

# Characterizing Arbitrarily Slow Convergence in the Method of Alternating Projections

Heinz H. Bauschke<sup>1</sup>, Frank Deutsch, and Hein Hundal

April 8, 2008 Version

<sup>1</sup>Supported in part by the Natural Sciences and Engineering Research Council of Canada and by the Canada Research Chair Program.

## **Abstract**

Bauschke, Borwein, and Lewis have stated a trichotomy theorem [4, Theorem 5.7.16] that characterizes when the convergence of the method of alternating projections can be arbitrarily slow. However, there are two errors in their proof of this theorem. In this note, we show that although one of the errors is critical, the theorem itself is correct. We give a different proof that uses the multiplicative form of the spectral theorem, and the theorem holds in any real or complex Hilbert space, not just in a real Hilbert space.

1991 Mathematics Subject Classification: 41A65, 46C05, 46N10, 47H09.

Key Words and Phrases: alternating projections, cyclic projections, orthogonal projections, angle between subspaces, rate of convergence of the method of alternating projections.

# 1 Introduction

For the notation and basic Hilbert space results necessary to read this paper, the book [6] is a good source, especially chapter 9.

Let  $H$  be a (real or complex) Hilbert space with inner product  $\langle x, y \rangle$  and norm  $\|x\| = \sqrt{\langle x, x \rangle}$ . If  $M$  is any closed (linear) subspace of  $H$ , let  $P_M$  denote the orthogonal projection onto  $M$ . That is,  $P_M : H \rightarrow M$  is defined by

$$\|x - P_M(x)\| = \inf_{y \in M} \|x - y\|.$$

Let  $M_1$  and  $M_2$  be closed subspaces in  $H$  and  $M := M_1 \cap M_2$ . It is well-known that  $P_{M_1}P_{M_2} = P_M$  if and only if  $P_{M_1}$  and  $P_{M_2}$  commute:  $P_{M_1}P_{M_2} = P_{M_2}P_{M_1}$ . Von Neumann established the following result which yields an interesting analogue in the non-commuting case.

**Theorem 1.1 (von Neumann [14])** *For each  $x \in H$ , there holds*

$$\lim_{n \rightarrow \infty} \|(P_{M_2}P_{M_1})^n(x) - P_M(x)\| = 0. \quad (1.1)$$

The method of constructing the sequence  $(P_{M_2}P_{M_1})^n(x)$  by alternately projecting onto one subspace and then the other is called the *method of alternating projections*. While von Neumann's theorem shows that the sequence of iterates  $(P_{M_2}P_{M_1})^n(x)$ , *always* converges to  $P_M(x)$  for every  $x$ , it does not say anything about the speed or rate of convergence. To say something about this, we will use the notion of angle between subspaces. Recall that the (Friedrichs) **angle** between the subspaces  $M_1$  and  $M_2$  is defined to be the angle in  $[0, \pi/2]$  whose cosine is given by

$$c(M_1, M_2) := \sup\{|\langle x, y \rangle| \mid x \in M_1 \cap M^\perp \cap B_H, y \in M_2 \cap M^\perp \cap B_H\},$$

where  $B_H := \{x \in H \mid \|x\| \leq 1\}$  is the unit ball in  $H$ . It is easy to see that  $0 \leq c(M_1, M_2) \leq 1$ .

**Theorem 1.2 (Aronszajn [1])** *For each  $x \in H$  and  $n \geq 1$ , we have*

$$\|(P_{M_2}P_{M_1})^n(x) - P_M(x)\| \leq c(M_1, M_2)^{2n-1} \|x\|. \quad (1.2)$$

Kayalar and Weinert [13] showed that the constant in Aronszajn's theorem is smallest possible independent of  $x$ . More precisely, they proved that

$$\|(P_{M_2}P_{M_1})^n - P_M\| = c(M_1, M_2)^{2n-1} \quad \text{for each } n \geq 1. \quad (1.3)$$

The usefulness of the bound in (1.2) depends on knowing when the cosine of the angle between  $M_1$  and  $M_2$  is less than one, i.e., when the angle is positive. A useful characterization of when this happens is the following.

**Lemma 1.3**  $c(M_1, M_2) < 1$  if and only if  $M_1^\perp + M_2^\perp$  is closed.

This lemma is a consequence of results of Deutsch [5] and Simonic, whose result appeared in [2, Lemma 4.10] (see also [6, Theorem 9.35, p. 222]).

Recall that a sequence  $(x_n)$  is said to converge to  $x$  **linearly** provided there exists an  $\alpha < 1$  and a constant  $c$  such that

$$\|x_n - x\| \leq c\alpha^n \quad \text{for each } n \geq 1.$$

In this case, we say that the rate of convergence is  $\alpha$ .

Using Lemma 1.3 and Theorem 1.2, we see that there is *linear convergence* for the method of alternating projections whenever the sum of the subspaces is closed. What can be said when the sum is not closed?

Franchetti and Light [11] gave the first example of a Hilbert space and two closed subspaces whose sum was not closed such that: given any sequence of reals decreasing to zero, there exists a point in the space with the property that the convergence in the von Neumann theorem was at least as slow as this sequence of reals. But this still left open the question of whether such a construction could be made in *any* Hilbert space whenever  $M_1$  and  $M_2$  were *any* closed subspaces whose sum was not closed.

