*Set-Valued Analysis* **1:** 185–212, 1993.



# On the Convergence of von Neumann's Alternating Projection Algorithm for Two Sets

H. H. BAUSCHKE
*Department of Mathematics, Statistics and Computing Science, Dalhousie University, Halifax, Nova Scotia, Canada, B3H 3J5*

and

J. M. BORWEIN
*Department of Mathematics and Statistics, Simon Fraser University, Burnaby, V5A 1S6 British Columbia, Canada*

(Received: 1 December 1992)

**Abstract.** We give several unifying results, interpretations, and examples regarding the convergence of the von Neumann alternating projection algorithm for two arbitrary closed convex nonempty subsets of a Hilbert space. Our research is formulated within the framework of Fejér monotonicity, convex and set-valued analysis. We also discuss the case of finitely many sets.

**Mathematics Subject Classifications (1991):** Primary 47H09, 65J05; secondary 41A25, 41A50, 46C99, 47N10, 49M45, 65Kxx, 90C25.

**Key words:** Algorithm, von Neumann's algorithm, method, alternating method, iterative method, projection, cyclic projections, successive projections, Hilbert space, convex sets, linear convergence, norm convergence, weak convergence, open mapping theorem, multifunctions, convex relations, convex feasibility problem, least-squares approximation, angle between two subspaces.

## 1. Introduction and Notation

Suppose $X$ is a Hilbert space and $A$, $B$ are closed convex nonempty subsets. Finding a point in $A \cap B$ or — if $A \cap B$ is empty — a good substitute for it, is a basic problem in various areas of mathematics.

Define the *distance* between two nonempty subsets $M$, $N$ by

$$d(M,N) := \inf \| M - N \| := \inf\{\| m - n \| : \ m \in M, n \in N\}.$$

Clearly, a good generalization of $A \cap B$ is

$$\begin{aligned} E &:= \{a \in A \ : \ d(a,B) = d(A,B)\}, \\ F &:= \{b \in B \ : \ d(b,A) = d(B,A)\}, \end{aligned}$$

because if $A \cap B \neq \emptyset$, then $E = F = A \cap B$. Recall that the *projection* onto a closed convex nonempty subset $C$ sends any point $x$ to its nearest point in $C$, denoted $P_C x$, and is characterized by

*Kolmogorov's criterion*:

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

Define further the *displacement vector* $v$ by

$$v = P_{\overline{B-A}}(0).$$

Given a mapping $Q$ from $X$ to $X$, the set of *fixed points of* $Q$, $\{x \in X : Qx = x\}$, is denoted by Fix $Q$. We then have further information on $E$, $F$ (see [1, Section 2] for proofs):

FACTS 1.1.
(i) $\| v \| = d(A, B), \qquad E + v = F,$
(ii) $E = \mathrm{Fix}(P_A P_B) = A \cap (B - v),$
$F = \mathrm{Fix}(P_B P_A) = B \cap (A + v),$
(iii) $P_B e = P_F e = e + v \ (e \in E),$
$P_A f = P_E f = f - v \ (f \in F).$

The following elegant and highly successful algorithm for finding a point in $A \cap B$ is at least 60 years old:

Given a starting point $x \in X$, define the terms of the sequences $(a_n)$, $(b_n)$ by

$$b_0 := x, \qquad a_n := P_A b_{n-1}, \qquad b_n := P_B a_n$$

for every integer $n \geq 1$.

While developing modern operator theory, von Neumann [25, Theorem 13.7] proved that both sequences converge to $P_{A\cap B}(x)$ in norm *whenever $A$, $B$ are closed subspaces!*

Not surprisingly, the algorithm is called *von Neumann's alternating projection algorithm.* We will refer to the sequences $(a_n)$, $(b_n)$ as *von Neumann sequences* and to the sequence $(b_0, a_1, b_1, a_2, b_2, \cdots)$ as *alternating von Neumann sequence.*

His result is truly remarkable, since the algorithm not only yields the *nearest* point in $A \cap B$, but also converges *in norm!*

What happens in our setting, where $A$, $B$ are arbitrary closed convex (possibly nonintersecting) sets? Well, firstly we cannot expect to find $P_{A\cap B}x$, as simple examples in $\mathbb{R}^2$ show (take e.g. the unit disc and the $x$-axis). Secondly, we do not know if the von Neumann sequences actually converge in norm. However, some positive results are known (see [1, Section 4]):

FACTS 1.2.

(i) $b_n - a_n, b_n - a_{n+1} \longrightarrow v.$

(ii) If $E$, $F$ are empty, then $\| a_n \|, \| b_n \| \longrightarrow \infty$.
If $E$, $F$ are nonempty, then

$$a_n \rightharpoonup e^*, b_n \rightharpoonup f^* = e^* + v$$

for some $e^* E$, $f^* \in F$.

(iii) If $E$, $F$ are nonempty, then the von Neumann sequences converge in norm, whenever
- $A$ or $B$ is locally compact or
- $A$, $B$ are affine closed subspaces.

It is a famous open problem as to whether or not the convergence can actually be only weak!

We would like to mention that Dykstra has invented an ingenious, but more complicated algorithm that either (i) converges in norm to the projection of the starting point onto $E$, $F$ (if $E, F \neq \emptyset$) or (ii) tends to infinity (if $E, F = \emptyset$); see [1] for details and more references.

Nonetheless, von Neumann's algorithm is very much alive, because its simplicity is striking and it has found numerous applications. In fact, it has been rediscovered frequently! We recommend Deutsch's recent survey article [10] as an excellent pointer to the relevant literature.

Here, we provide several results on norm and linear convergence of the von Neumann sequences. The paper is organized as follows:

Section 2 gives a new interpretation of the displacement vector which we exploit for the case where $A$, $B$ are translates of cones.

In Section 3, we regard the von Neumann sequences as *Fejér monotone* sequences; this viewpoint leads to some natural sufficient conditions for norm and linear convergence.

Our main results are in Section 4, where we give an easy-to-verify sufficient condition for linear convergence for the case when $A$, $B$ intersect. In particular, we obtain linear convergence whenever $0 \in \operatorname{int}(A - B)$ or $A$, $B$ are subspaces with closed sum. Variations of these results are known and due to Gubin *et al.* [14] and Deutsch [9]. However, we obtain them here *simultaneously* by using the *Open Mapping Theorem for Convex Relations*; thus we discover an underlying common structure of two ostensibly different looking results. We finish the section by proving that if $A$, $B$ are *affine* subspaces with closed sum, then the von Neumann sequences converge linearly with a rate independent of the starting point.

In Section 5, we discuss *en detail* the case where $A$ is an affine subspace of finite codimension and $B$ is a Hilbert lattice cone — a special case of a moment problem. Those problems appear in certain applications and are widely studied. We give a result on linear convergence under a mild regularity condition.

The last section deals with the case where we have $N$ closed convex nonempty subsets. A (now almost classical) product space formalization used by Pierra [18] allows us to transform the $N$-set problem into a 2-set problem. Thus our results apply and we obtain several linear convergence results relating to or improving theorems by Pierra [18] and Gubin *et al.* [14].

Throughout the paper, we make use of the definitions and facts listed above as well as the following:

If $C$ is a closed convex nonempty subset and $z \in X$, then the projection onto the translate $z + C$ is given by $P_{z+C}(x) = z + P_C(x - z)$ for any $x \in X$. The projection $P_C$ is also *firmly nonexpansive*, i.e.

$$\| P_C x - P_C y \|^2 + \| (\mathrm{Id} - P_C)x - (\mathrm{Id} - P_C)y \|^2 \le \| x - y \|^2$$

for every $x, y \in X$; so in particular it is *Lipschitz continuous with constant* 1, or equivalently, *nonexpansive*.

We denote the norm dual space of a Banach space $X$ by $X^*$. If $M$, $Z$ are any subsets of $X$, then $\overline{\mathrm{conv}}M$, $\overline{M}$, $\mathrm{int}_Z M$, icore $M$ denotes the *closed convex hull* of $M$, the *closure* of $M$, the *interior* of $M$ w.r.t. $Z$, the *intrinsic core* of $M (= \mathrm{int}_{\overline{\mathrm{aff}}M} M$, where $\overline{\mathrm{aff}}M$ is the *closed affine span* of $M$), respectively. If $\bar{x} \in X$, $\bar{r} > 0$, then define

$$B(\bar{x}, \bar{r}) := \{x \in X \ : \| x - \bar{x} \| \le \bar{r}\} \quad \text{and} \quad B_X := B(0, 1).$$

Suppose $S$ is a *cone*, i.e. a convex nonempty subset closed under nonnegative scalar multiplication. Then we define $S^{\oplus} := \{x^* \in X^* \ : \ x^*(S) \ge 0\}$ to be the *positive dual cone*; analogously, $S^{\ominus} := \{x^* \in X^* \ : \ x^*(S) \le 0\}$ is the *negative dual cone*. Of course, if $S$ is a subspace, then $S^{\oplus} = S^{\ominus} = S^{\perp}$, where $S^{\perp}$ is the *orthogonal complement* of $S$. If $S = S^{\oplus}$, then we say that $S$ is *self-adjoint*. If $S - S = X$, then $S$ is *generating*. If $X$ is a Hilbert space vector lattice, then the *positive cone* is denoted by $X^+$. Should $(X, X^+)$ actually be a *Hilbert lattice*, i.e. a Banach lattice in the Hilbert space norm, then $X^+$ is self-adjoint. This actually *characterizes* Hilbert lattices; see [7, Theorem 8] for a proof and more.

