Electronic Transactions on Numerical Analysis.
Volume 45, pp. 330–341, 2016.

ISSN 1068–9613.

ETNA
Kent State University
http://etna.math.kent.edu

A NEW GEOMETRIC ACCELERATION OF THE VON NEUMANN-HALPERIN PROJECTION METHOD*

WILLIAMS LÓPEZ†

Abstract. We develop a geometrical acceleration scheme for the von Neumann-Halperin alternating projection method, when applied to the problem of finding the projection of a point onto the intersection of a finite number of closed subspaces of a Hilbert space. We study the convergence properties of the new scheme. We also present some encouraging preliminary numerical results to illustrate the performance of the new scheme when compared with a well-known geometrical acceleration scheme, and also with the original von Neumann-Halperin alternating projection method.

Key words. von Neumann-Haperin algorithm, alternating projection methods, orthogonal projections, acceleration schemes

AMS subject classifications. 52A20, 46C07, 65H10, 47J25

1. Introduction. Let H be a Hilbert space with inner product $\langle .,. \rangle$. For a given $x_0 \in H$ and many closed subspaces $M_1, \ldots, M_p$ in H, we consider the best approximation problem: find the closest point to x_0 in $M = \cap_{i=1}^{p} M_i$, which can be stated as an optimization problem as follows:

$$\text{minimize } \|x_0 - x\| \quad \text{subject to} \quad x \in M, \tag{1.1}$$

where, for any $z \in H$, $\|z\|^2 = \langle z, z \rangle$. The unique solution x^* of problem (1.1) is called the projection of x_0 onto M and is denoted as $P_M(x_0)$.

In 1933, von Neumann solved problem (1.1) for the particular case of two closed subspaces.

THEOREM 1.1 (von Neumann [28]). *If M_1 and M_2 are closed subspaces in H, then for each $x_0 \in H$,*

$$\lim_{k\to\infty} (P_{M_2} P_{M_1})^k (x_0) = P_{M_1 \cap M_2}(x_0).$$

Figure 1.1 shows the geometric interpretation of Theorem 1.1. The extension to more than two subspaces was established in 1962 by Halperin.

THEOREM 1.2 (I. Halperin [23]). *If $M_1, \ldots, M_p$ are closed subspaces in H, then for each $x_0 \in H$,*

$$\lim_{k\to\infty} (P_{M_p} P_{M_{p-1}} \cdots P_{M_1})^k (x_0) = P_{\cap_{i=1}^{p} M_i}(x_0).$$

Theorem 1.2 suggests an algorithm, called the method of alternating projections (or MAP for short); see [11, 16], which can be described as follows: for any $x_0 \in H$, set

$$\begin{aligned} x_0^k &= x_p^{k-1} \\ x_i^k &= P_{M_i}(x_{i-1}^k) \qquad i = 1, 2, \ldots, p, \end{aligned} \tag{1.2}$$

for $k \in Z^+$, with initial value $x_p^0 = x_0$. The MAP is closely related to Kaczmarz alternating projection method [25] for solving linear systems of equations.

*Received October 27, 2015. Accepted June 9, 2016. Published online on September 2, 2016. Recommended by Fiorella Sgallari.

†Departamento de Física y Matemáticas, Núcleo Universiraio "Rafael Rangel", Universidad de Los Andes, Trujillo, Venezuela (`wlopez@ula.ve`).

ETNA
Kent State University
http://etna.math.kent.edu

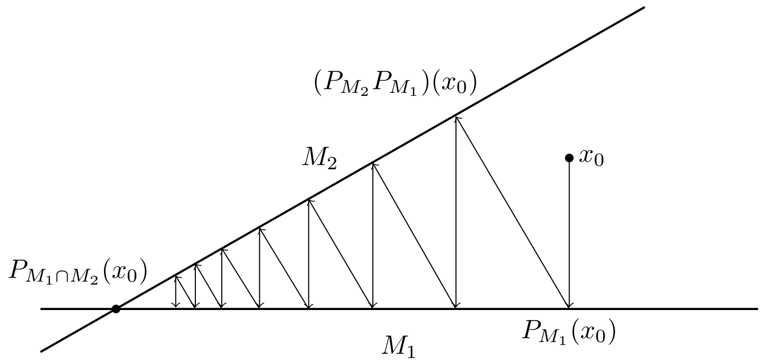

FIG. 1.1. $\lim_{k\to\infty}(P_{M_2}P_{M_1})^k(x_0) = P_{M_1\cap M_2}(x_0)$.

THEOREM 1.3. *For any $i = 1, 2, \ldots, p$, the sequence $\{x_i^k\}$ generated by (1.2) converges to $P_M(x_0)$.*

Proof. Let $i \in \{1, 2, \ldots, p\}$. Then $x_i^{k+1} = (P_{M_i}P_{M_{i-1}}\cdots P_{M_1})(x_p^k)$, for all $k \geq 0$, where $x_p^k = (P_{M_p}P_{M_{p-1}}\cdots P_{M_1})^k(x_0)$. From Theorem 1.2 we have that $x_p^k \to P_M(x_0)$ when $k \to \infty$ and since $P_{M_i}, P_{M_{i-1}}, \ldots, P_{M_1}$ are bounded linear operators, then

$$x_i^{k+1} = (P_{M_i}P_{M_{i-1}}\cdots P_{M_1})(x_p^k) \to (P_{M_i}P_{M_{i-1}}\cdots P_{M_1})(P_M(x_0)) = P_M(x_0)$$

when $k \to \infty$. ∎

Applications of MAP in various areas of mathematics can be found in [16, 9, 10, 3, 7, 8, 22]. The MAP has an r-linear rate of convergence that can be very slow when the angles between the subspaces (in the sense of Friedrichs [20]) are small; see, e.g., [16]. Franchetti and Light [19] showed that the convergence in Theorem 1.1 may be arbitrarily slow if $M_1 + M_2$ is not closed (i.e., if the angle between M_1 and M_2 is zero). Consequently, several acceleration schemes have been proposed; see, e.g., [4, 21, 24, 12]. Motivated by a recent work of López and Raydan [26], in this work we present and analyze a geometrical acceleration scheme for MAP in (1.2), which solves problem (1.1). The new scheme involves finding points closest to a solution (x^*, x^*) in the product space $H \times H$, where x^* is the solution of (1.1).

