King Fahd University of Petroleum & Minerals

DEPARTMENT OF MATHEMATICAL SCIENCES

Technical Report Series

TR 172

April 1995

Alternating Projection and Decomposition with Respect to Two Convex Sets

Alberto Seeger

DHAHRAN 31261 ● SAUDI ARABIA ● www.kfupm.edu.sa/math/ ● E-mail: mathdept@kfupm.edu.sa

ALTERNATING PROJECTION AND DECOMPOSITION WITH RESPECT TO TWO CONVEX SETS

Alberto Seeger
King Fahd University of Petroleum and Minerals
Department of Mathematical Sciences
Dhahran 31261, Saudi Arabia.

Abstract

We apply von Neumann's alternating projection algorithm to find a decomposition, with respect to two convex sets X and Y, of a given element in the Minkowski sum $X + Y$. We consider also the problem of finding the best decomposition, that is to say, the one that has least deviation. The latter problem is solved by adapting Dykstra's alternating projection method.

Key words: Least deviation decomposition, alternating projection, von Neumann's algorithm, Dykstra's algorithm.

1991 Mathematics subject classification: 41 A 65, 49 M 45, 52 A 05, 65 J 05.

1 Introduction.

This note is a complement to previous works by Martinez–Legaz and Seeger [MS], and Luc, Martinez–Legaz and Seeger [LMS]. We are concerned with the design of algorithms that decompose a given vector $z \in H$ in the form

$$z = x + y, \quad \text{with } x \in X \text{ and } y \in Y, \tag{1.1}$$

where X and Y are two nonempty closed convex sets in a Hilbert space $(H, \langle \cdot, \cdot \rangle)$. Such kind of decomposition problem arises in optimization, linear algebra, statistics, and other areas.

The element z is decomposable in the form (1.1) if and only if z belongs to the Minkowski sum

$$X + Y := \{x + y : x \in X, \quad y \in Y\}.$$

In such a case, however, the decomposition (1.1) need not be unique. For this reason we address also the question of finding the best decomposition in the set

$$D(z) := \{(x, y) \in X \times Y : z = x + y\}. \tag{1.2}$$

As in [LMS], we understand the term "best" in the sense of least deviation.

Definition 1.1 [LMS]. *The pair* $(\overline{x}, \overline{y}) \in H \times H$ *is called a* least deviation decomposition *of* z *if*

$$\begin{cases} (\overline{x}, \overline{y}) \in D(z), \\ \|\overline{x} - \overline{y}\| \leq \|x - y\| \text{ for all } (x, y) \in D(z). \end{cases} \tag{1.3}$$

In the above definition $\|\cdot\|$ is the norm associated to the inner product $\langle \cdot, \cdot \rangle$. Thus, each $z \in X + Y$ admits a unique least deviation decomposition. Least deviation is a natural criterion to use when it comes to selecting one particular decomposition in

$D(z)$. The concept of least deviation decomposition enjoys many interesting theoretical properties; for instance, it can be seen as an extension of the celebrated Moreau orthogonal decomposition with respect to a pair of mutually polar cones.

As explained in Section 3, one way to construct the least deviation decomposition $(\bar{x}, \bar{y})$ of $z \in X + Y$ is by finding the projection of a given point in a Hilbert space onto the nonempty intersection of two closed convex sets. Algorithms for solving this abstract projection problem have been extensively discussed in the literature. For the sake of completeness in the exposition, we record in the Appendix two results taken from Bauschke and Borwein [BB2]. The first result deals with von Neumann's alternating projection algorithm, and the second one deals with Dykstra's algorithm.

2 Finding an Admissible Decomposition.

2.1 Introducing Algorithm A1. It is always assumed that X and Y are two nonempty closed convex sets in a Hilbert space $(H, \langle\cdot,\cdot\rangle)$. The class of such sets is denoted by $P(H)$.

As indicated in (1.2), the set $D(z)$ contains all the (admissible) decompositions of z. It is clear that $D(z)$ can be written also in the form

$$D(z) = \{(x, z - x) : x \in X \cap (z - Y)\}. \tag{2.1}$$

Thus, the problem of decomposing z is equivalent to that of finding an element in the intersection of X and $z - Y$. Such an element can be found by adapting von Neumann's alternating projection algorithm. Recall that the projection $\Pi_X(f)$ of $f \in H$ onto X

is the unique solution of the minimization problem

$$d_X(f) := \min\{\|f - x\| : x \in X\}.$$

In this work it is assumed that the projection operators $\Pi_X, \Pi_Y : H \to H$ can be easily evaluated.