In the framework of convex functions (see, e.g., [11, 16, 20, 24]), we use lsc for *lower semi-continuous*, $\iota_M$ for the *indicator function* of a subset $M$ ($\iota_M(x) := 0$, if $x \in M$, $+\infty$ otherwise), $f_1 \Box f_2$ for the *infimal convolution* of two convex functions $f_1, f_2$

$$(f_1 \Box f_2)(x) := \inf\{f_1(x_1) + f_2(x_2) \ : \ x_1 + x_2 = x\}.$$

If $f$ is a convex function, then $\partial f$ denotes the *subdifferential* of $f$

$$\partial f(x) := \{x^* \in X^* \ : \ \langle x^*, h \rangle \le f(x + h) - f(x) \quad \text{for all} \quad h\},$$

$f^*$ denotes the *convex conjugate* of $f$

$$f^*(x^*) := \sup\{\langle x^*, x \rangle - f(x) \ : \ x \in X\};$$

if $g$ is a concave function, then $g_*$ denotes the *concave conjugate* of $g$

$$g_*(x^*) := \inf\{\langle x^*, x\rangle - g(x) \ : \ x \in X\}.$$

A sequence $(x_n)_{n\geq 0}$ in $X$ is said to *converge linearly with rate* $\kappa < 1$, if there is some constant $\beta$ s.t.

$$\| \ x_n - x \ \| \leq \beta\kappa^n$$

for all $n \geq 0$ and $x := \lim x_n$.

Finally, we will see a few times the quantifiers $\forall$ (= for all) and $\exists$ (= there exists) although we usually oppose the use of quantifiers. We feel, however, it is justified whenever the corresponding statement in plain language becomes a monster.

## 2. On the Displacement Vector

Viewing the distance between two sets as a convex program is very instructive: We will see that the displacement vector is nothing but the (unique) solution of the dual program. We define $f := \frac{1}{2}d^2(\cdot, A)$, $g := -\iota_B$, then we have the primal program

$$\text{(P)} \quad \mu := \inf_{x\in X} f(x) - g(x) = \frac{1}{2}d^2(A, B),$$

so that the corresponding Fenchel dual program is

$$\text{(D)} \quad \sigma := \sup_{x^*\in X^*} g_*(x^*) - f^*(x^*).$$

Since $g$ is proper and $f$ is continuous everywhere, we have *strong duality*, i.e. $\sigma = \mu$ and $\sigma$ is attained [24, Theorem 7.15].

THEOREM 2.1. *The displacement vector $v$ is the unique solution of the dual program* (D).

*Proof.* Clearly, $f = 1/2 \ \| \cdot \|^2 \ \Box \iota_A$, so by [24, Example 6.4.(a), Theorem 6.6] $f^* = \frac{1}{2} \ \| \cdot \|^2 + \iota_A^*$ and also $-g_* = \iota_{-B}^*$. Now

$$\text{(D)} \quad \sigma := -\min_{x^*\in X^*}(f^*(x^*) - g_*(x^*)) = -\min_{x^*\in X^*}\left(\frac{1}{2} \ \| \cdot \|^2 + \iota_A^* + \iota_{-B}^*\right);$$

thus if $x^*$ solves (D), then

$$0 \in \partial\left(\frac{1}{2} \ \| \cdot \|^2 + \iota_A^* + \iota_{-B}^*\right)(x^*).$$

But $\frac{1}{2} \ \| \cdot \|^2$ is continuous everywhere and $\iota_A^* + \iota_{-B}^*$ is proper (0 is in its domain), so the *sum formula* [24, Theorem 5.38] applies and yields

$$-x^* \in \partial(\iota_A^* + \iota_{-B}^*)(x^*).$$

Letting $h := \iota_A^* + \iota_{-B}^* = (\iota_A \Box \iota_{-B})^*$, we see that $h$ is a conjugate function, hence lsc. Thus $h = h^{**}$ by *Hörmander's Theorem* [24, Theorem 6.18] and

$$-x^* \in \partial h(x^*) \Longleftrightarrow x^* \in \partial h^*(-x^*)$$

by [11, Corollary 5.2]. It is easy to check that $\iota_A \Box \iota_{-B} = \iota_{A-B}$, hence, using [24, Theorem 6.20. (e)],

$$h^* = (\iota_A \Box \iota_{-B})^{**} = \iota_{A-B}^{**} = \iota_{\overline{\text{conv}}(A-B)} = \iota_{\overline{A-B}}.$$

Therefore,

$$\begin{aligned} x^* \in \partial h^*(-x^*) &\Longleftrightarrow x^* \in \partial \iota_{\overline{A-B}}(-x^*) \\ &\Longleftrightarrow x^* \in \overline{B-A} \quad \text{and} \quad \langle x^*, A-B+x^*\rangle \leq 0 \\ &\Longleftrightarrow x^* = P_{\overline{B-A}}(0) = v. \end{aligned}$$

□

If the sets $A$, $B$ are translated cones, then we are able to obtain an alternative description of the displacement vector:

EXAMPLE 2.2. *If $A$, $B$ are translated closed cones, say $A = a + K$, $B = b + L$, where $a, b \in X$ and $K$, $L$ are closed cones, then*

$$v = P_{K^\ominus \cap L^\oplus}(b-a).$$

*If furthermore $K$, $L$ are subspaces, then*

$$v = P_{K^\perp \cap L^\perp}(b-a).$$

*Proof.* It is easy to verify that if $C$ is a closed cone, then $\iota_C^* = \iota_{C^\ominus}$. Thus,

$$f^*(x^*) = \frac{1}{2} \parallel x^* \parallel^2 + \iota_{a+K}^*(x^*) = \frac{1}{2} \parallel x^* \parallel^2 + \langle x^*, a\rangle + \iota_{K^\ominus}(x^*)$$

and

$$g_*(x^*) = \langle x^*, b\rangle - \iota_{L^\oplus}(x^*).$$

Hence

$$\begin{aligned} \sigma &= \max_{x^* \in X^*} (g_*(x^*) - f^*(x^*)) = - \min_{x^* \in X^*} (f^*(x^*) - g_*(x^*)) \\ &= - \min_{x^* \in X^*} \left(\tfrac{1}{2} \parallel x^* \parallel^2 + \langle x^*, a\rangle + \iota_{K^\ominus}(x^*) - \langle x^*, b\rangle + \iota_{L^\oplus}(x^*)\right) \\ &= - \min_{x^* \in X^*} \left(\iota_{K^\ominus \cap L^\oplus}(x^*) + \tfrac{1}{2} \parallel x^* + (a-b) \parallel^2 - \tfrac{1}{2} \parallel a - b \parallel^2\right) \\ &= \tfrac{1}{2} \parallel a - b \parallel^2 - \min_{x^* \in K^\ominus \cap L^\oplus} \tfrac{1}{2} \parallel (b-a) - x^* \parallel^2; \end{aligned}$$

therefore, the unique solution of the dual is

$$x^* = P_{K^\ominus \cap L^\oplus}(b-a).$$

By the last theorem, the proof is complete. □

## 3. Fejér Monotonicity and Regularity

DEFINITION 3.1. Given a closed convex nonempty set $C$, a sequence $(x_n)_{n\geq 0}$ is called *Fejér monotone w.r.t.* $C$ if

$$\| x_{n+1} - c \| \leq \| x_n - c \|$$

for all $n \geq 0$ and $c \in C$.

This notion goes back at least to [17, Section 4]. The nonexpansivity of projections immediately implies the following:

EXAMPLE 3.2. *The von Neumann sequence* $(a_n)$ *(resp.* $(b_n)$*) is Fejér monotone w.r.t.* $E$ *(resp.* $F$*). If* $A \cap B \neq \emptyset$*, then the alternating von Neumann sequence is Fejér monotone w.r.t.* $A \cap B$.

The next theorem summarizes basic properties of Fejér monotone sequences. Most of them are well-known; property (iv) is due to Gubin *et al.* [14].

THEOREM 3.3. *Suppose* $(x_n)_{n\geq 0}$ *is Fejér monotone w.r.t.* $C$*. Then*

(i) $(x_n)$ *is bounded and* $d(x_{n+1}, C) \leq d(x_n, C)$.

(ii) *If* $\operatorname{int}(C) \neq \emptyset$*, then* $(x_n)$ *has at most one norm cluster point.*

(iii) $(x_n)$ *converges weakly to some point in* $C$ *if and only if all weak cluster points lie in* $C$.

(iv) $(x_n)$ *converges in norm to some point in* $C$*, say* $x$*, if and only if* $d(x_n, C)$ *tends to* 0*. In this case,*

$$\| x_n - x \| \leq 2d(x_n, C)$$

*for every nonnegative integer* $n$.

(v) *If there is some* $\alpha > 0$ *s.t. for every nonnegative integer* $n$

$$\alpha d^2(x_n, C) \leq d^2(x_n, C) - d^2(x_{n+1}, C),$$

*then there is some point in* $C$*, say* $x$*, s.t.*

$$\| x_n - x \| \leq 2(1-\alpha)^{n/2} d(x_0, C),$$

*i.e.* $(x_n)$ *converges linearly to* $x$ *with rate* $\sqrt{1-\alpha}$.

