, if $\left.\begin{array}{c}(x_n) \text{ in } C,\ x \in C,\\ d((x_n+x)/2, \operatorname{bd}(C)) \to 0\end{array}\right\}$ implies $x_n \to x$.

(iv) *strictly convex*, if $\left.\begin{array}{c}x, y \in C,\\ (x+y)/2 \in \operatorname{bd}(C)\end{array}\right\}$ implies $x = y$.

PROPOSITION 3.5.2. Suppose C is a closed convex nonempty subset of H. Then:

(i) C uniformly convex $\Rightarrow$ C locally uniformly convex $\Rightarrow$ C strictly convex.
(ii) C strictly convex and compact $\Rightarrow$ C uniformly convex.
(iii) If $\operatorname{int}(C) = \emptyset$, then: C is strictly convex $\Leftrightarrow$ C is singleton.
(iv) If $\operatorname{int}(C) = \emptyset$, then: C is Kadec/Klee $\Leftrightarrow$ C is boundedly compact.
(v) C boundedly compact $\Rightarrow$ C Kadec/Klee $\Rightarrow$ $\operatorname{int}(C) \neq \emptyset$ or C boundedly compact.
(vi) C locally uniformly convex $\Rightarrow$ C Kadec/Klee.
(vii) The function

$$\delta_C(t) := \sup\{r \geq 0 : B((x+y)/2, r) \subseteq C, \text{ for all } x, y \in C \text{ with } \|x-y\| = t\}$$

is increasing. Moreover: C is uniformly convex $\Leftrightarrow$ $\delta_C(t) > 0$, for all $t > 0$.

PROOF. We prove here only (vi); the rest is straightforward. Suppose C is locally uniformly convex. In view of (iii) and (iv), we can assume that $\operatorname{int}(C) \neq \emptyset$. Pick a sequence (x_n) in C that converges weakly to some $x \in \operatorname{bd}(C)$. Then x is a support point of C and there is a support functional f with $f(x) \geq f(C)$ and $\|f\| = 1$. Let $t_n := 1/n + (f(x) - f(x_n))/2$, for all n; then $t_n \to 0^+$. Also, $f(t_n f + (x_n + x)/2) = 1/n + f(x) > f(x)$ so that $t_n f + (x_n+x)/2 \notin C$, for all n. Hence $d((x_n+x)/2, \operatorname{bd}(C)) \leq \|(x_n+x)/2 - ((x_n+x)/2 + t_n f)\| = t_n \to 0$. Therefore, since C is locally uniformly convex, $x_n \to x$ and (vi) holds. ☐

REMARKS 3.5.3.

- Let $H = \mathbb{R}^2$ and suppose $C = \{(x,y) \in H : y \geq 1/x > 0\}$. Then C is locally uniformly convex but not uniformly convex. Hence the assumption of compactness in (ii) is important.
- Note that every non-singleton boundedly compact closed convex nonempty set is Kadec/Klee; thus there are many Kadec/Klee sets which are not strictly convex.
- (Borwein and Fitzpatrick's [**11**, Example following Corollary 3.5]) Let $H = \ell_2 \times \mathbb{R}$ and define a new norm $|||\cdot|||$ on H by $|||(x,r)|||^2 := \sum_n (x_n/n)^2 + r^2 + \max\{\sum_n x_n^2, r^2\}$. Suppose $C := \{(x,r) \in H : |||(x,r)||| \leq 1\}$. Then C is strictly convex. Denote the n^{th} unit vector of ℓ_2 by e_n. Then the sequence $((e_n,1)/\sqrt{2+1/n^2})$ lies in the boundary of C and converges weakly to $(0, 1/\sqrt{2}) \in \operatorname{bd}(C)$, but not in norm. Therefore, C is an strictly convex set

with nonempty interior but C is neither Kadec/Klee nor locally uniformly convex.

- Borwein and Lewis [**12**] showed that the notions of locally uniform convexity and strict convexity coincide in Euclidean spaces.

DEFINITION 3.5.4. Suppose C is a uniformly convex subset of H. If there is some $\gamma > 0$ such that $\delta_C(t) \geq \gamma t^2$, then C is called *strongly convex.*

REMARKS 3.5.5.

- If $C = \rho B_H$, where $\rho > 0$, then $\delta_C(t) \geq t^2/(8\rho)$ (use the Parallelogram Law) and hence C is strongly convex.
- Let $H = \mathbb{R}^2$ and suppose $C = \{(x,y) \in H : x^4 + y^4 \leq 1\}$. Fix an arbitrary $0 < s < 1$ and set $\mathbf{x}_s = (s, (1-s^4)^{(1/4)})$, $\mathbf{y}_s = (s, -(1-s^4)^{(1/4)})$. Then $\mathbf{x}_s, \mathbf{y}_s \in C$, $(\mathbf{x}_s + \mathbf{y}_s)/2 = (s, 0)$, and $d((\mathbf{x}_s+\mathbf{y}_s)/2, \mathrm{bd}(C)) \leq \|(\mathbf{x}_s+\mathbf{y}_s)/2 - (1,0)\| = 1-s$. Also, $\|\mathbf{x}_s - \mathbf{y}_s\| = 2(1-s^4)^{(1/4)}$, and thus $\delta_C(2(1-s^4)^{(1/4)}) \leq 1-s$. Changing variables, we obtain $\delta_C(t) \leq 1-(1-(t/2)^4)^{(1/4)} = t^4/64 + \mathcal{O}(t^8)$. Therefore, $\delta_C(t)/t^2 \leq t^2/64 + \mathcal{O}(t^6)$ and C is not strongly convex.

3.6. Regularities.

The two set case.

DEFINITION 3.6.1. Suppose $N = 2$ and $d_1 := P_{\overline{C_2 - C_1}}(0) \in C_2 - C_1$. We say that (C_1, C_2) is:

(i) *boundedly regular*, if

$$\left.\begin{array}{c}(x_n) \text{ is a bounded sequence in } H,\\ \max\{d(x_n, C_1), d(x_n, C_2 - d_1)\} \to 0\end{array}\right\} \text{ implies } d(x_n, C_1 \cap (C_2 - d_1)) \to 0.$$

(ii) *regular*, if

$$\left.\begin{array}{c}(x_n) \text{ is a sequence in } H,\\ \max\{d(x_n, C_1), d(x_n, C_2 - d_1)\} \to 0\end{array}\right\} \text{ implies } d(x_n, C_1 \cap (C_2 - d_1)) \to 0.$$

Obviously, regularity implies bounded regularity. These conditions have turned out to be sufficient for norm convergence of orbits; see Subsection 5.5.

FACT 3.6.2. Suppose $N = 2$.

(i) ([**9**, Theorem 3.9]) If C_1 or C_2 is boundedly compact and $d_1 \in C_2 - C_1$, then (C_1, C_2) is boundedly regular.

(ii) ([**9**, Corollary 3.14]) If C_1 or C_2 is bounded and (C_1, C_2) is boundedly regular, then (C_1, C_2) is regular.

REMARK 3.6.3. The first two authors wish to point out the following lacunae in their paper [**9**]: the definition of bounded regularity [**9**, Definiton 3.6] needs in addition the assumption $P_{\overline{C_2 - C_1}}(0) \in C_2 - C_1$ and is then equivalent to the present, more handy Definition 3.6.1; similarly, [**9**, Theorems 3.7, 3.8, and Corollary 3.10] need in addition the assumption $P_{\overline{C_2 - C_1}}(0) \in C_2 - C_1$.

The consistent case.

DEFINITION 3.6.4. Suppose $\bigcap_{i=1}^N C_i \neq \emptyset$. We say that $(C_1, \ldots, C_N)$ is:

(i) *boundedly regular*, if $\max_i d(x_n, C_i) \to 0$ implies $d(x_n, \bigcap_i C_i) \to 0$, for all bounded sequences (x_n) in H.

(ii) *regular*, if $\max_i d(x_n, C_i) \to 0$ implies $d(x_n, \bigcap_i C_i) \to 0$, for all sequences (x_n) in H.

(iii) *boundedly linearly regular*, if for every bounded subset S of H, there exists $\kappa_S > 0$ such that $d(x, \bigcap_i C_i) \leq \kappa_S \max_i d(x, C_i)$, for all $x \in S$.
(iv) *linearly regular*, if there exists $\kappa > 0$ such that $d(x, \bigcap_i C_i) \leq \kappa \max_i d(x, C_i)$, for all $x \in H$.

Clearly,

$$\begin{array}{ccc} \text{linearly regular} & \Longrightarrow & \text{boundedly linearly regular} \\ \Downarrow & & \Downarrow \\ \text{regular} & \Longrightarrow & \text{boundedly regular,} \end{array}$$

and again we prefer the more intuitive definition of (bounded) regularity to the equivalent [**8**, Definition 5.1].

THEOREM 3.6.5.

(i) ([**8**, Proposition 5.4.(i)]) If some C_i is boundedly compact, then $(C_1, \ldots, C_N)$ is boundedly regular.
(ii) ([**8**, Proposition 5.4.(iii)]) If H is finite-dimensional, then $(C_1, \ldots, C_N)$ is always boundedly regular.
(iii) If every C_i, except possibly one, is Kadec/Klee, then $(C_1, \ldots, C_N)$ is boundedly regular.
(iv) ([**8**, Proposition 5.4.(ii)]) If $(C_1, \ldots, C_N)$ is boundedly regular and some C_i is bounded, then $(C_1, \ldots, C_N)$ is regular.
(v) (Gubin et al.'s [**51**, Proof of Lemma 5]) If every C_i, except possibly one, is uniformly convex, then $(C_1, \ldots, C_N)$ is regular.

PROOF. (iii): After relabeling if necessary, we assume that $C_1, \ldots, C_{N-1}$ are Kadec/Klee. Now pick a bounded sequence with $\max_{i=1}^N d(x_n, C_i) \to 0$. After passing to a subsequence if necessary, we assume that $(d(x_n, \bigcap_{i=1}^N C_i))$ converges to some limit, say L. It suffices to show that L equals 0. After passing once more to a subsequence, we also assume that (x_n) converges weakly to some x, where x lies necessarily in $\bigcap_{i=1}^N C_i$ (Proposition 2.1.2.(i)). If $x \in \mathrm{bd}(C_i)$, for some $i \in \{1, \ldots, N-1\}$, then $(P_i x_n)$ converges in norm to x (because $x_n - P_i x_n \to 0$ and C_i is Kadec/Klee) and so does (x_n) implying $L = 0$, as desired. Otherwise, $x \in C_N \cap \mathrm{int}(\bigcap_{i=1}^{N-1} C_i)$. Then, by Fact 3.6.9, $(C_1, \ldots, C_N)$ is even boundedly linearly regular. The proof is complete. ☐

REMARK 3.6.6.

- Gubin et al. ([**51**, Lemma 5]) claimed only bounded regularity in (iv); their proof, however, works for regularity.
- For an example of two intersecting sets C_1 and C_2 where (C_1, C_2) is not boundedly regular, see Bauschke and Borwein's [**9**, Example 5.5]. Hence the assumption of bounded compactness in (i) is important.

We now give some useful results on (bounded) linear regularity; see also [**8**].

FACT 3.6.7. ([**8**, Theorem 5.19]) Suppose $L_1, \ldots, L_N$ are closed subspaces of H. Then $L_1^\perp + \cdots + L_N^\perp$ is closed if and only if $(L_1, \ldots, L_N)$ is regular in any of of the four senses. Consequently, if each C_i is a closed affine subspace, where $C_i = c + L_i$ and L_i is linear, then $(C_1, \ldots, C_N)$ is regular in any of the four senses if and only if $L_1^\perp + \cdots + L_N^\perp$ is closed.

FACT 3.6.8. ([**8**, Corollary 5.26]) If each C_i is a convex polyhedron, then the N-tuple $(C_1, \ldots, C_N)$ is linearly regular.

FACT 3.6.9. ([**8**, Corollaries 5.13 and 5.14]) If $0 \in \bigcap_{i=1}^{N-1} \operatorname{icr}((\bigcap_{j=1}^{i} C_j) - C_{i+1})$, then $(C_1, \ldots, C_N)$ is boundedly linearly regular. In particular, this happens whenever $C_N \cap \operatorname{int}(\bigcap_{i=1}^{N-1} C_i) \neq \emptyset$.

REMARK 3.6.10. Since the definition of (bounded) (linear) regularity does not depend on the order of the sets, the last fact has corresponding formulations for all other orders of the sets.

The general case. In general, without understanding the geometry of the problem (see Subsection 5.1), it seems to be hard even to suggest useful notions of regularity. We feel that a positive answer to Conjecture 5.1.6 would yield the appropriate general notion of regularity.

3.7. The angle of a tuple of subspaces. In this subsection, we generalize the notion of the angle between two subspaces to finitely many subspaces. The results are also interesting from an operator-theoretic point of view and will lead to some new linear convergence results in Subsection 5.7.

Throughout this subsection, we assume that

$$\boxed{L_1, L_2, \ldots, L_N \text{ are closed subspaces and } L := \bigcap_{i=1}^{N} L_i.}$$

FACT 3.7.1. (Kayalar and Weinert's [**58**, Theorem 1])

$$P_{L_N} \cdots P_{L_1} P_{L^\perp} = P_{L_N \cap L^\perp} \cdots P_{L_1 \cap L^\perp}.$$

The next proposition is easily verified.

PROPOSITION 3.7.2. Suppose B is a closed subspace and A is a subset of $B^\perp$. Then A is closed if and only if $A \oplus B$ is.

PROPOSITION 3.7.3. The following are equivalent:

(i) $(L_1 \cap L^\perp, \ldots, L_N \cap L^\perp)$ is regular in any of the four senses.
(ii) $L_1^\perp + \cdots + L_N^\perp + L$ is closed.
(iii) $(L_1, \ldots, L_N)$ is regular in any of the four senses.
(iv) $L_1^\perp + \cdots + L_N^\perp$ is closed.

PROOF. In view of Fact 3.6.7, it suffices to establish the equivalence of (ii) and (iv). By Proposition 3.7.2, $(L_i \cap L^\perp)^\perp = \overline{L_i^\perp + L} = L_i^\perp + L$, for each i. Using Proposition 3.7.2 once more, we get: $L_1^\perp + \cdots + L_N^\perp$ is closed $\Leftrightarrow (L_1^\perp + \cdots + L_N^\perp) + L$ is closed $\Leftrightarrow (L_1^\perp + L) + \cdots + (L_N^\perp + L)$ is closed $\Leftrightarrow (L_1 \cap L^\perp)^\perp + \cdots + (L_N \cap L^\perp)^\perp$ is closed. $\square$

THEOREM 3.7.4. The Hilbert space operator norm $\|P_{L_N} \cdots P_{L_1} P_{L^\perp}\|$ is strictly less than 1 if and only if the sum $L_1^\perp + \cdots + L_N^\perp$ is closed.

PROOF. "$\Longrightarrow$": We prove the contrapositive and assume that $L_1^\perp + \cdots + L_N^\perp$ is not closed. By Proposition 3.7.3, $(L_1 \cap L^\perp, \ldots, L_N \cap L^\perp)$ is not boundedly regular. Hence there is some bounded sequence (x_n) with $\max_i d(x_n, L_i \cap L^\perp) \to 0$, but $d(x_n, \bigcap_i L_i \cap L^\perp) = \|x_n\| \not\to 0$. After passing to a subsequence and normalizing if necessary, we may assume $\|x_n\| = 1$, for all n. Now $x_n - P_{L_1 \cap L^\perp} x_n \to 0$, which implies, by nonexpansivity of $P_{L_2 \cap L^\perp}$, that $P_{L_2 \cap L^\perp} x_n - P_{L_2 \cap L^\perp} P_{L_1 \cap L^\perp} x_n \to 0$.

