ALTERNATING PROJECTIONS ON NON-TANGENTIAL MANIFOLDS.

FREDRIK ANDERSSON AND MARCUS CARLSSON

Abstract. We consider sequences $(B_k)_{k=0}^{\infty}$ of points obtained by projecting back and forth between two manifolds $\mathcal{M}_1$ and $\mathcal{M}_2$, and give conditions guaranteeing that the sequence converge to a limit $B_\infty \in \mathcal{M}_1 \cap \mathcal{M}_2$. Our motivation is the study of algorithms based on finding the limit of such sequences, which have proven useful in a number of areas. The intersection is typically a set with desirable properties, but for which there is no efficient method of finding the closest point B_{opt} in $\mathcal{M}_1 \cap \mathcal{M}_2$. We prove not only that the sequence of alternating projections converges, but that the limit point is fairly close to B_{opt}, in a manner relative to the distance $\|B_0 - B_{opt}\|$, thereby significantly improving earlier results in the field. A concrete example with applications to frequency estimation of signals is also presented.

1. Introduction

Let $\mathcal{K}$ be a finite dimensional Hilbert space over $\mathbb{R}$ and let $\mathcal{M}_1,\ \mathcal{M}_2 \subset \mathcal{K}$ be manifolds. Suppose that for any $B \in \mathcal{K}$ the closest point on $\mathcal{M}_j$, $j = 1, 2$, is well defined and lets denote it by $\pi_j(B)$. Let the corresponding projection onto the intersection $\mathcal{M}_1 \cap \mathcal{M}_2$ be denoted by $\pi(B)$. Suppose that we are interested in finding the closest point on $\mathcal{M}_1 \cap \mathcal{M}_2$ and that the "projection operators" π_1 and π_2 can be efficiently computed, whereas π can not. The issue treated here is how to use π_1 and π_2 to obtain an approximation of π. A classical result by von Neumann [18] says that if $\mathcal{M}_1$ and $\mathcal{M}_2$ are affine linear manifolds, then the sequence of alternating projections

$$\pi_1(B),\ \pi_2(\pi_1(B)),\ \pi_1(\pi_2(\pi_1(B))),\ \pi_2(\pi_1(\pi_2(\pi_1(B)))), \dots \tag{1.1}$$

converges to $\pi(B)$. Moreover, the convergence rate is determined by the angle between $\mathcal{M}_1$ and $\mathcal{M}_2$. This paper is concerned with extensions of this result to non-linear manifolds.

Hence, given $B \in \mathcal{K}$, let $B_1 = \pi_1(B)$ and

$$B_{k+1} = \begin{cases} \pi_1(B_k) & k, \text{ is even,} \\ \pi_2(B_k) & k, \text{ is odd.} \end{cases} \tag{1.2}$$

In contrast to the case where $\mathcal{M}_j$ is an affine linear manifold, $B_\infty \neq \pi(B)$. However, given that $\mathcal{M}_j$ behave nicely, we may expect $\mathcal{B}_\infty \approx \pi(B)$.

Alternating projection schemes of this kind have been used in a number of applications, cf. [7, 10, 12, 14, 15, 17, 16, 19]. For instance, $\mathcal{K}$ can be the set $\mathbb{M}_{m,n}$ of $m \times n$-matrices, and the manifolds $\mathcal{M}_j$ are subsets with a certain structure, e.g. matrices with a certain rank, self-adjoint matrices, Hankel or Toeplitz matrices

2010 *Mathematics Subject Classification.* 41A65, 49Q99, 53B25.

Key words and phrases. Alternating projections, convergence, Non-convexity.

etc. Alternating projection schemes between several linear subsets in the infinite dimensional setting was recently investigated in [2]. Much emphasis has been put towards the usage of alternating projections for the case of complex manifolds, see for instance [3, 4]. Connections to the EM algorithm are given in [5].

In [7], Zangwill's Global Convergence Theorem [20] is used to motivate the convergence of the alternating projection scheme above. From Zangwill's theorem it is possible to deduce that if the sequence $(B_k)_{k=1}^{\infty}$ is bounded and the distance to $\mathcal{M}_1 \cap \mathcal{M}_2$ is strictly decreasing, then $(B_k)_{k=1}^{\infty}$ has a convergent subsequence to a point $B_\infty \in \mathcal{M}_1 \cap \mathcal{M}_2$, i.e., a result related to the fact that any bounded sequence in a compact set has a convergent subsequence. Thus, the use of Zangwill's theorem in this context does not provide any information about whether the limit point B_∞ exists, or if so, whether it is close to $\pi(B)$.

Recently, A. Lewis and J. Malick presented stronger results, valid under more restrictive conditions on $\mathcal{M}_1$ and $\mathcal{M}_2$. Before discussing their results in more detail, we give two simple examples that illustrate some of the difficulties that may arise when using alternating projections on non-linear manifolds.

Example 1.1. Let $\mathcal{K} = \mathbb{R}^2$ and set $\mathcal{M}_1 = \{(t, (t+1)(3-t)/4) : t \in \mathbb{R}\}$ and $\mathcal{M}_2 = \mathbb{R} \times \{0\}$. It is easily seen that $\pi_1((1,0)) = (1,1)$ and $\pi_2((1,1)) = (1,0)$, and hence the sequence of alternating projections does not converge, cf. Figure 1. On the other hand, if $(1+\epsilon, 0) \in \mathcal{M}_2$, $\epsilon > 0$, is used as a starting point, the sequence of alternating projections will converge to $(3,0) \in \mathcal{M}_1 \cap \mathcal{M}_2$.

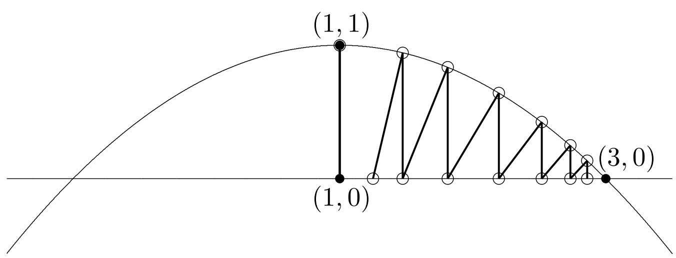


FIGURE 1. An example demonstrating that the algorithm can get stuck in loops where even no subsequence converges to a point in the intersection.