*Proof.* (i) is trivial, (iii) is proved exactly as [1, Lemma 4.7]. (iv) follows from [14, Proof of Lemma 6] and implies (v). We prove (ii): We may assume that $B_X \subseteq C$ (after translation of $C$ and scaling the norm, if necessary). Suppose $y_1$, $y_2$

are two norm cluster points of $(x_n)$. Then $\| y_1 - b' \| = \| y_2 - b' \|$ for all $b' \in B_X$. Denoting the projection onto $B_X$ by $P$, we conclude

$$\| y_1 - Py_1 \| = \| y_2 - Py_1 \| \geq \| y_2 - Py_2 \| = \| y_1 - Py_2 \|,$$

hence $Py_1 = Py_2 =: b$. In particular,

$$y_1, y_2 \in P^{-1}(b) \qquad \text{and} \qquad \| y_1 - b \| = \| y_2 - b \| =: r.$$

However,

$$P^{-1}(b) = \begin{cases} \{b\}, & \text{if} \quad \| b \| < 1, \\ \{tb : t \geq 1\}, & \text{if} \quad \| b \| = 1; \end{cases}$$

so $P^{-1}(b)$ contains at most one point with distance $r$ to $b$. Hence norm cluster points are unique whenever they exist. □

An immediate consequence of (ii) is:

COROLLARY 3.4. *If $X$ is finite-dimensional and* $\mathrm{int}(A \cap B) \neq \emptyset$, *then the alternating von Neumann sequence converges in norm.*

REMARK 3.5. Of course, we know this already from Facts 1.2 (iii). Nevertheless, it indicates that the interiority condition $\mathrm{int}(A \cap B) \neq \emptyset$ forces 'good' behaviour of the von Neumann sequences. This is indeed true, as we will see in the next section.

Theorem 3.3 gives us sufficient conditions for norm or even linear convergence of Fejér sequences. The program is now clear: We must adjust those conditions to our setting. We start with a sufficient condition for norm convergence.

DEFINITION 3.6. We say that $(A, B)$ is *boundedly regular* if

$$\underset{X \supseteq S \text{ bounded}}{\forall} \quad \underset{\epsilon > 0}{\forall} \quad \underset{\delta > 0}{\exists} \quad \underset{\substack{x \in S: \\ \max\{d(x,A),\, d(x, B - v)\} \leq \delta}}{\forall} \quad d(x, E) \leq \epsilon.$$

Loosely speaking, we can say that if a point is close to $A$ and to $B - v$, then it cannot be too far away from $A \cap (B - v) = E$. Note that if $(A, B)$ is boundedly regular, then $E, F \neq \emptyset$, since $\inf \emptyset = \infty$! In some sense, the definition of bounded regularity is symmetric in $A$ and $B$, because it is equivalent to

$$\underset{X \supseteq T \text{ bounded}}{\forall} \quad \underset{\epsilon > 0}{\forall} \quad \underset{\delta > 0}{\exists} \quad \underset{\substack{y \in T \\ \max\{d(y,B),\, d(y, A + v)\} \leq \delta}}{\forall} \quad d(y, F) \leq \epsilon.$$

THEOREM 3.7. *If $(A, B)$ is boundedly regular, then the von Neumann sequences converge in norm.*

*Proof.* The sequence $(a_n)$ is Fejér monotone w.r.t. $E$, so by Theorem 3.3 (iv) it is enough to show that $d(a_n, E)$ tends to 0. Now $(a_n)$ lies in $A$ and

$$d(a_n, B - v) \leq \| a_n - (b_n - v) \| \to 0$$

by Facts 1.2(i). Thus the sequence $\max\{d(a_n, A), d(a_n, B - v)\}$ tends to 0. Consequently, the sequence $(a_n)$ converges in norm to some point in $E$, because $(a_n)$ is bounded and $(A, B)$ is boundedly regular. Also, $(b_n)$ converges in norm by observing that either the definition of bounded regularity is 'symmetric' in $A$ and $B$ or that $(b_n)$ is the sum of two norm convergent sequences, namely $(b_n - a_n)$ and $(a_n)$. □

REMARK 3.8. It should be noted that not all pairs $(A, B)$ are boundedly regular; see Example 5.5, which shows that bounded regularity is sufficient but not necessary for norm convergence of the von Neumann sequences!

THEOREM 3.9. *If $A$ or $B$ is boundedly compact, then $(A, B)$ is boundedly regular.*

*Proof.* Without loss, assume $A$ is boundedly compact. Suppose to the contrary that $(A, B)$ is not boundedly regular. Then

$$\underset{X \supseteq S \text{ bounded}}{\exists} \quad \underset{\epsilon > 0}{\exists} \quad \underset{\delta > 0}{\forall} \quad \underset{\substack{x \in S: \\ \max\{d(x,A),\, d(x,B-v)\} \leq \delta}}{\exists} \quad d(x, E) > \epsilon.$$

In particular, we obtain a bounded sequence $(x_n)$ s.t.

$$x_n - P_A x_n, \qquad x_n - P_{B-v} x_n \to 0 \qquad \text{and} \qquad d(x_n, E) > \epsilon.$$

After passing to a subsequence if necessary, we may assume that $(x_n)$ converges weakly to some point $\bar{x}$. Because distance functions are weakly lsc and $d(x_n, A)$, $d(x_n, B - v) \to 0$, we conclude (with Facts 1.1 (ii)) that $\bar{x} \in A \cap (B - v) = E$. Hence, the sequence $(P_A x_n)$ converges weakly to $\bar{x}$ and has (by bounded compactness) *at least* one norm cluster point. On the other hand, every weakly convergent sequence has *at most* one norm cluster point. Altogether, $(P_A x_n)$ converges to $\bar{x}$ in norm. Therefore, $(x_n)$ converges to $\bar{x}$ in norm implying $d(x_n, E) \to 0$, a contradiction to $d(x_n, E) > \epsilon$. □

COROLLARY 3.10. *If $A$ or $B$ is boundedly compact, then the von Neumann sequences converge in norm.*

Again, the last corollary is well known (see, e.g., Facts 1.2 (iii)), but we deduced it here directly from the geometrical property of bounded regularity. Now we formulate a sufficient condition for linear convergence.

DEFINITION 3.11. We say that $(A, B)$ is *boundedly linearly regular* if

$$\underset{X \supseteq S \text{ bounded}}{\forall} \quad \underset{\kappa > 0}{\exists} \quad \underset{x \in S}{\forall} \quad d(x, E) \leq \kappa \max\{d(x, A), d(x, B - v)\}.$$

It is obvious that bounded linear regularity implies bounded regularity. Again we have 'symmetry' in the sense that the definition is equivalent to

$$\underset{X \supseteq T \text{ bounded}}{\forall} \quad \underset{\kappa > 0}{\exists} \quad \underset{y \in T}{\forall} \quad d(y, F) \leq \kappa \max\{d(y, B), d(y, A + v)\}.$$

THEOREM 3.12. *If $(A, B)$ is boundedly linearly regular, then the von Neumann sequences converge linearly.*

*Proof.* Since the sequence $(a_n)$ is bounded, we obtain $\kappa > 0$ s.t.

$$d(a_n, E) \leq \kappa \max\{d(a_n, A), d(a_n, B - v)\} = \kappa d(a_n + v, B).$$

Fix an arbitrary $e \in E$. Because $P_B$ is firmly nonexpansive, $P_B e = e + v$ and Fix $(P_A P_B) = E$ (see Facts 1.1), we estimate

$$\begin{aligned}
d^2(a_n + v, B) &\leq \| a_n + v - b_n \|^2 = \| (a_n - e) - (P_B a_n - P_B e) \|^2 \\
&\leq \| a_n - e \|^2 - \| P_B a_n - P_B e \|^2 \\
&\leq (\| a_n - e \|^2 - \| P_B a_n - P_B e \|^2) + \\
&\quad + (\| P_B a_n - P_B e \|^2 - \| P_A P_B a_n - P_A P_B e \|^2) \\
&= \| a_n - e \|^2 - \| a_{n+1} - e \|^2
\end{aligned}$$

for all $n \geq 0$. In particular, if $e = P_E a_n$, we obtain

$$\begin{aligned}
\frac{1}{k^2} d^2(a_n, E) &\leq \| a_n - P_E a_n \|^2 - \| a_{n+1} - P_E a_n \|^2 \\
&\leq d^2(a_n, E) - d^2(a_{n+1}, E).
\end{aligned}$$

By Theorem 3.3 (v), the sequence $(a_n)$ converges linearly. By symmetry, the sequence $(b_n)$ also converges linearly. □

The reader will note that the rate of convergence possibly depends on the starting point. For applications, one is more interested in results on linear convergence with a rate independent of the starting point. This can be achieved by requiring an even stronger regularity condition. So let us define a stronger, global version of bounded regularity and of linear bounded regularity.

DEFINITION 3.13. We say that $(A, B)$ is *regular* if

$$\underset{\epsilon > 0}{\forall} \quad \underset{\delta > 0}{\exists} \quad \underset{\substack{x \in X : \\ \max\{d(x, A), d(x, B - v)\} \leq \delta}}{\forall} \quad d(x, E) \leq \epsilon.$$

Analogously, we say that $(A, B)$ is *linearly regular with rate* $\kappa$, if

$$\underset{x \in X}{\forall} \quad d(x, E) \leq \kappa \max\{d(x, A), d(x, B - v)\}.$$

Note that we have once more 'symmetry' and that these regularities imply their corresponding bounded versions. The proof of the last theorem in combination with Theorem 3.3 (v) immediately leads to the following corollary.

COROLLARY 3.14. *If* $(A, B)$ *is linearly regular with rate* $\kappa$*, then the von Neumann sequences converge linearly with rate*

$$\sqrt{1 - \frac{1}{\kappa^2}}.$$

Next, we explore some relations among the various 'regularities'.

THEOREM 3.15. If $(A, B)$ is boundedly regular and $A$ or $B$ is bounded, then $(A, B)$ is regular.