Our first result concerns the following alternating projection method.

Algorithm A1: Choose any $y_0 \in H$. For $n \geq 1$, compute

$$\begin{cases} x_n = \Pi_X(z - y_{n-1}), \\ y_n = \Pi_Y(z - x_n). \end{cases} \tag{2.2}$$

Theorem 2.1. *Let $X, Y \in P(H)$, and $\{(x_n, y_n)\}_{n\geq 1}$ be a sequence generated by Algorithm A1. Then one has:*

(a) $\{(x_n, y_n)\}_{n\geq 1} \subset X \times Y$;

(b) $x_n + y_n \to \Pi_{\overline{X+Y}}(z)$, *i.e. the sequence* $\{x_n + y_n\}_{n\geq 1}$ *converges (strongly) to the projection of z onto the closure of $X + Y$. In particular,*

$$\|x_n + y_n - z\| \to d_{X+Y}(z);$$

(c) If $z \in X + Y$, then $\{(x_n, y_n)\}_{n\geq 1}$ converges weakly to some $(x, y) \in D(z)$.

Proof. We apply Lemma I (cf. Appendix) to the particular case $A = X$ and $B = z - Y$. If one writes

$$a_n = x_n \quad \text{and} \quad b_n = z - y_n,$$

then von Neumann's algorithm takes the form

$$\begin{cases} x_n = \Pi_X(z - y_{n-1}), \\ z - y_n = \Pi_{z-Y}(x_n). \end{cases} \tag{2.3}$$

But the second equality in (2.3) can be written as

$$y_n = z - \Pi_{z-Y}(x_n),$$

or, equivalently,

$$y_n = \Pi_Y(z - x_n).$$

From the convergence result

$$(z - y_n) - x_n \rightarrow \Pi_{\overline{z-Y-X}}(0),$$

one gets

$$x_n + y_n \rightarrow z - \Pi_{\overline{z-Y-X}}(0) = \Pi_{\overline{X+Y}}(z).$$

Finally, if $z \in X + Y$, then the infimum

$$\delta(X, z - Y) = \inf\{\|x - b\| : x \in X, \quad b \in z - Y\}$$

is attained. Moreover,

$$\delta(X, z - Y) = d_{X+Y}(z) = 0,$$

and

$$E = \{x \in X : d_{z-Y}(x) = 0\} = X \cap (z - Y).$$

Thus, the sequence $\{x_n\}_{n\geq 1}$ converges weakly to some $x \in X \cap (z - Y)$, and $\{y_n\}_{n\geq 1}$ converges weakly to $y = z - x$. □

Remark: Part (b) of Theorem 2.1 is proven by Franchetti and Light [FL, Lemma 3.2] in the particular case in which X and Y are closed subspaces of H, and $\Pi_{\overline{X+Y}}(z) \in X + Y$.

2.2 Further Analysis on the Convergence of Algorithm A1.

In the context of this note, the most important part of Theorem 2.1 is the last one.

Indeed, part (c) tells us how to construct an admissible decomposition for the vector $z \in X + Y$. Despite its astonishing simplicity, the Algorithm A1 generates a sequence $\{(x_n, y_n)\}_{n\geq 1}$ that converges weakly to some decomposition $(x, y) \in D(z)$. A natural question is whether or not the convergence can actually be only weak. Conditions for strong convergence can be obtained from the general theory of von Neumann sequences. By way of example, one has

Proposition 2.1. *The convergence in Theorem 2.1(c) is strong if any of the following conditions hold:*

(a) X or Y is locally compact;

(b) X and Y are closed affine subspaces.

The conditions (a) and (b) mentioned in the above proposition are quite stringent in practice. They ensure the strong convergence of $\{(x_n, y_n)\}_{n\geq 1}$, no matter which is the vector $z \in X + Y$ that we want to decompose. A much finer analysis is carried out in the next proposition, where the particular choice of z is bring into consideration. Before stating our result, it is convenient to recall first some notation and prove a technical lemma. Given a convex set C in some Banach space V, the ***strong quasi-interior*** of C, denoted by $\mathrm{sqi}(C)$, is the set of those $c \in C$ for which

$$\mathrm{cone}(C - c) := \bigcup_{\lambda \geq 0} \lambda\,(C - c)$$

is a closed subspace in V. This concept plays an important role in the duality analysis of abstract variational problems (cf. [BL], [Vo], [AT]). Here it is used to state an open mapping principle which extends the one due to Robinson [Ro, Theorem 1].