In their study of the method of alternating projections, Bauschke, Borwein, and Lewis [4] stated the following dichotomy. (Actually, they stated their result as a trichotomy since they were considering the more general setting of closed *affine* sets, i.e., translates of subspaces, rather than subspaces. In this situation, unlike the subspace case, one must also consider the possibility that the intersection of the affine sets is empty. However, when the intersection is nonempty, the affine sets case easily reduces to the subspace case by a simple translation.) Roughly speaking, it states that in the method of alternating projections, either there is linear convergence for each starting point, or there exists a point which converges arbitrarily slowly.

**Theorem 1.4** (dichotomy) *Let  $M_1$  and  $M_2$  be closed subspaces in a Hilbert space  $H$  and  $M = M_1 \cap M_2$ . Then exactly one of the following alternatives holds.*

- (1)  $M_1^\perp + M_2^\perp$  is closed. Then for each  $x \in H$ , the sequence  $(P_{M_2}P_{M_1})^n(x)$  converges linearly to  $P_M(x)$  with a rate  $[c(M_1, M_2)]^2$ .
- (2)  $M_1^\perp + M_2^\perp$  is not closed. Then for each  $x \in H$ , the sequence  $(P_{M_2}P_{M_1})^n(x)$  converges to  $P_M(x)$ . But convergence is “arbitrarily slow” in the following sense: for each sequence  $(\lambda_n)$  of positive real numbers with  $1 > \lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n \rightarrow 0$ , there exists a point  $x_\lambda \in H$  such that

$$\|(P_{M_2}P_{M_1})^n(x_\lambda) - P_M(x_\lambda)\| \geq \lambda_n \quad \text{for all } n.$$

**Remarks** (1) Clearly, the first statement of Theorem 1.4 is an immediate consequence of Theorem 1.2 and Lemma 1.3. Thus we need only verify the second statement. We will do this in Section 3 below.

(2) After this paper was submitted, two of us [10] established a result on powers of linear operators in Banach spaces, with a completely different proof, that implied that the dichotomy theorem holds for *any finite* number of subspaces, not just two.

(3) The dichotomy theorem also holds in the more general situation of a finite collection of closed *affine sets*, i.e., translates of linear subspaces, whose intersection is not empty. This can be deduced from the subspace case by a translation argument and a repeated application of the fact that, for any closed convex set  $C$  in a Hilbert space  $H$ , we have  $P_C(x) = P_{C+y}(x+y) - y$  for each  $x, y \in H$ .

(4) Since parallel projections correspond to alternating projections in the product space, the following variant of Theorem 1.4 is a consequence of Theorem 1.4 and [3, Lemma 5.18 and Theorem 5.19].

Let  $M_1, \dots, M_r$  be  $r$  closed subspaces of  $H$ ,  $M = \bigcap_1^r M_i$ , and for any  $x \in H$ , consider the sequence generated by the parallel projections method:

$$x_0 = x, \quad x_n = \left( \frac{1}{r} P_{M_1} + \dots + \frac{1}{r} P_{M_r} \right) (x_{n-1}) = \left( \frac{1}{r} \sum_1^r P_{M_i} \right)^n (x), \quad n \geq 1.$$

If  $\sum_1^r M_i^\perp$  is closed, then this sequence converges linearly to  $P_M(x)$ . Otherwise, the convergence is arbitrarily slow.

**Acknowledgement** We are indebted to two anonymous referees for some very useful and insightful suggestions that helped us to improve the paper.

## 2 Multiplicative form of the spectral theorem

The main fact that we will use in the proof of Theorem 1.4 is the multiplicative form of the spectral theorem (see Halmos [12] or Reed-Simon [15, Corollary on p. 227]). Recall that a bounded linear operator  $U : H_1 \rightarrow H_2$  between Hilbert spaces  $H_1$  and  $H_2$  is called *unitary* if  $U$  is invertible and  $U^* = U^{-1}$ . It follows that a unitary operator is isometric:  $\|Ux\| = \|x\|$  for each  $x \in H$ . Since the inverse of a unitary operator is unitary, it too is isometric. (We will use these facts in a few places below without explicit mention.)

**Theorem 2.1 (Spectral Theorem; multiplicative form)** *Let  $H$  be a (real or complex) Hilbert space, and let  $T$  be a self-adjoint bounded linear*

operator on  $H$ . Then there exists a finite measure space  $(\Omega, \mu)$ , a measurable essentially bounded real-valued function  $F$  on  $\Omega$ , and a unitary map  $U : H \rightarrow L_2(\Omega, \mu)$  such that

$$UTU^{-1}f = F \cdot f \quad \text{for all } f \in L_2(\Omega, \mu). \quad (2.1)$$

Defining  $D : L_2(\Omega, \mu) \rightarrow L_2(\Omega, \mu)$  to be the operator “multiplication by  $F$ ”,  $(Df)(t) := F(t)f(t)$ , this can be expressed in operator notation as

$$UTU^{-1} = D. \quad (2.2)$$

Moreover,  $\|T\| = \|F\|_\infty$ .

Actually, the last statement,  $\|T\| = \|F\|_\infty$ , is not in either [12] or [15], but it can be readily deduced from the first part of the theorem. Moreover, in both [12] and [15], the theorem is stated for a *complex* Hilbert space only, and [15] even assumes separability. However, it is easy to check that each of the tools used in the proof in [12], for example, has a corresponding real space analogue.