In general, it seems reasonable to assume that if the starting point is sufficiently close to an intersection point, then the sequence does converge to a point in the intersection. The next example shows that this is not the case of limited smoothness.

Example 1.2. Without going in to the details of the construction, we note that one can construct a C^1-function f such that, with $\mathcal{M}_1 = \mathbb{R} \times \{0\}$ and $\mathcal{M}_2 = \{(t, f(t)) : t \in \mathbb{R}\}$, the sequence of alternating projections can get stuck in projecting back and forth between the same two points. Figure 2 explains the idea.

However, if we have $f \in C^1$ and $f'(0) \neq 0$, it is hard to imagine how to make a similar construction work. This is indeed impossible if we assume additional smoothness, which follows both from the present paper and [13]. In the terminology of the latter, the condition $f'(0) \neq 0$ implies that $\mathcal{M}_1$ and $\mathcal{M}_2$ are *transversal* at

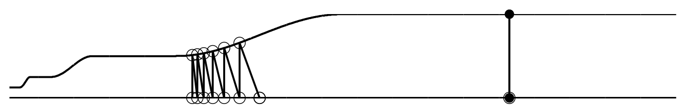


FIGURE 2. Alternating projections stuck in a loop.

$(0,0)$. In general, given C^1-manifolds $\mathcal{M}_1,\ \mathcal{M}_2$ and a point $A \in \mathcal{M}_1 \cap \mathcal{M}_2$, we say that A is *transversal* if

$$T_{\mathcal{M}_1}(A) + T_{\mathcal{M}_2}(A) = \mathcal{K}, \tag{1.3}$$

where $T_{\mathcal{M}_j}(A)$ denotes the tangent-space of $\mathcal{M}_j$ at A, $j = 1, 2$. The main result in [13] is roughly the following:

Theorem 1.3. *Let $\mathcal{M}_1$ and $\mathcal{M}_2$ be C^3-manifolds and let $A \in \mathcal{M}_1 \cap \mathcal{M}_2$ be transversal. If B is close enough to A, then the sequence of alternating projections $(B_k)_{k=1}^{\infty}$ given by (1.2) converges to a point B_∞ in $\mathcal{M}_1 \cap \mathcal{M}_2$. Moreover,*

$$\|B_\infty - \pi(B)\| \leq 2\|A - B\|.$$

The improvement over Zangwill's theorem is thus that the entire sequence converges, and that the assumption of boundedness no longer is necessary. Moreover, the limit B_∞ is not too far off from $\pi(B)$, (although in relative terms, i.e. comparing with $\mathsf{dist}(B, \mathcal{M}_1 \cap \mathcal{M}_2) = \|B - \pi(B)\|$, it need not be particularly close either).

On the other hand, the assumption of transversalsality is rather restrictive. To demonstrate the essence of the transversality assumption, we now present two cases which are not covered by the above result, but where the conclusion still holds.

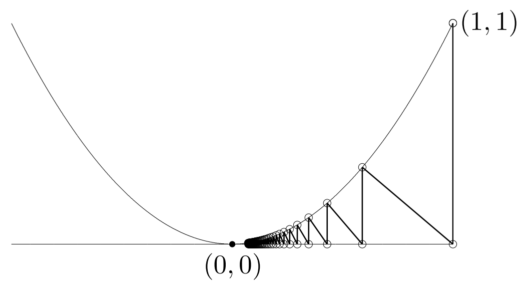


FIGURE 3. Lack of transversality due to tangential curves.

Example 1.4. With $\mathcal{M}_1 = \mathbb{R} \times \{0\}$ and $\mathcal{M}_2 = \{(t, t^2) : t \in \mathbb{R}\}$, transversality is not satisfied at $A = (0,0)$ (since $T_{\mathcal{M}_1}(A) = T_{\mathcal{M}_2}(A) = \mathcal{M}_1$), but it is not hard to see that the sequence of alternating projections still converges to $(0,0)$, (see Figure 3).

The difference between Example 1.2 and 1.4 is that in the latter case the manifolds are more regular.

Example 1.5. With $\mathcal{K} = \mathbb{R}^3$ and $\mathcal{M}_1 = \mathbb{R} \times \{0\}^2$ and $\mathcal{M}_2 = \{(t, t, t^2) : t \in \mathbb{R}\}$, transversality is not satisfied at $A = (0,0)$, but again it seems plausible that the sequence of alternating projections converges to $(0,0)$. See Figure 4.

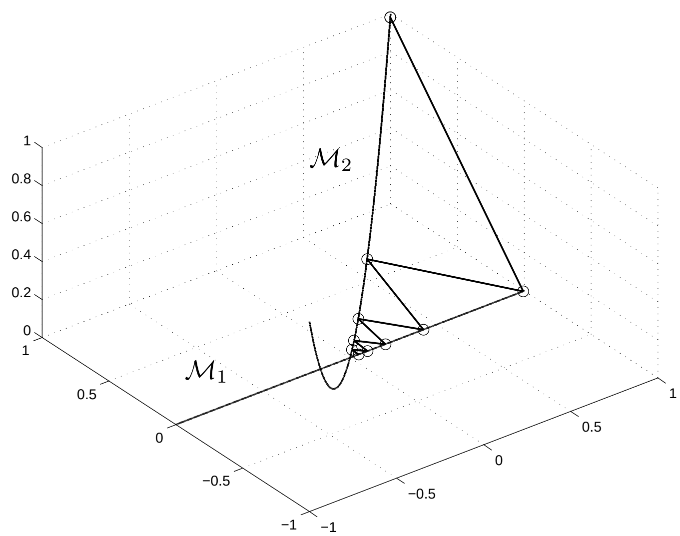


FIGURE 4. Lack of transversality due to too low dimensionality.