Since $x_n - P_{L_2 \cap L^\perp} x_n \to 0$, we obtain $x_n - P_{L_2 \cap L^\perp} P_{L_1 \cap L^\perp} x_n \to 0$. Repeating this line of thought yields eventually

$$x_n - P_{L_N \cap L^\perp} \cdots P_{L_1 \cap L^\perp} x_n \to 0.$$

The triangle inequality implies $\|x_n\| - \|P_{L_N \cap L^\perp} \cdots P_{L_1 \cap L^\perp} x_n\| \to 0$; thus $\|P_{L_N \cap L^\perp} \cdots P_{L_1 \cap L^\perp} x_n\| \to 1$. Therefore, by Fact 3.7.1, $\|P_{L_N \cap L^\perp} \cdots P_{L_1 \cap L^\perp}\| = \|P_{L_N} \cdots P_{L_1} P_{L^\perp}\| = 1$.

"$\Longleftarrow$": By Proposition 3.7.3, $(L_1 \cap L^\perp, \ldots, L_N \cap L^\perp)$ is linearly regular. Hence there is some $\kappa > 0$ such that

$$\|x\| = d(x, \textstyle\bigcap_i L_i \cap L^\perp) \leq \kappa \max_i d(x, L_i \cap L^\perp), \text{ for all } x \in H.$$

On the other hand, by Fact 4.5.2 below,

$$\begin{aligned} d^2(x, L_i \cap L^\perp) &\leq \|x - P_{L_i \cap L^\perp} \cdots P_{L_1 \cap L^\perp} x\|^2 \\ &\leq 2^{i-1}(\|x\|^2 - \|P_{L_i \cap L^\perp} \cdots P_{L_1 \cap L^\perp} x\|^2) \\ &\leq 2^{N-1}(\|x\|^2 - \|P_{L_N \cap L^\perp} \cdots P_{L_1 \cap L^\perp} x\|^2), \end{aligned}$$

for all $x \in H$ and each i. Altogether,

$$\|x\|^2 \leq \kappa^2 2^{N-1}(\|x\|^2 - \|P_{L_N \cap L^\perp} \cdots P_{L_1 \cap L^\perp} x\|^2), \text{ for all } x \in H,$$

which implies $\|P_{L_N \cap L^\perp} \cdots P_{L_1 \cap L^\perp}\| \leq (1 - \kappa^{-2} 2^{-(N-1)})^{1/2} < 1$. The theorem now follows from Fact 3.7.1. ☐

DEFINITION 3.7.5. We define the *angle* $\gamma := \gamma(L_1, \ldots, L_N) \in [0, \pi/2]$ *of the N-tuple* $(L_1, \ldots, L_N)$ by

$$\cos \gamma = \|P_{L_N} \cdots P_{L_1} P_{L^\perp}\|.$$

REMARKS 3.7.6.

(i) The classical definition of the angle $\bar{\gamma} := \bar{\gamma}(L_1, L_2)$ between *two* subspaces was given by Friedrichs [**43**] in 1937:

$$\cos \bar{\gamma} = \sup\{\langle l_1, l_2 \rangle : l_1 \in L_1 \cap L^\perp;\ l_2 \in L_2 \cap L^\perp;\ \|l_1\|, \|l_2\| \leq 1\}.$$

Deutsch ([**27**, Proof of Lemma 2.5.(iii)]) showed that $\|P_{L_2} P_{L_1} P_{(L_2 \cap L_1)^\perp}\|$ is equal to $\cos \bar{\gamma}$; therefore, $\gamma = \bar{\gamma}$ and our Definition 3.7.5 of the angle is consistent with Friedrichs' original one.

(ii) The importance of the angle between two subspaces stems from the fact that a positive angle guarantees linear convergence of the sequence of alternating projections; see, for instance, [**3**, pages 378ff.], [**9**, Theorem 4.11], [**27**, Theorem 2.3], and [**58**, Theorem 2]. In Subsection 5.7, we will show among other things that a positive angle of an N-tuple of subspaces yields linear convergence of the sequence of cyclic projections as well.
Very recently, Deutsch and Hundal [**29, 31**] studied the angle (in the sense of Friedrichs) and another notion of an angle (due to Dixmier) in great detail; they explored the relationships between these angles and error bounds for the method of cyclic projections for closed subspaces. Deutsch and Hundal [**32**] also carefully analyzed Dykstra's algorithm (see subsection 5.8) for intersecting halfspaces and obtained finite and linear convergence results. It is interesting to note that they provide an upper bound for the rate of convergence which depends on the angle of tuples of subspaces as in Definition 3.7.5.

(iii) By (i), the angle between two subspaces is independent of the order, that is $\gamma(L_1, L_2) = \gamma(L_2, L_1)$. For $N \geq 3$, the angle of the N-tuple $(L_1, \ldots, L_N)$ may well depend on the order of the subspaces; to see this, we pick three unit vectors a, b, c with $\langle b, c\rangle \neq 0$ and $|\langle a, b\rangle| \neq |\langle a, c\rangle|$. (This requires H to be of dimension at least 2.) Let $L_1 := L_2 := \cdots := L_{N-2} := \operatorname{span}(a)$, $L_{N-1} := \operatorname{span}(b)$, $L_N := \operatorname{span}(c)$; then $L_{N-1} \cap L_N = \{0\}$. Hence $L = \{0\}$ and one easily calculates:
$\|P_{L_N} P_{L_{N-1}} P_{L_{N-2}} \cdots P_{L_1} P_{L^\perp}\| = \|P_{L_N} P_{L_{N-1}} P_{L_{N-2}} \cdots P_{L_1}\| = |\langle a, b\rangle\langle b, c\rangle|$
$\|P_{L_{N-1}} P_{L_N} P_{L_{N-2}} \cdots P_{L_1} P_{L^\perp}\| = \|P_{L_{N-1}} P_{L_N} P_{L_{N-2}} \cdots P_{L_1}\| = |\langle a, c\rangle\langle c, b\rangle|$.
Thus $\gamma(L_1, \ldots, L_{N-2}, L_{N-1}, L_N) \neq \gamma(L_1, \ldots, L_{N-2}, L_N, L_{N-1})$.
In general, it follows from Theorem 3.7.4 that $\gamma(L_1, \ldots, L_N)$ is positive if and only if $\gamma(L_{\tau(1)}, \ldots, L_{\tau(N)})$ is, for every permutation τ of $\{1, \ldots, N\}$.

(iv) Although the angle of the N-tuple $(L_1, \ldots, L_N)$ depends on the order, it is true that $\gamma(L_1, \ldots, L_N) = \gamma(L_N, \ldots, L_1)$, because (see Fact 3.7.1) $\|P_{L_N} \cdots P_{L_1} P_{L^\perp}\| = \|P_{L_N \cap L^\perp} \cdots P_{L_1 \cap L^\perp}\| = \|(P_{L_N \cap L^\perp} \cdots P_{L_1 \cap L^\perp})^*\| = \|P_{L_1 \cap L^\perp} \cdots P_{L_N \cap L^\perp}\| = \|P_{L_1} \cdots P_{L_N} P_{L^\perp}\|$. This was observed by Kayalar and Weinert ([**58**, Corollary 1]) and it explains once more why $\gamma(L_1, L_2)$ equals $\gamma(L_2, L_1)$.

Proposition 3.7.7. The angle of the N-tuple $(L_1, \ldots, L_N)$ is positive if and only if the sum $L_1^\perp + \cdots + L_N^\perp$ is closed. In particular, this holds whenever one of the following conditions is satisfied:

(i) Some $L_i \cap L^\perp$ is finite-dimensional.
(ii) Some L_i is finite-dimensional.
(iii) H is finite-dimensional.
(iv) All L_i, except possibly one, are finite-codimensional.
(v) Each L_i is a hyperplane.
(vi) Each angle $\gamma(L_i, L_i \cap \cdots \cap L_N)$ is positive.

Proof. The first statement is just a reformulation of Theorem 3.7.4. (i): By assumption, $L_i \cap L^\perp$ is boundedly compact. Hence Theorem 3.6.5.(i) implies bounded regularity of the N-tuple $(L_1 \cap L^\perp, \ldots, L_N \cap L^\perp)$ which is equivalent to the closedness of $L_1^\perp + \cdots + L_N^\perp$ (Proposition 3.7.3). (ii) and (iii) clearly imply (i). Condition (iv) yields linear regularity of $(L_1, \ldots, L_N)$ and hence closedness of $L_1^\perp + \cdots + L_N^\perp$ by [**8**, Corollary 5.21.(ii)] and Proposition 3.7.3. Obviously, (v) implies (iv). Finally, (vi) follows from [**8**, Example 6.18] . □

Remarks 3.7.8.

- Gaffke and Mathar [**44**] suggested condition (i) in their study of Dykstra's algorithm. Condition (vi) is popular in computed tomography; see Smith et al.'s [**76**, Theorem 2.2].
- For $N = 2$, the angle between L_1 and L_2 is positive if and only if $L_1^\perp + L_2^\perp$ is closed which in turn is equivalent to the closedness of $L_1 + L_2$ (see, for instance, [**56**, Corollary 35.6]). For $N \geq 3$, however, the closedness of the sum $L_1 + \cdots + L_N$ is *independent* of the closedness of the sum $L_1^\perp + \cdots + L_N^\perp$ ([**8**, Remarks 5.20]).

3.8. Fejér monotone sequences.

DEFINITION 3.8.1. Suppose C is a closed convex nonempty subset of H. A sequence $(x_n)_{n\geq 0}$ is called *Fejér monotone w.r.t.* C, if

$$\|x_{n+1} - c\| \leq \|x_n - c\|, \text{ for every } c \in C \text{ and all } n \geq 0.$$

Fejér monotone sequences are a useful concept when dealing with projection algorithms or sequences of iterates of nonexpansive mappings. In our context, the most important property of Fejér monotone sequences is the following (others can be found in [**8**]):

PROPOSITION 3.8.2. Suppose $(x_n)_{n\geq 0}$ is Fejér monotone w.r.t. C. Then:

(i) (x_n) converges weakly to some point in C if and only if all weak cluster points of (x_n) lie in C. In this case, the weak limit of (x_n) belongs to $\mathcal{Q}_C x_0$.
(ii) (x_n) converges in norm to some point in C if and only if (x_n) has at least one norm cluster point in C.

PROOF. (i): Combine [**8**, Theorem 2.16.(ii)] and [**7**, Proposition 2.6.(vi)]. (ii) is [**8**, Theorem 2.16.(v)]. □

4. Nonexpansive mappings

4.1. Some bits of general theory.

DEFINITION 4.1.1. A selfmapping T of H is called *nonexpansive*, if

$$\|Tx - Ty\| \leq \|x - y\|, \text{ for all } x, y \in H.$$

A point $x \in H$ is a *fixed point of* T, if $x = Tx$; the set of all fixed points of T is denoted Fix T.

Obviously, the class of nonexpansive mappings is closed under composition. In view of Proposition 2.1.2, we thus have:

COROLLARY 4.1.2. Each P_i and each Q_i is nonexpansive.

PROPOSITION 4.1.3. (minimal displacement vector) Suppose T is a nonexpansive selfmapping of H. Then:

(i) (Pazy's [**68**, Lemma 4]) The set $\overline{\text{range}}(T - I)$ is convex and consequently has a unique element of minimal norm (namely the projection of 0 onto $\overline{\text{range}}(T-I)$) which we call the *minimal displacement vector of* T and denote $v(T)$.
(ii) The set $\{x \in H : Tx - x = v(T)\}$ is closed and convex. In particular, Fix T is a closed convex (possibly empty) subset of H.
(iii) (Pazy's [**68**, Corollary 2]) For every $x \in H$, the sequence $((1/n)T^n x)$ converges in norm to $v(T)$.
(iv) The minimal displacement vector $v(T)$ is contained in $\text{rec}(\overline{\text{conv}}(\text{range}(T)))$.

PROOF. (ii): Let $F := (I + T)/2$. Then F is *firmly nonexpansive* (see the following subsection). Since $\|Fx - x\| = \|Tx - x\|/2$ for all x, we have $\{x \in H : Tx - x = v(T)\} = \{x \in H : Fx - x = v(F)\}$. Now the latter set is convex by Corollary 4.2.4.
(iv): $T^n x \in \overline{\text{conv}}(\text{range}(T))$ for all n, thus $v(T) \in \text{rec}(\overline{\text{conv}}(\text{range}(T)))$ by (iii) and Proposition 3.2.2.(ii). □

PROPOSITION 4.1.4. Suppose T is a nonexpansive selfmapping of H with Fix $T \neq \emptyset$. Then the sequence of iterates $(T^n x)_{n\geq 0}$ is Fejér monotone w.r.t. Fix T.

PROOF. By Proposition 4.1.3.(ii), Fix T is closed and convex. If $f \in$ Fix T and $n \geq 0$, then $\|T^{n+1}x - f\| = \|T^{n+1}x - Tf\| \leq \|T^n x - f\|$; the result follows. $\square$

COROLLARY 4.1.5. If Fix Q_i is nonempty, then the sequence (x_i^n) is Fejér monotone w.r.t. Fix Q_i.

FACT 4.1.6. (Browder [**17**]/Göhde [**47**]/Kirk [**59**]) Every nonexpansive selfmapping of a bounded closed convex nonempty subset of the Hilbert space H has at least one fixed point.

REMARK 4.1.7. Numerous generalizations of this result have been established (see, for instance, Goebel and Kirk's [**45**] or Goebel and Reich's [**46**]); however, here we need only the Hilbert space case for which Pazy [**68**, Corollary 5] gave a very instructive proof.

4.2. Firmly nonexpansive mappings.

DEFINITION 4.2.1. A nonexpansive selfmapping T of H is called *firmly nonexpansive*, if

$$\langle Tx - Ty, x - y\rangle \geq \|Tx - Ty\|^2, \text{ for all } x, y \in H.$$

FACT 4.2.2. (Goebel and Kirk's [**45**, Theorem 12.1], Goebel and Reich's [**46**, Chapter 1, Section 11], Rockafellar's [**74**, Proposition 1], Zarantonello's [**84**, Lemma 1.3]) Suppose T is a selfmapping of H. Then the following are equivalent:

(i) T is firmly nonexpansive.
(ii) $I - T$ is firmly nonexpansive.
(iii) $\langle (I-T)x - (I-T)y, Tx - Ty\rangle \geq 0$, for all $x, y \in H$.
(iv) $2T - I$ is nonexpansive.

COROLLARY 4.2.3. Projections are strongly nonexpansive. Consequently, each P_i is firmly nonexpansive.

PROOF. Immediate from Proposition 2.1.2.(iii) and Fact 4.2.2. $\square$

COROLLARY 4.2.4. Suppose T is a firmly nonexpansive selfmapping of H and x is a point in H. Then $Tx - x = v(T)$ if and only if $\langle y - Ty + v(T), Ty - v(T) - x\rangle \geq 0$, for all $y \in H$. Consequently, the set $\{x \in H : Tx - x = v(T)\}$ is closed and convex.