The rest of this paper is organized as follows. In Section 2 we provide information about already existing acceleration schemes for MAP. In Section 2.1 we develop the new acceleration scheme and discuss its theoretical properties. In Section 3 we present encouraging preliminary numerical results.

2. Accelerations for MAP.

The sequences generated by the method of alternating projections often converge very slowly. Such a slow convergence is observed, e.g., if M_1 and M_2 are two closed subspaces with angles close to 0 (see, e.g., [19]). Since the method of alternating projections has many practical applications, any acceleration technique seems to be important.

Several acceleration schemes with a geometrical flavor have been proposed to improve the performance of MAP; see, e.g., De Pierro and Iusem [13], Dos Santos [14], Gearhart and Koshy [21], Bauschke et al. [4], Gubin, Polyak and Raik [22], Martínez [27], Appleby and Smolarski [1] and Censor [6]. In Section 2.1, we describe a geometrically appealing acceleration scheme following the presentation in [21]. Some other different acceleration ideas have also been developed (see, e.g., Echebest et al. [17, 18] and Scolnik et al. [29, 30]) based on the so-called projected aggregation method (PAM). Other specialized acceleration scheme ideas have been developed by Hernández-Ramos et. al. [24] based on the use of the conjugate gradient method for minimizing a related convex quadratic map. Usually vector extrapolation algorithms are used for accelerating the convergence of MAP; see, e.g., [3, 8, 7].

Dykstra's algorithm [15, 5] solves problem (1.1) when the involved sets are closed and convex (not necessarily closed subspaces). Dykstra's algorithm can be viewed as a natural extension to convex sets of von Neumann-Halperin's MAP for subspaces in a Hilbert space. The sequences generated by Dykstra's algorithm often converge very slowly. Such a slow convergence is observed, e.g., if M_1 is a hyperplane and M_2 is a ball which is tangent to M_1; see, e.g., [2, Example 5.3]. An acceleration scheme has been developed for Dykstra's algorithm [26]. In Section 2.1 we develop a new acceleration scheme for von Neumann-Halperin's MAP on closed subspaces, which is related to the scheme given in [26].

2.1. A new acceleration scheme. We now discuss an acceleration scheme for MAP to solve problem (1.1). For that we need to consider an auxiliary sequence in the product space $H \times H$ that will be denoted as H^2. If H is a Hilbert space with inner product $\langle ., . \rangle$, we will use the inner product in H^2 defined by $\langle (x,y),(w,z) \rangle = \langle x,w \rangle + \langle y,z \rangle$, for all $(x,y),(w,z) \in H^2$ and norm $\|(x,y)\|^2 = \langle (x,y),(x,y) \rangle = \|x\|^2 + \|y\|^2$, for all $(x,y) \in H^2$. Thus H^2 is a Hilbert space. For each $k \in \mathbb{Z}^+$, let us define

$$\widehat{x}_k = (x_{p-1}^k, x_p^k) \in H^2, \tag{2.1}$$

where x_{p-1}^k and x_p^k are defined in (1.2). Let us also denote

$$\widehat{x}^* = (P_M(x_0), P_M(x_0)) \in H^2.$$

It follows that $\{\widehat{x}_k\}$ is a sequence in H^2 that converges to $\widehat{x}^*$. Our idea to accelerate the convergence of MAP consists in building another sequence in H^2 which converges faster to $\widehat{x}^*$ than $\{\widehat{x}_k\}$. For that we need to define a suitable subspace Π of H^2 that contains $\widehat{x}^*$, as follows:

$$\Pi = \{(x,x) : x \in H\}.$$

Clearly, Π is a closed subspace in H^2 and $\widehat{x}^* \in \Pi$. Let us denote by $P_\Pi(\widehat{x})$ the orthogonal projection of $\widehat{x} \in H^2$ onto Π. For each $k \in \mathbb{Z}^+$, let us consider $P_\Pi(\widehat{x}_k) \in \Pi$, where $\{\widehat{x}_k\}$ is given by (2.1).

THEOREM 2.1. *For all $k \geq 1$, $\|P_\Pi(\widehat{x}_k) - \widehat{x}^*\| \leq \|\widehat{x}_k - \widehat{x}^*\|$.*

Proof. For all $k \geq 1$, since $\widehat{x}_k - P_\Pi(\widehat{x}_k)$ is orthogonal to Π, and since $P_\Pi(\widehat{x}_k)$ and $\widehat{x}^*$ belong to Π, then $\widehat{x}_k - P_\Pi(\widehat{x}_k)$ is orthogonal to $P_\Pi(\widehat{x}_k) - \widehat{x}^*$. Hence

$$\|\widehat{x}_k - \widehat{x}^*\|^2 = \|\widehat{x}_k - P_\Pi(\widehat{x}_k) + P_\Pi(\widehat{x}_k) - \widehat{x}^*\|^2 = \|\widehat{x}_k - P_\Pi(\widehat{x}_k)\|^2 + \|P_\Pi(\widehat{x}_k) - \widehat{x}^*\|^2.$$

Consequently, $\|\widehat{x}_k - \widehat{x}^*\|^2 \geq \|P_\Pi(\widehat{x}_k) - \widehat{x}^*\|^2$, thus $\|P_\Pi(\widehat{x}_k) - \widehat{x}^*\| \leq \|\widehat{x}_k - \widehat{x}^*\|$. ∎

The sequence $\{P_\Pi(\widehat{x}_k)\}$ will play a key role in the development of the acceleration scheme for MAP.

REMARK 2.2. If $(x,y) \in H^2$, then (see, e.g., [26])

$$P_\Pi(x,y) = 1/2(x+y, x+y). \tag{2.2}$$

Notice that computing the projection onto Π only requires us to compute the average of the two involved vectors, i.e., it requires a very inexpensive calculation.