*Proof.* Without loss, assume the set $A$ is bounded. Then so is $S := A + B_X$. Given a fixed $\epsilon > 0$, we get $\delta' > 0$ s.t.

$$\underset{\substack{x \in S: \\ \max\{d(x, A), d(x, B - v)\} \leq \delta'}}{\forall} \qquad d(x, E) \leq \epsilon.$$

If we let $\delta := \min\{1, \delta'\}$, then it is easy to see that this $\delta$ does not required job (in the definition of regularity). □

EXAMPLE 3.16. *Suppose*

$$X = \mathbb{R}^2, \qquad \rho > 0, \qquad A := \{(x, y) \in X : y = \rho\}, \qquad B := B_X.$$

*By* [1, Example 5.3], *the von Neumann sequence* $(a_n)$ *converges linearly if and only if* $\rho \neq 1$*. However,* $(A, B)$ *is regular for all* $\rho > 0$*. Hence, norm convergence is the best we can hope for when the sets are regular! We now show that* $(A, B)$ *is not boundedly linearly for* $\rho \geq 1$*. Indeed,*

$$F = \{(0, 1)\}, \qquad d((x, y), F) = \sqrt{x^2 + (y - 1)^2}$$

*and*

$$\begin{aligned} d((x, y), A + v) &= |y - 1|, \\ d((x, y), B) &= \begin{cases} 0, & \text{if } x^2 + y^2 \leq 1, \\ \sqrt{x^2 + y^2} - 1, & \text{otherwise.} \end{cases} \end{aligned}$$

*Consider* $(x,1)$ *for* $x > 0$. *Clearly,* $(x,1) \in (A+v)\backslash B$. *Hence*

$$\max\{d((x,1),B), d((x,1),A+v)\} = \sqrt{x^2+1}-1, \quad d((x,1),F) = x.$$

*However,*

$$\frac{x}{\sqrt{x^2+1}-1} \to \infty \ (x \to 0),$$

*so* $(A,B)$ *is not boundedly linearly regular. Hence, bounded linear regularity is a sufficient but not a necessary condition for linear convergence of a von Neumann sequence. Also, regularity can occur when bounded linear regularity does not.*

For cones, however, we can say more:

THEOREM 3.17. *If* $A$, $B$ *are nonempty closed convex cones, then TFAE:*
(i) $(A,B)$ *is regular.*
(ii) $(A,B)$ *is linearly regular.*
(iii) $(A,B)$ *is boundedly linearly regular.*

*Proof.* Clearly, (ii) implies (i) and (iii). Let us prove that (i) implies (ii)! For $\epsilon = 1$, we get $\delta > 0$ s.t. $d(x, A\cap B) \leq 1$ whenever

$$m(x) := \max\{d(x,A), d(x,B)\} \leq \delta.$$

Pick $y \in X\backslash(A\cap B)$ and let $x := (\delta/m(y))y$. Since projections onto cones are positively homogeneous, we conclude $m(x) = \delta$ and further

$$d(x, A\cap B) = \frac{\delta}{m(y)} d(y, A\cap B) \leq 1 \ \text{ so } \ d(y, A\cap B) \leq \frac{1}{\delta} m(y).$$

But the last inequality is true for all $y \in X$ and therefore (ii) holds. It remains to show that (iii) implies (ii). For $B_X$, we get $\kappa > 0$ s.t.

$$d(x, A\cap B) \leq \kappa \max\{d(x,A), d(x,B)\}$$

for all $x \in B_X$. Using the positive homogeneity of the projection onto a cone once more, we see that $(A,B)$ is linearly regular with rate $\kappa$. □

REMARK 3.18. Franchetti and Light [12, Section 4] constructed two subspaces with nonclosed sum and an alternating von Neumann sequence which does not converge linearly. By Corollary 3.14 and Theorem 3.17, this pair of subspaces is neither regular nor boundedly linearly regular. We do not know if it is boundedly regular. Their examples also shows that regularity is only sufficient but not necessary for norm convergence of the von Neumann sequences.

## 4. Regularity and the Open Mapping Theorem

LEMMA 4.1. *Suppose $M$, $N$ are closed convex sets with nonempty intersection. If*

$$\underset{\bar{x}\in M}{\exists}\quad \underset{\kappa,\bar{r}>0}{\exists}\quad \underset{x\in M\cap B(\bar{x},2\bar{r})}{\forall}\quad d(x,M\cap N)\leq \kappa d(x,N),$$

*then*

$$\underset{x\in B(\bar{x},\bar{r})}{\forall}\quad d(x,M\cap N)\leq (2\kappa+1)\max\{d(x,M),d(x,N)\}.$$

*Proof. Claim 1:* For every $x \in X$, we have

$$d(x,M\cap N)\leq (2\kappa+1)\max\{d(x,M\cap B(\bar{x},2\bar{r})),d(x,N)\}.$$

Indeed,

$$f(x):=\kappa\max\{d(x,M),d(x,N)\}-d(x,M\cap N)$$

is $(\kappa+1)$–Lipschitz on $X$ and, hence, for any $x \in X$

$$(\kappa+1)d(x,M\cap B(\bar{x},2\bar{r}))+f(x)\geq \inf f(M\cap B(\bar{x},2\bar{r})),$$

implying

$$\begin{aligned} &d(x,M\cap N)\\ &\quad\leq (\kappa+1)d(x,M\cap B(\bar{x},2\bar{r}))+\kappa\max\{d(x,M\cap B(\bar{x},2\bar{r})),d(x,N)\}\end{aligned}$$

and Claim 1 is verified.

*Claim 2:* $d(x,M)=d(x,M\cap B(\bar{x},2\bar{r}))$ for every $x\in B(\bar{x},\bar{r})$.

Clearly,

$$\| x-P_Mx \| \leq \| x-\bar{x} \| \leq \bar{r},$$

so

$$\| P_Mx-\bar{x} \| \leq \| P_Mx-x \| + \| x-\bar{x} \| \leq 2\bar{r}$$

and, hence, $P_Mx\in M\cap B(\bar{x},2\bar{r})$ implying

$$d(x,M\cap B(\bar{x},2\bar{r}))\leq d(x,M).$$

Thus, Claim 2 is verified, since obviously $d(\cdot,M)\leq d(\cdot,M\cap B(\bar{x},2\bar{r}))$.

Both claims together finish the proof. □

The next lemma is a very special version of an inversion theorem for convex multifunctions given by Robinson [19, Theorem 2]. Since the proof is short, we include it here.

LEMMA 4.2. *Suppose $X$, $Y$ are Hilbert spaces and $T : X \to Y$ is linear, continuous. Suppose, further, that $M$ is a closed convex subset of $X$ and $N$ is a closed convex subset of $Y$. Suppose, finally,*

$$\underset{\bar{x} \in M \cap T^{-1}(N)}{\exists} \quad \underset{\eta > 0}{\exists} \quad \underset{x \in M \backslash T^{-1}(N)}{\forall}$$
$$\eta \, \frac{P_N Tx - Tx}{\| P_N Tx - Tx \|} \in T(B(\bar{x}, 1) \cap M) - N.$$

*Then*

$$\text{(i)} \quad \underset{y \in TM}{\forall} \quad d(y, (TM) \cap N) \leq \frac{\| T\bar{x} - y \| + \| T \|}{\eta} d(y, N).$$

*If, additionally, $N \subseteq$ range $T$, then*

$$\text{(ii)} \quad \underset{x \in M}{\forall} \quad d(x, M \cap T^{-1}(N)) \leq \frac{\| T \|}{\eta}(1 + \| x - \bar{x} \|) d(x, T^{-1}(N)).$$

*Proof.* Without loss, assume $x \notin T^{-1}(N)$ and $y = Tx$. Select $m \in M \cap B(\bar{x}, 1)$, $n \in N$ s.t.

$$\eta \frac{P_N Tx - Tx}{\| P_N Tx - Tx \|} = Tm - n,$$

and define

$$\lambda := \frac{\| P_N Tx - Tx \|}{\| P_N Tx - Tx \| + \eta} \in ]0, 1[, \bar{m} := (1 - \lambda)x + \lambda m \in M.$$

Then, one checks that

$$T\bar{m} = (1 - \lambda)Tx + \lambda Tm = (1 - \lambda)P_N Tx + \lambda n \in N.$$

Hence, $\bar{m} \in M \cap T^{-1}(N)$, $T\bar{m} \in TM \cap N$. Thus (i) is proved by

$$\begin{aligned}
&d(y, (TM) \cap N) \\
&\quad \leq \| y - T\bar{m} \| = \lambda \| Tx - Tm \| \leq \lambda(\| Tx - T\bar{x} \| + \| T\bar{x} - Tm \|) \\
&\quad \leq \lambda(\| Tx - T\bar{x} \| + \| T \|) \\
&\quad \leq \frac{\| P_N Tx - Tx \|}{\eta}(\| Tx - T\bar{x} \| + \| T \|) \\
&\quad = \frac{\| Tx - T\bar{x} \| + \| T \|}{\eta} d(Tx, N) \\
&\quad = \frac{\| y - T\bar{x} \| + \| T \|}{\eta} d(y, N).
\end{aligned}$$

We always have

$$\begin{aligned}
&d(x, M \cap T^{-1}(N)) \\
&\quad \leq \| x - \bar{m} \| = \lambda \| x - m \| \leq \lambda(\| x - \bar{x} \| + \| \bar{x} - m \|) \\
&\quad \leq (1 + \| x - \bar{x} \|)\lambda \\
&\quad \leq (1 + \| x - \bar{x} \|)\frac{\| P_N T x - T x \|}{\eta} \\
&\quad = \frac{1 + \| x - \bar{x} \|}{\eta} d(Tx, N).
\end{aligned}$$

If range $T \supseteq N$, then

$$d(Tx, N) \leq \| T \| \, d(x, T^{-1}(N)).$$

Therefore, (ii) follows. □

THEOREM 4.3. *Suppose $X$, $Y$ are Hilbert spaces and $T : X \to Y$ is linear, continuous. Suppose furthermore $M$ is a closed convex subset of $X$, $N$ is a closed convex subset of $Y$. Suppose finally* $0 \in$ icore $(TM - N)$. *Then:*