**Acknowledgements** We are greatly indebted to Joel Anderson, Nigel Higson, and Barry Simon for personally transmitting some very useful comments to us related to the multiplicative form of the spectral theorem.

A self-adjoint operator  $T$  on  $H$  is called *positive* if  $\langle Tx, x \rangle \geq 0$  for each  $x \in H$ . A simple, but important, example of a positive operator is the orthogonal projection  $P_S$  onto any closed subspace  $S \subset H$  (see, e.g., [6, p. 79]).

**Corollary 2.2** *Assume the hypothesis of Theorem 2.1. If  $T$  is also positive, then the function  $F$  of Theorem 2.1 is also nonnegative a.e.  $(\mu)$ .*

*Proof.* Let  $f \in L_2(\Omega, \mu)$  be arbitrary and  $y = U^{-1}f$ . Since  $T$  is positive, we have that

$$\begin{aligned} \int_{\Omega} F|f|^2 d\mu &= \langle Ff, f \rangle = \langle Df, f \rangle = \langle UTU^{-1}f, f \rangle \\ &= \langle TU^{-1}f, U^*f \rangle = \langle Ty, y \rangle \geq 0. \end{aligned}$$

Briefly,  $\int_{\Omega} F|f|^2 d\mu \geq 0$  for each  $f \in L_2(\Omega, \mu)$ . We readily deduce that  $F \geq 0$  a.e.  $(\mu)$ . ■

### 3 Proof of Theorem 1.4

In this section we will prove the second statement of Theorem 1.4. Our proof is along the same general lines as in [4] in that we proceed by a series of small steps that are each easily digested. However, there are subtle errors in steps 2 and 3 of [4] (see Section 4 for the details). We will avoid these errors by using Theorem 2.1 and following a somewhat different path.

**Proof of the second statement in Theorem 1.4.** Suppose  $M_1 + M_2$  is not closed, and let  $(\lambda_n)$  be a sequence with  $1 > \lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n > 0$ , and  $\lambda_n \rightarrow 0$ . By Lemma 1.3,  $c(M_1, M_2) = 1$ . Let

$$A = M_1 \cap M^\perp \quad \text{and} \quad B = M_2 \cap M^\perp. \quad (3.1)$$

Note that  $A$  and  $B$  are closed subspaces with  $A \cap B = \{0\}$ . Clearly,

$$c(A, B) = c(M_1, M_2) = 1 \quad (3.2)$$

and hence, by Lemma 1.3 again,  $A+B$  is not closed. Since  $c(A, B) = \|P_B P_A\|$  by [5] (see also [6, Lemma 9.5(7), p. 197]), it follows that  $\|P_B P_A\| = 1$ .

**Lemma 3.1** *The operator  $T := P_A P_B P_A$  is a bounded self-adjoint linear operator on  $H$  which is positive and  $\|T\| = 1$ . Hence there exists a finite measure space  $(\Omega, \mu)$ , a nonnegative essentially bounded measurable function  $F$  on  $\Omega$ , and a unitary operator  $U : H \rightarrow L_2 := L_2(\Omega, \mu)$  such that*

$$UTU^{-1} = D \quad \text{and} \quad \|F\|_\infty = 1, \quad (3.3)$$

where  $D : L_2 \rightarrow L_2$  is defined by  $Df := Ff$  for each  $f \in L_2$ .

*Proof of Lemma 3.1.* By Corollary 2.2, it suffices to verify the first statement of the lemma. Clearly,  $T$  is self-adjoint and bounded. Moreover, using [9, Corollary 5.17],  $\|T\| = \|P_A P_B P_A\| = \|P_B P_A\|^2 = 1$ . Fix any  $x \in H$  and set  $y = P_A x$ . Since  $P_B$  is positive, we have that

$$\langle Tx, x \rangle = \langle P_A P_B P_A x, x \rangle = \langle P_B P_A x, P_A x \rangle = \langle P_B y, y \rangle \geq 0.$$

This shows that  $T$  is positive on  $H$  and completes the proof of Lemma 3.1.

By inserting more terms into the sequence  $(\lambda_k)$  (e.g., 1/1.1, 1/2.1, 1/3.1, 1/4.1, ...), we may assume without loss of generality that the sequence  $(s_k)$  defined by

$$s_k := \lfloor 1/\lambda_k \rfloor$$

contains the set of all natural numbers  $\mathbb{N} := \{1, 2, 3, \dots\}$ . (Here  $\lfloor \cdot \rfloor : \mathbb{R} \rightarrow \mathbb{N}$  denotes the greatest integer function, i.e.,  $\lfloor x \rfloor$  is the greatest integer  $\leq x$ .)

Moreover, by increasing slightly some of the  $\lambda_k$ , we may further assume without loss of generality, that the sequence  $(1/\lambda_k)$  contains *no* integers. The following claim is now obvious.

**Claim 1.**  $s_k \lambda_k < 1 \leq (s_k + 1) \lambda_k$  for all  $k \in \mathbb{N}$ ,  $s_1 \leq s_2 \leq s_3 \leq \dots$ , and each  $s_k$  occurs only finitely often.