Let $\{\widehat{x}_k\}$ be the sequence defined by (2.1). For any $k \geq 1$, with $P_\Pi(\widehat{x}_{k+1}) \neq P_\Pi(\widehat{x}_k)$, let $\widehat{o}_k \in \Pi$ be the projection of $\widehat{x}^*$ onto the line that goes through $P_\Pi(\widehat{x}_{k+1})$ and $P_\Pi(\widehat{x}_k)$ (see

ETNA
Kent State University
http://etna.math.kent.edu

Figure 2.1). Then $\langle \widehat{o}_k - \widehat{x}^*, P_\Pi(\widehat{x}_{k+1}) - P_\Pi(\widehat{x}_k)\rangle = 0$ and there exists a unique $\alpha_k \in \mathbb{R}$ such that

$$\widehat{o}_k = P_\Pi(\widehat{x}_k) + \alpha_k(P_\Pi(\widehat{x}_{k+1}) - P_\Pi(\widehat{x}_k)).$$

Therefore

$$\langle P_\Pi(\widehat{x}_k) - \widehat{x}^* + \alpha_k(P_\Pi(\widehat{x}_{k+1}) - P_\Pi(\widehat{x}_k)), P_\Pi(\widehat{x}_{k+1}) - P_\Pi(\widehat{x}_k)\rangle = 0. \tag{2.3}$$

Solving for α_k, from (2.3), gives

$$\alpha_k = \frac{\langle P_\Pi(\widehat{x}_k) - \widehat{x}^*, P_\Pi(\widehat{x}_k) - P_\Pi(\widehat{x}_{k+1})\rangle}{\|P_\Pi(\widehat{x}_{k+1}) - P_\Pi(\widehat{x}_k)\|^2}. \tag{2.4}$$

In the following result we will give a formula for α_k in (2.4) that does not require knowledge of $\widehat{x}^*$.

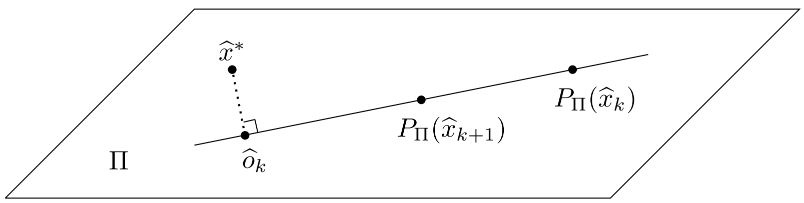

FIG. 2.1. $\langle \widehat{o}_k - \widehat{x}^*, P_\Pi(\widehat{x}_{k+1}) - P_\Pi(\widehat{x}_k)\rangle = 0$.

THEOREM 2.3. *Let $\{\widehat{x}_k\}$ be defined by (2.1). Then $\langle \widehat{x}^*, P_\Pi(\widehat{x}_k) - P_\Pi(\widehat{x}_{k+1})\rangle = 0$ for all $k \geq 1$.*

Proof. Let $k \geq 1$. Since $\widehat{x}^* \in \Pi$ and the projection P_Π is self-adjoint, we obtain

$$\begin{aligned}\langle \widehat{x}^*, P_\Pi(\widehat{x}_k) - P_\Pi(\widehat{x}_{k+1})\rangle &= \langle \widehat{x}^*, \widehat{x}_k - \widehat{x}_{k+1}\rangle \\ &= \left\langle (x^*, x^*), (x^k_{p-1} - x^{k+1}_{p-1}, x^k_p - x^{k+1}_p)\right\rangle \\ &= \left\langle x^*, x^k_{p-1} - x^{k+1}_{p-1}\right\rangle + \left\langle x^*, x^k_p - x^{k+1}_p\right\rangle.\end{aligned}$$

On the other hand, since $x^* \in M = \cap_{i=1}^p M_i$ and the projections P_{M_i} are self-adjoint, we have

$$\begin{aligned}\left\langle x^*, x^k_p - x^{k+1}_p\right\rangle &= \left\langle x^*, x^k_p\right\rangle - \left\langle x^*, x^{k+1}_p\right\rangle \\ &= \left\langle x^*, x^k_p\right\rangle - \left\langle x^*, (P_{M_p}P_{M_{p-1}}\dots P_{M_1})(x^k_p)\right\rangle \\ &= \left\langle x^*, x^k_p\right\rangle - \left\langle x^*, x^k_p\right\rangle = 0.\end{aligned}$$

Similarly, we also obtain $\left\langle x^*, x^k_{p-1} - x^{k+1}_{p-1}\right\rangle = 0$. □

COROLLARY 2.4. *Let α_k, for some k, be given by (2.4). Then*

$$\alpha_k = \frac{\langle P_\Pi(\widehat{x}_k), P_\Pi(\widehat{x}_k) - P_\Pi(\widehat{x}_{k+1})\rangle}{\|P_\Pi(\widehat{x}_{k+1}) - P_\Pi(\widehat{x}_k)\|^2}. \tag{2.5}$$

Proof. This is a consequence of Theorem 2.3 and (2.4). □

ETNA
Kent State University
http://etna.math.kent.edu

Let us define the sequence $\{\widehat{o}_k\}$ in Π, for $k \geq 1$, as follow

$$\widehat{o}_k = \begin{cases} P_\Pi(\widehat{x}_{k+1}) & \text{if } \|P_\Pi(\widehat{x}_{k+1}) - P_\Pi(\widehat{x}_k)\| = 0, \\ P_\Pi(\widehat{x}_k) + \alpha_k(P_\Pi(\widehat{x}_{k+1}) - P_\Pi(\widehat{x}_k)) & \text{if } \|P_\Pi(\widehat{x}_{k+1}) - P_\Pi(\widehat{x}_k)\| \neq 0, \end{cases} \tag{2.6}$$

where α_k is given by (2.5).

