C. R. Acad. Sci. Paris, Ser. I 348 (2010) 53–56

Contents lists available at ScienceDirect

C. R. Acad. Sci. Paris, Ser. I

www.sciencedirect.com

Functional Analysis

# A generalization of the Friedrichs angle and the method of alternating projections

## *Une généralisation de l'angle de Friedrichs et la méthode des projections alternées*

Catalin Badea [a], Sophie Grivaux [a], Vladimir Müller [b]

[a] *Laboratoire Paul Painlevé, Université Lille 1, CNRS UMR 8524, 59655 Villeneuve d'Ascq, France*
[b] *Institute of Mathematics AV CR, Zitna 25, 115 67 Prague 1, Czech Republic*

ARTICLE INFO

*Article history:*
Received 26 October 2009
Accepted 25 November 2009
Available online 23 December 2009

Presented by Jean-Pierre Kahane

ABSTRACT

We present a generalization to an arbitrary number of subspaces of the cosine of the Friedrichs angle between two subspaces of a Hilbert space. This parameter is used to analyze the rate of convergence in the von Neumann–Halperin method of alternating projections.



RÉSUMÉ

On considère une généralisation à plusieurs espaces du cosinus de l'angle de Friedrichs entre deux sous-espaces d'un espace de Hilbert. On utilise ce paramètre pour analyser la vitesse de convergence dans la méthode des projections alternées de von Neumann–Halperin.



### 1. Introduction

Let $M_1, \ldots, M_N$ be $N \geqslant 2$ closed subspaces of a complex Hilbert space $H$, whose intersection is denoted by $M = M_1 \cap M_2 \cap \cdots \cap M_N$. Throughout this Note we will denote by $P_S$ the orthogonal projection onto the closed subspace $S$ of $H$. The method of considering the convergence of the iterates of the product $T = P_{M_N} \cdots P_{M_2} P_{M_1}$ in order to obtain $T^\infty = P_M$ is called the *method of alternating projections*. It was proved for $N = 2$ by J. von Neumann and in general for $N \geqslant 3$ by I. Halperin [7] that for each $x \in H$

$$\lim_{n\to\infty} \left\| (P_{M_N} \cdots P_{M_2} P_{M_1})^n x - P_M x \right\| = 0.$$

For $N = 2$ we have a quite complete description of the rate of convergence involving the *angle* of the two subspaces (see [5] and [9, Lecture VIII]).

**Definition 1.1** *(Friedrichs angle).* Let $H$ be a Hilbert space whose closed unit ball we denote by $B_H$, and let $M_1$ and $M_2$ be two closed subspaces of $H$ with intersection $M = M_1 \cap M_2$. 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

*E-mail addresses:* badea@math.univ-lille1.fr (C. Badea), grivaux@math.univ-lille1.fr (S. Grivaux), muller@math.cas.cz (V. Müller).

1631-073X/$ – see front matter 
doi:10.1016/j.crma.2009.11.018

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

It was proved by Aronszajn [1] (upper bound) and by Kayalar and Weinert [8] (equality) that

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

This formula shows that the sequence $(T^n)$ of iterates of $T = P_{M_2}P_{M_1}$ converges *uniformly* to $T^\infty = P_M$ if and only if $c(M_1, M_2) < 1$, that is if and only if the Friedrichs angle between $M_1$ and $M_2$ is positive. In this case the iterates of $T = P_{M_2}P_{M_1}$ converge "quickly" (as fast as a geometrical progression) to $T^\infty = P_M$, in the following sense:

(QUC) (*quick uniform convergence*) There exist a positive constant $C$ and an $\alpha \in ]0, 1[$ such that $\|T^n - T^\infty\| \leqslant C\alpha^n$ for any $n \geqslant 1$.

It is also known [5] that $c(M_1, M_2) < 1$ if and only if $M_1 + M_2$ is closed, if and only if $M_1^\perp + M_2^\perp$ is closed, if and only if $(M_1 \cap M^\perp) + (M_2 \cap M^\perp)$ is closed. When $M_1 + M_2$ is not closed, we have strong, but not uniform convergence of the iterates of $T$ to $T^\infty$. It was recently proved by Bauschke, Deutsch and Hundal (see [3] for the history of this result) that in this case we have arbitrarily slow convergence of the iterates of $T = P_{M_2}P_{M_1}$ to $T^\infty$:

(ASC) (*arbitrarily slow convergence*) For every $\varepsilon > 0$ and every sequence $(a_n)_{n\geqslant 1}$ of positive numbers such that $\lim_{n\to\infty} a_n = 0$, there exists a vector $x \in X$ such that $\|T^n x - T^\infty x\| \geqslant a_n$ for every $n \geqslant 1$.