For each  $n \in \mathbb{N}$ , define

$$k_0(n) := \min\{k \mid s_k = n\} \text{ and } k_1(n) := \max\{k \mid s_k = n\}. \quad (3.4)$$

Clearly,  $k_1(n) = k_0(n+1) - 1$ ,  $k_0(n) \rightarrow \infty$ , and

$$s_{k_0(n)-1} = n - 1 < n = s_{k_0(n)} = s_{k_0(n)+1} = \dots = s_{k_0(n+1)-1} < n + 1 = s_{k_0(n+1)}.$$

Set

$$\alpha_n := (\lambda_{k_0(n)} n)^{\frac{1}{2k_1(n)}} \text{ for each } n \in \mathbb{N}. \quad (3.5)$$

**Claim 2.** For each  $n \in \mathbb{N}$ ,

$$1 > \lambda_{k_0(n)} s_{k_0(n)} = \lambda_{k_0(n)} n \geq 1 - \lambda_{k_0(n)}, \quad (3.6)$$

$$0 < \alpha_n < 1, \quad \text{and} \quad \alpha_n \rightarrow 1. \quad (3.7)$$

To see this, note that by definition and Claim 1,  $\lambda_{k_0(n)} n = \lambda_{k_0(n)} s_{k_0(n)} < 1$ , and  $1 \leq \lambda_{k_0(n)} (s_{k_0(n)} + 1)$ . But the latter inequality implies that  $1 - \lambda_{k_0(n)} \leq \lambda_{k_0(n)} s_{k_0(n)} = \lambda_{k_0(n)} n$ . Also,  $\lambda_{k_0(n)} n < 1$  implies that  $0 < \alpha_n < 1$ . Since  $\lambda_{k_0(n)} \rightarrow 0$ , relation (3.6) implies that  $\lambda_{k_0(n)} n \rightarrow 1$ . This, along with  $k_1(n) \rightarrow \infty$ , shows that  $\alpha_n \rightarrow 1$ , which completes the proof of Claim 2.

We note that the first two claims follow similarly to the proof given in [4]. However, at this point our approach will deviate significantly from that of [4].

**Claim 3.**  $\mu\{F^{-1}([1, \infty))\} = 0$ .

To see this, let  $S := F^{-1}[1, \infty)$  and  $y = U^{-1}(\chi_S)$ , where  $\chi_S$  denotes the characteristic function of  $S$ :  $\chi_S(t) = 1$  if  $t \in S$  and 0 otherwise. We must show that  $\mu(S) = 0$ . Since

$$\|y\| = \|U^{-1}(\chi_S)\| = \|\chi_S\| = \left( \int_S 1 d\mu \right)^{1/2} = [\mu(S)]^{1/2}, \quad (3.8)$$

it suffices to show that  $y = 0$ . Using (3.8), we have

$$\begin{aligned} \|Ty\| &= \|U^{-1}DUy\| = \|U^{-1}D(\chi_S)\| = \|U^{-1}(F\chi_S)\| = \|F\chi_S\| \\ &= \left[ \int_S F^2 d\mu \right]^{\frac{1}{2}} \geq \left[ \int_S 1 d\mu \right]^{\frac{1}{2}} = \|y\|. \end{aligned} \quad (3.9)$$

This shows that  $\|Ty\| \geq \|y\|$ . But since  $\|T\| = 1$  by Lemma 3.1, we must have  $\|Ty\| = \|y\|$ . We deduce that

$$\|y\| = \|Ty\| = \|P_AP_BP_Ay\| \leq \|P_BP_Ay\| \leq \|P_Ay\| \leq \|y\|. \quad (3.10)$$

Thus we must have equality holding throughout the string of inequalities (3.10). It follows (see, e.g., [6, Theorem 5.8(5), p. 76]) that  $y \in A \cap B = \{0\}$  and hence  $y = 0$ . This proves Claim 3.

During the course of proving (3.10), we should note that we have actually verified the following more general fact.

**Lemma 3.2** *If  $A, B$  are closed subspaces with  $A \cap B = \{0\}$ , then the non-expansive mapping  $P_BP_A$  is “strict,” i.e.,*

$$\|P_BP_A(x) - P_BP_A(y)\| < \|x - y\| \text{ for each } x, y \text{ with } x \neq y.$$

Consequently,  $P_AP_BP_A$  is also strict.

**Claim 4.** *For each  $\varepsilon > 0$ ,  $\mu\{F^{-1}((1 - \varepsilon, 1))\} > 0$ .*

Since  $F^{-1}(\{1\}) \subset F^{-1}([1, \infty))$ , it follows by Claim 3 that  $\mu\{F^{-1}(\{1\})\} = 0$ . Thus  $\mu\{F^{-1}((1 - \varepsilon, 1))\} = \mu\{F^{-1}((1 - \varepsilon, 1])\}$  for each  $\varepsilon > 0$ . But since  $\|F\|_\infty = 1$ , we must have  $\mu\{F^{-1}((1 - \varepsilon, 1])\} > 0$  for each  $\varepsilon > 0$ . This proves Claim 4.

**Claim 5.** *For each  $\varepsilon > 0$ , there exists  $\varepsilon_1 \in (0, \varepsilon)$  such that*

$$\mu\{F^{-1}((1 - \varepsilon, 1 - \varepsilon_1))\} > 0.$$

To verify this, it suffices to use Claim 4 and observe that  $F^{-1}((1 - \varepsilon, 1))$  is the countable union of the sets  $F^{-1}((1 - \varepsilon, 1 - \varepsilon/2^k))$ .