PROOF. Follows easily from Fact 4.2.2.(iii). $\square$

Among all nonexpansive mappings, the class of firmly nonexpansive mappings is probably the closest to the class of projections. It possesses, however, a serious drawback:

EXAMPLE 4.2.5. The class of firmly nonexpansive mapping is not closed under composition: consider in $H := \mathbb{R}^2$ the mappings $T_1(x,y) := P_{\text{span}\{(1,0)\}}(x,y) = (x,0)$ and $T_2(x,y) := P_{\text{span}\{(1,1)\}}(x,y) = (x+y, x+y)/2$. Then $T_2T_1(x,y) = (x,x)/2$ and $\langle T_2T_1(1,-2) - T_2T_1(0,0), (1,-2) - (0,0)\rangle = -1/2 \not\geq 0$; thus T_2T_1 is not firmly nonexpansive.
Conclusion: The composition of two projections can fail even to be firmly nonexpansive.

Useful substitutes to "repair this defect" were provided by Bruck and Reich, by De Pierro and Iusem, and by Bauschke and Borwein: they introduced classes of nonexpansive mappings which are closed under composition and preserve characteristic features of projections.

4.3. Bruck and Reich's strongly nonexpansive mappings.

DEFINITION 4.3.1. (Bruck and Reich [**19**]) A nonexpansive selfmapping T of H is called *strongly nonexpansive*, if for every pair of sequences $(x_n), (y_n)$,

$$\left.\begin{array}{c}(x_n - y_n) \text{ is bounded,} \\ \\ \|x_n - y_n\| - \|Tx_n - Ty_n\| \to 0\end{array}\right\} \text{ implies } (x_n - y_n) - (Tx_n - Ty_n) \to 0.$$

PROPOSITION 4.3.2. Projections are strongly nonexpansive. Consequently, each P_i is strongly nonexpansive.

PROOF. Let P be a projection. If $(x_n - y_n)$ is bounded, then so is $(\|Px_n - Py_n\|)$ and hence $(\|x_n - y_n\| + \|Px_n - Py_n\|)$. Thus if $\|x_n - y_n\| - \|Px_n - Py_n\| \to 0$, then $\|x_n - y_n\|^2 - \|Px_n - Py_n\|^2 \to 0$ and Proposition 2.1.2 yields $(x_n - y_n) - (Px_n - Py_n) \to 0$. $\square$

REMARK 4.3.3. In fact, Bruck and Reich [**19**, Proposition 2.1] proved that every firmly nonexpansive mapping is strongly nonexpansive whenever the underlying space is *uniformly convex* (which is the case in our Hilbert space setting).

FACT 4.3.4. (Bruck and Reich's [**19**, Lemma 2.1]) Suppose $T_1, \dots, T_N$ are strongly nonexpansive selfmappings of H. Then $T_N T_{N-1} \cdots T_1$ is strongly nonexpansive. Moreover, $\operatorname{Fix} T_N T_{N-1} \cdots T_1 = \bigcap_{j=1}^N \operatorname{Fix} T_j$ whenever the latter intersection is nonempty.

COROLLARY 4.3.5. Each Q_i is strongly nonexpansive. Also, $\operatorname{Fix} Q_i = \bigcap_{j=1}^N C_j$ provided the latter intersection is nonempty.

REMARK 4.3.6. In view of the last remark and Example 4.2.5, we see that the class of firmly nonexpansive mappings is a proper subclass of the class of strongly nonexpansive mappings.

FACT 4.3.7. (Bruck and Reich's [**19**, Corollaries 1.3, 1.4, and 1.5]) Suppose T is a strongly nonexpansive selfmapping of H. Then:

(i) $\operatorname{Fix} T \neq \emptyset$ if and only if $(T^n x)$ converges weakly to some fixed point of T, for all $x \in H$.
(ii) $\operatorname{Fix} T = \emptyset$ if and only if $\lim_n \|T^n x\| = +\infty$, for all $x \in H$.
(iii) $\lim_n T^{n+1}x - T^n x = v(T)$.

THEOREM 4.3.8. The minimal displacement vectors $v(Q_1) = \lim_n x_1^{n+1} - x_1^n, \dots, v(Q_N) = \lim_n x_N^{n+1} - x_N^n$ all coincide and are contained in

$$\bigcap_{i=1}^N \operatorname{rec}(C_i) \cap \left(\bigcap_{i=1}^N \operatorname{rec}(C_i)\right)^{\oplus}.$$

PROOF. By Fact 4.3.7.(iii), $x_i^{n+1} - x_i^n = Q_i^n x_i^1 - Q_i^{n-1} x_i^1 \longrightarrow v_i$, for each i. Corollary 2.1.3 implies (after choosing $y^0 := x_N^1$)

$$0 \longleftarrow (x_i^{n+1} - x_i^n) - (x_{i+1}^{n+1} - x_{i+1}^n), \text{ for each } i;$$

consequently, $v(Q_1) = \cdots = v(Q_N)$. On the one hand, by Proposition 4.1.3.(iv) and Proposition 3.2.2.(iii), $v(Q_i) \in \operatorname{rec}(\overline{\operatorname{conv}}(\operatorname{range}(Q_i))) \subseteq \operatorname{rec}(\overline{\operatorname{conv}}(\operatorname{range}(P_i))) = \operatorname{rec}(C_i)$, for each i. On the other hand,

$$\begin{array}{rcl} x_N^{n+1} - x_N^n & = & (x_N^{n+1} - x_{N-1}^{n+1}) + (x_{N-1}^{n+1} - x_{N-2}^{n+1}) + \cdots + (x_1^{n+1} - x_N^n) \\ & \in & \operatorname{range}(P_N - I) + \operatorname{range}(P_{N-1} - I) + \cdots + \operatorname{range}(P_1 - I), \end{array}$$

hence $v(Q_N) \in \overline{\mathrm{range}(P_N - I) + \mathrm{range}(P_{N-1} - I) + \cdots + \mathrm{range}(P_1 - I)}$. Therefore, by Fact 3.2.3.(ii) and Proposition 3.1.2.(ii),

$$v(Q_N) \in \overline{(\mathrm{rec}(C_N))^{\oplus} + \cdots + (\mathrm{rec}(C_1))^{\oplus}} \subseteq \left(\bigcap_{i=1}^{N} \mathrm{rec}(C_i)\right)^{\oplus}.$$

The proof is complete. □

Definition 4.3.9. We abbreviate $v(Q_1)$ $(= v(Q_2) = \cdots = v(Q_N))$ by v and refer to v as the *minimal displacement vector*. The corresponding nonnegative real number $\mu := \|v\| = \inf_x \|x - Q_1 x\|$ $(= \inf_x \|x - Q_2 x\| = \cdots = \inf_x \|x - Q_N x\|)$ is called the *minimal displacement*. We say that *the minimal displacement* μ *is attained* (resp. *unattained*), if the infimum $\inf_x \|x - Q_1 x\|$ is (resp. is not) attained. Finally, we write shorter F_i instead of Fix Q_i, for each i.

Proposition 4.3.10.

(i) Each F_i is a closed convex subset of C_i. If $\bigcap_{j=1}^{N} C_j \neq \emptyset$, then each F_i is equal to this intersection.
(ii) μ is attained if and only if each infimum $\inf_x \|x - Q_i x\|$ is.
(iii) Either each F_i is nonempty or each F_i is empty.

Proof. (i) follows from Proposition 4.1.3.(ii) and Corollary 4.3.5.
(ii), (iii): Suppose μ is attained; that is $\mu = v(Q_1) = \|\bar{x} - Q_1\bar{x}\|$, for some $\bar{x} \in H$. Then $\mu = v(Q_2) \leq \|Q_2 P_2 \bar{x} - P_2\bar{x}\| = \|P_2 Q_1 \bar{x} - P_2 \bar{x}\| \leq \|Q_1\bar{x} - \bar{x}\| = \mu$; hence $\inf_x \|x - Q_2 x\|$ is attained (at $P_2\bar{x}$) and so are the remaining infima. □

Remark 4.3.11. As we will see in the next subsection, the following is actually true: if μ is attained, then $\mu = 0$.

4.4. Nonexpansive mappings à la De Pierro and Iusem.

Definition 4.4.1. (De Pierro and Iusem [**71**]) We say a nonexpansive selfmapping T of H is *DPI-nonexpansive*, if for all $x, y \in H$:

$$\|Tx - Ty\| = \|x - y\| \text{ implies } \begin{cases} Tx - Ty = x - y, \\ \\ \langle x - y, Ty - y\rangle = 0. \end{cases}$$

Fact 4.4.2. (De Pierro and Iusem's [**71**, Proposition 1]) The class mappings that are DPI-nonexpansive is closed under composition.

Corollary 4.4.3. (De Pierro and Iusem's [**71**, Propositions 8 and 9]) Projections are DPI-nonexpansive. Consequently, each P_i and each Q_i is DPI-nonexpansive.

Proof. Combine Proposition 2.1.2.(iii) and the preceding fact. □

Fact 4.4.4. (De Pierro and Iusem's [**71**, Proof of Lemma 1]) Suppose T is a DPI-nonexpansive selfmapping of H. If $\inf_x \|x - Tx\|$ is attained, then it is 0.

Corollary 4.4.5. If the minimal displacement μ is positive, it is unattained.

Remark 4.4.6. De Pierro and Iusem formulated their results in Euclidean space — their proofs, nevertheless, work in general Hilbert space.

4.5. Strongly attracting mappings. Strongly attracting mappings are a useful tool for studying projection algorithms; see [**8**, Section 2].

DEFINITION 4.5.1. ([**8**]) A nonexpansive selfmapping T of H with $\operatorname{Fix} T \neq \emptyset$ is called *strongly attracting* or κ-attracting, if there exists $\kappa > 0$ such that

$$\kappa\|x - Tx\|^2 \leq \|x-y\|^2 - \|Tx - y\|^2, \text{ for all } x \in H,\ y \in \operatorname{Fix} T.$$

FACT 4.5.2.

(i) ([**8**, Corollary 2.5]) Projections are 1-attracting. In particular, each P_i is 1-attracting.

(ii) ([**8**, Proposition 2.10.(ii)]) If $T_1, \dots, T_N$ are strongly attracting with constants $\kappa_1, \dots, \kappa_N$ and if $\bigcap_{i=1}^N \operatorname{Fix} T_i \neq \emptyset$, then the composition $T_N \cdots T_1$ is $\min\{\kappa_1, \dots, \kappa_N\}/2^{N-1}$-attracting. Consequently, if $\bigcap_{i=1}^N C_i \neq \emptyset$, then each Q_i is $2^{-(N-1)}$-attracting.

5. Main results

5.1. Geometry of the problem.

DEFINITION 5.1.1. Suppose each F_i is nonempty. Fix any $f_1 \in F_1$ and recursively define $f_{i+1} := P_{i+1}f_i$ so that $f_i \in F_i$, for each i. We refer to the N-tuple $(f_1, \dots, f_N)$ as a *cycle*. The i^{th} *difference vector* d_i is defined by $d_i := f_{i+1} - f_i$, for each i.

THEOREM 5.1.2. Suppose each F_i is nonempty. Then the difference vectors are well-defined, i.e. they do not depend on the choice of f_1. They satisfy $\langle C_{i+1} - F_{i+1}, -d_i\rangle \leq 0$, $\langle F_{i+1} - F_{i+1}, d_i\rangle = 0$, for each i and $d_1 + d_2 + \cdots + d_N = 0$. The restriction $P_{i+1}|_{F_i}$ is a bijection between F_i and F_{i+1}; it is given by $d_i + I$ and the sets F_i are all translates of each other. Finally, $P_{F_{i+1}} = d_i + P_{F_i}$.

PROOF. Suppose the difference vectors d_i arise from a cycle (f_i). Let (g_i) be another cycle. Then Corollary 2.1.3 implies $d_i = g_{i+1} - g_i$ (the difference vectors are thus well-defined). Proposition 2.1.2.(ii) yields $\langle C_{i+1} - F_{i+1}, -d_i\rangle \leq 0$, hence $\langle F_{i+1} - F_{i+1}, d_i\rangle = 0$. The rest follows readily. ☐

REMARK 5.1.3. Youla and Velasco already showed that the difference vectors are well-defined; see [**83**, Theorem 4].

The two set case. In this situation, the geometry is well-understood:

FACT 5.1.4. Suppose $N = 2$. Then

(i) (Cheney and Goldstein's [**23**, Theorem 2]) $F_1 = \{c_1 \in C_1 : d(c_1, C_2) = d(C_1, C_2)\}$, $F_2 = \{c_2 \in C_2 : d(c_2, C_1) = d(C_2, C_1)\}$.

(ii) ([**10**, Lemmata 2.1 and 2.3]) $d_1 = P_{\overline{C_2 - C_1}}(0)$, $d_2 = P_{\overline{C_1 - C_2}}(0) = -d_1$. $F_1 = C_1 \cap (C_2 + d_2)$, $F_2 = C_2 \cap (C_1 + d_1)$.

REMARKS 5.1.5.

- Goldburg and Marks rediscovered (i); see [**48**, Theorem 3]. Gubin et al. stated (i) without proof; see [**51**, Note 1 to Theorem 2].
- It is important to note that (ii) gives a description of the difference vectors *which is independent of* F_1 *and* F_2. This alternative representation is the key to the understanding of the cyclic projection algorithm (and also of Dykstra's algorithm) for two sets.

The consistent case. Here Theorem 5.1.2 reduces to Proposition 4.3.10: if $\bigcap_{i=1}^N C_i \neq \emptyset$, then $F_1 = \cdots = F_N = \bigcap_{i=1}^N C_i$ and $d_1 = \cdots = d_N = 0$.

The general case again. The general case is likely to be much more complicated: consider, for instance, in $\mathbb{R}^2$ the case when $N = 3$ and the sets C_i are the sides of an equilateral triangle. Then each F_i is a singleton. If (f_1, f_2, f_3) denotes the unique cycle, then $\|f_2 - f_1\| + \|f_3 - f_2\| + \|f_1 - f_3\|$ is *not* the infimal length of a closed polygonal path with a single vertex in each set (and other similar measures) — in stark distinction from the two set case (Fact 5.1.4.(i))! A similar example was given by Kosmol [**60**].

We conclude this subsection with a conjecture which we saw holds true for the two set and for the consistent case:

Conjecture 5.1.6. (geometry conjecture) The difference vectors can be described without referring to the fixed point sets F_i and thus they exist even when each F_i is empty. Moreover: $d_1 + \cdots + d_N = 0$ and

$$F_N = C_N \cap (C_{N-1} + d_{N-1}) \cap \cdots \cap (C_1 + d_1 + \cdots d_{N-1});$$

corresponding formulae hold for $F_1, \ldots, F_{N-1}$.

Remark 5.1.7. By Theorem 5.1.2, we always have $F_N \subseteq C_N \cap (C_{N-1} + d_{N-1}) \cap \cdots \cap (C_1 + d_1 + \cdots d_{N-1})$.

5.2. Dichotomy.

Theorem 5.2.1. Exactly one of the following alternatives holds:

(i) Each F_i is empty. Then $\lim_n \|x_i^n\| = +\infty$, for all i.
(ii) Each F_i is nonempty. Then the N-tuple of sequence $(x_1^n, x_2^n, \ldots, x_N^n)$ converges weakly to some cycle $(f_1, f_2, \ldots, f_N)$ and each f_i is an element of $\mathcal{Q}_{F_i}(x_i^m)$, for all m. The sequence of differences $(x_{i+1}^n - x_i^n)$ converges in norm to the i^{th} difference vector d_i, for each i.