Thus the iterates of the product of two projections converge either quickly, or arbitrarily slowly. We call this alternative the (QUC)/(ASC) dichotomy.

The results concerning the rate of convergence in Halperin's theorem for $N \geqslant 3$ are far less complete than the results described above for $N = 2$. We refer to [5,6], [4, Chapter 9], [10] and their references for several results concerning the rate of convergence in the method of alternating projections for $N \geqslant 3$. An interesting fact is pointed out in [6, Example 3.7]: for $N \geqslant 3$ the error of the method of cyclic alternating projections is not a function of the various Friedrichs angles $c(M_i, M_j)$ between pairs of subspaces.

The aim of this Note is to announce several results concerning the rate of convergence in Halperin's theorem for $N \geqslant 3$, including that the dichotomy (QUC)/(ASC) holds for $N \geqslant 2$, with several possible meanings for (ASC). We introduce a generalization of the cosine of the Friedrichs angle to several subspaces, $c(M_1, \ldots, M_N)$. Like for $N = 2$, the condition (QUC) holds if and only if $c(M_1, \ldots, M_N) < 1$. Estimates for the error $\|(P_{M_N}\cdots P_{M_2}P_{M_1})^n - P_M\|$ are given in this case as well as several statements equivalent to the condition $c(M_1, \ldots, M_N) < 1$. We refer the reader to [2] for complete proofs and for other related results.

## 2. Main results

The rate of convergence in the method of alternating projections for $N \geqslant 3$ is obtained in terms of the following parameter:

**Definition 2.1.** Let $M_1, \ldots, M_N$ be $N \geqslant 2$ closed subspaces of $H$ with intersection $M = M_1 \cap \cdots \cap M_N$. The *Friedrichs number* $c(M_1, \ldots, M_N)$ associated to these $N$ subspaces is defined as

$$c(M_1, \ldots, M_N) = \sup\left\{\frac{2}{N-1}\frac{\sum_{j<k}\operatorname{Re}\langle m_j, m_k\rangle}{\|m_1\|^2 + \cdots + \|m_N\|^2}: m_j \in M_j \cap M^\perp,\ \|m_1\|^2 + \cdots + \|m_N\|^2 \neq 0\right\}.$$

It is not hard to show that this definition coincides with the classical one for two subspaces, and that the number $c(M_1, \ldots, M_N)$ always lies between 0 and 1. The Friedrichs number of $M_1, \ldots, M_N$ is related to the projections $P_{M_1}, \ldots, P_{M_N}$ by the following formula:

$$c(M_1, \ldots, M_N) = \frac{N}{N-1}\left\|\frac{P_{M_1} + \cdots + P_{M_N}}{N} - P_M\right\| - \frac{1}{N-1}.$$

We announce the following results which provide a fairly complete description of the rate of convergence in the method of alternating projections.

**Theorem 2.2.** *Let $M_1, \ldots, M_N$ be $N \geqslant 2$ closed subspaces of a complex Hilbert space $H$ with intersection $M = M_1 \cap M_2 \cap \cdots \cap M_N$. Let $T = P_{M_N}P_{M_{N-1}}\cdots P_{M_1}$ denote the product of the projections on these successive subspaces. Then the following alternative holds: either the range* $\operatorname{Ran}(T - I)$ *is closed and the iterates $T^n$ converge uniformly to $T^\infty = P_M$, and in this case (QUC) holds, or* $\operatorname{Ran}(T - I)$ *is not closed and then the convergence of $T^n$ to $T^\infty$ is arbitrarily slow in each one of the following senses:*

(ASC1: *arbitrarily slow convergence, variant* 1) *For every* $\varepsilon > 0$ *and every sequence* $(a_n)_{n\geqslant 1}$ *of positive numbers such that* $\lim_{n\to\infty} a_n = 0$, *there exists a vector* $x \in X$ *such that* $\|x\| < \sup_n a_n + \varepsilon$ *and* $\|T^n x - T^\infty x\| \geqslant a_n$ *for all* $n$;