**Claim 6.** *There exists a sequence of reals  $(\beta_n) \subset (0, 1)$  such that  $\alpha_n^2 \leq \beta_n < \beta_{n+1} < 1$  and  $\mu\{F^{-1}([\beta_n, \beta_{n+1}))\} > 0$  for each  $n \in \mathbb{N}$ .*

We prove Claim 6 by induction. For  $n = 1$ , take  $\beta_1 = \alpha_1^2$ . Then  $\beta_1 < 1$ . Assume next that  $\beta_1, \dots, \beta_m$  have been chosen so that  $\beta_1 < \beta_2 < \dots < \beta_m < 1$ ,  $\beta_k \geq \alpha_k^2$  for  $k = 1, 2, \dots, m$ , and  $\mu\{F^{-1}([\beta_k, \beta_{k+1}))\} > 0$  for  $k = 1, 2, \dots, m - 1$ . Let  $\varepsilon := \min\{1 - \alpha_{m+1}^2, 1 - \beta_m\}$ . Then  $\varepsilon > 0$  and Claim 5 implies the existence of  $\varepsilon_1 \in (0, \varepsilon)$  such that  $\mu\{F^{-1}((1 - \varepsilon, 1 - \varepsilon_1))\} > 0$ . Let  $\beta_{m+1} := 1 - \varepsilon_1$ . Then  $\beta_{m+1} > 1 - \varepsilon \geq \beta_m$ . Also,  $\beta_{m+1} > 1 - \varepsilon \geq \alpha_{m+1}^2$ . Finally,  $\mu\{F^{-1}([\beta_m, \beta_{m+1}))\} \geq \mu\{F^{-1}([1 - \varepsilon, 1 - \varepsilon_1))\} > 0$ . This completes the induction step and hence the proof.

**Definition 3.3** With  $\beta_n$  given as in Claim 6, for each  $n \in \mathbb{N}$ , let  $S_n := F^{-1}([\beta_n, \beta_{n+1}))$  and define the vector  $e_n \in H$  and its image  $f_n$  by

$$e_n := \frac{1}{\sqrt{\mu(S_n)}} U^{-1}(\chi_{S_n}) \text{ and } f_n = Ue_n.$$

Note that

$$f_n = Ue_n = \frac{1}{\sqrt{\mu(S_n)}} \chi_{S_n}.$$

**Claim 7.**  $\|e_n\| = \|f_n\| = 1$  for each  $n \in \mathbb{N}$ .

This follows from

$$\|e_n\| = \|Ue_n\| = \|f_n\| = \frac{1}{\sqrt{\mu(S_n)}} \|\chi_{S_n}\| = 1. \quad (3.11)$$

It is convenient to list next a few basic and easily verified facts concerning powers of  $T$  and  $D$ .

**Claim 8.**

- (1)  $T^k = (U^{-1}DU)^k = U^{-1}D^kU$ .
- (2)  $D^k f = F^k f$  for all  $f \in L_2(\Omega, \mu)$ .
- (3) If  $f, g \in L_2(\Omega, \mu)$  and  $f(t)g(t) = 0$  for  $\mu$  almost all  $t$ , then  $\langle D^j f, D^k g \rangle = 0$  for every  $j, k \in \mathbb{N} \cup \{0\}$ .

**Claim 9.** For all integers  $j, k \in \mathbb{N} \cup \{0\}$  and  $m, n \in \mathbb{N}$  with  $m \neq n$ , we have

$$\langle T^j e_m, T^k e_n \rangle = 0.$$

In particular (taking  $j = k = 0$ ),  $\{e_n \mid n \in \mathbb{N}\}$  is an orthonormal set.

To verify this, note that (3.11) implies  $f_r = Ue_r = \frac{1}{\sqrt{\mu(S_r)}} \chi_{S_r}$  for each  $r \in \mathbb{N}$ . Then  $\chi_{S_n} \chi_{S_m} = \chi_{S_n \cap S_m} = 0$  since  $S_n \cap S_m = \emptyset$ . Thus  $f_n f_m = 0$ . Using statements (1) and (3) of Claim 8, we get that

$$\langle T^j e_m, T^k e_n \rangle = \langle U^{-1}D^j Ue_m, U^{-1}D^k Ue_n \rangle = \langle D^j f_m, D^k f_n \rangle = 0.$$

The last statement of the Claim follows from the first and Claim 7.

**Claim 10.**  $\beta_{n+1}^k \geq \|T^k e_n\| \geq \beta_n^k \geq \alpha_n^{2k}$  for all  $k, n \in \mathbb{N}$ .

The last inequality follows from Claim 6. Next observe that

$$\begin{aligned}\|T^k e_n\|^2 &= \|U^{-1} D^k U e_n\|^2 = \left\| D^k \left( \frac{\chi_{S_n}}{\sqrt{\mu(S_n)}} \right) \right\|^2 = \int \left[ D^k \left( \frac{\chi_{S_n}}{\sqrt{\mu(S_n)}} \right) \right]^2 d\mu \\ &= \frac{1}{\mu(S_n)} \int F^{2k} \chi_{S_n}^2 d\mu = \frac{1}{\mu(S_n)} \int_{S_n} F^{2k} d\mu.\end{aligned}$$

Also, by the definition of  $S_n$  (in Definition 3.3), it is clear that

$$\beta_n^{2k} \leq \frac{1}{\mu(S_n)} \int_{S_n} F^{2k} d\mu \leq \beta_{n+1}^{2k}.$$

Taking square roots completes the proof of Claim 10.