Proof. (i): Use Corollary 4.3.5, Fact 4.3.7.(ii), and Proposition 4.3.10.(iii). (ii): If (i) doesn't hold, then each F_i is nonempty, by Proposition 4.3.10.(iii). Comparison of the orbit (x_i^n) with the orbit of any cycle yields $x_{i+1}^n - x_i^n - d_i \to 0$, for all i (Corollary 2.1.3). Also, by Fact 4.3.7.(i), $x_i^n \rightharpoonup f_i$ for some $f_i \in F_i$, for all i. Hence $f_{i+1} = f_i + d_i$, for each i and thus $(f_i)_{i=1}^N$ is identified as a cycle. Finally, $(x_i^n)_{n \geq m}$ is Fejér monotone w.r.t. F_i (Corollary 4.1.5) and converges weakly to $f_i \in F_i$; thus $f_i \in \mathcal{Q}_{F_i}(x_i^m)$ by Proposition 3.8.2.(i). $\square$

The two set case.

Fact 5.2.2. ([**10**, Theorem 4.8]) Suppose $N = 2$. Then

$$x_2^n - x_1^n \to d_1 = P_{\overline{C_2 - C_1}}(0), \; x_1^{n+1} - x_2^n \to d_2 = P_{\overline{C_1 - C_2}}(0) = -d_1.$$

Moreover, exactly one of the following alternatives holds:

(i) F_1 and F_2 are empty. Then $\|x_1^n\|, \|x_2^n\| \to +\infty$.
(ii) F_1 and F_2 are nonempty. Then $x_1^n \rightharpoonup f_1$ and $x_2^n \rightharpoonup f_2$, where $f_1 \in F_1$ and $f_2 = f_1 + d_1 \in F_2$.

The consistent case. Here Theorem 5.2.1.(ii) results in the following:

THEOREM 5.2.3. Suppose $\bigcap_{i=1}^{N} C_i \neq \emptyset$. Then the entire orbit (x_i^n) converges weakly to some point in $\bigcap_{i=1}^{N} C_i$. Also, $x_{i+1}^n - x_i^n \to 0$ which implies $d(x_i^n, C_j) \to 0$, for all i, j.

REMARK 5.2.4. The weak convergence of (x_i^n) was already known to Bregman ([**16**, Theorem 1]).

The general case again. Nurtured by the results for the two set case and the consistent case, we formulate two conjectures:

CONJECTURE 5.2.5. (convergent differences conjecture) The sequence of differences $(x_{i+1}^n - x_i^n)$ is norm convergent for all starting points and every i.

REMARK 5.2.6. If the convergent differences conjecture is true, then it follows that $\lim_n x_{i+1}^n - x_i^n$ is independent of the starting point (Corollary 2.1.3).

The second conjecture is formally weaker that the first one:

CONJECTURE 5.2.7. (bounded differences conjecture) The sequences of differences $(x_{i+1}^n - x_i^n)$ are bounded for all starting points and every i.

We would like to mention the following long-standing open problem:

OPEN PROBLEM 5.2.8. Even when $N = 2$ and $C_1 \cap C_2 \neq \emptyset$, it is still not known whether the convergence of the entire orbit (x_i^n) to some point in $C_1 \cap C_2$ can actually be only weak.

5.3. Minimal displacement.

THEOREM 5.3.1.

(i) $\mu = 0$ whenever $\bigcap_{i=1}^{N} \operatorname{rec}(C_i)$ is a subspace.

(ii) $\mu = 0$ and is attained if and only if there exists a bounded closed convex nonempty subset of H that is left invariant by some Q_j. In particular, this holds whenever some C_j is bounded.

(iii) $\mu = 0$ and is attained whenever $\bigcap_{i=1}^{N} \operatorname{rec}(C_i)$ is a subspace and there exists some j such that the "normalization" $\hat{C}_j := \{c_j/\|c_j\| : c_j \in C_j \setminus \{0\}\}$ of C_j is relatively compact. The latter "normalization" condition holds whenever H is finite-dimensional.

PROOF. (i): By Proposition 3.1.3, $\bigcap_{i=1}^{N} \operatorname{rec}(C_i)$ is a subspace if and only if $\bigcap_{i=1}^{N} \operatorname{rec}(C_i) \cap (\bigcap_{i=1}^{N} \operatorname{rec}(C_i))^{\oplus} = \{0\}$. (i) is now obvious from Theorem 4.3.8.
(ii): For the "only if" part, simply choose $f_j \in F_j$ so that $\{f_j\}$ is left invariant by Q_j. The "if" and "In particular" parts follow from Corollary 4.1.2 and Fact 4.1.6.
(iii): Suppose to the contrary μ is not simultaneously 0 and attained. In view of Corollary 4.4.5 and Theorem 5.2.1, each Q_i is fixed point free and all orbits (x_i^n) tend in norm to infinity. We may also assume (after relabeling, if necessary) that $j = N$. Fix an orbit (x_i^n) and extract a subsequence (n_k) of (n) such that

$$\|x_N^{n_k+1}\| = \|Q_N x_N^{n_k}\| > \|x_N^{n_k}\| =: r_k, \text{ for all } k.$$

Then (r_k) is strictly increasing and tends to infinity. Let us abbreviate

$$y_0^k := x_N^{n_k} \text{ and } y_i^k := P_i y_{i-1}^k, \text{ for all } k \text{ and every } i.$$

Now choose a positive integer m so large that $y_i^n \neq y_i^m$, for all $n \geq m$ and every i. (If this weren't possible, then we would conclude $y_N^n = y_N^m$ infinitely often, which would be absurd because $\|y_N^n\| > r_n \to +\infty$.)

Claim: $\lim_n \|y_i^n - y_i^m\|/\|y_{i-1}^n - y_{i-1}^m\| = 1$, for all i.

Since all of these quotients are at most 1, it suffices to show that $\lim_n \|y_N^n - y_N^m\|/\|y_0^n - y_0^m\| = 1$. This, however, follows from

$$1 \geq \frac{\|y_N^n - y_N^m\|}{\|y_0^n - y_0^m\|} \geq \frac{\|y_N^n\| - \|y_N^m\|}{\|y_0^n\| + \|y_0^m\|} > \frac{r_n - \|y_N^m\|}{r_n + \|y_0^m\|} \to 1;$$

the claim is thus verified.

Since $\|y_0^n\| = r_n \to +\infty$, we reach the following:

Conclusion: $\lim_n \|y_i^n\| = +\infty$, for all i.

After passing to another subsequence if necessary, we can also assume that the sequence $((y_i^n - y_i^m)/\|y_i^n - y_i^m\|)$ converges weakly to some point q_i, for each i. The Claim and Proposition 2.1.2.(iii) imply

$$\frac{y_{i-1}^n - y_{i-1}^m}{\|y_{i-1}^n - y_{i-1}^m\|} - \frac{y_i^n - y_i^m}{\|y_i^n - y_i^m\|}\frac{\|y_i^n - y_i^m\|}{\|y_{i-1}^n - y_{i-1}^m\|} \to 0, \text{ for all } i.$$

Hence the weak limits q_i all coincide, say $q := q_1 = \cdots = q_N$. Next, by the Conclusion and Proposition 3.2.2,

$$q_i = \text{weak}\lim_n \frac{y_i^n - y_i^m}{\|y_i^n - y_i^m\|} = \text{weak}\lim_n \frac{y_i^n}{\|y_i^n\|} \in \text{rec}(C_i), \text{ for all } i.$$

Thus

$$q \in \textstyle\bigcap_{i=1}^N \text{rec}(C_i).$$

Now $\hat{C}_N$ is relatively compact (by hypothesis), hence we can assume that $(y_N^n/\|y_N^n\|)$ converges in norm to q (after passing to yet another subsequence, if necessary) which implies

$$\|q\| = 1.$$

On the other hand, for the orbit (x_i^n), we certainly have (Fact 3.2.3.(ii) and Proposition 3.1.2.(ii)) $x_N^{n+1} - x_N^n \in \sum_{i=1}^N \text{range}(P_i - I) \subseteq (\bigcap_{i=1}^N \text{rec}(C_i))^{\oplus}$, for all n. Hence $x_N^{n+l} - x_N^n \in (\bigcap_{i=1}^N \text{rec}(C_i))^{\oplus}$, for all positive integers n, l. Remembering the special structure of (y_i^n), we conclude $(y_N^n - y_N^m)/\|y_N^n - y_N^m\| \in (\bigcap_{i=1}^N \text{rec}(C_i))^{\oplus}$, for all n, which yields

$$q \in (\textstyle\bigcap_{i=1}^N \text{rec}(C_i))^{\oplus}.$$

Altogether, we obtain

$$0 \neq q \in (\textstyle\bigcap_{i=1}^N \text{rec}(C_i)) \cap (\bigcap_{i=1}^N \text{rec}(C_i))^{\oplus}.$$

Finally, $\bigcap_{i=1}^N \text{rec}(C_i)$ is a subspace so that $(\bigcap_{i=1}^N \text{rec}(C_i)) \cap (\bigcap_{i=1}^N \text{rec}(C_i))^{\oplus} = \{0\}$ (Proposition 3.1.3). Therefore, q has to be 0 which is the desired contradiction. ☐

COROLLARY 5.3.2. If each C_i is a closed affine subspace, then $\mu = 0$.

REMARK 5.3.3. Theorem 5.3.1.(iii) is sharp in the following sense:

- Kosmol and Zhou [**62**, Theorem] and Bauschke and Borwein [**10**, Example 4.3] constructed examples where $N = 2$ and each C_i is a closed affine subspace but $\mu = 0$ and is *unattained.* Hence the preceding corollary cannot be improved and the assumption of relative compactness in Theorem 5.3.1.(iii) is important.

- The following simple example shows that the assumption on the recession cones in Theorem 5.3.1.(iii) is important: suppose $H = \mathbb{R}^2$, $N = 2$, and let $C_1 = \{(r, 0) : r \geq 0\}$, $C_2 = \{(x, y) : y \geq 1/x > 0\}$. Then $\operatorname{rec}(C_1) \cap \operatorname{rec}(C_2) = C_1$ is not linear, $\mu = 0$ (see the subsubsection on the two set case below) and unattained.

REMARKS 5.3.4. Suppose C is a closed convex nonempty subset of H and let $\hat{C} := \{c/\|c\| : c \in C \setminus \{0\}\}$.

- $\hat{C}$ need not be closed: consider in $\mathbb{R}^2$ the epigraph of the function $1/x$ restricted to the positive reals, i.e. $C = \{(x, y) \in \mathbb{R}^2 : y \geq 1/x > 0\}$. Then $\hat{C} = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1;\ x > 0;\ y > 0\}$ is not closed.
- It is easy to check that if $\hat{C}$ is relatively compact, then C is boundedly compact.
- In general, if C is boundedly compact, then $\hat{C}$ need not be relatively compact: consider in $H = \ell_2$ the compact convex nonempty set $C := \overline{\operatorname{conv}}\{e^n/n : n \geq 1\}$, where e^n denotes the n^{th} unit vector. Then $\hat{C}$ contains $\{e^1, e^2, \dots\}$ and thus cannot be relatively compact.

The two set case.

THEOREM 5.3.5. If $N = 2$, then $\mu = 0$.

PROOF. By Fact 5.2.2, $x_2^n - x_1^n \to d_1$ and $x_1^{n+1} - x_2^n \to d_2 = -d_1$. Adding both sequences yields $x_1^{n+1} - x_1^n = Q_1 x_1^n - x_1^n \to 0$ and therefore $\mu = \inf_x \|Q_1 x - x\| = 0$. □

The general case again. In view of the results of this subsection, the following conjecture is very plausible (and true for the two set and the consistent cases):

CONJECTURE 5.3.6. (zero displacement conjecture) The minimal displacement μ always equals 0.

THEOREM 5.3.7. The bounded differences conjecture is stronger than the zero displacement conjecture

PROOF. Suppose the bounded differences conjecture is true. Comparison of the orbits generated by x^0 and x_N^1 via Corollary 2.1.3 yields $\lim_n \langle x_i^n - x_{i+1}^n, x_{i+1}^n - x_{i+1}^{n+1} \rangle = 0$, for all i. Hence, by assumption and Theorem 4.3.8, $\lim_n \langle x_i^n - x_{i+1}^n, -v \rangle = 0$, for all i. Addition yields $0 = \lim_n \langle x_1^n - x_1^{n+1}, -v \rangle = \langle -v, -v \rangle = \|v\|^2$; therefore, $\mu = \|v\| = 0$ and the proof is complete. □

5.4. Convergence results I.

THEOREM 5.4.1.
(weak convergence) If some set C_j is bounded, then $(x_1^n, \dots, x_N^n)$ converges weakly to some cycle.
(norm convergence) $(x_1^n, \dots, x_N^n)$ converges in norm to some cycle whenever (at least) one of the following conditions holds:

(i) Each F_i is nonempty and some C_j is boundedly compact.
(ii) Some C_i is bounded and some C_j is boundedly compact.
(iii) Some C_j is compact.
(iv) Some C_j is bounded and H is finite-dimensional.

(v) Some normalization $\hat{C}_j = \{c_j/\|c_j\| : c_j \in C_j \setminus \{0\}\}$ is relatively compact and $\bigcap_{i=1}^N \operatorname{rec}(C_i)$ is a subspace.

(vi) H is finite-dimensional and $\bigcap_{i=1}^N \operatorname{rec}(C_i)$ is a subspace.

In all situations, $\lim_n x_{i+1}^n - x_i^n = d_i$, for each i.

Proof. (weak convergence): Combine Theorems 5.3.1.(ii) and 5.2.1.(ii).
(norm convergence).(i): On the one hand, by Theorem 5.2.1.(ii), (x_j^n) is weakly convergent to some $f_j \in F_j$. Since (x_j^n) is bounded and C_j is boundedly compact, f_j is actually a norm cluster point of (x_j^n). On the other hand, (x_j^n) is Fejér monotone w.r.t. F_j (Corollary 4.1.5). Thus, by Proposition 3.8.2.(ii), (x_j^n) converges in norm to f_j. Therefore, by continuity of each P_i, each (x_i^n) is norm convergent. The rest follows since: (iv) $\Rightarrow$ (iii) $\Rightarrow$ (ii) $\Rightarrow$ (i) $\Leftarrow$ (v) $\Leftarrow$ (vi). (For "(v) $\Rightarrow$ (i)", use Theorem 5.3.1.(iii) and Remarks 5.3.4.)
"In all situations" is nothing but Theorem 5.2.1.(ii). $\square$

Remarks 5.4.2.

- A slightly less general form of the weak convergence part of the last theorem is Gubin et al.'s [**51**, weak convergence part of Theorem 2]; they also observed (iv).
- The two set version of Theorem 5.4.1.(i) and (iii) is due to Cheney and Goldstein; see [**23**, Theorem 4]
- Consider the second example in the remark following Theorem 5.3.1. Since F_1 and F_2 are empty, the sequences (x_1^n) and (x_2^n) tend to infinity. Therefore, the assumption on the recession cones in Theorem 5.4.1.(v) is important.
- The following consistent case is interesting: if each set C_i is *symmetric* ($C_i = -C_i$), then Bruck and Reich's [**19**, Theorem 2.2] or Baillon et al.'s [**6**, Corollary 2.4] implies that the entire orbit (x_i^n) converges in norm to some point in $\bigcap_i C_i$.

The two set case. In this case, there is another sufficient condition for weak convergence:

Theorem 5.4.3. If $C_2 - C_1$ is closed, then (x_1^n, x_2^n) converges weakly to some cycle.

Proof. By Fact 5.1.4.(ii), F_1 and F_2 are nonempty. The result now follows from Theorem 5.2.1.(ii). $\square$

5.5. Convergence results II.

Theorem 5.5.1. Suppose each C_i, except possibly one, is Kadec/Klee and each F_i is nonempty. Then $(x_1^n, \dots, x_N^n)$ converges in norm to some cycle.