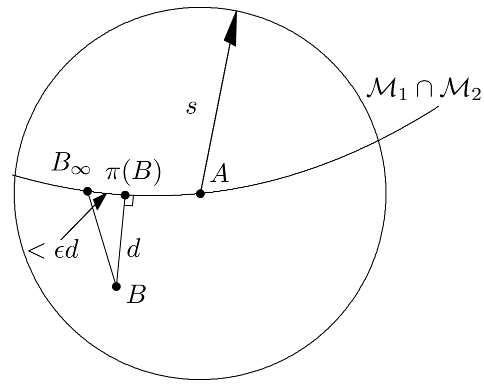
In Example 1.5, the two manifolds clearly sit at a positive angle, but this situation is not covered by Theorem 1.3, since the manifolds are of too low dimension to satisfy the transversality assumption. In fact, if $\mathcal{K}$ has dimension n and $\mathcal{M}_j$ has dimension m_j, $j = 1, 2$, the transversality (1.3) can never be satisfied if $m_1 + m_2 < n$. But for many applications of practical interest, one has $m_1 + m_2 << n$. We will introduce a concept which we call *non-tangential*, which loosely speaking says that the manifolds should have a positive angle in directions perpendicular to $\mathcal{M}_1 \cap \mathcal{M}_2$. The point $(0, 0)$ in Example 1.5 is non-tangential, whereas the same point in Example 1.4 is not. A proper definition of non-tangentiality is given in Definition 4.3. A simplified version of our main result is given in what follows.

Theorem 1.6. *Given a non-tangential point $A \in \mathcal{M}_1 \cap \mathcal{M}_2$ there exists an $s > 0$ such that the sequence of alternating projections (1.2) converges to a point $B_\infty \in \mathcal{M}_1 \cap \mathcal{M}_2$, given that $\|B - A\| < s$. Moreover, given any $\epsilon > 0$ one can take s such that*

$$\|B_\infty - \pi(B)\| < \epsilon \|B - \pi(B)\|.$$

The full version of the main theorem is given in Section 6.

The improvement over Theorem 1.3 mainly consists of two items. Primarily, the assumption that the surfaces be non-tangential is not at all restrictive, and in particular there is no implication on the dimensions of $\mathcal{M}_1$ and $\mathcal{M}_2$. For the applications we are aware of, the set of tangential points is very small, if it exists at all. Secondly, as has been highlighted before, we are usually interested not just in any point of $\mathcal{M}_1 \cap \mathcal{M}_2$, but the closest point $\pi(B)$. Here the theorem says that in relative terms, i.e. after dividing with the distance to $\mathcal{M}_1 \cap \mathcal{M}_2$, the error is small if the distance to $\mathcal{M}_1 \cap \mathcal{M}_2$ is small. It is also worth mentioning that we only assume that the manifolds are C^2, although the manifolds are C^∞ in the applications that we are aware of.



In order to motivate the need for certain technicalities in the abstract setting, as well as the need for a theorem concerning low-dimensional manifolds, we will develop the theory in parallel with an example, namely that when $\mathcal{K}$ is the space $\mathbb{M}_{n,n}(\mathbb{C})$ of complex $n \times n$-matrices, $\mathcal{M}_1$ is the set of Hankel matrices and $\mathcal{M}_2$ the set of matrices of rank at most k. The former is a subspace of real dimension $2(n-1)$, whereas the latter is not actually a manifold. However, it is "locally" a manifold of real dimension $2(2nk-k^2)$ at all matrices A with rank precisely k, (which clearly constitutes the majority of matrices in the set). If k is small and A is a non-tangential intersection point, then $\dim(T_{\mathcal{M}_1}) + \dim(T_{\mathcal{M}_2}) = 2(2nk - k^2 + n - 1)$, which is much less than $\dim \mathcal{K} = 2n^2$, unless k is close to n. Another thing that does not match between this example and the theory outlined above is that the projection π_2 is not uniquely defined at certain points. This is common in algorithmic theory and can be dealt with by using point-to-set maps, following Zangwill [20]. However, in our case, this is not necessary since we show that the projections are locally well defined near non-tangential points, (Proposition 3.3). Moreover, from a practical perspective this is unnecessary since such points are rarely encountered in applications.

For the above example, the interest in $\mathcal{M}_1 \cap \mathcal{M}_2$ lies in the fact that such sequences are samplings of functions which are sums of k exponential functions. Given a function f on an interval the alternating projections algorithm can then be used to find approximations of f by sums of k exponential functions, which is a problem of great practical interest. We demonstrate the idea in Figure 5. More thorough examples are conducted in [1].

2. Case study; rank k matrices versus Hankel matrices

We include this section for the reader to get the picture of a typical application. However, this particular application is treated in detail in [1], and therefore we will be very brief here. In the setting considered by Lewis and Malick, the fact that $\mathcal{M}_1 \cap \mathcal{M}_2$ is itself a manifold (around a point A, say) follows by the tranversality assumption (1.3) and standard differential geometry. In our (more general) setting, $\mathcal{M}_1 \cap \mathcal{M}_2$ may fail to be a manifold. However, in the example developed in this section, $\mathcal{M}_1 \cap \mathcal{M}_2$ will be a manifold locally. In fact, it is not hard to see that it will always happen for algebraic manifolds, which is the case for all of the various applications presented in [7].

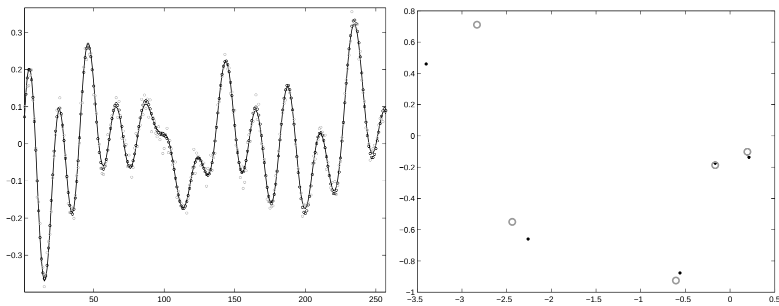


FIGURE 5. Signal with five exponentials. Left panel: original signal (black solid line); signal with white noise (gray circles); and the reconstruction using the alternating Hankel projection method (black circes). Right panel: original exponential nodes (black dots), estimated nodes from noisy signal (gray circles.)