Now we can define the element which will converge slower than the sequence  $(\lambda_n)$ .

**Definition 3.4** *Set*

$$x_\lambda := \sum_1^\infty \frac{1}{n} e_n.$$

Since  $\sum_1^\infty 1/n^2 < \infty$  and  $\{e_n\}$  is an orthonormal set, it follows that  $x_\lambda$  is a well-defined element of  $H$ .

**Claim 11.**  $\|T^k x_\lambda\| \geq \alpha_n^{2k}/n$  for all  $n, k \in \mathbb{N}$ .

We deduce

$$\begin{aligned}\|T^k x_\lambda\|^2 &= \langle T^k x_\lambda, T^k x_\lambda \rangle = \left\langle T^k \left( \sum_{n=1}^\infty e_n/n \right), T^k \left( \sum_{m=1}^\infty e_m/m \right) \right\rangle \\ &= \sum_{n=1}^\infty \frac{1}{n} \sum_{m=1}^\infty \frac{1}{m} \langle T^k e_n, T^k e_m \rangle \\ &= \sum_{n=1}^\infty \frac{1}{n^2} \|T^k e_n\|^2 \quad (\text{by Claim 9}) \\ &\geq \frac{1}{n^2} \|T^k e_n\|^2 \quad \text{for each } n \\ &\geq \frac{\alpha_n^{4k}}{n^2} \quad (\text{by Claim 10}).\end{aligned}$$

Thus  $\|T^k x_\lambda\| \geq \alpha_n^{2k}/n$  as claimed.

**Claim 12.**  $\|(P_B P_A)^k x_\lambda\| \geq \lambda_k$  for each  $k \in \mathbb{N}$ .

Fix any  $k \in \mathbb{N}$  and let  $n = s_k$ . Using Claim 11, we get that

$$\|(P_B P_A)^k x_\lambda\| \geq \|P_A (P_B P_A)^k x_\lambda\| = \|T^k x_\lambda\| \geq \frac{\alpha_n^{2k}}{n} \geq \frac{\alpha_n^{2k_1(n)}}{n} = \lambda_{k_0(n)} \geq \lambda_k,$$

which proves Claim 12.

**Claim 13.** For each  $k \in \mathbb{N}$ ,  $(P_{M_2} P_{M_1})^k - P_M = (P_B P_A)^k$ .

This is a special case of [6, Lemma 9.30, p. 218].

Combining Claims 12 and 13, we immediately obtain

**Claim 14.**  $\|(P_{M_2} P_{M_1})^k(x_\lambda) - P_M(x_\lambda)\| \geq \lambda_k$  for each  $k \in \mathbb{N}$ .

This completes the proof of the second statement of Theorem 1.4.

## 4 Two errors in [4]

In this section, we point out two errors in [4]. We shall use the notation of [4]. (Note that this is the same as the notation of the present paper except that here we have used  $M_1, M_2$  instead of  $C_1, C_2$ .)

**First error.** The proof of the Claim in Step 2 of the proof of Theorem 5.7.16 in [4] has a mistake. The Claim itself is correct, only the proof of this claim is incorrect.

We recall the hypothesis.  $M_1$  and  $M_2$  are closed subspaces with  $M := M_1 \cap M_2$  such that  $M_1^\perp + M_2^\perp$  is not closed. Thus  $c(M_1, M_2) = 1$ . Let  $A = M_1 \cap M^\perp$  and  $B = M_2 \cap M^\perp$ . Note that  $A$  and  $B$  are closed subspaces with  $A \cap B = \{0\}$ . Moreover,

$$c(A, B) = c(M_1, M_2) = 1. \quad (4.1)$$

By Claim 2 in the proof of Lemma 3.1, we saw that  $0 < \alpha_n < 1$  and  $\alpha_n \rightarrow 1$ .

**Claim in Step 2 of [4, Theorem 5.7.16]** *There exist sequences  $(e'_n)$  in  $A$  and  $(f'_n)$  in  $B$  such that for all  $n, m$  with  $n \neq m$ :*

$$\langle e'_n, e'_m \rangle = \langle e'_n, f'_m \rangle = \langle f'_n, f'_m \rangle = 0, \quad (4.2)$$

$$\|e'_n\| \leq 1, \|f'_n\| \leq 1, \text{ and } \langle e'_n, f'_n \rangle \geq \alpha_n. \quad (4.3)$$

*Proof of Claim.* By (4.1), we deduce that

$$\sup\{\langle a, b \rangle \mid a \in A, b \in B, \|a\| \leq 1, \|b\| \leq 1\} = 1. \quad (4.4)$$

We proceed inductively. Clearly,  $e'_1$  and  $f'_1$  exist. Now suppose that we already found  $e'_1, \dots, e'_m$  and  $f'_1, \dots, f'_m$ . Set  $E := \text{span}\{e'_1, \dots, e'_m\}$  and

$F := \text{span}\{f'_1, \dots, f'_m\}$ . By (4.4) we can pick sequences  $(a_n)$  in  $A$  and  $(b_n)$  with  $\|a_n\| = \|b_n\| = 1$  and  $\langle a_n, b_n \rangle \rightarrow 1$ . By passing to a subsequence if necessary, we may assume that  $a_n \rightarrow a \in A$  weakly and  $b_n \rightarrow b \in B$  weakly. By expanding  $\|a_n - b_n\|^2$ , we see that  $a_n - b_n \rightarrow 0$  and hence that  $a = b = 0$ . Because  $E + F$  is *finite-dimensional*, the sum  $A^\perp + (E + F)$  is closed. Hence  $\{A^\perp, E + F\}$  is regular (by [3, Proposition 5.16]) and so is  $\{A^{\perp\perp}, (E + F)^\perp\} = \{A, E^\perp \cap F^\perp\}$  (again by [3, Proposition 5.16]). The following observation is now a consequence of the definition of regularity.