(ASC2: *arbitrarily slow convergence, variant* 2) *For every sequence* $(a_n)_{n\geqslant 1}$ *of positive numbers such that* $\lim_{n\to\infty} a_n = 0$, *there exists a* dense *subset of points* $x \in X$ *such that* $\|T^n x - T^\infty x\| \geqslant a_n$ *for all but a* finite *number of* $n$*'s*;

(ASCH: *arbitrarily slow convergence, Hilbertian version*) *For every* $\varepsilon > 0$ *and every sequence* $(a_n)_{n\geqslant 1}$ *of positive numbers such that* $\lim_{n\to\infty} a_n = 0$, *there exists a vector* $x \in H$ *such that* $\|x\| < \sup_n a_n + \varepsilon$ *and* $\operatorname{Re}\langle T^n x - T^\infty x, x\rangle \geqslant a_n$ *for all* $n \geqslant 1$;

(ASCHR: *arbitrarily slow convergence for random products*) *For every* $\varepsilon > 0$, *every sequence* $(a_n)_{n\geqslant 0}$ *of positive reals with* $\lim_{n\to\infty} a_n = 0$, *and every sequence of indices* $(i_k)_{k\geqslant 1}$ *in* $\{1, 2, \ldots, N\}$, *there exists* $x \in H$ *with* $\|x\| < \sup_n a_n + \varepsilon$ *such that* $\operatorname{Re}\langle P_{M_{i_n}} \cdots P_{M_{i_1}} - P_M x, x\rangle \geqslant a_n$ *for each* $n \geqslant 1$.

More precisely, the next theorem characterizes in several ways when the above dichotomy occurs. We denote by $\sigma(A)$ the spectrum of $A$, by $\|\cdot\|_e$ the *essential norm* and by $\sigma_e$ the *essential spectrum*.

**Theorem 2.3.** *With the same notation as in the previous theorem, the following assertions are equivalent*:

(1) *the range* $\operatorname{Ran}(T - I)$ *of* $T - I$ *is not closed*;

(1′) *for every* $k \geqslant N$ *and every sequence of indices* $(i_k)_{k\geqslant 1}$ *such that* $\{i_1, \ldots, i_k\} = \{1, 2, \ldots, N\}$, $\operatorname{Ran}(P_{M_{i_k}} \cdots P_{M_{i_1}} - I)$ *is not closed*;

(2) *one/all of the conditions* (ASC1), (ASC2), (ASCH) *hold for* $T$;

(2′) *the condition* (ASCHR) *holds*;

(3) $c(M_1, \ldots, M_N) = 1$;

(4) *for every* $\varepsilon > 0$, *every closed subspace* $K \subset M^\perp$ *of finite codimension* (*in* $M^\perp$), *there exists a vector* $x \in K$ *such that* $\|x\| = 1$ *and* $\max\{\operatorname{dist}(x, M_j)\colon j = 1, \ldots, N\} < \varepsilon$;

(5) $1 \in \sigma(T - P_M)$;

(5′) *for every* $k$ *and every* $i_1, \ldots, i_k \in \{1, 2, \ldots, N\}$ *we have* $1 \in \sigma(P_{M_{i_k}} \cdots P_{M_{i_1}} - P_M)$;

(6) $\|T - P_M\| = 1$;

(6′) *for every* $k$ *and every sequence of indices* $(i_k)_{k\geqslant 1}$ *with* $\{i_1, \ldots, i_k\} = \{1, 2, \ldots, N\}$ *we have* $\|P_{M_{i_k}} \cdots P_{M_{i_1}} - P_M\| = 1$;

(7) $\|T - P_M\|_e = 1$;

(7′) *for every* $k$ *and every* $i_1, \ldots, i_k \in \{1, 2, \ldots, N\}$ *we have* $\|P_{M_{i_k}} \cdots P_{M_{i_1}} - P_M\|_e = 1$;

(8) $1 \in \sigma_e(T - P_M)$;

(8′) *for every* $k$, *every* $i_1, \ldots, i_k \in \{1, 2, \ldots, N\}$ *we have* $1 \in \sigma_e(P_{M_{i_k}} \cdots P_{M_{i_1}} - P_M)$;