Proof. By Theorem 5.2.1.(ii), $(x_1^n, \dots, x_N^n)$ converges weakly to some cycle $(f_1, \dots, f_N)$. *Case 1:* $\bigcap_{i=1}^N C_i = \emptyset$. Then there is some i such that $f_{i+1} = P_{i+1} f_i \neq f_i$ and C_{i+1} is Kadec/Klee. Hence $f_{i+1} \in \operatorname{bd}(C_{i+1})$. Since $x_{i+1}^n \rightharpoonup f_{i+1}$, we conclude $x_{i+1}^n \to f_{i+1}$ and thus $(x_1^n, \dots, x_N^n) \to (f_1, \dots, f_N)$. *Case 2:* $\bigcap_{i=1}^N C_i \neq \emptyset$. Then $(C_1, \dots, C_N)$ is boundedly regular (Theorem 3.6.5.(iii)) and the result follows from Fact 5.5.5.(iii) below. $\square$

Corollary 5.5.2. (Gubin et al.'s [**51**, part of Theorem 2]) Suppose each C_i, except possibly one, is uniformly convex and some C_j is bounded. Then $(x_1^n, \dots, x_N^n)$ converges in norm to some cycle.

Proof. Clear from Proposition 3.5.2, Theorem 5.4.1, and Theorem 5.5.1. $\square$

The two set case.

FACT 5.5.3. ([**9**, Theorem 3.7 and Corollary 3.10]) Suppose $N = 2$. If (C_1, C_2) is boundedly regular, then (x_1^n, x_2^n) converges in norm to some cycle. In particular, this happens whenever C_1 or C_2 is boundedly compact and each F_i is nonempty.

REMARK 5.5.4. The first two authors' [**9**, Theorem 3.7 and Corollary 3.10] lack the assumption of nonemptiness of each F_i; see Remark 3.6.3.

The consistent case. Our observations on (linear) regularity in Subsection 3.6 combined with convergence results in [**8**, Section 5] yield the following two facts:

FACT 5.5.5. ([**8**, Theorem 5.2]) Suppose $\bigcap_i C_i \neq \emptyset$ and $(C_1, \dots, C_N)$ is boundedly regular. Then the entire orbit (x_i^n) converges in norm to some point in $\bigcap_i C_i$. In particular, this happens if (i) some C_i is boundedly compact; (ii) H is finite-dimensional; (iii) every C_i, except possibly one, is Kadec/Klee; or (iv) every C_i, except possibly one, is uniformly convex.

REMARK 5.5.6. Fact 5.5.5.(i),(ii) give a different explanation for the consistent case of Theorem 5.4.1.(i). Fact 5.5.5.(iv) is due to Gubin et al.; see [**51**, Theorem 1.(b)].

FACT 5.5.7. ([**8**, Theorem 5.7]) Suppose $\bigcap_i C_i \neq \emptyset$ and $(C_1, \dots, C_N)$ is boundedly linearly regular. Then: (a) the entire orbit (x_i^n) converges linearly to some point in $\bigcap_i C_i$; moreover, (b) the rate of convergence is independent of the starting point whenever $(C_1, \dots, C_N)$ is linearly regular. In particular, (a) happens whenever (i) $0 \in \bigcap_{i=1}^{N-1} \operatorname{icr}((\bigcap_{j=1}^{i} C_j) - C_{i+1})$ or (ii) $C_N \cap \operatorname{int}(\bigcap_{i=1}^{N-1} C_i) \neq \emptyset$.

REMARKS 5.5.8. Fact 5.5.7.(a).(ii) is also due to Gubin et al.; see [**51**, Theorem 1.(a)]. Corresponding formulations of Fact 5.5.7.(a),(i),(ii) with different orders hold, too; see Remark 3.6.10. We discuss some more special cases of Fact 5.5.7 in Subsections 5.6 and 5.7.

FACT 5.5.9. ([**9**, Theorem 5.3.(iv)]) Suppose $N = 2$ and H is a Hilbert lattice with lattice cone H^+. If C_1 is a hyperplane, $C_2 = H^+$, and $C_1 \cap C_2 \neq \emptyset$, then the entire orbit (x_i^n) converges in norm to some point in $C_1 \cap C_2$.

REMARKS 5.5.10.

- Classical examples are $H = \ell_2$ with $H^+ = \{(s_n) \text{ in } H : s_n \geq 0, \text{ for all } n\}$ and $H = L_2[0,1]$ with $H^+ = \{f \in H : f \geq 0 \text{ almost everywhere}\}$.
- Fact 5.5.9 relies heavily on the order structure of H. It would be interesting to know whether this result holds when C_1 is a closed affine finite-codimensional subspace instead of a hyperplane.

The inconsistent case.

FACT 5.5.11. (Gubin et al.'s [**51**, part of Theorem 2]) Suppose each C_i, except possibly one, is strongly convex, some C_j is bounded, and $\bigcap_{i=1}^{N} C_i = \emptyset$. Then $(x_1^n, \dots, x_N^n)$ converges linearly to some cycle.

REMARK 5.5.12. The simple example of two discs in $\mathbb{R}^2$ with precisely one common point shows that the assumption of emptiness of the intersection is important.

5.6. The convex polyhedral case.

THEOREM 5.6.1. If each C_i is a convex polyhedron, then each F_i is nonempty and the N-tuple of sequences $(x_1^n, \dots, x_N^n)$ *converges in norm* to some cycle.

PROOF. Corollary 3.4.8, Corollary 4.4.5, and Theorem 5.2.1.(ii) imply the weak convergence of $(x_1^n, \dots, x_N^n)$ to some cycle. Now suppose that each C_i is given as $\{x \in H : \langle a_{i,j}, x\rangle \leq b_{i,j}\}$. Let $K := \bigcap_{i,j} \operatorname{kernel}(a_{i,j})$ and let $D_i := C_i \cap K^\perp$, for each i. Then, by Proposition 3.3.2,

$$x_{i+1}^n - x_i^n = P_K x_i^n + P_{D_{i+1}} P_{K^\perp} x_i^n - x_i^n = -P_{K^\perp} x_i^n + P_{D_{i+1}} P_{K^\perp} x_i^n \in K^\perp,$$

for all n and each i. Telescoping yields

$$x_i^n \in x^0 + K^\perp, \text{ for all } n \text{ and every } i;$$

hence the entire orbit (x_i^n) lies in an affine finite-dimensional subspace. Here, however, weak and norm convergence coincide and the result follows. □

REMARK 5.6.2. Theorem 5.6.1 is related to a result by De Pierro and Iusem [**71**, Proposition 13 and Theorem 1] which states that (x_N^n) converges in norm to some point in F_N when H is finite-dimensional and each C_i is a halfspace; their algorithm, however, is more general.

The consistent case. Combining Fact 3.6.8 with Fact 5.5.7.(b) yields:

THEOREM 5.6.3. Suppose each C_i is a convex polyhedron and $\bigcap_{i=1}^N C_i \neq \emptyset$. Then the entire orbit (x_i^n) converges linearly to some point in $\bigcap_{i=1}^N C_i$ with a rate independent of the starting point.

REMARK 5.6.4. In the setting of the problem, special cases of this theorem were obtained by Gubin et al. ([**51**, Theorem 1.(d)]) and by Herman et al. ([**53**, Theorem 1]); see also Mandel's [**65**, Theorem 3.1] for an upper bound for the rates of convergence and [**8**, Section 6] for more general algorithms.

The general case again. Encouraged by the last fact, we believe the following is true:

CONJECTURE 5.6.5. If each C_i is a convex polyhedron, then $(x_1^n, \dots, x_N^n)$ *converges linearly* to some cycle *with a rate independent* of the starting point.

5.7. The affine subspace case. Throughout this subsection, we assume that

$$\boxed{C_1 = c_1 + L_1,\ \dots,\ C_N = c_N + L_N \text{ are closed affine subspaces,}}$$

where, without loss of generality,

$$\boxed{L_i \text{ is linear, } L := \bigcap_{i=1}^N L_i, \text{ and } c_i \in L_i^\perp, \text{ for each } i.}$$

We investigate the behaviour of the method of cyclic projections in this context and restrict ourselves to the subsequence (x_N^n) of an arbitrary but fixed orbit. Proposition 2.2.2.(i) yields $P_i x = c_i + P_{L_i} x$, for each i and all x; thus if we let $\mathbf{c} := (c_1, \dots, c_N)$ and define the operator

$$\boxed{\mathbf{T} : L_1^\perp \times \cdots \times L_N^\perp \to H : \mathbf{y} = (y_1, \dots, y_N) \mapsto \textstyle\sum_{i=1}^N P_{L_N} \cdots P_{L_{i+1}} y_i,}$$

it then follows that $Q_N x = P_{L_N} \cdots P_{L_1} x + \mathbf{Tc}$, and hence inductively

$$x_N^n = Q_N^n x^0 = (P_{L_N} \cdots P_{L_1})^n x^0 + \textstyle\sum_{k=0}^{n-1} (P_{L_N} \cdots P_{L_1})^k (\mathbf{Tc}), \text{ for all } n \geq 1.$$

In passing, we note that the first term on the right hand side of the last equation, $(P_{L_N}\cdots P_{L_1})^n x^0$, tends to $P_L x^0$ by the Fact 1.1.1.

PROPOSITION 5.7.1.

(i) $L^\perp$ is an invariant subspace of P_L and $P_{L_N}\cdots P_{L_i}$, for each i.
(ii) $\mathbf{Tc}\in L^\perp$.
(iii) $(P_{L_N}\cdots P_{L_1})^k = P_L\oplus(P_{L_N}\cdots P_{L_1}P_{L^\perp})^k$, for all $k\geq 1$.
(iv) $P_L = P_L P_{L_i}$, for each i.

PROOF. (i): Obviously, $L^\perp$ is an invariant subspace of P_L. Now fix $l^\perp\in L^\perp$ and $l\in L$. Then $\langle l, P_{L_N}\cdots P_{L_i}l^\perp\rangle = \langle (P_{L_N}\cdots P_{L_i})^* l, l^\perp\rangle = \langle P_{L_i}\cdots P_{L_N} l, l^\perp\rangle = \langle l, l^\perp\rangle = 0$, which implies that $L^\perp$ is an invariant subspaces of $P_{L_N}\cdots P_{L_i}$.
(ii): Since $c_i\in L_i^\perp\subseteq L^\perp$, for all i, (i) and the definition of $\mathbf{T}$ imply $\mathbf{Tc}\in L^\perp$.
(iii) is proved by induction on k: $P_{L_N}\cdots P_{L_1} = P_{L_N}\cdots P_{L_1}(P_L\oplus P_{L^\perp}) = P_L\oplus P_{L_N}\cdots P_{L_1}P_{L^\perp}$; hence (iii) holds for $k=1$. If it holds for some $k\geq 1$, then $(P_{L_N}\cdots P_{L_1})^{k+1} = (P_{L_N}\cdots P_{L_1})(P_L\oplus(P_{L_N}\cdots P_{L_1}P_{L^\perp})^k) =$
$P_L\oplus(P_{L_N}\cdots P_{L_1}P_{L^\perp})^{k+1}$, because $L^\perp$ is an invariant subspace of $P_{L_N}\cdots P_{L_1}$ by (i).
(iv): Fix $x\in H$ and $l\in L$. Then $\langle l-P_LP_{L_i}x, x-P_LP_{L_i}x\rangle = \langle l-P_LP_{L_i}x, P_{L_i}x - P_LP_{L_i}x\rangle + \langle l-P_LP_{L_i}x, P_{L_i^\perp}x\rangle = 0$; since this is true for all $l\in L$, we conclude (Proposition 2.1.2.(ii)) that $P_Lx = P_LP_{L_i}x$ and the result follows. ☐

COROLLARY 5.7.2. The sequence (x_N^n) satisfies

$$x_N^n = P_Lx^0 + (P_{L_N}\cdots P_{L_1}P_{L^\perp})^n x^0 + \textstyle\sum_{k=0}^{n-1}(P_{L_N}\cdots P_{L_1}P_{L^\perp})^k(\mathbf{Tc}),\ \text{for all } n\geq 1.$$

This observation is useful, since the operator norm of $P_{L_N}\cdots P_{L_1}P_{L^\perp}$ is less than or equal to the operator norm of $P_{L_N}\cdots P_{L_1}$. Moreover, $\|P_{L_N}\cdots P_{L_1}P_{L^\perp}\|$ is nothing but the cosine of the angle of the N-tuple $(L_1,\ldots,L_N)$; see Definition 3.7.5.

The next proposition provides information on the set of fixed points F_N:

PROPOSITION 5.7.3.

(i) $F_N = (I-P_{L_N}\cdots P_{L_1})^{-1}(\mathbf{Tc})$.
(ii) (Kosmol and Zhou's [**62**, Proposition 4]) If F_N is nonempty, then $F_N = f_N^* + L$, for all $f_N^*\in F_N$.

PROOF. (i): $x\in F_N \Leftrightarrow x - Q_Nx = 0 \Leftrightarrow x-(P_{L_N}\cdots P_{L_1})x = \mathbf{Tc}$.
(ii): Pick $f_N, f_N^*\in F_N$. Then, using (i), the linearity of $P_{L_N}\cdots P_{L_1}$, and Corollary 4.3.5, we get: $f_N - P_{L_N}\cdots P_{L_1}f_N = \mathbf{Tc} = f_N^* - P_{L_N}\cdots P_{L_1}f_N^* \Leftrightarrow f_N - f_N^* = (P_{L_N}\cdots P_{L_1})(f_N-f_N^*) \Leftrightarrow f_N - f_N^*\in \mathrm{Fix}\,(P_{L_N}\cdots P_{L_1}) \Leftrightarrow f_N\in f_N^*+L$. ☐

The following useful identity is due to Kosmol and Zhou; it is proved inductively using Proposition 5.7.3.(i).

FACT 5.7.4. ([**61**, Proof of Theorem 3]) If F_N is nonempty and $f_N^*\in F_N$, then

$$x_N^n = (P_{L_N}\cdots P_{L_1})^n(x^0-f_N^*) + f_N^*,\ \text{for all } n\geq 1.$$

THEOREM 5.7.5. (norm dichotomy) Exactly one of the following alternatives holds:

(i) F_N is empty. Then $(\|x_N^n\|)$ tends to $+\infty$.
(ii) F_N is nonempty. Then the sequence (x_N^n) converges in norm:

$$\lim_n x_N^n = P_{F_N}x^0 = P_Lx^0 + \textstyle\sum_{k=0}^{\infty}(P_{L_N}\cdots P_{L_1}P_{L^\perp})^k(\mathbf{Tc}).$$

Consequently, $\sum_{k=0}^\infty(P_{L_N}\cdots P_{L_1}P_{L^\perp})^k(\mathbf{Tc})$ is equal to $P_{F_N}0$, the minimum norm fixed point of Q_N.

PROOF. In view of Theorem 5.2.1, we assume that F_N is nonempty and fix $f_N^* \in F_N$. By Fact 5.7.4 and Fact 1.1.1, $x_N^n \to P_L(x^0 - f_N^*) + f_N^* = P_L x^0 + P_{L^\perp} f_N^*$. Also, by Proposition 2.2.2.(i) and Proposition 5.7.3.(ii), $P_L(x^0 - f_N^*) + f_N^* = P_{f_N^*+L}(x^0) = P_{F_N} x^0$ and $P_{L^\perp} f_N^* = f_N^* + P_L(0 - f_N^*) = P_{f_N^*+L}(0) = P_{F_N} 0$. Hence, by Corollary 5.7.2, $P_{L^\perp} f_N^* = P_{F_N} 0 = \sum_{k=0}^{\infty} (P_{L_N} \cdots P_{L_1} P_{L^\perp})^k (\mathbf{Tc})$. □

COROLLARY 5.7.6. If all F_i are empty, then $\lim_n \|x_i^n\| = +\infty$, for each i. Otherwise, all F_i are nonempty, each sequence (x_i^n) converges in norm to $P_{F_i} x^0$, and $d_i = P_{L^\perp}(F_{i+1} - F_i)$.