**Observation.** *If  $(z_n)$  is a bounded sequence with  $\max\{d(z_n, A), d(z_n, E^\perp \cap F^\perp)\} \rightarrow 0$ , then  $d(z_n, A \cap E^\perp \cap F^\perp) \rightarrow 0$ . (And analogously when  $A$  is replaced by  $B$ .)*

Note that  $P_{E+F}$  is a compact operator. (In [4], the error was in assuming only that  $P_E$  and  $P_F$  were compact, which was not sufficient.) Since  $a_n \rightarrow 0$  weakly and  $b_n \rightarrow 0$  weakly, we deduce that

$$P_{E+F}a_n \rightarrow 0 \quad \text{and} \quad P_{E+F}b_n \rightarrow 0. \quad (4.5)$$

Since  $(E + F)^\perp = E^\perp \cap F^\perp$ , this implies

$$a_n - P_{E^\perp \cap F^\perp}a_n \rightarrow 0 \quad \text{and} \quad b_n - P_{E^\perp \cap F^\perp}b_n \rightarrow 0; \quad (4.6)$$

or equivalently,

$$d(a_n, E^\perp \cap F^\perp) \rightarrow 0 \quad \text{and} \quad d(b_n, E^\perp \cap F^\perp) \rightarrow 0. \quad (4.7)$$

The above Observation now implies  $d(a_n, A \cap E^\perp \cap F^\perp) \rightarrow 0$  and  $d(b_n, B \cap E^\perp \cap F^\perp) \rightarrow 0$ ; equivalently,

$$a_n - P_{A \cap E^\perp \cap F^\perp}a_n \rightarrow 0 \quad \text{and} \quad b_n - P_{B \cap E^\perp \cap F^\perp}b_n \rightarrow 0. \quad (4.8)$$

Since  $\langle a_n, b_n \rangle \rightarrow 1$ , we deduce that

$$\langle P_{A \cap E^\perp \cap F^\perp}a_n, P_{B \cap E^\perp \cap F^\perp}b_n \rangle \rightarrow 1. \quad (4.9)$$

Thus, for all  $n$  sufficiently large, we have  $\|P_{A \cap E^\perp \cap F^\perp}a_n\| \leq 1$ ,  $\|P_{B \cap E^\perp \cap F^\perp}b_n\| \leq 1$ ,  $P_{A \cap E^\perp \cap F^\perp}a_n \in A \cap E^\perp \cap F^\perp$ ,  $P_{B \cap E^\perp \cap F^\perp}b_n \in B \cap E^\perp \cap F^\perp$ , and  $\langle P_{A \cap E^\perp \cap F^\perp}a_n, P_{B \cap E^\perp \cap F^\perp}b_n \rangle$  is as close to 1 (from below) as we like. Then for  $n$  sufficiently large, we can take  $e'_{m+1} = P_{A \cap E^\perp \cap F^\perp}a_n$  and  $f'_{m+1} = P_{B \cap E^\perp \cap F^\perp}b_n$ .

This proves the claim.

**Second error.** The second error is on the third line on page 32 of [4], where it is claimed that

$$C_1 = (C_1 \cap C_2) \oplus E \oplus (A \cap E^\perp \cap F^\perp), \quad C_2 = (C_1 \cap C_2) \oplus F \oplus (B \cap E^\perp \cap F^\perp). \quad (4.10)$$

Unfortunately, only

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

is true. This invalidates the rest of the proof in [4].

Here is a counterexample to (4.10). Let  $\{u_n \mid n \in \mathbb{N}\}$  be an orthonormal basis of a separable Hilbert space. Set

$$C_1 := \overline{\text{span}}\{u_{2n} + \frac{1}{n}u_{2n-1} \mid n \in \mathbb{N}\} \quad \text{and} \quad C_2 := \overline{\text{span}}\{u_{2n} + \frac{1}{n}u_{2n+1} \mid n \in \mathbb{N}\}.$$

Then

$$C_1 \cap C_2 = \{0\}. \quad (4.11)$$

(Sketch: the spanning vectors are orthogonal. Normalize and use Fourier expansions. Equate coefficients, compare odd and even ones. Deduce that they are all equal; thus they must be equal to 0.) Hence  $A = C_1$  and  $B = C_2$ . Set

$$e_n = e'_n = \rho_n(u_{4n} + \frac{1}{2n}u_{4n-1}) \quad \text{and} \quad f_n = f'_n = \rho_n(u_{4n} + \frac{1}{2n}u_{4n+1}), \quad (4.12)$$

where  $\rho_n := (1 + \frac{1}{4n^2})^{-1/2}$ . Since  $\langle e'_n, f'_n \rangle = (1 + \frac{1}{4n^2})^{-1}$ , the sequences  $(e'_n)$  and  $(f'_n)$  are as in the Claim of Step 2, and the sequences  $(e_n)$  and  $(f_n)$  are as in Step 3. Set

$$E = \overline{\text{span}}\{e_n \mid n \in \mathbb{N}\} \quad \text{and} \quad F = \overline{\text{span}}\{f_n \mid n \in \mathbb{N}\}. \quad (4.13)$$