(9) *for every* $\varepsilon > 0$, *every closed subspace* $K \subset M^\perp$ *of finite codimension* (*in* $M^\perp$), *there exists a vector* $x \in K$ *such that* $\|Tx - x\| \leqslant \varepsilon$;

(9′) *for every* $\varepsilon > 0$, *every closed subspace* $K \subset M^\perp$ *of finite codimension* (*in* $M^\perp$), *there exists a vector* $x \in K$ *such that* $\|P_{M_{i_k}} \cdots P_{M_{i_1}} x - x\| \leqslant \varepsilon$ *for every* $k$ *and every* $i_1, \ldots, i_k \in \{1, 2, \ldots, N\}$;

(10) *the sum of* $\operatorname{diag}(M_1) := \{(y, \ldots, y)\colon y \in M_1\} \subset H^{N-1}$ *and* $M_2 \oplus \cdots \oplus M_N \subset H^{N-1}$ *is not closed in* $H^{N-1}$ (*and equivalent statements for* $\operatorname{diag}(M_j) \subset H^{N-1}$, $2 \leqslant j \leqslant N$);

(11) $M_1^\perp + \cdots + M_N^\perp$ *is not closed in* $H$.

The conditions (1), (2), (5), (6), (7), (8) and (9) (most of which are of spectral nature) are conditions about the product $T = P_{M_N} \cdots P_{M_1}$, while the corresponding conditions denoted with primes are analog conditions concerning *random products* $P_{M_{i_N}} \cdots P_{M_{i_1}}$. The conditions (3), (4), (10) and (11) concern the geometry of subspaces. The last theorem gives an estimate on the rate of convergence when we have (QUC) in terms of the Friedrichs parameter:

**Theorem 2.4.** *Let* $M_1, \ldots, M_N$ *be* $N \geqslant 2$ *closed subspaces of* $H$ *with intersection* $M = M_1 \cap M_2 \cap \cdots \cap M_N$, *and* $T = P_{M_N} \cdots P_{M_1}$. *Suppose that* $c := c(M_1, \ldots, M_N) < 1$. *Then we have quick uniform convergence of the powers* $T^n$ *to* $P_M$, *and more precisely*

$$\|T^n - P_M\| \leqslant \left(1 - \left(\frac{1-c}{4N}\right)^2\right)^{n/2} \quad \text{for any } n \geqslant 1.$$

## Acknowledgements

This work has been partially supported by ANR Project Blanc DYNOP, Grant No. 201/06/0128 of GA ČR and Institutional Research Plan AV 0Z 10190503.

## References

[1] N. Aronszajn, Theory of reproducing kernels, Trans. Amer. Math. Soc. 68 (1950) 337–404.
[2] C. Badea, S. Grivaux, V. Müller, The rate of convergence in the method of alternating projections, preprint.
[3] H.H. Bauschke, F. Deutsch, H. Hundal, Characterizing arbitrarily slow convergence in the method of alternating projections, arXiv:0710.2387.
[4] F. Deutsch, Best Approximation in Inner Product Spaces, CMS Books in Mathematics, vol. 7, Springer, New York, 2001.

[5] F. Deutsch, Rate of convergence of the method of alternating projections, in: Parametric Optimization and Approximation, in: Internat. Schriftenreihe Numer. Math., vol. 72, Birkhäuser, Basel, 1985, pp. 96–107.
[6] F. Deutsch, H. Hundal, The rate of convergence for the method of alternating projections. II, J. Math. Anal. Appl. 205 (1997) 381–405.
[7] I. Halperin, The product of projection operators, Acta Sci. Math. (Szeged) 23 (1962) 96–99.
[8] S. Kayalar, H. Weinert, Error bounds for the method of alternating projections, Math. Control Signals Systems 1 (1988) 43–59.
[9] N. Nikolski, Treatise on the Shift Operator. Spectral Function Theory (Translated from the Russian by Jaak Peetre), Grundlehren der Mathematischen Wissenschaften, vol. 273, Springer-Verlag, Berlin, 1986.
[10] J. Xu, L. Zikatanov, The method of alternating projections and the method of subspace corrections in Hilbert space, J. Amer. Math. Soc. 15 (2002) 573–597.