THEOREM 2.5. *Let $\{\widehat{o}_k\}$ be given by* (2.6). *Then $\|\widehat{o}_k - \widehat{x}^*\| \leq \|\widehat{x}_{k+1} - \widehat{x}^*\|$ for all $k \geq 1$.*

Proof. Let $k \geq 1$. If $\|P_\Pi(\widehat{x}_{k+1}) - P_\Pi(\widehat{x}_k)\| = 0$, then from Theorem 2.1 and (2.6) it follows that $\|\widehat{o}_k - \widehat{x}^*\| = \|P_\Pi(\widehat{x}_{k+1}) - \widehat{x}^*\| \leq \|\widehat{x}_{k+1} - \widehat{x}^*\|$. On the other hand, if $\|P_\Pi(\widehat{x}_{k+1}) - P_\Pi(\widehat{x}_k)\| \neq 0$, then $\widehat{o}_k$ is chosen to minimize the distance to $\widehat{x}^*$ along the line connecting $P_\Pi(\widehat{x}_{k+1})$ and $P_\Pi(\widehat{x}_k)$. Therefore $\|\widehat{o}_k - \widehat{x}^*\| \leq \|P_\Pi(\widehat{x}_{k+1}) - \widehat{x}^*\|$ and, from Theorem 2.1, it follows that $\|\widehat{o}_k - \widehat{x}^*\| \leq \|\widehat{x}_{k+1} - \widehat{x}^*\|$. $\square$

Theorem 2.5 guarantees than the sequence $\{\widehat{o}_k\}$ defined by (2.6) accelerates the convergence of the sequence $\{\widehat{x}_k\}$ defined by (2.1).

REMARK 2.6. For the particular case of closed subspaces, the sequence $\{\widehat{o}_k\}$ defined by (2.6) accelerates the convergence of the sequence $\{\widehat{z}_k\}$ defined in [26]. For any $k \geq 1$, $\widehat{z}_k$ defined in [26] is the intersection of the subspace Π with the line that goes through $\widehat{x}_k$ and $\widehat{x}_{k+1}$ (see Figure 2.2). Consequently

$$\widehat{z}_k = \widehat{x}_k + \alpha'_k(\widehat{x}_{k+1} - \widehat{x}_k) \in \Pi,$$

for some $\alpha'_k \in \mathbb{R}$. Since $\widehat{z}_k \in \Pi$ and P_Π is a linear operator,

$$\widehat{z}_k = P_\Pi(\widehat{z}_k) = P_\Pi(\widehat{x}_k) + \alpha'_k(P_\Pi(\widehat{x}_{k+1}) - P_\Pi(\widehat{x}_k)).$$

Therefore we have that $\widehat{z}_k$ belongs to the line connecting $P_\Pi(\widehat{x}_{k+1})$ and $P_\Pi(\widehat{x}_k)$. Since $\widehat{o}_k$ is chosen to minimize the distance to $\widehat{x}^*$ along the line connecting $P_\Pi(\widehat{x}_{k+1})$ and $P_\Pi(\widehat{x}_k)$, $\|\widehat{o}_k - \widehat{x}^*\| \leq \|\widehat{z}_k - \widehat{x}^*\|$ (see Figure 2.2).

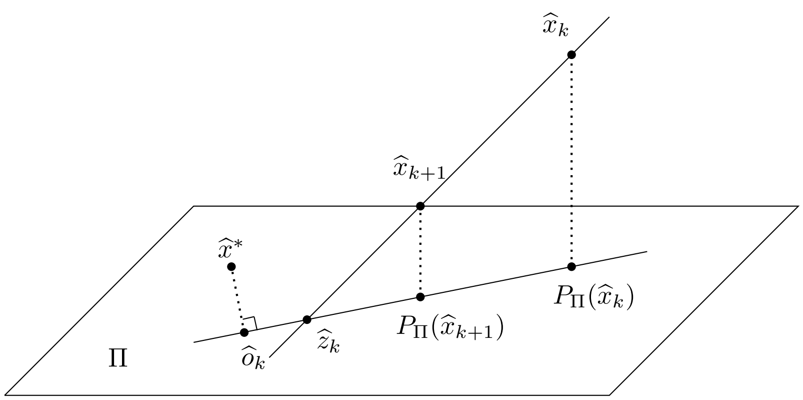

FIG. 2.2. $\widehat{o}_k$ *vs.* $\widehat{z}_k$.

REMARK 2.7. Since the sequence $\{\widehat{o}_k\}$ defined by (2.6) belongs to Π, for any $k \geq 1$, there exists a unique $o_k \in H$ such that

$$\widehat{o}_k = (o_k, o_k) \in \Pi. \tag{2.7}$$

THEOREM 2.8. *Let* $\{o_k\}$ *the sequence defined by* (2.7). *Then, for each* $k \geq 1$,

$$\|o_k - P_M(x_0)\| \leq \max\{\|x_{p-1}^{k+1} - P_M(x_0)\|, \|x_p^{k+1} - P_M(x_0)\|\}.$$

Proof. Let $k \geq 1$. From Theorem 2.5, we have that

$$\begin{aligned}
\sqrt{2}\|o_k - P_M(x_0)\| &= \sqrt{\|o_k - P_M(x_0)\|^2 + \|o_k - P_M(x_0)\|^2} \\
&= \|\widehat{o}_k - \widehat{x}^*\| \\
&\leq \|\widehat{x}_{k+1} - \widehat{x}^*\| \\
&= \sqrt{\|x_{p-1}^{k+1} - P_M(x_0)\|^2 + \|x_p^{k+1} - P_M(x_0)\|^2} \\
&\leq \sqrt{2\max\{\|x_{p-1}^{k+1} - P_M(x_0)\|^2, \|x_p^{k+1} - P_M(x_0)\|^2\}} \\
&= \sqrt{2}\max\{\|x_{p-1}^{k+1} - P_M(x_0)\|, \|x_p^{k+1} - P_M(x_0)\|\},
\end{aligned}$$

and the result is established. ☐

COROLLARY 2.9. *The sequence* $\{o_k\}$ *defined by* (2.7) *converges to* $P_M(x_0)$.

Proof. This is a consequence of Theorem 2.8 and Theorem 1.3. ☐

REMARK 2.10. Observe that the sequence $\{\widehat{o}_k\}$ in H^2 accelerates the convergence of the sequence $\{\widehat{x}_k\}$ in H^2 (see Theorem 2.5). However, Theorem 2.8 does not guarantee acceleration when the sequence $\{o_k\}$ in H is compared to the original MAP sequence $\{x_p^k\}$ in H, since a "maximal" is involved. Nevertheless, based on numerical experimentation, the sequence $\{o_k\}$ almost always accelerates the convergence of the original MAP sequence $\{x_p^k\}$.