Lemma 2.1. *Let U and V be Banach spaces, and let $F : U \to V$ be a set-valued mapping whose graph is a closed convex set in the product space $U \times V$. Suppose 0 is* *a strong quasi-interior point of the range $R(F)$ of F. Then, for all $u \in F^{-1}(0)$, there exists a positive η such that*

$$\eta S_W \subset F(u + S_U),$$

where S_U and S_W denote the closed unit balls in the spaces U and $W :=$ cone $(R(F))$, respectively.

Proof. First of all, it is important to observe that W is a closed subspace of V. This is because $0 \in \text{sqi}(R(F))$. Since $R(F) \subset W$, we regard F as a set-valued mapping from U into W. This trick allows us to adjust Robinson's proof [Ro, Theorem 1] in a rather straightforward manner. For notational convenience, we set $u = 0$. In this case, the convex set $F(S_U) \subset W$ contains the origin $0 \in W$. As in [Ro], one needs to prove that $F(S_U)$ is absorbing (not in V, but in the space W!). But, if w is any point in $W \backslash \{0\}$, then the very definition of W implies that

$$\mu w \in R(F) \quad \text{for some } \mu > 0.$$

So, we are in fact in the same situation as in [Ro]. □

The concept of linear convergence will be used also in the sequel. Recall that a sequence $\{h_n\}_{n \geq 1}$ in H is said to *converge linearly* to $h \in H$ if there are constants $\beta > 0$ and $\delta \in]0, 1[$ such that

$$\|h_n - h\| \leq \beta \delta^n \quad \text{for all } n \geq 1.$$

The vector h is of course the (strong) limit of $\{h_n\}_{n \geq 1}$.

Proposition 2.2. *Let $X, Y \in P(H)$. If $z \in sqi(X+Y)$, then the sequence $\{(x_n, y_n)\}_{n \geq 1}$ generated by Algorithm A1 converges linearly to some $(x, y) \in D(z)$. Moreover, $\{x_n\}_{n \geq 1}$ converges linearly to x, and $\{y_n\}_{n \geq 1}$ converges linearly to y.*

Proof. As mentioned in the proof of Theorem 2.1, the sequence $\{(x_n, y_n)\}_{n\geq 1}$ is obtained by adjusting von Neumann's algorithm to the sets $A = X$ and $B = z - Y$. According to Bauschke and Borwein [BB1, Theorem 3.12], the von Neumann sequences converge linearly if the pair (A, B) is boundedly linearly regular in the sense that for each bounded set $\Theta \subset H$, there exists a positive α such that

$$d_{A\cap B}(v) \leq \alpha \max\{d_A(v), d_B(v)\} \quad \text{for all } v \in \Theta.$$

To check the above condition, we apply a technique developed in [BB1, Theorem 4.3]. However, some adjustments to this technique are in order. Consider the set-valued mapping $F : H \rightarrow H$ defined by

$$F(u) := \begin{cases} u - B & \text{if } u \in A, \\ \phi & \text{otherwise .} \end{cases}$$

The graph of F is clearly a closed convex set in the product space $H \times H$, and the range of F is

$$R(F) = A - B.$$

The "regularity" assumption $z \in \text{sqi}(X + Y)$ amounts to saying that

$$W := \bigcup_{\lambda\geq 0} \lambda(X + Y - z) = \bigcup_{\lambda\geq 0} \lambda(A - B)$$

is a closed subspace in H. In other words, 0 is a strong quasi-interior point of $R(F)$. Thus, we can apply Lemma 2.1 (contrary to [BB1], we do not rely on Borwein's open mapping theorem) to show that, for all $u \in F^{-1}(0) = A \cap B$, there exists a positive η such that

$$\eta S_W \subset F(u + S_H) = A \cap (u + S_H) - B.$$

Now, we are going to apply [BB1, Lemma 4.2]. To do this, one needs first to observe that

$$a - \Pi_B(a) \in A - B \subset W \quad \text{for all } a \in A,$$

and, therefore,

$$\frac{\Pi_B(a) - a}{\|\Pi_B(a) - a\|} \in S_W \quad \text{for all } a \in A\backslash B.$$