Let $n \in \mathbb{N}$ be fixed, let $\mathcal{K} = \mathbb{M}_{n,n}(\mathbb{C})$ and let $\mathcal{H} \subset \mathcal{K}$ be the set of Hankel matrices, e.g., matrices of the form

$$\begin{pmatrix} a_1 & a_2 & a_3 & \cdots & a_n \\ a_2 & a_3 & \iddots & a_n & a_{n+1} \\ a_3 & \iddots & \iddots & \iddots & \vdots \\ \vdots & a_n & \iddots & \iddots & a_{2n-2} \\ a_n & a_{n+1} & \cdots & a_{2n-2} & a_{2n-1} \end{pmatrix} \tag{2.1}$$

$\mathcal{H}$ is a linear subspace and, hence, a manifold at each point, of (real) dimension $2(2n-1)$. Denoting the matrix in (2.1) by $H(a)$, where $a = (a_1, \dots, a_{2n-1})$, we obtain a natural chart for $\mathcal{H}$, (by identifying $\mathbb{C}$ with $\mathbb{R}^2$ in the obvious way).

Given $k < n$, we denote by $\mathcal{R}_k \subset \mathcal{K}$ the set of matrices of rank less than or equal to k. We wish to do alternating projections as outlined in the introduction between $\mathcal{H}$ and $\mathcal{R}_k$, which is slightly complicated by the fact that $\mathcal{R}_k$ is not a manifold. However, it turns out that $\mathcal{R}_k$ is locally a manifold of (real) dimension $2(2nk - k^2)$, apart from at some exceptional points. More precisely, suppose $A \in \mathcal{R}_k$, and use the singular value decomposition of A to find $\sigma_A \in (\mathbb{R}^+)^k$ and unitary matrices $U_A, V_A \in \mathbb{M}_{n,k}$ such that

$$A = V_A \begin{pmatrix} \sigma_{A,1} & 0 & \cdots & 0 \\ 0 & \sigma_{A,2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \sigma_{A,k} \end{pmatrix} U_A^* = V_A I_\sigma U_A^*. \tag{2.2}$$

The vast majority of matrices in $\mathcal{R}_k$ satisfy

$$\sigma_{A,1} > \sigma_{A,2} > \ldots > \sigma_{A,k} > 0, \tag{2.3}$$

and an arbitrarily small numerical perturbation will yield distinct singular values. The subset of $\mathcal{R}_k$ satisfying (2.3) will be denoted $\mathcal{R}_k^d$, where d stands for distinct. If $\mathcal{M}$ is a manifold and its closure $\overline{\mathcal{M}}$ is such that $\overline{\mathcal{M}} \setminus \mathcal{M}$ is a union of manifolds of lower dimension than $\mathcal{M}$, we will say that $\overline{\mathcal{M}} \setminus \mathcal{M}$ is *thin*. The proof of the following proposition can be found in [1].

Proposition 2.1. *$\mathcal{R}_k^d$ is a manifold of (real) dimension $2(2nk - k^2)$. Moreover, $\mathcal{R}_k = \overline{\mathcal{R}_k^d}$ and $\mathcal{R}_k \setminus \mathcal{R}_k^d$ is thin.*

The structure of the set $\mathcal{R}_k$ around the exceptional points in $\mathcal{R}_k \setminus \mathcal{R}_k^d$ can be rather complicated, and although the sequence generated by the alternating projection scheme could theoretically approach such a point, we have not done any analysis of convergence properties in this setting. This seems to be a hard problem. The rate of convergence is however very poor around such points (compare Figure 1 with Figure 3). The first contains 12 iterations and the second 100.

We note that by the Eckart-Young theorem [11], it is easy to project any matrix $B \in \mathbb{M}_{n,n}$ onto its closest point in $\mathcal{R}_k$ while using the Frobenius norm

$$\|B\|_2 = \sqrt{\sum_{i,j} |b_{i,j}|^2}.$$

Specifically, the theorem says that if $B = V_B I_\sigma U_B$ is a singular value decomposition of B, then the closest point in $\mathcal{R}_k$ is given by replacing I_σ with I_τ where $\tau = (\sigma_1, \ldots, \sigma_k, 0, \ldots, 0)$. The closest point is thus unique as long as the singular values are distinct, which is always the case when working with "real numerical" data, so we will for simplicity treat the projection onto $\mathcal{R}_k$ as a well defined map which we denote by $\pi_{\mathcal{R}_k}$. More stringently one could work with "point to set"-maps, as in [7] and [20].

The last manifold that needs to be discussed is

$$\mathcal{H}_k = \mathcal{R}_k \cap \mathcal{H},$$

i.e., the set of Hankel matrices with rank $\leq k$. It is easily seen that, given any $\alpha \in \mathbb{C}$, the matrix

$$H(\alpha) = \begin{pmatrix} 1 & \alpha & \alpha^2 & \cdots & \alpha^{n-1} \\ \alpha & \alpha^2 & \iddots & \alpha^{n-1} & \alpha^n \\ \alpha^2 & \iddots & \iddots & \iddots & \vdots \\ \vdots & \alpha^{n-1} & \iddots & \iddots & \alpha^{2n-3} \\ \alpha^{n-1} & \alpha^n & \cdots & \alpha^{2n-3} & \alpha^{2n-2} \end{pmatrix} \tag{2.4}$$

defines a rank 1 Hankel matrix, and thus

$$\mathfrak{H}(c, \alpha) = \sum_{j=1}^{k} c_j H(\alpha_j) \tag{2.5}$$

has rank k. This is the typical case, as we will show below. Let us denote the image of $\mathfrak{H}$ intersected with $\mathcal{R}_k^d$ by $\mathcal{H}_k^n$, where n stands for "nice". Again, the analysis is complicated by some exceptional points. For example, $\frac{d^j}{d\alpha^j} H(\alpha)$ is a Hankel matrix of rank j which is not covered by $\mathfrak{H}$. Moreover, rank k Hankel matrices containing summands of this type are usually inside $\mathcal{R}_k^d$, but despite that, some investigations show that the manifold structure of $\mathcal{H}_k$ collapses around such points. The sup-script "$n''(=$ *nice*) is thus really more restrictive than "$d''(=$ *distinct*). Nevertheless, $\mathcal{H}_k \setminus \mathcal{H}_k^n$ is thin and never encountered in practice, so we omit a study of such cases. The proof of the following result can be found in [1].