We will now describe explicitly the sequence $\{o_k\}$ defined by (2.7). If $\|P_\Pi(\widehat{x}_{k+1}) - P_\Pi(\widehat{x}_k)\| = 0$, then from (2.1) and (2.2) we have that

$$\widehat{o}_k = P_\Pi(\widehat{x}_{k+1}) = P_\Pi\left(x_{p-1}^{k+1}, x_p^{k+1}\right) = 1/2\left(x_{p-1}^{k+1} + x_p^{k+1}, x_{p-1}^{k+1} + x_p^{k+1}\right) = (o_k, o_k) \in \Pi.$$

In this case $o_k = 1/2\left(x_{p-1}^{k+1} + x_p^{k+1}\right) \in H$. If, on the other hand, $\|P_\Pi(\widehat{x}_{k+1}) - P_\Pi(\widehat{x}_k)\| \neq 0$, then

$$\widehat{o}_k = P_\Pi(\widehat{x}_k) + \alpha_k(P_\Pi(\widehat{x}_{k+1}) - P_\Pi(\widehat{x}_k)) = (o_k, o_k) \in \Pi,$$

where

$$o_k = 1/2\left(x_p^k + \alpha_k\left(x_p^{k+1} - x_p^k\right) + x_{p-1}^k + \alpha_k\left(x_{p-1}^{k+1} - x_{p-1}^k\right)\right) \in H.$$

From Corollary 2.9 it follows that

$$o_k = 1/2\left(x_p^k + \alpha_k\left(x_p^{k+1} - x_p^k\right) + x_{p-1}^k + \alpha_k\left(x_{p-1}^{k+1} - x_{p-1}^k\right)\right) \to P_M(x_0)$$

when $k \to \infty$. Since P_{M_p} is a bounded linear operator (where P_{M_p} is the projection operator onto M_p),

$$P_{M_p}\left[1/2\left(x_p^k + \alpha_k\left(x_p^{k+1} - x_p^k\right) + x_{p-1}^k + \alpha_k\left(x_{p-1}^{k+1} - x_{p-1}^k\right)\right)\right] \to P_{M_p}(P_M(x_0))$$

when $k \to \infty$. Consequently

$$1/2\left[x_p^k + \alpha_k\left(x_p^{k+1} - x_p^k\right) + x_p^k + \alpha_k\left(x_p^{k+1} - x_p^k\right)\right] \to P_M(x_0)$$

when $k \to \infty$. Therefore

$$x_p^k + \alpha_k \left(x_p^{k+1} - x_p^k\right) \to P_M(x_0) \tag{2.8}$$

when $k \to \infty$. Equation (2.8) suggests how to define an accelerated sequence in H and a specialized algorithm for which convergence to the solution $P_M(x_0)$ will be later established. Nevertheless, to achieve this it is convenient to write α_k given by (2.5) as a function of the original MAP iterations in H.

LEMMA 2.11. *Let α_k, for some k, be given by* (2.5)*. Then*

$$\alpha_k = \frac{\left\langle x_{p-1}^k + x_p^k, x_{p-1}^k - x_{p-1}^{k+1}\right\rangle + \left\langle x_{p-1}^k + x_p^k, x_p^k - x_p^{k+1}\right\rangle}{\|x_{p-1}^{k+1} + x_p^{k+1} - x_{p-1}^k - x_p^k\|^2}. \tag{2.9}$$

Proof. Let $k \geq 1$. The numerator in (2.5) can be written as

$$\begin{aligned}\langle P_\Pi(\widehat{x}_k),\ P_\Pi(\widehat{x}_k) \ &- \ P_\Pi(\widehat{x}_{k+1})\rangle \\ &= \ \langle P_\Pi(\widehat{x}_k), \widehat{x}_k - \widehat{x}_{k+1}\rangle \\ &= \ \left\langle 1/2\left(x_{p-1}^k + x_p^k, x_{p-1}^k + x_p^k\right), \left(x_{p-1}^k - x_{p-1}^{k+1}, x_p^k - x_p^{k+1}\right)\right\rangle \\ &= \ 1/2\left(\left\langle x_{p-1}^k + x_p^k, x_{p-1}^k - x_{p-1}^{k+1}\right\rangle + \left\langle x_{p-1}^k + x_p^k, x_p^k - x_p^{k+1}\right\rangle\right).\end{aligned}$$

On the other hand

$$\begin{aligned}P_\Pi(\widehat{x}_{k+1}) - P_\Pi(\widehat{x}_k) &= 1/2\left(x_{p-1}^{k+1} + x_p^{k+1}, x_{p-1}^{k+1} + x_p^{k+1}\right) - 1/2\left(x_{p-1}^k + x_p^k, x_{p-1}^k + x_p^k\right) \\ &= 1/2\left(x_{p-1}^{k+1} + x_p^{k+1} - x_{p-1}^k - x_p^k, x_{p-1}^{k+1} + x_p^{k+1} - x_{p-1}^k - x_p^k\right).\end{aligned}$$

Therefore the denominator in (2.5) can be written as

$$\begin{aligned}\|P_\Pi(\widehat{x}_{k+1}) - P_\Pi(\widehat{x}_k)\|^2 &= 1/4\left(2\|x_{p-1}^{k+1} + x_p^{k+1} - x_{p-1}^k - x_p^k\|^2\right) \\ &= 1/2\|x_{p-1}^{k+1} + x_p^{k+1} - x_{p-1}^k - x_p^k\|^2,\end{aligned}$$

and the result holds. □

Let $\{x_i^k\}$ be the MAP's iterates given by (1.2). Let us now define a new sequence $\{o_k'\}$ in H, for $k \geq 1$, as follows

$$o_k' = \begin{cases} x_p^{k+1} & \text{if } \|x_{p-1}^{k+1} + x_p^{k+1} - x_{p-1}^k - x_p^k\| = 0, \\ x_p^k + \alpha_k(x_p^{k+1} - x_p^k) & \text{if } \|x_{p-1}^{k+1} + x_p^{k+1} - x_{p-1}^k - x_p^k\| \neq 0, \end{cases} \tag{2.10}$$

where α_k is given by (2.9).