By applying [BB1, Lemma 4.2] one arrives at the following conclusion: for each $u \in A \cap B$ there is a positive η such that

$$d_{A\cap B}(v) \leq \frac{\|u - v\| + 1}{\eta} d_B(v) \quad \text{for all } v \in A.$$

As shown in the proof of [BB1, Theorem 4.3], the above condition leads us to the bounded linear regularity of the pair (A, B). In this way we have shown that the sequences $\{x_n\}_{n\geq 1}$ and $\{y_n\}_{n\geq 1}$ converge linearly to $x \in X$ and $y \in Y$ respectively. The pair (x, y) is necessarily a decomposition of $z \in X+Y$, and $\{(x_n, y_n)\}_{n\geq 1}$ converges linearly to (x, y). □

Some remarks concerning Proposition 2.2 are in order:

(i) The assumption $z \in \text{sqi}(X+Y)$ plays an important role in the proof of Proposition 2.2. If z belongs to the interior of $X + Y$, then one does not need Lemma 2.1. Indeed, it suffices to invoke the classical open mapping principle of Robinson [Ro, Theorem 1]. If z belongs to the intrinsic core of $X + Y$, then the proof of [BB1, Theorem 4.3] does not require major adjustments.

(ii) An interesting question which remains open is the following one: Is it possible to find sets $X, Y \in P(H)$ and a vector $z \in (X + Y)\backslash\text{sqi}(X + Y)$ such that the weakly convergent sequence $\{(x_n, y_n)\}_{n\geq 1}$ generated by Algorithm A1 does not converge strongly?

(iii) The assumption $z \in \text{sqi}(X+Y)$ is sufficient for $\{(x_n, y_n)\}_{n\geq 1}$ to converge linearly. Strong convergence may occur, however, even if $z \notin \text{sqi}(X+Y)$. To see this, just take two closed affine subspaces X and Y whose sum $X+Y$ is not closed. In this case the set $\mathrm{sqi}(X+Y)$ is empty, but one has strong convergence whenever $z \in X+Y$.

2.3 Introducing Algorithm A2.

The set $D(z)$ can be written also as the intersection of the closed convex set $X \times Y$ and the closed affine subspace

$$\begin{aligned}\Delta(z) &:= \{(x,y) \in H \times H : x+y=z\} \\ &= (0,z) + \{(x,-x) : x \in H\}.\end{aligned}$$

This way of looking at $D(z)$ leads us to the following

Algorithm A2: Choose any $(x_0, y_0) \in H \times H$. For $n \geq 1$, compute

$$\begin{cases} e_n &= \frac{1}{2}\left[\Pi_X(x_{n-1}) - \Pi_Y(y_{n-1})\right], \\ x_n &= \frac{z}{2} + e_n, \\ y_n &= \frac{z}{2} - e_n. \end{cases} \tag{2.4}$$

Before stating the main properties of the above algorithm, it is convenient to record some technical results.

Lemma 2.2. *Let $X, Y \in P(H)$. Then,*

(a) The projection operator $\Pi_{\Delta(z)} : H \times H \to H \times H$ is given by

$$\Pi_{\Delta(z)}(c,d) = \left(\frac{z}{2}, \frac{z}{2}\right) + \left(\frac{c-d}{2}, \frac{d-c}{2}\right); \tag{2.5}$$

(b) The shortest distance $\delta(X \times Y, \Delta(z))$ between the sets $X \times Y$ and $\Delta(z)$ is given by

$$\delta^2(X \times Y, \Delta(z)) = \min_{x+y=z} \left\{d_X^2(x) + d_Y^2(y)\right\}, \tag{2.6}$$

or, equivalently,

$$\delta^2(X \times Y, \Delta(z)) = \frac{1}{2} d_{X+Y}^2(z); \tag{2.7}$$

(c) The pair $(\tilde{x}, \tilde{y})$ is a solution to the minimization problem (2.6) if and only if

$$\begin{cases} \tilde{x} + \tilde{y} = z \\ [I - \Pi_X](\tilde{x}) = [I - \Pi_Y](\tilde{y}), \end{cases} \tag{2.8}$$

where $I : H \to H$ stands for the identity operator;

(d) Let $(\tilde{x}, \tilde{y})$ be a solution to (2.6), and let $p = [I - \Pi_X](\tilde{x})$. Then

$$\Pi_{\overline{\Delta(z) - X \times Y}}(0,0) = \Pi_{\Delta(z) - X \times Y}(0,0) = (p,p).$$