(i) $(TM, N)$ *is boundedly linearly regular.*

*If, additionally,* $N \subseteq$ range $T$, *then*

(ii) $(M, T^{-1}(N))$ *is boundedly linearly regular.*

*Proof.* Consider

$$\Omega : X \rightrightarrows Z := \overline{\text{span}}(TM - N) : x \mapsto \begin{cases} Tx - N, & x \in M, \\ \emptyset, & \text{else.} \end{cases}$$

Then $\Omega$ is a closed convex relation with range $\Omega = TM - N$. The interiority assumption together with the *Open Mapping Theorem* (see [2, Theorem 8(c)]) imply that $\Omega$ is open at 0. Thus, we can pick any $\bar{x} \in \Omega^{-1}(0) = M \cap T^{-1}(N)$ and obtain

$$0 \in \text{int}_Z \Omega(B(\bar{x}, 1)) = \text{int}_Z \Omega(T(B(\bar{x}, 1) \cap M) - N).$$

Hence, there is some constant $\eta > 0$ s.t. $\eta B_Z \subseteq T(B(\bar{x}, 1) \cap M) - N$. Applying Lemma 4.2. (i) yields

$$\forall_{y \in TM} \quad d(y, (TM) \cap N) \leq \frac{\| T\bar{x} - y \| + \| T \|}{\eta} d(y, N).$$

If $\bar{r}$ is any positive number, then

$$\forall_{y \in TM \cap B(T\bar{x}, 2\bar{r})} \quad d(y, (TM) \cap N) \leq \frac{2\bar{r} + \| T \|}{\eta} d(y, N).$$

Therefore, by Lemma 4.1,

$$\begin{aligned}\forall_{y \in B(T\bar{x},\bar{r})} \quad & d(y,(TM)\cap N) \\ & \leq \left(2\frac{2\bar{r}+\parallel T \parallel}{\eta}+1\right)\max\{d(y,TM),d(y,N)\}.\end{aligned}$$

So (i) is verified. By using Lemma 4.2.(ii) we can prove (ii) similarly. □

REMARK 4.4. Borrowing notations from the last proof, we see that the convex relation $\Omega$ is regular at $(\bar{x}, 0)$ in the sense of Borwein [3, Corollary 4.1] and that 0 is a regular value of $\Omega$ in the sense of Robinson [19, Section 2]. We hope that the reader may accept this as an a posteriori justification for choosing the notations 'regularity' in the definitions of Section 3.

The last theorem gives a nice sufficient condition for linear convergence when $A$, $B$ have nonempty intersection:

COROLLARY 4.5. *If* $0 \in \operatorname{icore}(A-B)$, *then* $(A, B)$ *is boundedly linearly regular and the von Neumann sequences converge linearly. In particular, this happens whenever*

(i) $0 \in \operatorname{int}(A-B)$ *or*

(ii) $A-B$ *is a closed subspace.*

*Proof.* Apply the last theorem with $Y = X, T = \mathrm{Id}$, $M = A$ and $N = B$; then $(A, B)$ is boundedly linearly regular and the result follows from Theorem 3.12. □

In view of Corollary 3.14 and Theorem 3.17, we also obtain the following corollary.

COROLLARY 4.6. *If $A$, $B$ are two closed subspaces with closed sum, then the alternating von Neumann sequence converges linearly with a rate independent of the starting point.*

REMARKS 4.7.

- We emphasize that the last theorem and its corollaries apply only to the case where $A$, $B$ have nonempty intersection. For if $v \neq 0$, then interiority arguments will not work since the displacement vector $v$ is in the boundary of $B - A$.
- Corollary 4.5.(i) strengthens a result by Gubin *et al.* [14, Theorem 1.(a)] who — in the case of two sets — assumed the stronger

$$(\mathrm{intA} \cap \mathrm{B}) \cup (\mathrm{intB} \cap \mathrm{A}) \neq \emptyset.$$

- Corollary 4.6 is in the flavour of a result by Deutsch [9, Theorem 2.3], who is even able to specify the rate of convergence in terms of the *angle* between $A$ and $B$. We also known that Corollary 4.6 is sharp in the sense that linear

convergence can fail if the sum of the two subspaces is not closed, see again Franchetti and Light's example [12, Section 4].
- It seems to be new that these results can be obtained simultaneously using techniques from set-valued analysis!

For the remaining part of this section, we assume that $A$, $B$ are closed affine subspaces, say $A = a + K$, $B = b + L$ for vectors $a, b \in X$ and closed subspaces $K$, $L$. The following definition dates back to Friedricks (1937) [13]:

DEFINITION 4.8. The *angle* between $K$ and $L$ is defined to be the angle $\gamma(K, L)$ between 0 and $\pi/2$ whose cosine is given by

$$\cos\gamma := \sup\{\langle k, l\rangle : k \in B_X \cap K \cap (K \cap L)^{\perp}, l \in B_X \cap L \cap (K \cap L)^{\perp}\}.$$

Deutsch [9, Lemma 2.5.(4)] proved the following nice result regarding the angle:

FACT 4.9.

$$\gamma(K, L) > 0 \Longleftrightarrow [K \cap (K \cap L)^{\perp}] + [L \cap (K \cap L)^{\perp}] \textit{ is closed.}$$

Fortunately, the latter condition has a nicer equivalent formulation; its proof is due to A. Simonič [22]:

LEMMA 4.10.

$$[K \cap (K \cap L)^{\perp}] + [L \cap (K \cap L)^{\perp}] \textit{ is closed} \Longleftrightarrow K + L \textit{ is closed.}$$

*Proof.* Denote $[K \cap (K \cap L)^{\perp}] + [L \cap (K \cap L)^{\perp}]$ by $S$.

$$\textit{Claim 1}: \quad S = (K + L) \cap (K \cap L)^{\perp}.$$

'$\subseteq$': is trivial. Conversely, fix $x \in (K + L) \cap (K \cap L)^{\perp}$. Then $x = k + l$ for some vectors $k \in K$, $l \in L$ and

$$x = (\mathrm{Id} - P_{K\cap L})k + (\mathrm{Id} - P_{K\cap L})l,$$

since $x \in (K \cap L)^{\perp}$. This shows '$\supseteq$'. Claim 1 is verified and immediately implies '$\Longleftarrow$'. Similarly, we see

$$\textit{Claim 2}: \quad K + L = [(K + L) \cap (K \cap L)^{\perp}] + (K \cap L).$$

These claims together imply

$$K + L = S \oplus (K \cap L).$$

Since $K + L$ is the *orthogonal* sum of $S$ and $K \cap L$, we conclude that $K + L$ is closed whenever $S$ is. Thus '$\Longrightarrow$' holds. $\square$

The set up is complete. With the help of some other results by Deutsch [9], we can now prove the following:

THEOREM 4.11. *If $K+L$ is closed, then the von Neumann sequences converge linearly with rate* $\cos\gamma(K,L)$ independent *of the starting point. In particular, this happens whenever one of the following conditions holds:*

(i) *$K$ or $L$ has finite dimension.*

(ii) *$K$ or $L$ has finite codimension.*

*Proof.* The assumption that $K+L$ is closed is equivalent to $\gamma(K,L)>0$ and $E,F\neq\emptyset$ (by Fact 4.9 and Lemma 4.10). Recalling

$$P_A x = P_{a+K}x = a + P_K(x-a) = P_{K^\perp}a + P_K x$$

and

$$P_B x = P_{L^\perp}b + P_L x$$

for every $x\in X$, and defining

$$r := P_{K^\perp}a + P_K P_{L^\perp}b, \qquad s := P_{L^\perp}b + P_L P_{K^\perp}a,$$

we can 'decompose' the von Neumann sequences as follows:

$$\begin{aligned} b_{n+1} &= P_B a_{n+1} = P_{L^\perp}b + P_L a_{n+1} = P_{L^\perp}b + P_L P_A b_n \\ &= P_{L^\perp}b + P_L(P_{K^\perp}a + P_K b_n) \\ &= s + P_L P_K b_n \\ &= s + P_L P_K(s + P_L P_K b_{n-1}) \\ &= \vdots \\ &= \sum_{k=0}^{n}(P_L P_K)^k s + (P_L P_K)^{n+1} b_0, \end{aligned}$$

and similarly

$$a_{n+1} = \sum_{k=0}^{n}(P_K P_L)^k r + (P_K P_L)^n a_1$$

for all integers $n\geq 0$. By [9, Theorem 2.3], we know that

$$(P_L P_K)^n b_0 \longrightarrow P_{K\cap L}b_0, \qquad (P_K P_L)^n a_1 \longrightarrow P_{K\cap L}a_1$$

linearly with rate $c^2$, where $c:=\cos\gamma(K,L)$. In fact, the limits coincide, since

$$\begin{aligned} (P_K P_L)^n a_1 &= (P_K P_L)^n (P_A b_0) = (P_K P_L)^n (P_{K^\perp}a + P_K b_0) \\ &= (P_K P_L)^n (P_{K^\perp}a) + P_K (P_L P_K)^n b_0 \\ &\to P_{K\cap L}P_{K^\perp}a + P_K P_{K\cap L} b_0 \\ &= P_{K\cap L} b_0. \end{aligned}$$

So we must discuss the sums in the decomposition of the von Neumann sequences! Firstly, it is easy to check that $r, s \in (K \cap L)^\perp$ and that

$$\begin{aligned} P_L P_K &= P_{K\cap L} + P_L P_K P_{(K\cap L)^\perp}, \\ P_K P_L &= P_{K\cap L} + P_K P_L P_{(K\cap L)^\perp}; \end{aligned}$$

hence secondly for any integer $n \geq 0$:

$$\begin{aligned} a_n &= \left( \sum_{k=0}^{n-1} (P_K P_L P_{(K\cap L)^\perp})^k \right) r + (P_K P_L)^{n-1} a_1, \\ b_n &= \left( \sum_{k=0}^{n-1} (P_L P_K P_{(K\cap L)^\perp})^k \right) s + (P_L P_K)^n b_0. \end{aligned}$$