THEOREM 2.12. *The sequence $\{o_k'\}$ defined by* (2.10) *converges to $P_M(x_0)$.*

Proof. This is a consequence of Corollary 2.9, equation (2.8), and Theorem 1.3. □

We are now ready to present our acceleration scheme for MAP.

Algorithm 1. Let M_i, $i = 1, \ldots, p$, be p closed subspaces in H. Given $x_0 \in H$; set $k = 1$.
set $x_p^0 = x_0$

$x_{p-1}^k = (P_{M_{p-1}} \ldots P_{M_1})(x_p^{k-1})$

$x_p^k = P_{M_p}(x_{p-1}^k)$

for $k = 1, 2, \ldots$ **do**

$x_{p-1}^{k+1} = (P_{M_{p-1}} \ldots P_{M_1})(x_p^k)$

$x_p^{k+1} = P_{M_p}(x_{p-1}^{k+1})$

if $\|x_{p-1}^{k+1} + x_p^{k+1} - x_{p-1}^k - x_p^k\| \neq 0$ **then**

compute α_k using (2.9), and **set** $o'_k = x_p^k + \alpha_k(x_p^{k+1} - x_p^k)$

else

set $o'_k = x_p^{k+1}$

end if

end for

In the following, we will compare our scheme, given in Algorithm 1, with the acceleration scheme given by Gearhart and Koshy [21]. Therefore, we will give a brief explanation of the Gearhart-Koshy scheme. Let us denote by x_0 the given starting point and by Q the composition of the projection operators, i.e, $Q = P_{M_p} P_{M_{p-1}} \cdots P_{M_1}$, where P_{M_i} is the projection operator onto M_i for all i. Let x_k be the k th iterate, and let Qx_k be the next iterate after applying a sweep of MAP. The idea is to search along the line through the points x_k and Qx_k to obtain the point closest to the solution $x^* = P_{\cap_{i=1}^p M_i}(x_0)$. Let us represent any point on this line as

$$x^k(t) = tQx_k + (1-t)x_k = x_k + t(Qx_k - x_k),$$

for some real number t. Let t_k be the value of t for which this point is the closest to x^*. Then,

$$\left\langle x^k(t_k) - x^*, x_k - Qx_k \right\rangle = 0. \tag{2.11}$$

Now, since $x^* \in \cap_{i=1}^p M_i$ and the projections P_{M_i} are self-adjoint, it follows

$$\langle x^*, Qx_k \rangle = \left\langle P_{M_1} P_{M_2} \cdots P_{M_p} x^*, x_k \right\rangle = \langle x^*, x_k \rangle .$$

Consequently, $\langle x^*, x_k - Qx_k \rangle = 0$, and so x^* can be eliminated from (2.11) to obtain

$$\left\langle x^k(t_k), x_k - Qx_k \right\rangle = 0.$$

Solving for t_k gives

$$t_k = \frac{\langle x_k, x_k - Qx_k \rangle}{\|x_k - Qx_k\|^2}.$$

To summarize, we have the following steps.

Algorithm 2. (Gearhart-Koshy) Let M_i, $i = 1, \ldots, p$, be p closed subspaces in H. Given $x_0 \in H$.

for $k = 1, 2, \ldots$ **do**

$Q_k = Qx_k$

$t_k = \langle x_k, x_k - Qx_k \rangle / \|x_k - Qx_k\|^2$

$x_{k+1} = t_k Q_k + (1 - t_k) x_k$

end for

3. Numerical experiments. We compare our acceleration scheme, given in Algorithm 1, with the acceleration scheme given by Gearhart and Koshy (Algorithm 2), and with the original MAP with no acceleration given by (1.2). All computations were performed in MATLAB. For all experiments we know the exact solution x^*, and we stop each algorithm when the absolute error (the distance from the k-th iterate to x^*) is less than or equal to 10^{-6}.

ETNA
Kent State University
http://etna.math.kent.edu

For our first experiment, we consider the following three subspaces of the space of square real matrices $\mathbb{R}^{3\times 3}$, with the Frobenius norm $\|A\|_F^2 = \langle A, A\rangle = \text{trace}(A^T A)$:

$$M_1 = \left\{A \in \mathbb{R}^{3\times 3} : A^T = A\right\} ,$$
$$M_2 = \left\{A \in \mathbb{R}^{3\times 3} : a_{i,j+1} = 0,\ i = 1, 2,\ j = i+1, \dots, 3\right\} ,$$
$$M_3 = \left\{A \in \mathbb{R}^{3\times 3} : a_{1,1} = a_{1,3} = a_{3,1} = a_{3,3}\right\} .$$

If $A = (A_{ij}) \in \mathbb{R}^{3\times 3}$, then $P_{M_1}(A) = (A^T + A)/2$,

$$P_{M_2}(A) = \begin{bmatrix} A_{11} & 0 & 0 \\ A_{21} & A_{22} & 0 \\ A_{31} & A_{32} & A_{33} \end{bmatrix}, \text{ and } P_{M_3}(A) = \begin{bmatrix} P & A_{12} & P \\ A_{21} & A_{22} & A_{23} \\ P & A_{32} & P \end{bmatrix},$$

where $P = (A_{11} + A_{13} + A_{31} + A_{33})/4$.
We choose

$$A_0 = \begin{bmatrix} 10 & 20 & 30 \\ 40 & 50 & 60 \\ 70 & 80 & 90 \end{bmatrix}$$

as the initial given point. Then

$$P_{M_1\cap M_2\cap M_3}(A_0) = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 50 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

since

$$A \in \left\{ \begin{bmatrix} 0 & 0 & 0 \\ 0 & x & 0 \\ 0 & 0 & 0 \end{bmatrix} : x \in \mathbb{R} \right\} = M_1 \cap M_2 \cap M_3,$$

$$\langle A, A_0 - P_{M_1\cap M_2\cap M_3}(A_0)\rangle = \text{trace}\left(\begin{bmatrix} 0 & 0 & 0 \\ 0 & x & 0 \\ 0 & 0 & 0 \end{bmatrix}^T \begin{bmatrix} 10 & 20 & 30 \\ 40 & 0 & 60 \\ 70 & 80 & 90 \end{bmatrix} \right) = 0.$$

The performance of each method is shown in Figure 3.1.