PROOF. In view of Theorem 5.2.1, we only have to verify the "otherwise" part. Theorem 5.7.5.(ii) implies $x_N^n \to P_{F_N} x^0$; if we apply the theorem once more (this time to (x_1^n)), we get $x_1^n \to P_{F_1} x_1^1$. Fix $f_1 \in F_1$. Then, by Proposition 5.7.3.(ii), Proposition 2.2.2.(i), and Proposition 5.7.1, $P_{F_1} x_1^1 = P_{f_1+L}(P_1 x^0) = f_1 + P_L(c_1 + P_{L_1} x^0 - f_1) = f_1 + P_L(x^0 - f_1) = P_{f_1+L}(x^0) = P_{F_1} x^0$. Also, $x_2^n \to P_{F_2} x_2^1$, and, similarly to what we just did, $P_{F_2} x_2^1 = P_{F_2} x_1^1 = P_{F_2} x^0$. Analogously, $P_{F_3} x_3^1 = P_{F_3} x^0$, …, up to $P_{F_{N-1}} x_{N-1}^1 = P_{F_{N-1}} x^0$. Finally, by Theorem 5.1.2, Proposition 5.7.3.(ii), and Proposition 2.2.2.(i), $d_i = P_{F_{i+1}}(x^0) - P_{F_i}(x^0) = P_{f_{i+1}+L}(x^0) - P_{f_i+L}(x^0) = f_{i+1} + P_L(x^0 - f_{i+1}) - (f_i + P_L(x^0 - f_i)) = P_{L^\perp}(f_{i+1} - f_i)$, for all $f_i \in F_i$ and $f_{i+1} \in F_{i+1}$. □

REMARKS 5.7.7.

- Kosmol and Zhou obtained the "otherwise" part of Corollary 5.7.6; see [**61**, Theorem 3].
- The fact that the limit of the entire orbit (x_i^n) has to be $P_{F_i} x^0$ (provided it exists) can be explained differently by using Fejér monotone sequences and quasi-projections. Indeed, the sequence (x_n^i) is Fejér monotone w.r.t. F_i (Corollary 4.1.5), hence its limit lies in $\mathcal{Q}_{F_i} x_i^1$ (Proposition 3.8.2) which is — in this setting — the singleton $\{P_{F_i} x_i^1\}$ (Proposition 2.3.2.(iii)). However, as we saw in the proof of Corollary 5.7.6, $P_{F_i} x_i^1 = P_{F_i} x^0$.

THEOREM 5.7.8. (angle and linear convergence) If $L_1^\perp + \cdots + L_N^\perp$ is closed, then F_N is nonempty and (x_N^n) converges linearly with a rate no worse than the cosine of the angle of the N-tuple $(L_1, \ldots, L_N)$:

$$\lim_n x_N^n = P_{F_N} x^0 = P_L x^0 + P_{F_N} 0 = P_L x^0 + (I - P_{L_N} \cdots P_{L_1} P_{L^\perp})^{-1} (\mathbf{Tc}).$$

In particular, this holds whenever:

(i) Some $L_i \cap L^\perp$ is finite-dimensional.
(ii) Some L_i is finite-dimensional.
(iii) H is finite-dimensional.
(iv) All L_i, except possibly one, are finite-codimensional.
(v) Each C_i is a hyperplane.
(vi) Each angle $\gamma(L_i, L_{i+1} \cap \cdots \cap L_N)$ is positive.

Consequently, each sequence (x_i^n) converges linearly to $P_{F_i} x^0$ with a rate no worse than the cosine of the angle $\gamma(L_{i+1}, \ldots, L_N, L_1, \ldots, L_i)$.

PROOF. Proposition 3.7.7 implies that the angle $\gamma(L_1, \ldots, L_N)$ of the N-tuple $(L_1, \ldots, L_N)$ is positive, in other words: $\|P_{L_N} \cdots P_{L_1} P_{L^\perp}\| < 1$. Hence, by Corollary 5.7.2, (x_N^n) converges linearly with a rate no worse than $\cos \gamma(L_1, \ldots, L_N)$; its limit is $P_L x^0 + (I - P_{L_N} \cdots P_{L_1} P_{L^\perp})^{-1} (\mathbf{Tc})$. The main statement of the theorem now follows with Theorem 5.7.5. Conditions (i) through (vi) ensure, again by Proposition 3.7.7, $\gamma(L_1, \ldots, L_N) > 0$. Finally, the "consequently" part follows from Corollary 5.7.6 and the first statement of this theorem. $\square$

REMARKS 5.7.9.

- If $N = 2$ and $L_1^\perp + L_2^\perp$ is not closed, then convergence can be "arbitrarily slow"; see Theorem 5.7.16.
- Kosmol proved a less general version of (iii); see [**60**, Satz 3].

We now specialize to the case when each C_i is a hyperplane:

$$\boxed{C_i = \{x \in H : \langle a_i, x\rangle = b_i\}, \text{ for some } a_i \in H \setminus \{0\} \text{ and } b_i \in \mathbb{R}.}$$

In the terminology of this subsection,

$$c_i = \frac{b_i}{\|a_i\|^2} a_i, \; L_i = \text{kernel}(a_i) = \{a_i\}^\perp, \; L_i^\perp = \text{span}\{a_i\}, \text{ for each } i.$$

Theorem 5.7.8.(v) implies that the angle $\gamma(L_1, \ldots, L_N)$ is positive; hence $\|P_{L_N} \cdots P_{L_1} P_{L^\perp}\| < 1$ and each subsequence (x_i^n) converges linearly.

The mapping

$$\boxed{\Psi : \mathbb{R}^N \to L_1^\perp \times \cdots \times L_N^\perp : \mathbf{y} = (y_1, \ldots, y_N) \mapsto \left(\frac{y_1}{\|a_1\|^2} a_1, \ldots, \frac{y_N}{\|a_N\|^2} a_N\right)}$$

not only defines a linear isomorphism between $\mathbb{R}^N$ and $L_1^\perp \times \cdots \times L_N^\perp$, but also allows us to think of $\mathbf{T}$ as being defined on $\mathbb{R}^N$; this will yield a nice connection to generalized inverses.

DEFINITION 5.7.10. Suppose Y is another Hilbert space and $A : H \to Y$ is a continuous linear operator with closed range. The *generalized inverse of* A, denoted $A^\dagger$, is the unique continuous linear operator from Y to H satisfying (i) $A^\dagger A = I$ on $(\text{kernel}(A))^\perp$ and (ii) $A^\dagger = 0$ on $(\text{range}(A))^\perp$.

REMARK 5.7.11. Here, we have chosen the definition of the generalized inverse due to Desoer and Whalen [**26**] which is equivalent to better known definitions by Moore [**67**] and by Penrose [**69**]. A proof of these equivalences and more can be found in Groetsch's monograph [**50**].

Now consider

$$\boxed{A : H \to \mathbb{R}^N : x \mapsto (\langle a_1, x\rangle, \ldots, \langle a_N, x\rangle).}$$

Then $A^* : \mathbb{R}^N \to H : \mathbf{y} = (y_1, \ldots, y_N) \mapsto \sum_{i=1}^N y_i a_i$. Because of $(\text{kernel}(A))^\perp = \text{range}(A^*)$ and $(\text{range}(A))^\perp = \text{kernel}(A^*)$, we have

$$\text{span}\{a_1, \ldots, a_N\} = \text{range}(A^*) = (\text{kernel}(A))^\perp = (\textstyle\bigcap_{i=1}^N \text{kernel}(a_i))^\perp = L^\perp,$$

and

$$(\text{range}(A))^\perp = \text{kernel}(A^*) = \{\mathbf{y} = (y_1, \ldots, y_N) \in \mathbb{R}^N : \textstyle\sum_{i=1}^N y_i a_i = 0\}.$$

PROPOSITION 5.7.12. If the vectors $a_1, \ldots, a_N$ are linearly independent, then $(I - P_{L_N} \cdots P_{L_1} P_{L^\perp})^{-1} \mathbf{T}\Psi$ is the generalized inverse $A^\dagger$ of A.

PROOF. We have to check conditions (i) and (ii) of Definition 5.7.10.
(i): Fix $x \in L^\perp = (\text{kernel}(A))^\perp$. Then

$$\begin{aligned}
\mathbf{T}\Psi A x &= \mathbf{T}\Psi(\langle a_1, x\rangle, \ldots, \langle a_N, x\rangle) \\
&= \mathbf{T}\left(\frac{\langle a_1, x\rangle}{\|a_1\|^2}a_1, \ldots, \frac{\langle a_N, x\rangle}{\|a_N\|^2}a_N\right) \\
&= \mathbf{T}(P_{L_1^\perp}x, \ldots, P_{L_N^\perp}x) \\
&= \textstyle\sum_{i=1}^N (P_{L_N}\cdots P_{L_{i+1}})P_{L_i^\perp}x \\
&= \textstyle\sum_{i=1}^N (P_{L_N}\cdots P_{L_{i+1}})(x - P_{L_i}x) \\
&= x - P_{L_N}\cdots P_{L_1}x \\
&= (I - P_{L_N}\cdots P_{L_1}P_{L^\perp})x,
\end{aligned}$$

so that $(I - P_{L_N}\cdots P_{L_1}P_{L^\perp})^{-1}\mathbf{T}\Psi A x = x$, as promised.
(ii): Since the vectors $a_1, \ldots, a_N$ are linearly independent, $(\text{range}(A))^\perp = \{0\}$ and hence (ii) clearly holds. $\square$

The following example shows that the assumption of linear independence in the last theorem is important:

EXAMPLE 5.7.13. Let $H = \mathbb{R}$, $N = 2$, $a_1 = a_2 = 1$. A simple calculation yields

$$A = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \ (I - P_{L_2}P_{L_1}P_{L^\perp})^{-1}\mathbf{T}\Psi = (0,1), \text{ but } A^\dagger = \left(\tfrac{1}{2}, \tfrac{1}{2}\right).$$

THEOREM 5.7.14. Suppose $a_1, \ldots, a_N$ are nonzero vectors in H and let $A : H \to \mathbb{R}^N : x \mapsto (\langle a_i, x\rangle)_{i=1}^N$. Given any vector $\mathbf{b} = (b_i)_{i=1}^N \in \mathbb{R}^N$, let $C_i = \{x \in H : \langle a_i, x\rangle = b_i\}$, for all i. Then:

(i) The subsequence (x_N^n) of the orbit with starting point x^0 converges linearly to $(I - P_{L_N}\cdots P_{L_1}P_{L^\perp})^{-1}\mathbf{T}\Psi\mathbf{b} + (I - A^\dagger A)x^0$.
(ii) (generalized inverses via cyclic projections) If the vectors $a_1, \ldots, a_N$ are furthermore linearly independent, then the subsequence (x_N^n) of the orbit with starting point $x^0 = 0$ converges linearly to $A^\dagger\mathbf{b}$.

PROOF. (i): On the one hand, (x_N^n) converges linearly to

$$P_L x^0 + (I - P_{L_N}\cdots P_{L_1}P_{L^\perp})^{-1}\mathbf{T}\mathbf{c} = P_L x^0 + (I - P_{L_N}\cdots P_{L_1}P_{L^\perp})^{-1}\mathbf{T}\Psi\mathbf{b}.$$

On the other hand, by Examples 2.2.1.(v),(vii), $P_L = I - P_{L^\perp} = I - P_{\text{span}\{a_1,\ldots,a_N\}} = I - A^\dagger A$. Hence (i) holds. (ii) follows from the last proposition and (i). $\square$

REMARKS 5.7.15.

- Related finite-dimensional versions of (i) were obtained by Censor et al. [**22**, Lemma 2], De Pierro and Iusem [**71**, Theorem 1 and Proposition 10], Eggermont et al. [**38**, Theorem 1.1], Tanabe [**77**, Corollary 9], and Trummer [**78**, Theorem 14].
- In view of Example 5.7.13, the assumption of linear independence in (ii) is important.

The two set case. We proved the two set version of Theorem 5.7.8 in [**9**, Theorem 4.11]. In addition, it should be noted that

$$d_1 = P_{L_1^\perp \cap L_2^\perp}(c_2 - c_1),\ d_2 = P_{L_1^\perp \cap L_2^\perp}(c_1 - c_2) = -d_1;$$

see [**9**, Example 2.2].

We conclude with a striking trichotomy result:

THEOREM 5.7.16. (trichotomy) Suppose $N = 2$. Exactly one of the following alternatives holds:

(i) F_1, F_2 are empty. Then $\gamma(L_1, L_2) = 0$ is zero and $\lim_n \|x_2^n\| = +\infty$.
(ii) F_1, F_2 are nonempty and $\gamma(L_1, L_2) = 0$. Then for every starting point $x^0 \in H$, the sequence (x_2^n) converges in norm to $P_{F_2}x^0$. Convergence is "arbitrarily slow" in the following sense: for each sequence (λ_k) of real numbers with $1 > \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_{k-1} \geq \lambda_k \to 0$, there exists some starting point $x^0 \in H$ such that $\|x_2^k - P_{F_2}x^0\| \geq \lambda_k$, for all k.
(iii) F_1, F_2 are nonempty and $\gamma(L_1, L_2) > 0$. Then for every starting point x^0, the sequence (x_2^n) converges linearly to $P_{F_2}x^0$ with a rate no worse than $\cos\gamma(L_1, L_2)$.

PROOF. By Theorems 5.7.5, 5.7.8, we have only to prove (ii). In view of Fact 5.7.4, we assume without loss of generality that $C_1 = L_1$, $C_2 = L_2$, so that $F_2 = F_1 = C_1 \cap C_2$. Set $A := C_1 \cap (C_1 \cap C_2)^\perp$, $B := C_2 \cap (C_1 \cap C_2)^\perp$.

Step 1: Select a sequence of integers (s_k) such that

$$s_k\lambda_k < 1 \leq (s_k + 1)\lambda_k, \text{ for all } k.$$

Then (s_k) is increasing and every term s_k occurs only finitely often. Let (t_n) be the strictly increasing sequence of integers with $\{t_1, t_2, \dots\} = \{s_1, s_2, \dots\}$. Note that $\sum_n 1/t_n^2$ is convergent.

Step 2: Define $k_0(n) := \min\{k : s_k = t_n\}$, $k_1(n) := \max\{k : s_k = t_n\}$ so that

$$s_{k_0(n)-1} = t_{n-1} < t_n = s_{k_0(n)} = s_{k_0(n)+1} = \cdots = s_{k_1(n)} < t_{n+1} = s_{k_1(n)+1};$$

and further set

$$\alpha_n := \left(\lambda_{k_0(n)} t_n\right)^{\frac{1}{2k_1(n)}}, \text{ for all } n.$$

Since $1 > \lambda_{k_0(n)} s_{k_0(n)} = \lambda_{k_0(n)} t_n \geq 1 - \lambda_{k_0(n)}$ and $\lambda_k \to 0$, we get $\alpha_n \to 1^-$.