Thirdly, by [9, Proof of Lemma 2.5.(3)],

$$\| P_L P_K P_{(K\cap L)^\perp} \| = \| P_K P_L P_{(K\cap L)^\perp} \| = c < 1.$$

Therefore,

$$\begin{aligned} \sum_{k=0}^{n-1} (P_K P_L P_{(K\cap L)^\perp})^k r &\to (\mathrm{Id} - P_K P_L P_{(K\cap L)^\perp})^{-1} r, \\ \sum_{k=0}^{n-1} (P_L P_K P_{(K\cap L)^\perp})^k s &\to (\mathrm{Id} - P_L P_K P_{(K\cap L)^\perp})^{-1} s \end{aligned}$$

linearly with rate $c$. Altogether: Every von Neumann sequences is — as a sum of two sequences, where one converges with rate $c^2$ and the other converges with rate $c$ — linearly convergent with rate $c$. Note that (i) follows since the sum of an arbitrary closed subspace and a finite-dimensional subspace is closed. Finally, (ii) follows since $K + L$ is closed if and only if $K^\perp + L^\perp$ is. The proof is complete. □

## 5. An Important Example

Throughout this section, we assume the following:

$X$ is a Hilbert lattice with lattice cone $S := X^+$, while $T : X \to Y := \mathbb{R}^N$ is linear, continuous and given by $x \mapsto (\langle t_i, x \rangle)_{i=1}^n$ for some vectors $t_1, \cdots, t_N \in X$. We also assume that $T$ is onto, equivalently (see, e.g., [21, Theorem 4.15])

$$T^* : \mathbb{R}^N \longrightarrow X : (\lambda_i)_{i=1}^n \longmapsto \sum_{i=1}^n \lambda_i t_i$$

is one-to-one, equivalently the $t_i$'s are linearly independent. Since kernel $(TT^*) =$ kernel $(T^*) = \{0\}$, we may define

$$Q := T^*(TT^*)^{-1}T.$$

It is straightforward to check that

$$TQ = T, \quad QT^* = T^*, \quad Q^2 = Q, \quad Q = Q^*, \quad P_{\text{range } T^*} = Q,$$

and, since always kernel $T = (\text{range } T^*)^{\perp}$,

$$P_{\text{kernel } T} = \text{Id} - Q = \text{Id} - T^*(TT^*)^{-1}T.$$

We now define the sets $A$, $B$. Given by $\bar{y} \in Y$, we let

$$A := T^{-1}(\bar{y}) \ \text{ and } \ B := S.$$

Thus we are interested in finding a solution to $Tx = \bar{y}$ subject to a nonnegativity constraint. Such problems arise frequently in applications and are often referred to as 'moment problems' (see [6] for a recent survey article). Often, one wants to find a nonnegative solution which minimizes a convex objective function. In our case, the objective function is identically zero. Nonetheless, even this feasibility problem has many applications!

Working with a Hilbert lattice, we can readily describe $P_B$; it is just taking the positive part (use [7, Theorem 8] to check Kolmogorov's criterion!). But how can we describe $P_A$? Well, fix an arbitrary $\bar{x} \in A$, then $A = \bar{x} + \text{kernel } T$ and, hence, for any $x \in X$:

$$\begin{aligned} P_A x &= P_{\bar{x}+\text{kernel } T}x = \bar{x} + P_{\text{kernel } T}(x - \bar{x}) \\ &= \bar{x} + (\text{Id} - Q)(x - \bar{x}) = Q\bar{x} + (\text{Id} - Q)x \\ &= x + Q(\bar{x} - x). \end{aligned}$$

Possessing a lattice cone, we can describe the relation between $a_n$ and $a_{n+1}$ nicely:

$$a_n \stackrel{P_B}{\longmapsto} b_n = P_B a_n = a_n^+ \stackrel{P_A}{\longmapsto} a_{n+1} = P_A b_n = Q\bar{x} + (\text{Id} - Q)a_n^+.$$

Since this is true for an *arbitrary* $\bar{x} \in A$, we can simply choose $\bar{x} = a_n$ and obtain the even nicer iteration:

$$a_{n+1} = Qa_n + (\text{Id} - Q)a_n^+ = a_n^+ - Qa_n^-.$$

Note, by the way, that the range $T^*$-part of $(a_n)$ is constant: $Qa_1 = Qa_2 = Qa_3 = \cdots$. What is the displacement vector in this setting? By Example 2.2, we get

$$v = P_{(\text{kernel } T)^{\ominus} \cap S^{\oplus}}(-\bar{x}) = P_{\text{range } T^* \cap S^{\oplus}}(-\bar{x}),$$

for an arbitrary $\bar{x} \in A$. And from Facts 1.2(i), we learn

$$a_n^- \longrightarrow v.$$

THEOREM 5.1. *If $Q(S) \subseteq S$, then $E, F \neq \emptyset$ and the von Neumann sequences converge in norm. In particular, this happens whenever $\{t_1, t_2, \cdots, t_N\}$ is an orthogonal subset of $S \cup (-S)$.*

*Proof.* Because of

$$a_{n+1}^{+} = (a_n^{+} - Qa_n^{-})^{+} \leq (a_n^{+})^{+} + (-Qa_n^{-})^{+} = a_n^{+},$$

the sequence $(a_n^{+})$ is decreasing and, hence, converges to some $s \in S$ in norm. But then

$$a_n = a_n^{+} - a_n^{-} \longrightarrow s - v$$

and

$$b_n = (b_n - a_n) + a_n \longrightarrow v + (s - v) = s.$$

It remains to verify the 'in particular' part. Using orthogonality of the $t_i$'s, we check that

$$Qx = \sum_{i=1}^{N} \langle \frac{t_i}{\| t_i \|}, x \rangle \frac{t_i}{\| t_i \|}$$

for every $x \in X$. However: If $x \in S$ and $s \in S \cup (-S)$, then $\langle s, x \rangle s \in S$. Hence $Qx \in S$ and the proof is complete. □

The following lemma is needed for the discussion of the case when $A$, $B$ have nonempty intersection. It is true for an *arbitrary* closed cone $S$.

LEMMA 5.2. *If $S$ is a closed cone, then TFAE:*

(i) $\displaystyle\forall_{y \in Y : T^* y \in S^{\oplus}} \quad y = 0.$

(ii) $TS = Y$.

(iii) $\operatorname{range} T^* \cap S^{\oplus} = \{0\}$.

*Proof.* '(i) $\Longrightarrow$ (ii)': Suppose to the contrary that $TS \subsetneqq Y$.

*Claim 1:* $\overline{TS} \subsetneqq Y$.
By [23, Theorem 3], $TS$ is a $CS$-closed subset since $TS$ is convex and $Y$ is finite-dimensional. Were $\overline{TS} = Y$, then, by [15, Theorem 22.4], $Y = \operatorname{int}(Y) = \operatorname{int}(\overline{TS}) = \operatorname{int}(TS) \subseteq TS \subsetneqq Y$ — a contradiction. Claim 1 is verified.

*Claim 2:* $\langle y^*, TS \rangle \geq 0 \quad$ for some $\quad y^* \in Y \backslash \{0\}$.
By Claim 1, $\overline{TS}$ is a proper closed convex subset of $Y$. Fix $z \in Y \backslash \overline{TS}$. We can separate and thus obtain $y^* \in Y \backslash \{0\}$ s.t.

$$\langle y^*, TS \rangle \geq \langle y^*, z \rangle.$$

Claim 2 follows, because $TS$ is a *cone*.

But Claim 2 also means that $T^*y^* \in S^\oplus$. By assumption, $y^*$ must equal zero – the desired contradiction. '(ii)$\Longrightarrow$(iii)': Suppose $T^*y \in S^\oplus$. Then

$$0 \leq \langle T^*y, S\rangle = \langle y, TS\rangle = \langle y, Y\rangle.$$

This implies $y = 0$. '(iii)$\Longrightarrow$(i)': Suppose $T^*y \in S^\oplus$. Then, by assumption, $T^*y = 0$ implying $y = 0$, since $T^*$ is one-to-one. The proof is complete. □

We return to our original setting, i.e. $S$ is a Hilbert lattice cone.

THEOREM 5.3. *Suppose $A \cap B \neq \emptyset$. Then the alternating von Neumann sequence converges* $\cdots$

(i) *linearly, whenever* $\bar{y} \in \operatorname{int} TS$.
(ii) *linearly, whenever* $TS = Y$.
(iii) *linearly, whenever* range $T^* \cap S^\oplus = \{0\}$.
(iv) *in norm, whenever* $N = 1$.

*Proof.* (i) follows by application of Theorem 4.3 (ii) to $M = S$, $N = \{\bar{y}\}$: We obtain bounded linear regularity of $(A, B)$ whenever $0 \in \operatorname{icore}(TS - \bar{y})$. Corollary 3.14 then implies linear convergence. By the last lemma, (ii) and (iii) are equivalent and clearly imply (i). It remains to prove (iv). If $N = 1$, then range $T^* = \operatorname{span}\{t_1\}$. Either $t_1 \in S \cup (-S)$ and the last theorem applies. Or $t_1 \notin S \cup (-S)$, then range $T^* \cap S = \{0\}$ and (iii) applies. □

REMARKS 5.4.