Consider the problem of solving the linear system of equations $Ax = 0$, where $A \in \mathbb{R}^{m\times n}$ and $x \in \mathbb{R}^n$. This problem can be generalized to any Hilbert space H by considering the following formulation: find x in the intersection of m closed subspaces given by $H_i = \{x \in H : \langle a_i, x\rangle = 0\}$, for every $i = 1, \dots, m$. Here a_i denotes the ith row of A or, in general, a fixed given vector in H. If $z \in H$, then $P_{H_i}(z) = z - (\langle a_i, z\rangle \,/\, \langle a_i, a_i\rangle)a_i$.

In the following two test problems we will solve the linear equations $Ax = 0$, where A is a square nonsingular matrix which is randomly generated in MATLAB (we verified that $\det(A) \neq 0$ after randomly generating the matrix A in MATLAB). Therefore the predetermined solutions is always $x^* = 0$. In both experiments we chose $x_0 = (100, 100, \dots, 100)$ as the initial given point. Figure 3.2 shows the performance for each methods for $A \in \mathbb{R}^{5\times 5}$ while Figure 3.3 shows the performance for each method for $A \in \mathbb{R}^{10\times 10}$.

Our preliminary numerical experiments seem to indicate that our acceleration scheme, when compared to the other two “competitors”, benefits from a reduction in the number of cycles, as well as in the computational work.

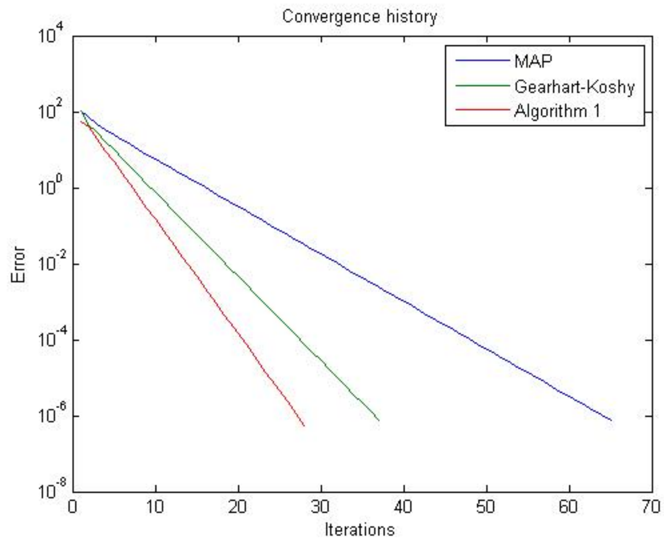

FIG. 3.1. *Acceleration for three subspaces in* $\mathbb{R}^{3\times 3}$.

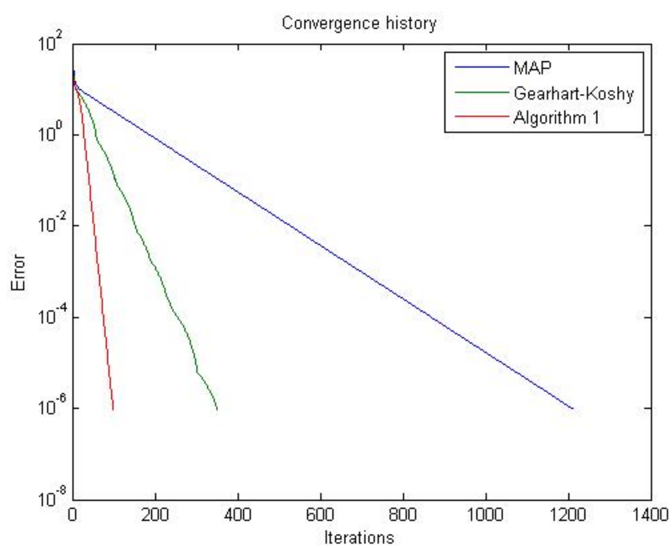

FIG. 3.2. *Acceleration for solving* $Ax = 0$*, where* $A \in \mathbb{R}^{5\times 5}$.

Acknowledgements. I thank the three anonymous referees for constructive remarks and additional references.

REFERENCES

[1] G. APPLEBY AND D. C. SMOLARSKI, *A linear acceleration row action method for projecting onto subspaces*, Electron. Trans. Numer. Anal, 20 (2005), pp. 253–275.
http://etna.ricam.oeaw.ac.at/vol.20.2005/pp253-275.dir/pp253-275.pdf

[2] H. H. BAUSCHKE AND J. M. BORWEIN, *Dykstra's alternating projection algorithm for two sets*, J. Approx.

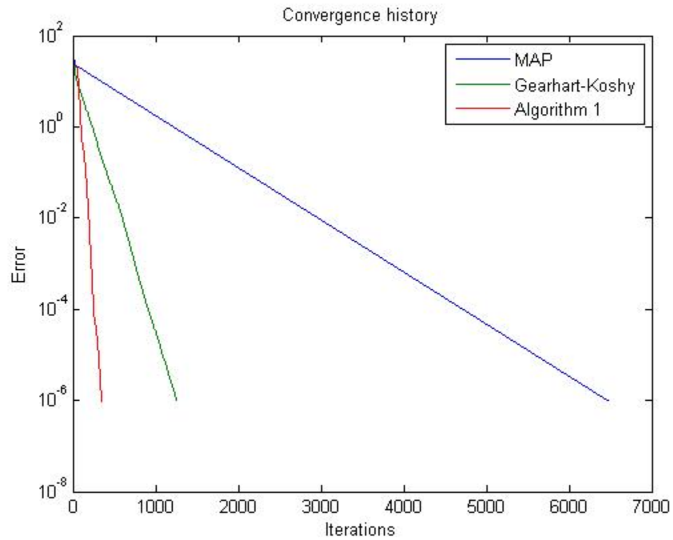

FIG. 3.3. *Acceleration for solving $Ax = 0$, where $A \in \mathbb{R}^{10\times 10}$.*

Theory, 79 (1994), pp. 418–443.