Claim: There exist sequences (e_n') in A, (f_n') in B such that for all n, m with $n \neq m$:

$$\langle e_n', e_m' \rangle = \langle e_n', f_m' \rangle = \langle f_n', f_m' \rangle = 0,\ \|e_n'\|, \|f_n'\| \leq 1,\ \langle e_n', f_n' \rangle \geq \alpha_n.$$

By Remarks 3.7.6.(i), $\sup\{\langle a, b \rangle : a \in A,\ b \in B,\ \|a\|, \|b\| \leq 1\} = 1$. We construct the sequences inductively. Clearly, e_1', f_1' exist. Now suppose we have already found $e_1', \dots, e_m'$ and $f_1', \dots, f_m'$. Set $E := \operatorname{span}\{e_1', \dots, e_m'\}, F := \operatorname{span}\{f_1', \dots, f_m'\}$. Pick sequences (a_n) in A, (b_n) in B with $\|a_n\| = \|b_n\| = 1$, $\langle a_n, b_n \rangle \to 1$, and $a_n \rightharpoonup a$, $b_n \rightharpoonup b$. Then $a_n - b_n \to 0$ and hence $a = b = 0$. Because P_E, P_F are compact operators, we conclude $P_E a_n, P_F b_n \to 0$ and further

$$1 = \lim_n \langle a_n, b_n \rangle = \lim_n \langle P_E a_n + P_{E^\perp} a_n, P_F b_n + P_{F^\perp} b_n \rangle = \lim_n \langle P_{E^\perp} a_n, P_{F^\perp} b_n \rangle.$$

Consequently, for n sufficiently large, $e_{m+1}' := P_{E^\perp} a_n, f_{m+1}' := P_{F^\perp} b_n$ do the job and the claim is verified.

Step 3: Define $e_n := e_n'/\|e'n\|$, $f_n := f_n'/\|f_n'\|$, for all n and $E := \overline{\operatorname{span}}\{e_1, e_2, \dots\}$, $F := \overline{\operatorname{span}}\{f_1, f_2, \dots\}$. Then the sequences (e_n), (f_n) have the same properties as the sequences (e'_n), (f'_n); moreover, their terms are norm one. Since $E = A \cap (E+F)$, $F = B \cap (E+F)$, $(E+F)^\perp = E^\perp \cap F^\perp$, we obtain

$$C_1 = (C_1 \cap C_2) \oplus E \oplus (A \cap E^\perp \cap F^\perp), C_2 = (C_1 \cap C_2) \oplus F \oplus (B \cap E^\perp \cap F^\perp).$$

If $z \in F$, then $P_1 z = P_{(C_1 \cap C_2)} z + P_E z + P_{(A \cap E^\perp \cap F^\perp)} z = P_E z$, because F is orthogonal to $C_1 \cap C_2$ and to $A \cap E^\perp \cap F^\perp$. Similarly, if $z \in F$, then $P_2 z = P_E z$. Thus $Q_2|_F \equiv P_2 P_1|_F \equiv P_F P_E|_F$.

Step 4: $P_{C_1 \cap C_2}|_F \equiv P_{E \cap F}|_F \equiv 0$. This follows from Fact 1.1.1 and Step 3: if $z \in F$, then $P_{C_1 \cap C_2} z = \lim_n (P_2 P_1)^n z = \lim_n (P_F P_E)^n z = P_{E \cap F} z = P_{\{0\}} z = 0$.

Step 5: Set $x^0 := \sum_n (1/t_n) f_n$. Then $P_{F_2} x^0 = P_{C_1 \cap C_2} x^0 = 0$ and one readily verifies that $(P_2 P_1)^k x^0 = \sum_n (1/t_n) \langle e_n, f_n \rangle^{2k} f_n$ which yields

$$\|x_2^k - P_{F_2} x^0\| \geq \frac{1}{t_n} \alpha_n^{2k}, \text{ for all } n, k.$$

Last Step: Fix an arbitrary k and get $\bar{n}$ such that $k_0(\bar{n}) \leq k \leq k_1(\bar{n})$. Then

$$\|x_2^k - P_{F_2} x^0\| \geq \frac{1}{t_{\bar{n}}} \alpha_{\bar{n}}^{2k} \geq \frac{\alpha_{\bar{n}}^{2k_1(\bar{n})}}{t_{\bar{n}}} \geq \lambda_{k_0(\bar{n})} \geq \lambda_k;$$

the proof is complete. □

REMARKS 5.7.17. For examples of case (i), see Kosmol and Zhou's [**62**, Theorem] and Bauschke and Borwein's [**10**, Example 4.3]. Franchetti and Light have an example of case (ii) in [**42**, Section 4]; we give another example below. Theorem 5.7.8 provides many examples of case (iii).

EXAMPLE 5.7.18. Let $H := \ell_2$ and denote the unit vectors by (u_n). Suppose (α_n) is a sequence of positive real numbers with $1 > \alpha_n > -1$ and $\lim_n \alpha_n = 1$. Set $e_n := u_{2n-1}$, $f_n := \alpha_n u_{2n-1} + \sqrt{1-\alpha_n^2} u_{2n}$, $C_1 := \overline{\mathrm{span}}\{e_1, e_2, \dots\}$, and $C_2 := \overline{\mathrm{span}}\{f_1, f_2, \dots\}$. Then $C_1 \cap C_2 = \{0\}$ and the sum $C_1^\perp + C_2^\perp$ is dense in H but not closed. By choosing (α_n) and (x^0) appropriately, we can arrange arbitrarily slow convergence of (x_2^n); see the proof of the last theorem.

The consistent case. Our journey is over: the consistent case of Theorem 5.7.5 is the von Neumann/Halperin result (Fact 1.1.1) we started with. A vast number of papers on this result have been appearing; we feel the reader is in good hands when we refer him or her once more to Deutsch's survey article [**28**]. Two special cases have received much attention:

- two subspaces with positive angle. Sharper upper bounds on the rate of convergence can be found in [**3, 27, 58**].
- intersecting hyperplanes. Here one obtains the well-known method of Kaczmarz [**57**] (see also [**49**]) for solving systems of linear equations.

For more general algorithmic schemes, see [**8**] and the references therein.

5.8. Related algorithms.

Random projections. Suppose r is a *random mapping* for $\{1,\ldots,N\}$, i.e. a surjective mapping from $\mathbb{N}$ to $\{1,\ldots,N\}$ that assumes every value infinitely often. Then generate a *random sequence* (x_n) by

$$x_0 \in H \text{ arbitrary}, \; x_{n+1} := P_{r(n+1)}x_n, \; \text{for all } n \geq 0.$$

This *method of random projections* is more general than the method of cyclic projections; indeed, if $r(\cdot) \equiv [\cdot]$, then the random sequence (x_n) and the orbit generated by x_0 (in the sense of Section 1) coincide.
It is not surprising that little is known on random projections for the inconsistent case with three or more sets, because the preceding material for the more restrictive case of cyclic projections has left some important questions open.

The two set case. Each projection P_i is *idempotent*, i.e. $P_i^2 = P_i$; thus the random sequence (x_n) is essentially the orbit (for cyclic projections) with starting point x_0.

The consistent case. The random sequence (x_n) converges weakly to some point in $\bigcap_{i=1}^N C_i$, when each C_i is a closed subspace (Amemiya and Ando [**2**]), or $N = 3$ (Dye and Reich [**37**]), or the sets C_i have a common "inner point" (Dye and Reich [**37**], Youla [**82**]; see also Bruck [**18**], Dye [**34**], Dye et al. [**35**], and Dye and Reich [**36**]). It is not known whether or not every random sequence converges weakly to some point in $\bigcap_{i=1}^N C_i$.

Norm convergence of the random sequence (x_n) to some point in $\bigcap_i C_i$ follows when some C_i is boundedly compact (Bauschke and Borwein [**8**, Example 4.6]; see also Aharoni and Censor [**1**], Bruck [**18**], Elsner et al. [**39**], Flåm and Zowe [**41**], and Tseng [**79**]) or when the N-tuple $(C_1,\ldots,C_N)$ is *innately boundedly regular*, i.e. $(C_j)_{j\in J}$ is boundedly regular, for every nonempty subset J of $\{1,\ldots,N\}$ (Bauschke [**7**]).

All of the above authors have established more general results.

Weighted projections. The method of (equally) *weighted projections* is given by

$$x_0 \in H \text{ arbitrary}, \; x_{n+1} := \textstyle\sum_{i=1}^N \frac{1}{N}P_i x_n, \; \text{for all } n \geq 0.$$

At first glance, this looks quite different from the method of cyclic projections; however Pierra's product space formalization ([**70**]; see also [**41, 55**]) allows an interpretation as the method of cyclic projections in a suitable product space:

Let $\mathbf{H} := \prod_{i=1}^N (H, \frac{1}{N}\langle\cdot,\cdot\rangle)$, $\mathbf{C}_1 := \{\mathbf{x} = (x_i)_{i=1}^N \in \mathbf{H} : x_1 = \cdots = x_N\}$, and $\mathbf{C}_2 := \prod_{i=1}^N C_i$.
The projection of $\mathbf{x} = (x_i)_{i=1}^N$ onto $\mathbf{C}_1$, $\mathbf{C}_2$ is given by

$$P_{\mathbf{C}_1}\mathbf{x} = \left(\textstyle\sum_{i=1}^N \frac{1}{N}x_i, \ldots, \sum_{i=1}^N \frac{1}{N}x_i\right), \; P_{\mathbf{C}_2}\mathbf{x} = (P_1x_1,\ldots,P_Nx_N).$$

Moreover, by [**10, 9**],

$$\begin{aligned} \mathbf{F}_1 &:= \text{Fix}\,(P_{\mathbf{C}_1}P_{\mathbf{C}_2}) \\ &= \{(x,\ldots,x) \in \mathbf{H} : x \in \text{Fix}\,(\tfrac{P_1+\cdots+P_N}{N})\} \\ &= \{(x,\ldots,x) \in \mathbf{H} : x \in \text{arginf}_{\,x\in H} \textstyle\sum_{i=1}^N d^2(x,C_i)\}, \end{aligned}$$

and

$$\mathbf{F}_2 := \text{Fix}\,(P_{\mathbf{C}_2}P_{\mathbf{C}_1}) = \text{arginf}_{\,(x_1,\ldots,x_N)\in\mathbf{C}_2} \textstyle\sum_{i,j} \|x_i - x_j\|^2.$$

If we now consider the method of cyclic projections for $\mathbf{C}_1$, $\mathbf{C}_2$ in $\mathbf{H}$ with starting point $\mathbf{x}^0 = (x_0, \ldots, x_0)$, then

$$\mathbf{x}_2^n = (x_{n+1}, \ldots, x_{n+1}), \text{ for all } n \geq 0.$$

Hence all results on the two set case can be "translated" to this particular setting; see, for instance, [**9**, Section 6].
Also, Reich proved norm convergence for weighted projections if each set C_i is symmetric; see [**72**] for this result and more.

Flexible weighted relaxed projections. The method of *flexible weighted relaxed projections* is a generalization of the two previously discussed methods: Given $x_0 \in H$, generate a sequence (x_n) by

$$x_{n+1} := \left(\sum_{i=1}^N \lambda_i^{(n)} \left[(1-\alpha_i^{(n)})I + \alpha_i^{(n)} P_i\right]\right) x_n, \text{ for all } n \geq 0.$$

The $\alpha_i^{(n)} \in [0,2]$ are called *relaxation parameters* and the $\lambda_i^{(n)}$ are nonnegative *weights* with $\sum_{i=1}^N \lambda_i^{(n)} = 1$; they might well depend on n and i.

Under additional assumptions on the relaxation parameters and on the weights, numerous (linear, norm, and weak) convergence results have been obtained for the consistent case; see [**8**] and the references therein for further information.

Moreover, we would like to mention that at least some of the results on the inconsistent case given in the previous sections generalize to the method of *relaxed projections*, which is the method of flexible weighted projections with $\alpha_i^{(n)} \equiv \alpha_i \in]0,2[$ and $\lambda_{[n]}^{(n-1)} \equiv 1$.

Dykstra's algorithm. Dykstra's algorithm is closely related to the method of cyclic projections; in fact, if each C_i is a closed affine subspace, then Dykstra's algorithm is essentially the method of cyclic projections. It is defined as follows:

Fix an arbitrary $x_0 \in H$ and set $e_{-(N-1)} := e_{-(N-2)} := \cdots := e_0 := 0$. Then define two sequences (x_n), (e_n) by

$$x_{n+1} := P_{[n+1]}(x_n + e_{n+1-N}), e_{n+1} := (I - P_{[n+1]})(x_n + e_{n+1-N}), \text{ for all } n \geq 0.$$

CONJECTURE 5.8.1. (on Dykstra's algorithm) If all F_i are empty, then $\lim_n \|x_n\| = +\infty$. Otherwise, all F_i are nonempty and

$$(x_n)_{[n]=i} \text{ converges in norm to } P_{F_i} x_0, \text{ for every } i.$$

The following fact not only supports this conjecture but also demonstrates the power of Dykstra's algorithm:

FACT 5.8.2. The previous conjecture is true for: (i) the two set case; (ii) the consistent case; and (iii) the affine subspace case.

PROOF. (i) is Bauschke and Borwein's [**10**, Theorem 3.8], (ii) is Boyle and Dykstra's [**15**, Theorem 2]. (iii): Denote the orbit (for the method of cyclic projections) with starting point x_0 by (x_i^n). It is not hard to check that $x_i^n := x_{(n-1)N+i}$, for all n and every i. The result now follows from Corollary 5.7.6. ☐

For further information on Dykstra's algorithm, the reader is referred to [**15, 10, 30, 32, 54**]

Acknowledgments. It is our pleasure to thank Frank Deutsch and Jon Vanderwerff for numerous helpful discussions and remarks. We also wish to thank Yair Censor, Simeon Reich, and an anonymous referee for careful reading and for helpful suggestions.

References

[1] R. AHARONI and Y. CENSOR. Block-iterative projection methods for parallel computation of solutions to convex feasibility problems. *Linear Algebra and its Applications*, 120:165–175, 1989.

[2] I. AMEMIYA and T. ANDO. Convergence of random products of contractions in Hilbert space. *Acta scientiarum mathematicarum (Szeged)*, 26:239–244, 1965.

[3] N. ARONSZAJN. Theory of reproducing kernels. *Transactions of the American Mathematical Society*, 68:337–404, May 1950.

[4] J.-P. AUBIN and I. EKELAND. *Applied Nonlinear Analysis.* Wiley-Interscience, New York, 1984.

[5] J.B. BAILLON and R.E. BRUCK. Ergodic theorems and the asymptotic behavior of contraction semigroups. In K.K. Tan, editor, *Fixed Point Theory and Applications*, pages 12–26, Singapore, 1992. World Scientific Publ. Proceedings of the second international conference held in Halifax, Nova Scotia, Canada, June 9–14, 1991.

[6] J.B. BAILLON, R.E. BRUCK, and S. REICH. On the asymptotic behaviour of nonexpansive mappings and semigroups in Banach spaces. *Houston Journal of Mathematics*, 4(1):1–9, 1978.

[7] H.H. BAUSCHKE. A norm convergence result on random products of relaxed projections in Hilbert space. *Transactions of the American Mathematical Society*, 347(4):1365–1374, April 1995.

[8] H.H. BAUSCHKE and J.M. BORWEIN. On projection algorithms for solving convex feasibility problems. *SIAM Review.* To appear.

[9] H.H. BAUSCHKE and J.M. BORWEIN. On the convergence of von Neumann's alternating projection algorithm for two sets. *Set-Valued Analysis*, 1(2):185–212, 1993.