Proposition 2.2.
$\mathcal{H}$ is a $2(2n-1)$-dimensional linear subspace.
$\mathcal{H}_k^n$ is a $4k$-dimensional manifold which is dense in $\mathcal{H}_k$, its complement is thin.
The map $\pi_{\mathcal{R}_k}$ is well defined on all points in $\mathbb{M}_{n,n}$ except a for thin subset.

We now explain why one would like to do alternating projections between $\mathcal{H}$ and $\mathcal{R}_k$. Let $\pi_{\mathcal{H}}$ be the orthogonal projection onto $\mathcal{H}$, and let $\pi(B)$ denote the closest point to B in $\mathcal{H} \cap \mathcal{R}_k$. Say we have a sampled function on an interval, e.g., the signal in Figure 5. Denote the corresponding sequence by f and note that

$$\sum_{j=1}^{2n-1} |f_j|^2(n - |j-n|) = \|H(f)\|_2^2. \tag{2.6}$$

We denote the weight sequence in (2.6) by w and the corresponding norm on $\mathbb{C}^{2n-1}$ by $\|f\|_w^2$. By Proposition 2.2, the function g defining $\pi(B) = H(g)$ is almost surely in the range of $\mathfrak{H}$, i.e. of the form

$$g = \sum_{j=1}^{k} c_j (\alpha_j^l)_{l=0}^{2n-2}.$$

Such functions are precisely what one encounters when sampling sums of k exponential functions, and hence g is the closest such function to f in the norm given by (2.6). Approximating a given function with sums of a predetermined number of exponentials is a problem with a large number of applications. There is, however, no computationally efficient method for computing g. In contrast, the projections $\pi_{\mathcal{R}_k}$ and $\pi_{\mathcal{H}}$ are relatively easy to compute, and hence, by the results of this paper, we can get a fairly good approximation of g by finding the limit of the sequence of alternating projections, starting at $H(f)$.

For the sake of efficient computations, care has to be taken concerning the implementation of $\pi_{\mathcal{R}_k}$ and $\pi_{\mathcal{H}}$. Constructing fast algorithms for this purpose is one of the topics of a companion paper [1], where we also discuss other aspects of this specific application of alternating projections as well as give proofs of the above claims. In the present paper, we will continue discussing this application in Section 7.

3. Preliminaries

Let $\mathcal{K}$ be a Hilbert space of dimension $n \in \mathbb{N}$. Given $A \in \mathcal{K}$ and $r > 0$ we write $\mathcal{B}(A, r)$ or $\mathcal{B}_{\mathcal{K}}(A, r)$ for the open ball centered at A with radius r. Since $\mathcal{K}$ is finite-dimensional it has a unique Euclidean topology. Any subset $\mathcal{M}$ of $\mathcal{K}$ will be given the induced topology from $\mathcal{K}$.

Definition 3.1. *We say that $\mathcal{M} \subset \mathcal{K}$ is locally an m-dimensional C^p-manifold around $A \in \mathcal{M}$ if there exists $r_1 > 0$ and a C^p-map $\phi : \mathcal{B}_{\mathbb{R}^m}(0, r_1) \to \mathcal{K}$ with the following properties:*

- *$d\phi(x)$ is injective for all $x \in \mathcal{B}_{\mathbb{R}^m}(0, r_1)$,*
- *$\phi(0) = A$,*
- *ϕ is a homeomorphism onto an open neighborhood of A in $\mathcal{M}$.*

This is in line with the standard definition of C^p-manifolds, see Theorem 2.1.2 [6] for a number of equivalent definitions. As a consequence of the homeomorphism condition, note that there exists an $s_1 > 0$ such that

$$\mathcal{M} \cap \mathcal{B}_{\mathcal{K}}(A, s_1) = \mathsf{Im}\phi \cap \mathcal{B}_{\mathcal{K}}(A, s_1), \tag{3.1}$$

where $\mathsf{Im}\phi$ denotes the image of ϕ. Given $B \in \mathsf{Im}\phi$, there exists a unique $x \in \mathcal{B}(0, r_1)$ such that $B = \phi(x)$. We will without further comment denote this x by x_B. All the manifolds considered in this paper are at least C^1, and hence we have

$$\phi(x) = C + d\phi(x_C)(x - x_C) + o(x - x_C) \tag{3.2}$$

where o stands for "little ordo".[1] We define the tangent space $T_{\mathcal{M}}(B)$ by $T_{\mathcal{M}}(B) = \mathsf{Ran}\ d\phi(x_B)$. It is a standard fact from differential geometry that this definition is independent of ϕ. Moreover, we set

$$\tilde{T}_{\mathcal{M}}(B) = B + T_{\mathcal{M}}(B),$$

i.e., $\tilde{T}_{\mathcal{M}}(B)$ is the affine linear manifold which is tangent to $\mathcal{M}$ at B. Throughout this section, $\mathcal{M}$ will be a locally C^p-manifold at A, where $p \geq 1$ and we associate with it r_1 and s_1 as in Definition 3.1 and (3.1). Moreover, there will follow a row of decreasing numbers $r_2,\ r_3$ and $s_2,\ s_3$ etc., related to the above A, and the reader has to bear in mind where these numbers were defined. The following proposition basically says that the affine tangent-spaces are close to $\mathcal{M}$ locally.