- Canonical examples for Hilbert lattices are $X = L_2[0, 1]$ and $\ell_2(\mathbb{N})$ with the point-wise ordering.
- Theorem 5.3 (i), (ii), (iii) hold if $S$ is any *vector* lattice cone. But one has to be careful since in general $P_S x \neq x^+$; so the iteration for the sequence $(a_n)$ becomes

  $$a_{n+1} = P_S a_n + Q(a_n - P_S a_n).$$

- If the *quasi-relative interior of S* is nonempty, then Borwein and Lewis proved that

  $$\operatorname{int} TS = T(\operatorname{qri} S).$$

  To apply (i), we must check that $\bar{y} \in T(\operatorname{qri} S)$ which is usually easier! For example, if $X = L_2[0, 1]$, $S = X^+$, then

  $$\operatorname{qri} S = \{x \in L_2[0, 1] : x > 0 \text{ a.e.}\},$$

  and it is often a priori clear that a strictly positive solution exists. For more on this concept see [4, 5].

- (ii) implies (i), but the converse is false in general: Indeed, let $Y := \mathbb{R}$ and fix $T \in S\backslash\{0\}$. Then $TS = [0, +\infty[$, $1 \in \operatorname{int} TS$, but $TS \subsetneqq Y$.
- We proved (i) by showing that $\bar{y} \in \operatorname{int} TS$ implies $(A, B)$ is boundedly linearly regular. The next example shows that we cannot drop the interiority assumption.

EXAMPLE 5.5. *Let*

$$X := \ell_2(\mathbb{N}), \qquad N := 1, \quad \bar{y} := 0, \qquad T := (T_1, T_2, \cdots) \in S.$$

*If* $\{n \in \mathbb{N} : T_n > 0\}$ *is infinite, then* $(A, B)$ *is not boundedly regular. Nevertheless, the alternating von Neumann sequence converges in norm. Consequently, bounded regularity is a sufficient but not a necessary condition for norm convergence of the von Neumann sequences.*

*Proof.* Define $I := \{n \in \mathbb{N} : T_n > 0\}$. Clearly,

$$A \cap B = \{x \in X : x_n = 0 \text{ if } n \in I; x_n \geq 0 \text{ otherwise}\}.$$

If $c \in A \cap B$, then $\| e_n - c \| \geq |1 - c_n|$; so for all $n \in I$, we conclude $\| e_n - c \| \geq 1$ and further

$$d(e_n, A \cap B) \geq 1.$$

In fact, we can pick $0 \in A \cap B$ to get equality. On the one hand, $e_n \in S = B$, so for all $n \geq 0$,

$$d(e_n, B) = 0.$$

On the other hand,

$$P_A x = P_{\text{kernel } T} x = P_{\{T\}^\perp} x = x - \frac{\langle T, x \rangle}{\| T \|^2} T,$$

thus

$$d(e_n, A) = \| e_n - P_A e_n \| = \frac{T_n}{\| T \|} \longrightarrow 0.$$

Altogether: $d(e_n, A)$, $d(e_n, B) \longrightarrow 0$, but $d(e_n, A \cap B) = 1$ infinitely often. Therefore, $(A, B)$ is not boundedly regular. The convergence of the alternating von Neumann sequence follows from Theorem 5.3 (iv). $\square$

## 6. Finitely Many Sets

In applications, it is often of interest to find a point in the intersection of finitely many sets (or, if the intersection is empty, a good substitute of it). This can be done by Pierra's clever product space formalization [18] which reduces the problem to one involving two sets only! Recently, this approach, among much more general iteration schemes, has become increasingly popular since it is 'parallelizable'; we refer the interested reader to a survey article by Censor [8].

*Setting.* Let $X$ be a Hilbert space, $C_1, C_2, \cdots, C_N$ $(N \geq 2)$ closed convex nonempty subsets with projections $P_1, P_2, \cdots, P_N$. Define

$$\begin{aligned}
\mathbf{X} &:= \Pi_{i=1}^{N}(X, \frac{1}{N}\langle\cdot,\cdot\rangle), \\
\mathbf{A} &:= \{(x_1, \cdots, x_N) \in \mathbf{X} : x_1 = x_2 = \cdots = x_N \in X\}, \\
\mathbf{B} &:= \{(x_1, \cdots, x_N) \in \mathbf{X} : x_i \in C_i \text{ for } i = 1 \cdots N\} = \Pi_{i=1}^{N} C_i.
\end{aligned}$$

As shown in [1, Section 6], the following is true: For two points $\mathbf{x} = (x_1, \cdots, x_N)$, $\mathbf{y} = (y_1, \cdots, y_N) \in \mathbf{X}$, we have

$$\| \mathbf{x} - \mathbf{y} \|^2 = \sum_{i=1}^{N} \frac{1}{N} \| x_i - y_i \|^2,$$

and the projection onto $\mathbf{A}$, $\mathbf{B}$ is given by

$$\begin{aligned}
P_{\mathbf{A}}(x_1, x_2, \cdots, x_N) &= \left( \sum_{i=1}^{N} \frac{1}{N} x_i, \sum_{i=1}^{N} \frac{1}{N} x_i, \cdots, \sum_{i=1}^{N} \frac{1}{N} x_i, \right), \\
P_{\mathbf{B}}(x_1, x_2, \cdots, x_N) &= (P_1 x_1, P_2 x_2, \cdots, P_N x_N).
\end{aligned}$$

The set $\mathbf{E} = \mathrm{Fix}(P_{\mathbf{A}} P_{\mathbf{B}})$ was found to have the alternative description $\mathbf{E} = \{(x, x, \cdots, x) \in \mathbf{X} : x \in E\}$, where

$$E = \mathrm{Fix}\, \frac{P_1 + P_2 + \cdots + P_N}{N} = \mathrm{arginf}_{x \in X} \sum_{i=1}^{N} d^2(x, C_i).$$

This time, we give a different description for $\mathbf{F} = \mathrm{Fix}(P_{\mathbf{B}} P_{\mathbf{A}})$:

THEOREM 6.1.

$$\mathbf{F} = \mathrm{arginf}_{(c_i) \in \mathbf{B}} \sum_{\mathrm{i=1 \cdots N}} \sum_{\mathrm{j=1 \cdots N: j \neq i}} \| c_i - c_j \|^2$$

*Proof.* We define the *right-shift operator* $R : \mathbf{X} \longrightarrow \mathbf{X} : (x_1, \cdots, x_N) \longmapsto (x_N, x_1, \cdots, x_{N-1})$, the *left-shift operator* $L := R^*$ and consider for an arbitrary but fixed $\lambda > 0$ the convex program

$$(P_\lambda) \qquad \inf_{x \in \mathbf{B}} \frac{\lambda}{2} \sum_{j=1}^{N-1} \| \mathbf{x} - R^j \mathbf{x} \|^2 = \inf_{\mathbf{x} \in \mathbf{X}} f_\lambda(\mathbf{x}) + \iota_{\mathbf{B}}(\mathbf{x}),$$

where $f_\lambda(\mathbf{x}) := (\lambda/2) \sum_{j=1}^{N-1} \| \mathbf{x} - R^j \mathbf{x} \|^2$. By [24, Example 6.4(a), Theorem 5.38] and the *chain rule* [11, Proposition 5.7],

$$\nabla f_\lambda(\mathbf{x}) = \lambda \sum_{j=1}^{N-1} (\mathrm{Id} - L^j)(\mathrm{Id} - R^j)\mathbf{x},$$

so the optimality condition for $(P_\lambda)$ becomes

$$0 \in \nabla f_\lambda(\mathbf{x}) + \partial\iota_{\mathbf{B}}(\mathbf{x}) \Longleftrightarrow -\nabla f_\lambda(\mathbf{x}) \in \partial\iota_{\mathbf{B}}(\mathbf{x}).$$

On the other hand,

$$\mathbf{x}^* \in \partial\iota_{\mathbf{B}}(\mathbf{x}) \Longleftrightarrow \mathbf{x} = P_{\mathbf{B}}(\mathbf{x}^* + \mathbf{x}).$$

Thus

$$\mathbf{x} \text{ solves } (P_\lambda) \Longleftrightarrow \mathbf{x} = P_{\mathbf{B}}(\mathbf{x} - \nabla f_\lambda(\mathbf{x})).$$

Hence

$$\operatorname{arginf}\,(P_\lambda) = \operatorname{Fix} P_{\mathbf{B}}(\mathrm{Id} - \nabla f_\lambda).$$

Since this is true for any positive $\lambda$, we can choose $\bar{\lambda} = 1/(2N)$, which — after an elementary calculation — yields

$$\mathbf{x} - \nabla f_{\bar{\lambda}}(\mathbf{x}) = \left(\sum_{i=1}^{N} \frac{1}{N} x_i, \sum_{i=1}^{N} \frac{1}{N} x_i, \cdots, \sum_{i=1}^{N} \frac{1}{N} x_i\right) = P_{\mathbf{A}}(\mathbf{x}),$$

for every $\mathbf{x} = (x_1, \cdots, x_N) \in \mathbf{X}$; so $\mathrm{Id} - \nabla f_{\bar{\lambda}} = P_{\mathbf{A}}$ and therefore

$$\operatorname{arginf}\,(P_\lambda) = \operatorname{arginf}\,(P_{\bar{\lambda}}) = \operatorname{Fix} P_{\mathbf{B}} P_{\mathbf{A}} = \mathbf{F}.$$

Finally, observing $\mathbf{F} = \operatorname{arginf}\,(P_2)$ finishes the proof. □

We believe that the alternative descriptions of $E$ and $\mathbf{F}$ are highly satisfying: Loosely speaking, the set $E$ consists of all points which are simultaneously close to all of the sets $C_1, \cdots, C_N$, whereas the set $\mathbf{F}$ consists of all $N$-tuples of points which are simultaneously close to each other. Of course, if the sets $C_1, \cdots, C_N$ have nonempty intersection, then all these meanings collapse to this intersection.