[10] H.H. BAUSCHKE and J.M. BORWEIN. Dykstra's alternating projection algorithm for two sets. *Journal of Approximation Theory*, 79(3):418–443, December 1994.

[11] J.M. BORWEIN and S. FITZPATRICK. Mosco convergence and the Kadec property. *Proceedings of the American Mathematical Society*, 106(3):843–851, July 1989.

[12] J.M. BORWEIN and A.S. LEWIS, 1993. Unpublished manuscript.

[13] J.M. BORWEIN and D.T. YOST. Absolute norms on vector lattices. *Proceedings of the Edinburgh Mathematical Society*, 27:215–222, 1984.

[14] J.E. BOYD and J.J. LITTLE. Complementary data fusion for limited-angle tomography. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 288–294, 1994.

[15] J.P. BOYLE and R.L. DYKSTRA. A method for finding projections onto the intersection of convex sets in Hilbert spaces. In R.L Dykstra, T. Robertson, and F.T. Wright, editors, *Advances in Order Restricted Statistical Inference*, pages 28–47. Springer-Verlag, 1985. Lecture Notes in Statistics; vol 37. Proceedings, Iowa City.

[16] L.M. BREGMAN. The method of successive projection for finding a common point of convex sets. *Soviet Mathematics Doklady*, 6:688–692, 1965.

[17] F.E. BROWDER. Nonexpansive nonlinear operators in a Banach space. *Proceedings of the National Academy of Science, USA*, 54:1041–1044, 1965.

[18] R.E. BRUCK. Random products of contractions in metric and Banach spaces. *Journal of Mathematical Analysis and Applications*, 88:319–332, 1982.

[19] R.E. BRUCK and S. REICH. Nonexpansive projections and resolvents of accretive operators in Banach spaces. *Houston Journal of Mathematics*, 3(4):459–470, 1977.

[20] C. CASTAING and M. VALADIER. *Convex Analysis and Measurable Multifunctions*, volume 580 of *Lecture notes in mathematics.* Springer-Verlag, Berlin, 1977.

[21] Y. CENSOR. On variable block algebraic reconstruction techniques. In G.T. Herman, A.K. Louis, and F. Natterer, editors, *Mathematical Methods in Tomography*, pages 133–140. Springer, 1990. Proceedings of a Conference held in Oberwolfach, Germany, June 11-15. LNM Vol 1497.

[22] Y. CENSOR, P.P.B. EGGERMONT, and D. GORDAN. Strong underrelaxation in Kaczmarz's method for inconsistent systems. *Numerische Mathematik*, 41:83–92, 1983.

[23] W. CHENEY and A.A. GOLDSTEIN. Proximity maps for convex sets. *Proceedings of the American Mathematical Society*, 10:448–450, 1959.

[24] P.L. COMBETTES. The foundations of set theoretic estimation. *Proceedings of the IEEE*, 81(2):182–208, February 1993.

[25] P.L. COMBETTES. *The Convex Feasibility Problem in Image Recovery*, volume 95 of *Advances in Imaging and Electron Physics*, pages 155–270. Academic Press, Inc., 1996.

[26] C.A. DESOER and B.H. WHALEN. A note on pseudoinverses. *SIAM Journal*, 11(2):442–447, June 1963.

[27] F. DEUTSCH. Rate of convergence of the method of alternating projections. In B. Brosowski and F. Deutsch, editors, *Parametric optimization and approximation*, pages 96–107. Birkhäuser, 1983. International Series of Numerical Mathematics Vol 72.

[28] F. DEUTSCH. The method of alternating orthogonal projections. In S.P Singh, editor, *Approximation Theory, Spline functions and Applications*, pages 105–121. Kluwer Academic Publ., 1992. Proceedings of a Conference held in the Hotel villa del Mare, Maratea, Italy between April 28, 1991 and May 9, 1991.

[29] F. DEUTSCH. The angle between subspaces of a Hilbert space. In S.P. Singh, editor, *Approximation Theory, Wavelets and Applications*, pages 107–130. Kluwer Academic Publishers, 1995.

[30] F. DEUTSCH. Dykstra's cyclic projections algorithm: the rate of convergence. In S.P. Singh, editor, *Approximation Theory, Wavelets and Applications*, pages 87–94. Kluwer Academic Publishers, 1995.

[31] F. DEUTSCH and H. HUNDAL. The rate of convergence for the method of alternating projections II. *Journal of Mathematical Analysis and Applications*. to appear.

[32] F. DEUTSCH and H. HUNDAL. The rate of convergence of Dykstra's cyclic projections algorithm: the polyhedral case. *Numerical Functional Analysis and Optimization*, 15(5-6):537–565, 1994.

[33] R. DEVILLE, G. GODEFROY, and V. ZIZLER. *Smoothness and Renormings in Banach Spaces*, volume 64 of *Monographs and Surveys in Pure and Applied Mathematics*. Pitnam, 1993.

[34] J.M. DYE. Convergence of random products of compact contractions in Hilbert space. *Integral Equations and Operator Theory*, 12:12–22, 1989.

[35] J.M. DYE, T. KUCZUMOW, P.-K. LIN, and S. REICH. Random products of nonexpansive mappings in spaces with the Opial property. In B.-L. Lin and W.B. Johnson, editors, *Banach Spaces*, pages 87–93, Providence, Rhode Island, 1993. American Mathematical Society. Proceedings of an International Workshop on Banach Space Theory, held January 6–17, 1992. Contemporary mathematics; v. 144.

[36] J.M. DYE and S. REICH. Random products of nonexpansive mappings. In A. Ioffe, M. Marcus, and S. Reich, editors, *Optimization and Nonlinear Analysis*, pages 106–118, Harlow, England, 1992. Longman Scientific & Technical. Proceedings of a Binational Workshop on Optimization and Nonlinear Analysis, held at Technion City, Haifa, 21–27 March 1990. Pitman Research Notes in Mathematics Series, Vol 244.

[37] J.M. DYE and S. REICH. Unrestricted iterations of nonexpansive mappings in Hilbert space. *Nonlinear Analysis*, 18(2):199–207, 1992.

[38] P.P.B. EGGERMONT, G.T. HERMAN, and A. LENT. Iterative algorithms for large partitioned linear systems, with applications to image reconstruction. *Linear Algebra and its Applications*, 40:37–67, 1981.

[39] L. ELSNER, I. KOLTRACHT, and M. NEUMANN. Convergence of sequential and asynchronous nonlinear paracontractions. *Numerische Mathematik*, 62:305–319, 1992.

[40] Ky FAN. On systems of linear inequalities. In H.W. Kuhn and A.W. Tucker, editors, *Linear inequalities and related systems*, pages 99–156. Princeton University Press, 1956. Annals of Mathematics Studies Number 38.

[41] S.D. FLÅM and J. ZOWE. Relaxed outer projections, weighted averages and convex feasibility. *BIT*, 30:289–300, 1990.

[42] C. FRANCHETTI and W. LIGHT. On the von Neumann alternating algorithm in Hilbert space. *Journal of Mathematical Analysis and Applications*, 114:305–314, 1986.

[43] K. FRIEDRICHS. On certain inequalities and characteristic value problems for analytic functions and for functions of two variables. *Transactions of the American Mathematical Society*, 41:321–364, 1937.

[44] N. GAFFKE and R. MATHAR. A cyclic projection algorithm via duality. *Metrika*, 36:29–54, 1989.

[45] K. GOEBEL and W.A. KIRK. *Topics in metric fixed point theory.* Cambridge University Press, 1990. Cambridge studies in advanced mathematics 28.

[46] K. GOEBEL and S. REICH. *Uniform convexity, hyperbolic geometry, and nonexpansive mappings*, volume 83 of *Monographs and textbooks in pure and applied mathematics.* Marcel Dekker, New York, 1984.

[47] D. GÖHDE. Zum Prinzip der kontraktiven Abbildung. *Mathematische Nachrichten*, 30:251–258, 1965.

[48] M. GOLDBURG and R.J. MARKS II. Signal synthesis in the presence of an inconsistent set of constraints. *IEEE Transactions on Circuits and Systems*, CAS-32(7):647–663, July 1985.

[49] R. GORDON, R. BENDER, and G.T. HERMAN. Algebraic reconstruction techniques (ART) for three-dimensional electron microscopy and X-ray photography. *Journal of theoretical Biology*, 29:471–481, 1970.

[50] C.W. GROETSCH. *Generalized inverses of linear operators.* Marcel Dekker, Inc., New York, 1977. Monographs and textbooks in pure and applied mathematics, Vol 37.

[51] L.G. GUBIN, B.T. POLYAK, and E.V. RAIK. The method of projections for finding the common point of convex sets. *U.S.S.R. Computational Mathematics and Mathematical Physics*, 7(6):1–24, 1967.

[52] I. HALPERIN. The product of projection operators. *Acta scientiarum mathematicarum (Szeged)*, 23:96–99, 1962.

[53] G.T. HERMAN, A. LENT, and P.H. LUTZ. Relaxation methods for image reconstruction. *Communications of the Association for Computing Machinery*, 21(2):152–158, February 1978.

[54] H. HUNDAL and F. DEUTSCH. Two generalizations of Dykstra's cyclic projection algorithm, 1995. Preprint.

[55] A.N. IUSEM and A.R. DE PIERRO. On the convergence of Han's method for convex programming with quadratic objective. *Mathematical Programming*, 52:265–284, 1991.

[56] G.J.O. JAMESON. *Topology and Normed Spaces.* Chapman and Hall, 1974.

[57] S. KACZMARZ. Angenäherte Auflösung von Systemen linearer Gleichungen. *Bulletin internationel de l'Académie Polonaise des Sciences et des Lettres. Classe des Sciences mathématiques et naturelles. Séries A: Sciences mathématiques*, pages 355–357, 1937. Cracovie, Imprimerie de l'Université.

[58] S. KAYALAR and H.L. WEINERT. Error bounds for the method of alternating projections. *Mathematics of Control, Signals, and Systems*, 1:43–59, 1988.

[59] W.A. KIRK. A fixed point theorem for mappings which do not increase distances. *American Mathematical Monthly*, 72:1004–1006, 1965.

[60] P. KOSMOL. Über die sukzessive Wahl des kürzesten Weges. In Opitz and Rauhut, editors, *Ökonomie und Mathematik*, pages 35–42, Berlin, 1987. Springer-Verlag.

[61] P. KOSMOL and X. ZHOU. The limit points of affine iterations. *Numerical Functional Analysis and Optimization*, 11(3&4):403–409, 1990.

[62] P. KOSMOL and X.-L. ZHOU. The product of affine orthogonal projections. *Journal of Approximation Theory*, 64:351–355, 1991.

[63] V. LAKSHMIKANTHAM and D. TRIGIANTE. *Theory of Difference Equations*, volume 181 of *Mathematics in Science and Engineering.* Academic Press, San Diego, 1987.

[64] J.P. LASALLE. *The stability and control of discrete processes*, volume 62 of *Applied Mathematical Sciences.* Springer-Verlag, New York, 1986.

[65] J. MANDEL. Convergence of the cyclic relaxation method for linear inequalities. *Mathematical Programming*, 30:218–228, 1984.

[66] M. MARTELLI. *Discrete dynamical sytems and chaos*, volume 62 of *Pitman Monographs and Surveys in Pure and Applied Mathematics.* Longman Scientific & Technical, Harlow, England, 1992.

[67] E.H. MOORE. Abstract. *Bulletin of the American Mathematical Society*, 26:394–395, 1920.

[68] A. PAZY. Asymptotic behavior of contractions in Hilbert spaces. *Israel Journal of Mathematics*, 9:235–240, 1971.
[69] R. PENROSE. A generalized inverse for matrices. *Proceedings of the Cambridge Philosophical Society*, 51(3):406–413, July 1955.
[70] G. PIERRA. Decomposition through formalization in a product space. *Mathematical Programming*, 28:96–115, 1984.
[71] A.R. DE PIERRO and A.N. IUSEM. On the asymptotic behaviour of some alternate smoothing series expansion iterative methods. *Linear Algebra and its Applications*, 130:3–24, 1990.
[72] S. REICH. A limit theorem for projections. *Linear and Multilinear Algebra*, 13:281–290, 1983.
[73] S.M. ROBINSON. Some continuity properties of polyhedral multifunctions. *Mathematical Programming Study*, 14:206–214, 1981.
[74] R.T. ROCKAFELLAR. Monotone operators and the proximal point algorithm. *SIAM Journal on Control and Optimization*, 14(5):877–898, August 1976.
[75] M.I. SEZAN. An overview of convex projections theory and its applications to image recovery problems. *Ultramicroscopy*, 40:55–67, 1992.
[76] K.T. SMITH, D.C. SOLMON, and S.L. WAGNER. Practical and mathematical aspects of the problem of reconstructing objects from radiographs. *Bulletin of the American Mathematical Society*, 83(6):1227–1270, November 1977.
[77] K. TANABE. Projection method for solving a linear system of linear equations and its applications. *Numerische Mathematik*, 17:203–214, 1971.
[78] M.R. TRUMMER. Rekonstruktion von Bildern aus ihren Projektionen. Diplomarbeit, Seminar für Angewandte Mathematik, ETH Zürich, 1979.
[79] P. TSENG. On the convergence of the products of firmly nonexpansive mappings. *SIAM Journal on Optimization*, 2(3):425–434, August 1992.
[80] M.A. VIERGEVER. Introduction to discrete reconstruction methods in medical imaging. In M.A. Viergever and A. Todd-Pokropek, editors, *Mathematics and Computer Science in Medical Imaging*, pages 43–65, Berlin, 1988. Springer-Verlag. NATO Advanced Science Institute Series F: Computer and Systems Sciences Vol. 39. Proceedings, held in Il Ciocco, Italy, September 21 – October 4, 1986.
[81] J. von NEUMANN. *Functional Operators*, volume II. The Geometry of Orthogonal Spaces. Princeton University Press, 1950. Annals of mathematics studies Vol. 22. Reprint of mimeographed lecture notes first distributed in 1933.
[82] D.C. YOULA. On deterministic convergence of iterations of relaxed projection operators. *Journal of Visual Communication and Image Representation*, 1(1):12–20, September 1990.
[83] D.C. YOULA and V. VELASCO. Extensions of a result on the synthesis of signals in the presence of inconsistent constraints. *IEEE Transactions on Circuits and Systems*, CAS-33(4):465–468, April 1986.
[84] E.H. ZARANTONELLO. Projections on convex sets in Hilbert space and spectral theory. In E.H. Zarantonello, editor, *Contributions to Nonlinear Functional Analysis*, pages 237–424, New York, 1971. Academic Press. University of Wisconsin. Mathematics Research Center; Publication No. 27.

Centre for Experimental and Constructive Mathematics, Department of Mathematics and Statistics, Simon Fraser University, Burnaby, British Columbia V5A 1S6, CANADA
Current address: Department of Combinatorics and Optimization, Faculty of Mathematics, University of Waterloo, Waterloo, Ontario N2L 3G1, CANADA
E-mail address: `bauschke@cecm.sfu.ca`

Centre for Experimental and Constructive Mathematics, Department of Mathematics and Statistics, Simon Fraser University, Burnaby, British Columbia V5A 1S6, CANADA
E-mail address: `jborwein@cecm.sfu.ca`

Department of Combinatorics and Optimization, Faculty of Mathematics, University of Waterloo, Waterloo, Ontario N2L 3G1, CANADA
E-mail address: `aslewis@orion.uwaterloo.ca`