Proposition 3.2. *Let $\mathcal{M}$ be a locally C^1-manifold at A. For each $\epsilon_2 > 0$ there exists s_2, $0 < s_2 < s_1$, such that for all $C \in \mathcal{B}(A, s_2) \cap \mathcal{M}$ we have*

(*i*) $\mathsf{dist}(B, \tilde{T}_{\mathcal{M}}(C)) \leq \epsilon_2 \|B - C\|, \quad B \in \mathcal{B}(A, s_2) \cap \mathcal{M}.$
(*ii*) $\mathsf{dist}(B, \mathcal{M}) \leq \epsilon_2 \|B - C\|, \qquad B \in \mathcal{B}(A, s_2) \cap T_{\mathcal{M}}(C).$

Proof. Given $r < r_1$, we first show that $\|\phi(x) - \phi(y)\| / \|x - y\|$ is uniformly bounded above and below for $x, y \in \mathcal{B}_{\mathbb{R}^m}(0, r)$. By (3.2) and the mean value theorem we have

$$\|\phi(y) - \phi(x)\| = \|d\phi(z)(y - x)\|$$

for some z on the line between x and y. Now, $d\phi(z)$ depends continuously on z and its singular values $\sigma_1(d\phi(z)), \ldots, \sigma_m(d\phi(z))$ depend continuously on the matrix entries [9, p191], hence

$$\inf_{z \in \mathcal{B}_{\mathbb{R}^m}(0,r)} \{\sigma_n(d\phi(z))\} \|y - x\| \leq \|\phi(y) - \phi(x)\| \leq \sup_{z \in \mathcal{B}_{\mathbb{R}^m}(0,r)} \{\sigma_1(d\phi(z))\} \|y - x\|. \tag{3.3}$$

By Definition 3.1, $\sigma_m(d\phi(x))$ is never zero and $cl(\mathcal{B}_{\mathbb{R}^m}(0, r))$ is compact, so both the inf and sup amount to finite positive numbers, as desired. (cl denotes the closure).

We now prove (*i*); by (3.2) and the mean value theorem we have

$$\mathsf{dist}(B, \tilde{T}_{\mathcal{M}}(C)) \leq \|B - (C + d\phi(x_C)(x_B - x_C))\| = \|\phi(x_B) - \phi(x_C) - d\phi(x_C)(x_B - x_C))\| \leq$$
$$\leq \sup\Big\{\|d\phi(y) - d\phi(x_C)\| : y \in \mathcal{B}_{\mathbb{R}^m}(0, r) \text{ such that } \|y - x_C\| \leq \|x_B - x_C\|\Big\} \|x_B - x_C\|.$$

Since $d\phi$ is continuous on the compact set $cl(\mathcal{B}_{\mathbb{R}^m}(0, r))$, it is also equicontinuous. It follows that for each $\epsilon > 0$ we can pick a $\delta > 0$ such that

$$\sup\Big\{\|d\phi(y) - d\phi(x)\| : x, y \in \mathcal{B}_{\mathbb{R}^m}(0, r) \text{ such that } \|y - x\| \leq \delta\Big\} < \epsilon.$$

[1] i.e. it stands for a function with the property that $o(x)/\|x\|$ extends by continuity to 0 and takes the value 0 there.

Set $r_2 < \min(\delta/2, r)$ and let s_2, $0 < s_2 < s_1$, be such that $B \in B_{\mathcal{K}}(A, s_2) \cap \mathcal{M}$ implies $\|x_B\| < r_2$, which we can do since ϕ is a homeomorphism with $\phi(0) = A$. For $B, C \in \mathcal{B}_{\mathcal{K}}(A, s_2)$, we then have (by equation (3.3)) that

$$\mathsf{dist}(B, \tilde{T}_{\mathcal{M}}(C)) \leq \epsilon\|x_B - x_C\| \leq \epsilon k\|\phi(x_B) - \phi(x_C)\| = \epsilon k\|B - C\|,$$

where $k = (\inf_{z\in\mathcal{B}_{\mathbb{R}^m}(0,r)}\{\sigma_n(d\phi(z))\})^{-1}$, (cf equation (3.3)). By letting $\epsilon = \epsilon_2/k$, the desired statement follows.

The proof of (ii) is similar. Given C let y be such that $B = C + d\phi(x_C)y$ and note that $\phi(x_C + y) \in \mathcal{M}$ so

$$\mathsf{dist}(B, \mathcal{M}) \leq \|\phi(x_C + y) - (C + d\phi(x_C)y)\| \leq \sup\Big\{\|d\phi(x_C + ty) - d\phi(x_C)\| : t \in [0, 1]\Big\}\|y\|.$$

Moreover, if s_2 is such that $\|x_C\| \leq r$, it is easily seen that

$$\|y\| \leq 2s_2 \Big/ \sup_{z\in\mathcal{B}_{\mathbb{R}^m}(0,r)} \{\sigma_1(d\phi(z))\}.$$

Thus, by picking s_2 small enough we can ensure that for all $B, C \in \mathcal{B}_{\mathcal{K}}(A, s_2)$, all points of the form $x_C + ty$ are bounded by some pregiven $\delta > 0$. From here the proof follows a similar path as (i), we omit the remaining details. □

The value of ϵ_2 will be determined later, so for the moment we will consider it as a constant and the corresponding value of s_2 will be kept for future use. However, r_2 was an internal variable in the above proof, and to avoid too many subindices, we let r_2 take a new value in the proof of the next proposition, (in contrast to the value of r_1 which was given in Definition 3.1 and connected to the fixed point A). The next proposition shows that projection on $\mathcal{M}$ is a *locally* well defined operation.

Proposition 3.3. *Let $\mathcal{M}$ be a locally C^p-manifold at A with $p \geq 1$. Then there exists $s_3 > 0$ and a C^{p-1} map*

$$\pi : \mathcal{B}_{\mathcal{K}}(A, s_3) \to \mathcal{M}$$

such that for all $B \in \mathcal{B}_{\mathcal{K}}(A, s_3)$ there exists a unique closest point in $\mathcal{M}$ which is given by $\pi(B)$. Moreover, $C \in \mathcal{M} \cap \mathcal{B}_{\mathcal{K}}(A, s_3)$ equals $\pi(B)$ if and only if $B - C \perp T_{\mathcal{M}}(C)$.