We want to study von Neumann's algorithm in this setting. It is here more elegant to 'ignore' the product space setting by following just the von Neumann sequence in $\mathbf{A}$ — that is: $X$ — and to start the algorithm at a point in $\mathbf{A}$. The von Neumann algorithm then becomes:

$$a_{n+1} := \frac{P_1 a_n + P_2 a_n + \cdots + P_N a_n}{N}, \quad a_1 \in X.$$

Combining [1, Theorem 6.4] with the last theorem yields the following theorem.

THEOREM 6.2.

$$\sum_{i=1}^{N} \| a_n - P_i a_n \|^2, \sum_{i=1}^{N} \| P_i a_n - a_{n+1} \|^2 \to N d^2(\mathbf{A}, \mathbf{B}).$$

*Moreover:*

*If $E$ is empty, then*

$$\| a_n \|, \max\{\| P_1 a_n \|, \| P_2 a_n \|, \cdots, \| P_N a_n \|\} \to \infty.$$

*If $E$ is nonempty, then*

$$a_n \rightharpoonup e^*, P_1 a_n \rightharpoonup P_1 e^*, \cdots, P_N a_n \rightharpoonup P_N e^*$$

*for some $e^* \in E$ and*

$$e^* \in \operatorname{arginf}_{x \in X} \sum_{i=1}^{N} d^2(x, C_i),$$
$$(P_i e^*) \in \operatorname{arginf}_{(c_i) \in \mathbf{B}} \sum_{i=1 \cdots N} \sum_{\substack{j=1 \cdots N: \\ j \neq i}} \| c_i - c_j \|^2 .$$

All our results of the previous sections may be reformulated in this product space setting. We think that the following is the most interesting:

THEOREM 6.3. *If* $0 \in \operatorname{icore}(\mathbf{A} - \mathbf{B})$, *then the von Neumann sequence* $(a_n)$ *converges linearly. This happens whenever one of the following conditions holds:*

(i) $0 \in \operatorname{int}(\mathbf{A} - \mathbf{B})$.
(ii) $C_i \cap \bigcap_{j=1 \cdots N: j \neq i} \operatorname{int} C_j \neq \emptyset$ *for some* $i$.
(iii) $\operatorname{int}(\bigcap_{j=1 \cdots N} C_j) \neq \emptyset$.
(iv) $\mathbf{A} - \mathbf{B}$ *is a closed subspace.*
(v) Every $C_i$ *is a finite-dimensional affine subspace.*
(vi) *Every* $C_i$ *is an affine subspace of finite codimension.*
(vii) *Every* $C_i$ *is a hyperplane.*

*Proof.* The main statement, (i) and (iv) follow directly from Corollary 4.5. If all $C_i$'s are finite-dimensional, then so is $\mathbf{B}$ and thus (iv) holds — this gives (v). Similarly, (vi) implies (iv). Clearly, (vii) implies (vi). Also, (iii) implies (ii). We prove the remaining part (ii) by showing that it implies (i). Without loss suppose $i = 1$ and $z \in C_1 \cap \bigcap_{j=2}^{N} \operatorname{int} C_j$. Next, we pick a neighborhood $W$ of 0 s.t. $z + W - W \subseteq C_j$ for $j = 2 \cdots N$. Fix $\mathbf{w} = (w_1, \cdots w_N) \in \prod_{j=1}^{N} =: \mathbf{W}$. Then $c_j := z + w_1 - w_j \in C_j$ for $j = 1 \cdots N$ and hence $w_j = (z + w_1) - c_j$. This implies $\mathbf{w} \in \mathbf{A} - \mathbf{B}$ and, since $\mathbf{w}$ was chosen arbitrarily, $\mathbf{W} \subseteq \mathbf{A} - \mathbf{B}$. The proof is complete. □

We saw in the proof that (ii) implies (i). The converse, however, is false in general:

EXAMPLE 6.4. *Let* $X = \ell_2(\mathbb{N})$, $C_1 := C_2 := C_3 := S = X^+$. *Then*

$$\operatorname{int} C_j = \emptyset$$

*for* $j = 1, 2, 3$ *but*

$$\mathbf{A} - \mathbf{B} = X.$$

*Proof.* Given $(x_1, x_2, x_3) \in \mathbf{X}$, we get, since $S$ is generating, $s_1, s_2, t_1, t_2 \in S$ s.t.

$$x_1 - x_2 = s_2 - s_1 \qquad \text{and} \qquad x_2 - x_3 = t_1 - t_2.$$

Now define

$$b_1 := s_1 + t_2, \qquad b_2 := s_2 + t_2, \qquad b_3 := s_2 + t_1, \qquad a := x_1 + s_1 + t_2$$

and check that

$$(x_1, x_2, x_3) = (a, a, a) - (b_1, b_2, b_3) \in \mathbf{A} - \mathbf{B}.$$

This completes the proof. □

REMARKS 6.5.
- As far as we know, condition (i) is one of the best conditions guaranteeing linear convergence to a point in $\bigcap_{j=1}^{N} C_j$. Indeed, as we just saw, it is more general than condition
  - (iii) which was proposed by Pierra [18, Theorem 1.1.(ii).(a)].
  - (ii) which was proposed by Gubin *et al.* [14, Theorem 1.(a)] who obtained linear convergence for the *cyclic projection algorithm*.
- If condition (iv) holds and the $C_j$'s are affine closed subspaces, then (by Theorem 4.11) the rate of convergence is independent of the starting point! Of course, the same is true if (v), (vi) or (vii) holds.

## Acknowledgement

It is our pleasure to thank Aleksander Simonič for discussing and proving Lemma 4.10.

## References

1. Bauschke, H. H. and Borwein, J. M.: Dykstra's alternating projection algorithm for two sets. Technical Report CORR 92-15, Faculty of Mathematics, University of Waterloo, May 1992, to appear in *J. Approx. Theory.*
2. Borwein, J. M.: Convex relations in analysis and optimization, in S. Schaible and W. T. Ziemba, (eds), *Generalized Concavity in Optimization and Economics*, Academic Press, New York, pp. 335–377.
3. Borwein, J. M.: Stability and regular points of inequality systems. J. Optim. Theory Appl. **48**(1) (1986), 9–52.
4. Borwein, J. M. and Lewis, A. S.: Partially finite convex programming Part I: Quasi relative interiors and duality theory, *Math. Programming* **57** (1992), 15–48.

5. Borwein, J. M. and Lewis, A. S.: A Partially finite convex programming Part II: Explicit lattice models. *Math. Programming* **57** (1992), 49–83.
6. Borwein, J. M. and Lewis, A. S.: A survey of convergence results for maximum entropy methods, Technical Report CORR 92-32, Faculty of Mathematics, University of Waterloo, September 1992.
7. Borwein, J. M., and Yost, D. T.: Absolute norms on vector lattices. *Proc. Edinburgh Math. Soc.*, **27** (1984), 215–222.
8. Censor, Y.: On variable block algebraic reconstruction techniques, in G. T. Herman, A. K. Louis, and F. Natterer (eds), *Mathematical Methods in Tomography*, Springer, New York, 1990, pp. 133–140.
9. Deutsch, F.: Rate of convergence of the method of alternating projections, in B. Brosowski and F. Deutsch (eds), *Parametric Optimization and Approximation*, Birkhäuser, Basel, 1983, pp. 96–107.
10. Deutsch, F.: The method of alternating orthogonal projections, in S. P. Singh (ed), *Approximation Theory, Spline Functions and Applications*, Kluwer Academic Publ., Dordrecht, 1992, pp. 105–121.
11. Ekeland, I. and Temam, R.: *Convex Analysis and Variational Problems*, North-Holland, Amsterdam, 1976.
12. Franchetti, C. and Light, W.: On the von Neumann alternating algorithm in Hilbert space, *J. Math. Anal. Appl.* **114** (1986), 305–314.
13. Friedricks, K.: On certain inequalities and characteristic value problems for analytic functions and for functions of two variables, *Trans. Amer. Math. Soc.* **41** (1937), 321–364.
14. Gubin, L. G., Polyak, B. T., and Raik, E. V.: The method of projections for finding the common point of convex sets, *U.S.S.R. Comput. Math. and Math. Phys.* **7** (6) (1967), 1–24.
15. Jameson, G. J. O.: *Topology and Normed Spaces*, Chapman and Hall, London, 1974.
16. Luenberger, D. G.: *Optimization by Vector Space Methods*, Wiley, New York, 1969.
17. Motzkin, T. S. and Schoenberg, I. J.: The relaxation method for linear inequalities, *Canad. J. Math.* **6** (1954), 393–404.
18. Pierra, G.: Decomposition through formalization in a product space, *Math. Programming* **28** (1984), 96–115.
19. Robinson, S. M.: Regularity and stability for convex multivalued functions. *Math. Oper. Res.* **1** (2) (1976), 130–143.
20. Rockafellar, R. T.: *Convex Analysis*, Princeton University Press, Princeton, NJ, 1970.
21. Rudin, W.: *Functional Analysis*, McGraw-Hill, New York, 1973.
22. Simonič, A.: Personal communication.
23. Tijs, S. H. and Borwein, J. M.: Some generalizations of Carathéodory's theorem via barycenters, with application to mathematical programming, *Canad. Math. Bull.* **23** (3) (1980), 339–346.
24. Van Tiel, J.: *Convex Analysis*, Wiley, New York, 1984.
25. von Neumann, J.: *Functional Operators, Vol. II*, Princeton University Press, 1950. (Reprint of mimeographed lecture notes first distributed in 1933.)