Proof. The projection $\Pi_{\Delta(z)}(c,d)$ is the solution to the minimization problem

$$d_{\Delta(z)}(c,d) := \min_{(x,y) \in \Delta(z)} \|(c,d) - (x,y)\|. \tag{2.9}$$

As a matter of computation one has

$$\begin{aligned} d^2_{\Delta(z)}(c,d) &= \min_{x \in H} \|(c,d) - (x, z - x)\|^2 \\ &= \min_{x \in H} \{\|c - x\|^2 + \|d - z + x\|^2\}. \end{aligned}$$

Since the above minimum is attained at

$$x = \frac{z}{2} + \frac{c - d}{2},$$

the solution to (2.9) is given by (2.5). To obtain the formula (2.6), we write

$$\begin{aligned} \delta^2(X \times Y, \Delta(z)) &= \inf_{\substack{(c,d) \in X \times Y \\ (x,y) \in \Delta(z)}} \|(c,d) - (x,y)\|^2 \\ &= \inf_{\substack{(c,d) \in X \times Y \\ x + y = z}} \{\|c - x\|^2 + \|d - y\|^2\} \\ &= \min_{x + y = z} \left\{ \min_{c \in X} \|c - x\|^2 + \min_{d \in Y} \|d - y\|^2 \right\}. \end{aligned}$$

Formula (2.7) is obtained by writing

$$\delta^2(X \times Y, \Delta(z)) = \inf_{(c,d) \in X \times Y} \inf_{x + y = z} \{\|c - x\|^2 + \|d - y\|^2\},$$

and by observing that the inner infimum is attained at

$$(x,y) = \Pi_{\Delta(z)}(c,d).$$

To prove part (c), observe that the term on the right-hand side of (2.6) corresponds to the infimal-convolution of the functions d_X^2 and d_Y^2. General results concerning the attainment of an infimal-convolution (cf. [La, Chapter 6]) lead to the characterization (2.8). Finally, consider part (d) of the lemma. The projection $\Pi_{\Delta(z)-X\times Y}(0,0)$ is the unique solution to the minimization problem

$$\alpha := \inf\{\|(p,q)\| : (p,q) \in \Delta(z) - X \times Y\}. \tag{2.10}$$

To check that the above infimum is attained, pick up any solution $(\tilde{x},\tilde{y})$ to (2.6). A simple calculus shows that

$$\begin{aligned} \alpha^2 &= \delta^2(X \times Y, \Delta(z)) \\ &= \|\Pi_X(\tilde{x}) - \tilde{x}\|^2 + \|\Pi_Y(\tilde{y}) - \tilde{y}\|^2, \end{aligned}$$

or, equivalently,

$$\alpha = \|(\tilde{x},\tilde{y}) - (\Pi_X(\tilde{x}), \Pi_Y(\tilde{y}))\|.$$

This means that the infimum in (2.10) is attained at

$$\begin{aligned} (p,q) &= (\tilde{x},\tilde{y}) - (\Pi_X(\tilde{x}), \Pi_Y(\tilde{y})) \\ &= ([I - \Pi_X](\tilde{x}),\ [I - \Pi_Y](\tilde{y})). \end{aligned}$$

It is fairly clear that the same pair (p,q) characterizes the projection of $(0,0)$ onto $\overline{\Delta(z) - X \times Y}$. To complete the proof it suffices to observe that $p = q$. □

Now we are ready to state:

Theorem 2.2. *Let $X, Y \in P(H)$ and let $\{(x_n, y_n)\}_{n\geq 1}$ be a sequence generated by Algorithm A2. Then,*

(a) $\{(x_n, y_n)\}_{n\geq 1} \subset \Delta(z)$;

(b) $\Pi_X(x_n) + \Pi_Y(y_n) \to z - 2[I - \Pi_X](\tilde{x}) = z - 2[I - \Pi_Y](\tilde{y})$, *where* $(\tilde{x}, \tilde{y})$ *is a pair at which the infimal-convolution (2.6) is attained. In particular,*

$$\|\Pi_X(x_n) + \Pi_Y(y_n) - z\| \to d_{X+Y}(z);$$

(c) If $z \in X + Y$, *then both sequences* $\{(x_n, y_n)\}_{n\geq 1}$ *and* $\{(\Pi_X(x_n),\ \Pi_Y(y_n))\}_{n\geq 1}$ *converge weakly to some* $(x, y) \in D(z)$.