Proof. Recall that n is the dimension of $\mathcal{K}$ and m the dimension of $\mathcal{M}$ at A. By standard differential geometry there exists an $r_2 < r_1$ and C^{p-1}-functions $f_1, \ldots, f_{n-m} : \mathcal{B}_{\mathbb{R}^m}(0, r_2)$ with the property that

$$\big(T_{\mathcal{M}}(\phi(x))\big)^{\perp} = \mathsf{Span}\ \{f_1(x), \ldots, f_{n-m}(x)\}$$

for all $x \in \mathcal{B}_{\mathbb{R}^m}(0, r_2)$, (see e.g. Theorem 2.7.7 in [6]). Define $\sigma : \mathcal{B}_{\mathbb{R}^m}(0, r_2) \times \mathbb{R}^{n-m} \to \mathcal{K}$ via

$$\sigma(x, y) = \phi(x) + \sum_{i=1}^{n-m} y_i f_i(x).$$

Consider the set $\mathcal{S} \subset \mathcal{K}$ of points whose multiplicity under σ is greater than 1, (i.e. all points hit more than once by σ). By the inverse function theorem, the set $\sigma^{-1}(\mathcal{S})$ can not have 0 as an accumulation-point, for it says that there exists an $r < r_2$ such that σ restricted to $\mathcal{B}_{\mathbb{R}^n}(0, r)$ is a diffeomorphism onto its image. Pick $r_3 < \min\big(r, \mathsf{dist}(\sigma^{-1}(\mathcal{S}), 0)\big)$ and pick s_3 such that

$$\mathcal{M} \cap \mathcal{B}_{\mathcal{K}}(A, 2s_3) = \phi(\mathcal{B}_{\mathbb{R}^m}(0, r_3)) \cap \mathcal{B}_{\mathcal{K}}(A, 2s_3), \tag{3.4}$$

and

$$\mathcal{B}_{\mathcal{K}}(A, s_3) \subset \sigma(\mathcal{B}_{\mathbb{R}^n}(0, r_3)). \tag{3.5}$$

Note that, given $B \in \mathcal{B}_{\mathcal{K}}(A, s_3)$ there exists a unique (x_B, y_B)such that $B = \sigma((x_B, y_B))$ and moreover $\|(x_B, y_B)\| \leq r_3$ by (3.5). We define

$$\pi(B) = \phi(x_B).$$

To see that π is a C^{p-1}-map, let $\theta : \mathbb{R}^m \times \mathbb{R}^{n-m} \to \mathbb{R}^m$ be given by $\theta((x, y)) = x$ and note that

$$\pi = \phi \circ \theta \circ (\sigma|_{\mathcal{B}_{\mathbb{R}^n}(0,r_3)})^{-1}.$$

We now show that $\pi(B)$ have the desired properties. By the construction,

$$\pi(B) - B \in \mathsf{Span}\ \{f_1(x_B), \ldots, f_{n-m}(x_B)\} \perp T_{\mathcal{M}}(\phi(x_B)) = T_{\mathcal{M}}(\pi(B)).$$

Now suppose $C \in \mathcal{M}$ is a closest point to B. Since $\|A - B\| < s_3$ we clearly must have $\|C - A\| < 2s_3$ so by (3.4) there exists a $x_C \in \mathcal{B}_{\mathbb{R}^m}(0, r_3)$ with $\phi(x_C) = C$. Since $r_3 < r_1$ we know that $\mathcal{M}$ is completely determined by ϕ in the vicinity of C. In particular, it makes sense to talk about $T_{\mathcal{M}}(C)$ and it is easily seen that $B - C \perp T_{\mathcal{M}}(C)$, for by (3.2) we have

$$\begin{aligned}\|\phi(x) - B\|^2 &= \|C + d\phi(x_C)(x - x_C) + o(x - x_C) - B\| = \\ &= \|C - B\|^2 + 2\langle C - B, d\phi(x_C)(x - x_C)\rangle + o(\|x - x_C\|)\end{aligned}$$

and hence the scalar product needs to be zero for all x's. Thus there is a y such that $B = \sigma((x_C, y))$. But since (x_B, y_B) is the unique point with this property, we deduce that $x_B = x_C$ and hence $\pi(B) = \phi(x_B) = \phi(x_C) = C$. This establishes the first part of the proposition. Now let C be as in the second part of the proposition. As above we have $C = \phi(x_C)$ with $\|x_C\| < r_3$ and the orthogonality implies that there exists a y with $B = \sigma((x_C, y))$ and again this implies $C = \pi(B)$, as desired. □

4. Non-tangentiality

Suppose now that we are given closed sets $\mathcal{M}_1$ and $\mathcal{M}_2$ which locally are manifolds around an intersection point $A \in \mathcal{M}_1 \cap \mathcal{M}_2$.

Definition 4.1. *An intersection point A will be called regular if there are numbers m_1, m_2, m and $p \geq 1$ such that*

- *$\mathcal{M}_j$ is locally an m_j-dimensional C^p-manifold at A, $j = 1, 2$.*
- *$\mathcal{M}_1 \cap \mathcal{M}_2$ is locally an m-dimensional C^p-manifold at A.*

Note that the set of regular points is clearly a relatively open set in $\mathcal{M}_1 \cap \mathcal{M}_2$. Next, we introduce angles. For more information on angles, we refer to [13].

Definition 4.2. *For any regular point A, we define the angle $\alpha(A)$ of $\mathcal{M}_1$ and $\mathcal{M}_2$ at A to be the $\cos^{-1}$ of the number*

$$\sigma(A) = \lim_{r \to 0} \sup \left\{ \frac{\langle B_1 - A, B_2 - A\rangle}{\|B_1 - A\|\|B_2 - A\|} : B_j \in \mathcal{M}_j, \|B_j - A\| < r \textit{ and } B_j - A \perp T_{\mathcal{M}_1 \cap \mathcal{M}_2}(A) \right\}.$$

Given a linear subspace $\mathcal{M}$ we let $P_{\mathcal{M}}$ denote the orthogonal projection onto $\mathcal{M}$. It is easily verified that