[3] H. H. BAUSCHKE, P. L. COMBETTES, AND S. G. KRUK, *Extrapolation algorithm for affine-convex feasibility problems*, Numer. Algorithms, 41 (2006), pp. 239–274.

[4] H. H. BAUSCHKE, F. DEUTSCH, H. HUNDAL, AND S. H. PARK, *Accelerating the convergence of the method of alternating projections*, Trans. Amer. Math. Soc., 355 (2003), pp. 3433–3461.

[5] J. P. BOYLE AND L. DYKSTRA, *A method for finding projections onto the intersection of convex sets in Hilbert spaces*, in Advances in Order Restricted Statistical Inference, R. Dykstra, T. Robertson, and F. T. Wright, eds., Lecture Notes in Statistics, 37, Springer, Berlin, 1986, pp. 28–47.

[6] Y. CENSOR, *Computational acceleration of projection algorithms for the linear best approximation problem*, Linear Algebra Appl., 416 (2006), pp. 111–123.

[7] P. L. COMBETTES, *Convex set theoretic image recovery by extrapolated iterations of parallel subgradient projections*, IEEE Trans. Image Process., 6 (1997), pp. 493–506.

[8] ———, *The convex feasibility problem in image recovery*, in Advances in Imaging and Electron Physics, 95 P. W. Hawkes, ed., Academic Press, New York, 1996, pp. 155–270.

[9] F. DEUTSCH, *Applications of von Neumann's alternating projections algorithm*, in Mathematical Methods in Operations Research, P. Kenderov, ed., Bulgarian Academy of Science, Sofia, 1983, pp. 44–51.

[10] ———, *The method of alternating orthogonal projections*, in Approximation Theory, Spline Functions and Applications, S. P. Singh, ed., Kluwer, Netherlands, 1992, pp. 105–121.

[11] ———, *Best Approximation in Inner Product Space*, Springer, New York, 2001.

[12] ———, *Accelerating the convergence of the method of alternating projections via a line search: a brief survey*, in Inherently Parallel Algorithms in Feasibility and Optimization and their Applications, D. Butnariu, Y. Censor, S. Reich, eds., Studies in Computational Mathematics, 8, North-Holland Publishing, Amsterdam, 2001, pp. 203–217.

[13] A. DE PIERRO AND A. N. IUSEM, *A parallel projection method for finding a common point of a family of convex sets*, Pesquisa Operacional, 5 (1985), pp. 1–20.

[14] L. T. DOS SANTOS, *A parallel subgradient projections method for the convex feasibility problem*, J. Comput. Appl. Math., 18 (1987), pp. 307–320.

[15] R. L. DYKSTRA, *An algorithm for restricted least-squares regression*, J. Amer. Statist. Assoc., 78 (1983), pp. 837–842.

[16] R. ESCALANTE AND M. RAYDAN, *Alternating Projection Methods*, SIAM, Philadelphia, 2011.

[17] N. ECHEBEST, M. T. GUARDARUCCI, H. SCOLNIK, AND M. C. VACCHINO, *An acceleration scheme for solving convex feasibility problems using incomplete projection algorithms*, Numer. Algorithms, 35 (2004), pp. 335–350.

[18] ———, *An accelerated iterative method with diagonally scaled oblique projections for solving linear feasibility problems*, Ann. Oper. Res., 138 (2005), pp. 235–257.

[19] C. FRANCHETTI AND W. LIGHT, *On the von Neumann alternating algorithm in Hilbert space*, J. Math. Anal.

Appl., 114 (1986), pp. 305–314.

[20] K. FRIEDRICHS, *On certain inequalities and characteristic value problems for analytic functions and for functions of two variables*, Trans. Amer. Math. Soc., 41 (1937), pp. 321–364.

[21] W. GEARHART AND M. KOSHY, *Acceleration schemes for the method of alternating projection*, J. Comput. Appl. Math., 26 (1989), pp. 235–249.

[22] L. G. GUBIN, B. T. POLYAK, AND E. V. RAIK, *The method of projections for finding the common point of convex sets*, USSR Comput. Math. Math. Phys., 7 (1967), pp. 1–24.

[23] I. HALPERIN, *The product of projection operators*, Acta Sci. Math. (Szeged), 23 (1962), pp. 96–99.

[24] L. M. HERNÁNDEZ-RAMOS, R. ESCALANTE, AND M. RAYDAN, *Unconstrained optimization techniques for the acceleration of alternating projection methods*, Numer. Funct. Anal. Optim., 32 (2011), pp. 1041–1066.

[25] S. KACZMARZ, *Angenäherte Auflösung von Systemen linearer Gleichungen*, Bull. Internat. Acad. Pol. Sci. Lett. Ser. A, 35 (1937), pp. 355–357.

[26] W. LÓPEZ AND M. RAYDAN, *An acceleration scheme for Dykstra's algorithm*, Comput. Optim. Appl., 63 (2016), pp. 29–44.

[27] J. M. MARTÍNEZ, *An accelerated successive orthogonal projection method for solving large-scale linear feasibility problems*, Comput. Math. Appl., 15 (1988), pp. 367–373.

[28] J. VON NEUMANN, *Functional operators. Vol II. The Geometry of Orthogonal Spaces*, Annals of Math. Studies, vol. 22, Princeton University Press, Princeton, 1951.

[29] H. SCOLNIK, N. ECHEBEST, M. T. GUARDARUCCI, AND M. C. VACCHINO, *Acceleration scheme for parallel projected aggregation methods for solving large linear systems*, Ann. Oper. Res., 117 (2002), pp. 95–115.

[30] ———, *A class of optimized row projection methods for solving large nonsymmetric linear systems*, Appl. Numer. Math., 41 (2002), pp. 499–513.