Proof. We apply again Lemma I (cf. Appendix), but this time we choose $A = X \times Y$ and $B = \Delta(z)$. The terms of the von Neumann sequences are generated by

$$\left\{ \begin{array}{lcl} (c_n, d_n) & = & \Pi_{X\times Y}(x_{n-1}, y_{n-1}), \\ (x_n, y_n) & = & \Pi_{\Delta(z)}(c_n, d_n). \end{array} \right. \tag{2.11}$$

The first equality in (2.11) splits into

$$c_n = \Pi_X(x_{n-1}) \text{ and } d_n = \Pi_Y(y_{n-1}), \tag{2.12}$$

and, according to Lemma 2.2(a), the second equality in (2.11) can be written in the form

$$\left\{ \begin{array}{lcl} x_n & = & \dfrac{z}{2} + \dfrac{1}{2}[c_n - d_n], \\ y_n & = & \dfrac{z}{2} - \dfrac{1}{2}[c_n - d_n]. \end{array} \right. \tag{2.13}$$

Algorithm A2 is, of course, a combination of (2.12) and (2.13). Part (a) of the theorem just says that

$$x_n + y_n = z \quad \text{for all } n \geq 1,$$

which is immediate from (2.13). Consider now part (b). According to Lemma I, $(x_n, y_n) - (c_n, d_n)$ converges to $\Pi_{\overline{\Delta(z)-X\times Y}}(0,0)$. But,

$$(x_n, y_n) - (c_n, d_n) = \left(\frac{z}{2}, \frac{z}{2}\right) - \left(\frac{c_n + d_n}{2}, \frac{c_n + d_n}{2}\right),$$

and

$$\Pi_{\overline{\Delta(z)-X\times Y}}(0,0) = ([I - \Pi_X](\tilde{x}), [I - \Pi_X](\tilde{x})).$$

Hence

$$c_n + d_n \to z - 2[I - \Pi_X](\tilde{x}).$$

We know also that

$$\|(x_n, y_n) - (c_n, d_n)\| \to \delta(X \times Y, \Delta(z)).$$

But, according to Lemma 2.2(b), this is equivalent to

$$2\left\|\frac{c_n + d_n}{2} - \frac{z}{2}\right\|^2 \to \frac{1}{2}d^2_{X+Y}(z).$$

This takes care of the part (b). Finally, if $z \in X + Y$, then $\delta(X \times Y, \Delta(z)) = 0$, and

$$E = F = X \times Y \cap \Delta(z).$$

Lemma I ensures the weak convergence of the sequences $\{(x_n, y_n)\}_{n\geq 1}$ and $\{(c_n, d_n)\}_{n\geq 1}$ toward some $(x, y) \in X \times Y \cap \Delta(z)$. □.

3 Finding The Least Deviation Decomposition.

This section deals with the problem of finding the least deviation decomposition of a given vector $z \in X + Y$. The first algorithm that we are going to consider is based on the following result.

Lemma 3.1. [LMS]. *Let $X, Y \in P(H)$ and $z \in X+Y$. Then the following statements are equivalent:*

(a) $(\bar{x}, \bar{y})$ is the least deviation decomposition of z;

(b) $\bar{x}$ and $\bar{y}$ are the projections of $z/2$ onto $X \cap (z-Y)$ and $Y \cap (z-X)$, respectively, i.e.

$$\begin{cases} \bar{x} = \text{Argmin}\left\{\left\|\frac{z}{2} - x\right\| : x \in X \cap (z-Y)\right\}, \\ \bar{y} = \text{Argmin}\left\{\left\|\frac{z}{2} - y\right\| : y \in Y \cap (z-X)\right\}, \end{cases} \tag{3.1}$$

As indicated in the above lemma, to find the component $\bar{x}$ one needs to compute the projection of $z/2$ onto the intersection of X and $z-Y$. This can be done by adapting Dykstra's alternating projection algorithm. One gets in this way:

Algorithm A3. Take $u_0 = v_0 = 0$ and $y_0 = z/2$. For $n \geq 1$, compute

$$\begin{cases} x_n = \Pi_X(z - y_{n-1} + u_{n-1}), \\ u_n = z - y_{n-1} - x_n + u_{n-1}, \\ y_n = \Pi_Y(z - x_n + v_{n-1}), \\ v_n = z - y_n - x_n + v_{n-1}. \end{cases} \tag{3.2}$$

Theorem 3.1. *Let $X, Y \in P(H)$, and let $\{(x_n, y_n)\}_{n \geq 1}$ be the sequence generated by Algorithm A3. Then,*

(a) $\{(x_n, y_n)\}_{n\geq 1} \subset X \times Y$;

(b) $x_n + y_n \to \Pi_{\overline{X+Y}}(z)$. In particular, $\|x_n + y_n - z\| \to d_{X+Y}(z)$;

(c) If $z \in X+Y$, then $\{(x_n, y_n)\}_{n\geq 1}$ converges (strongly) to the least deviation decomposition $(\bar{x}, \bar{y})$ of z.