Then

$$\overline{E + F} = \overline{\text{span}}\{2ne_{4n} + u_{4n-1}, 2ne_{4n} + u_{4n+1} \mid n \in \mathbb{N}\} \quad (4.14)$$

is a subspace of  $\overline{\text{span}}\{u_{4n-1}, u_{4n}, u_{4n+1} \mid n \in \mathbb{N}\}$ . Thus  $\{u_1, u_2, u_6, u_{10}, \dots\} \subset (E + F)^\perp$ . Since the orthogonal complement of  $\overline{\text{span}}\{2ne_{4n} + u_{4n-1}, 2ne_{4n} + u_{4n+1} \mid n \in \mathbb{N}\}$  in  $\overline{\text{span}}\{u_{4n-1}, u_{4n}, u_{4n+1} \mid n \in \mathbb{N}\}$  is  $\overline{\text{span}}\{-2nu_{4n-1} + u_{4n} - 2nu_{4n+1} \mid n \in \mathbb{N}\}$ , we obtain

$$\begin{aligned} E^\perp \cap F^\perp &= (E + F)^\perp \\ &= \overline{\text{span}}\{u_1, u_{4n-2}, -2nu_{4n-1} + u_{4n} - 2nu_{4n+1} \mid n \in \mathbb{N}\}. \end{aligned} \quad (4.15)$$

Consider the vector  $x := u_6 + \frac{1}{3}u_5$ . Then  $x$  belongs to  $C_1 = A$ . Since  $E \subset \overline{\text{span}}\{u_{4n-1}, u_{4n} \mid n \in \mathbb{N}\}$ , it follows that  $x \in E^\perp$  and hence  $P_E x = 0$ . Now consider the first term in the false statement (4.10), which in our present situation becomes

$$A = E \oplus (A \cap E^\perp \cap F^\perp). \quad (4.16)$$

This would imply that  $x$  belongs entirely to  $A \cap E^\perp \cap F^\perp$ . While it is true that  $x \in A \cap E^\perp$ , it is *not* true that  $x$  belongs to  $E^\perp \cap F^\perp$ . This can be verified using relation (4.15).

## References

- [1] N. Aronszajn, *Theory of reproducing kernels*, Trans. Amer. Math. Soc., **68**(1950), 337–403.
- [2] H. H. Bauschke and J. M. Borwein, *On the convergence of von Neumann's alternating projection algorithm for two sets*, Set-Valued Anal., **1**(1993), 185–212.
- [3] H. H. Bauschke and J. M. Borwein, *On projection algorithms for solving convex feasibility problems*, SIAM Review, **38**(1996), 367–426.
- [4] H. H. Bauschke, J. M. Borwein, and A. S. Lewis, *The method of cyclic projections for closed convex sets in Hilbert space*, Contemporary Mathematics, **204**, 1997, 1–38.
- [5] F. Deutsch, *Rate of convergence of the method of alternating projections*, ISNM, **72**, Birkhäuser Verlag, Basel, 1984, 96–107.
- [6] F. Deutsch, *Best Approximation in Inner Product Spaces*, Springer-Verlag, New York, 2001.
- [7] F. Deutsch and H. Hundal, *The rate of convergence for the method of alternating projections, II*, J. Math. Anal. Appl., **205**(1997), 381–405.
- [8] F. Deutsch and H. Hundal, *The rate of convergence for the cyclic projections algorithm I: Angles between convex sets*, J. Approx. Theory, **142**(2006), 36–55.
- [9] F. Deutsch and H. Hundal, *The rate of convergence for the cyclic projections algorithm II: Norms of nonlinear operators*, J. Approx. Theory, **142**(2006), 56–82.
- [10] F. Deutsch and H. Hundal, *Characterizing arbitrarily slow convergence for powers of linear operators*, submitted for publication, January 10, 2008, 15 pages.
- [11] C. Franchetti and W. Light, *On the von Neumann alternating algorithm in Hilbert space*, J. Math. Anal. Appl., **114**(1986), 305–314.
- [12] P. R. Halmos, *What does the spectral theorem say?*, Amer. Math. Monthly, **70**(1963), 241–247.
- [13] S. Kayalar and H. Weinert, *Error bounds for the method of alternating projections*, Math. Control Signals Systems, **1**(1988), 43–59.

- [14] J. von Neumann, *Functional Operators-Vol. II. The Geometry of Orthogonal Spaces*, Annals of Math. Studies #22, Princeton University Press, Princeton, NJ, 1950. [This is a reprint of mimeographed lecture notes first distributed in 1933.]
- [15] M. Reed and B. Simon, *Methods of Modern Mathematical Physics I: Functional Analysis*, Academic Press, New York, 1972.

Heinz H. Bauschke	Frank Deutsch	Hein Hundal
Mathematics	Department of Mathematics	146 Cedar Ridge Drive
UBC Okanagan	Penn State University	Port Matilda, PA 16870
Kelowna, British Columbia	University Park, PA 16802	
V1V 1V7, Canada	USA	USA
heinz.bauschke@ubc.ca	deutsch@math.psu.edu	hundalhh@yahoo.com