Proof. We apply Lemma II (cf. Appendix) to the particular case $A = X$ and $B = z - Y$. For convenience, one changes variables in the way indicated below:

$$a_n \leftarrow x_n \quad , \quad b_n \leftarrow z - y_n$$
$$p_n \leftarrow u_n \quad , \quad q_n \leftarrow -v_n.$$

The rest of the proof is similar to that of Theorem 2.1. □

Algorithm A3 has some advantages over Algorithm A1. First of all, the sequence $\{(x_n, y_n)\}_{n\geq 1}$ converges to a very special admissible decomposition of z. Indeed, it converges to the best one, namely, the least deviation decomposition of z. Second, the convergence is not only in the weak sense, but also in the strong one.

The algorithm discussed next is based on a different characterization of the concept of least deviation decomposition.

Lemma 3.2. *[LMS]. Let $X, Y \in P(H)$ and $z \in X+Y$. Then the following statements are equivalent:*

(a) $(\overline{x}, \overline{y})$ is the least deviation decomposition of z;

(b) $(\overline{x}, \overline{y})$ is the projection of $(0,0)$ onto $D(z)$.

So, to find the pair $(\overline{x}, \overline{y})$ one needs to project the origin $(0,0) \in H \times H$ onto the intersection of $X \times Y$ and $\Delta(z)$. This observation leads us to the following

Algorithm A4. Take $p_0 = q_0 = u_0 = v_0 = x_0 = y_0 = 0$. For $n \geq 1$, compute

$$\begin{cases} c_n = \Pi_X(x_{n-1} + p_{n-1}), \\ d_n = \Pi_Y(y_{n-1} + q_{n-1}), \\ p_n = x_{n-1} + p_{n-1} - c_n, \\ q_n = y_{n-1} + q_{n-1} - d_n, \\ e_n = (c_n + u_{n-1} - d_n - v_{n-1})/2, \\ x_n = \dfrac{z}{2} + e_n, \\ y_n = \dfrac{z}{2} - e_n, \\ u_n = c_n + u_{n-1} - x_n, \\ v_n = d_n + v_{n-1} - y_n. \end{cases} \tag{3.3}$$

The limiting behavior of the main sequences $\{(x_n, y_n)\}_{n\geq 1}$ and $\{(c_n, d_n)\}_{n\geq 1}$ is discussed in the next theorem. The other sequences in (3.3) are only auxiliary and do not deserve further discussion.

Theorem 3.2. *Let $X, Y \in P(H)$, and let $\{(x_n, y_n)\}_{n\geq 1}$ and $\{(c_n, d_n)\}_{n\geq 1}$ be the sequences generated by Algorithm A4. If $z \in X + Y$, then both sequences converge (strongly) to the least deviation decomposition $(\overline{x}, \overline{y})$ of z.*

Proof. We apply Lemma II (cf. Appendix) to the case $A = X \times Y$ and $B = \Delta(z)$. Dykstra's algorithm takes the form

$$\begin{aligned}
(c_n, d_n) &= \Pi_{X\times Y}((x_{n-1}, y_{n-1}) + (p_{n-1}, q_{n-1})),\\
(p_n, q_n) &= (x_{n-1}, y_{n-1}) + (p_{n-1}, q_{n-1}) - (c_n, d_n),\\
(x_n, y_n) &= \Pi_{\Delta(z)}((c_n, d_n) + (u_{n-1}, v_{n-1})),\\
(u_n, v_n) &= (c_n, d_n) + (u_{n-1}, v_{n-1}) - (x_n, y_n).
\end{aligned}$$

All these equalities occur in the product space $H \times H$. After decoupling and rearranging in a suitable way, one ends up with the relations given in (3.3). If $z \in X + Y$, then

$$E = F = X \times Y \cap \Delta(z),$$

and the sequences $\{(x_n, y_n)\}_{n\geq 1}$ and $\{(c_n, d_n)\}_{n\geq 1}$ converge to the projection of $(0, 0)$ onto $X \times Y \cap \Delta(z)$. □

Appendix

The following two results have played a key role in our work. The first one deals with von Neumann's alternating projection algorithm.

Lemma I. (BB2, Theorem 4.8]). *Let $\mathcal{H}$ be a Hilbert space, and let $A, B \in P(\mathcal{H})$. Choose any $b_0 \in \mathcal{H}$ and define the terms of the von Neumann sequences by*

$$a_n = \Pi_A(b_{n-1}), \quad b_n = \Pi_B(a_n) \quad \textit{for all } n \geq 1.$$

Then $b_n - a_n \to \Pi_{\overline{B-A}}(0)$. In particular,

$$\|b_n - a_n\| \to \delta(A, B) := \inf\{\|a - b\| : \ a \in A, \ b \in B\}.$$

If the above infimum is attained, then (a_n, b_n) converges weakly to some pair (a^*, b^*) such that

$$\begin{aligned} b^* - a^* &= \Pi_{\overline{B-A}}(0), \\ a^* \in E: &= \{a \in A: \ d_B(a) = \delta(A, B)\}, \\ b^* \in F: &= \{b \in B: \ d_A(b) = \delta(A, B)\}. \end{aligned}$$

The second result deals with Dykstra's alternating projection algorithm.

Lemma II. ([BB2, Theorem 3.8]). *Let $\mathcal{H}$ be a Hilbert space, and let $A, B \in P(\mathcal{H})$. Take $p_0 = q_0 = 0$ and $b_0 = w$, where w is any point in $\mathcal{H}$. For $n \geq 1$, define the terms of the Dykstra sequences by*

$$\begin{aligned} a_n &= \Pi_A(b_{n-1} + p_{n-1}), \\ p_n &= b_{n-1} + p_{n-1} - a_n, \\ b_n &= \Pi_B(a_n + q_{n-1}), \\ q_n &= a_n + q_{n-1} - b_n. \end{aligned}$$

Then, $b_n - a_n \to \Pi_{\overline{B-A}}(0)$. In particular,

$$\|b_n - a_n\| \to \delta(A, B) := \inf\{\|a - b\| : a \in A, \quad b \in B\}.$$

If the above infimum is attained, then

$$a_n \to \Pi_E(w) \text{ and } b_n \to \Pi_F(w),$$

where E and F are as in Lemma I.

Acknowledgement. I am indebted to my colleagues D.T. Luc and J.E. Martinez–Legaz for discussions on the least deviation decomposition problem. H.H. Bauschke and J.M. Borwein kindly provided me with copies of their work on alternating projection methods.

References

[AT] H. ATTOUCH and M. THERA, A general duality principle for the sum of two operators, Preprint, 1994, Laboratoire d'Analyse Convexe, Université de Montpellier II.

[BB1] H.H. BAUSCHKE and J.M. BORWEIN, On the convergence of von Neumann's alternating projection algorithm for two sets. Set–Valued Analysis 1(1993), 185–212.

[BB2] H.H. BAUSCHKE and J.M. BORWEIN, Dykstra's alternating projection algorithm for two sets, J. Approximation Theory 79(1994), 418–443.

[BL] J.M. BORWEIN and A.S. LEWIS, Practical conditions for Fenchel duality in infinite dimensions, Proceedings of the International Conference on Fixed Points and Applications, J.B. Baillon and M. Thera Editors, Pitman Research Notes in Mathematics 252(1990), 83–99.

[FL] C. FRANCHETTI and W. LIGHT, On the von Neumann alternating algorithm in Hilbert space, J. Math. Anal. Appl. 114(1986), 305–314.

[La] P.J. LAURENT, Approximation et optimisation, Hermann, Paris, 1972.

[LMS] D.T. LUC, J.E. MARTINEZ–LEGAZ and A. SEEGER, Least deviation decomposition with respect to a pair of convex sets, Preprint, 1995, Dept. of Math. Sciences, King Fahd University of Petroleum and Minerals, Dhahran, Saudi Arabia.

[MS] J.E. MARTINEZ-LEGAZ and A. SEEGER, A general cone decomposition theory based on efficiency, Math. Programming 65(1994), 1–20.

[Ro] S.M. ROBINSON, Regularity and stability for convex multivalued functions, Math. Oper. Res. 1(1976), 130–143.

[Vo] M. VOLLE, Sur quelques formules de dualité convexe et non convexe, Set–Valued Analysis 2(1994), 369–379.