(4.1)

$$\sigma(A) = \sup\left\{\frac{\langle B_1 - A, B_2 - A\rangle}{\|B_1 - A\|\|B_2 - A\|} : B_j \in T_{\mathcal{M}_j}(A) \ominus T_{\mathcal{M}_1\cap\mathcal{M}_2}(A)\right\} =$$
$$= \|P_{T_{\mathcal{M}_1}(A)\ominus T_{\mathcal{M}_1\cap\mathcal{M}_2}(A)} P_{T_{\mathcal{M}_2}(A)\ominus T_{\mathcal{M}_1\cap\mathcal{M}_2}(A)}\| = \|P_{T_{\mathcal{M}_1}(A)} P_{T_{\mathcal{M}_2}(A)} - P_{T_{\mathcal{M}_1\cap\mathcal{M}_2}(A)}\|.$$

If $\mathcal{M}_1$ and $\mathcal{M}_2$ are hyperplanes through the origin, then it is clear that the above definition coincides with the classical definition;

$$\cos\alpha_{clas}(A) = \sigma_{clas}(A) = \|P_{\mathcal{M}_1}P_{\mathcal{M}_2} - P_{\mathcal{M}_1\cap\mathcal{M}_2}\|. \tag{4.2}$$

However, it is important to note that $\alpha_{clas}(A, T_{\mathcal{M}_1(A)}, T_{\mathcal{M}_2(A)})$ and $\alpha(A)$ are not necessarily the same. For example, take $\mathcal{M}_1 = \{(x_1, x_2, 0, 0, x_3) : x \in \mathbb{R}^3\}$ and $\mathcal{M}_2 = \{(x_1, x_2, x_2^2, x_3, x_3) : x \in \mathbb{R}^3\}$. Then $\mathcal{M}_1 \cap \mathcal{M}_2 = \{(x, 0, 0, 0, 0) : x \in \mathbb{R}\}$ and 0 is a regular point. Moreover $\alpha_{clas}(0, T_{\mathcal{M}_1(0)}, T_{\mathcal{M}_2(0)}) = \pi/4$ whereas $\alpha(0) = 0$. What goes wrong above is clearly that the two surfaces are tangential to each other in the direction $(0, 0, 1, 0, 0)$, and it is intuitively clear that when this is not the case, the two concepts should coincide. To avoid such obstacles, we therefore introduce:

Definition 4.3. *$\mathcal{M}_1$ and $\mathcal{M}_2$ are said to be non-tangential at A if A is regular and they have a positive angle at A, i.e. if $\sigma(A) < 1$. We will often simply say that A is non-tangential.*

Note that the angle is always a number between 0 and $\pi/2$, so in an intuitive sense only exceptional points do not satisfy non-tangentiality. We have no intention of formalizing this statement, but point out already that for the example considered in Section 2, this is indeed the case, which will be proven in [1].

Theorem 4.4. *With the notation as in Section 2, we have that the set of tangential points between $\mathcal{R}_k$ and $\mathcal{H}$ is thin in $\mathcal{H}_k$.*

In the general setting, pathological examples do exist. For example, take $\mathcal{M}_1 = \{(x_1, x_2, 0) : x \in \mathbb{R}^2\}$ and $\mathcal{M}_2 = \{(x_1, x_2, x_1^2) : x \in \mathbb{R}^2\}$. However, when $\mathcal{M}_1$ and $\mathcal{M}_2$ are defined by polynomials, one can use similar methods, as in the proof of Theorem 4.4, to show that under mild conditions, if non-tangentiality holds at one point, then it holds everywhere except for a thin set. We will not pursue this.

By (4.1) it is easy to see that A is non-tangential if and only if

$$\Big(T_{\mathcal{M}_1}(A) \ominus T_{\mathcal{M}_1\cap\mathcal{M}_2}(A)\Big) \cap \Big(T_{\mathcal{M}_2}(A) \ominus T_{\mathcal{M}_1\cap\mathcal{M}_2}(A)\Big) = \{0\} \tag{4.3}$$

which in turn happens if and only if

$$T_{\mathcal{M}_1}(A) \cap T_{\mathcal{M}_2}(A) = T_{\mathcal{M}_1\cap\mathcal{M}_2}(A). \tag{4.4}$$

This latter condition is usually referred to as *clean intersection* in microlocal analysis. We have chosen the terminology non-tangential since it is more intuitive. We now show that non-tangentiality is a weaker concept than transversality, defined in (1.3).

Proposition 4.5. *Let $\mathcal{M}_1$ and $\mathcal{M}_2$ be two C^1-manifolds and $A \in \mathcal{M}_1 \cap \mathcal{M}_2$ a transversal point. A transversal point A is also non-tangential.*

Proof. Applying the implicit function theorem to $\phi_1-\phi_2$ easily yields that $\mathcal{M}_1\cap\mathcal{M}_2$ is a C^1-manifold of dimension m_1+m_2-n at A. Thus $T_{\mathcal{M}_1}(A)\ominus T_{\mathcal{M}_1\cap\mathcal{M}_2}(A)$ has dimension $n-m_2$ and $T_{\mathcal{M}_2}(A)\ominus T_{\mathcal{M}_1\cap\mathcal{M}_2}(A)$ has dimension $n-m_1$. Thus

$$\dim(T_{\mathcal{M}_1}(A)\ominus T_{\mathcal{M}_1\cap\mathcal{M}_2}(A))+\dim(T_{\mathcal{M}_2}(A)\ominus T_{\mathcal{M}_1\cap\mathcal{M}_2}(A))+\dim(T_{\mathcal{M}_1\cap\mathcal{M}_2}(A))=n.$$

But by transversality the sum of the above subspaces equals $\mathcal{K}$, which can only happen if (4.3) is satisfied. □

We will denote $\mathcal{M}_1\cap\mathcal{M}_2$ by $\mathcal{M}$, and the objects from Section 3 associated to $\mathcal{M}_1$, $\mathcal{M}_2$ and $\mathcal{M}$, e.g. ϕ, by ϕ_1, ϕ_2 and ϕ respectively. We thus omit subindex when dealing with $\mathcal{M}_1\cap\mathcal{M}_2$. We now prove that for non-tangential points, the angle as defined here and the classical angle of the respective tangent spaces coincide.

Proposition 4.6. *If A is non-tangential, then*

$$\alpha_{clas}(A,T_{\mathcal{M}_1(A)},T_{\mathcal{M}_2(A)})=\alpha(A).$$

Proof. Using (4.1) it is easy to see that

$$\sigma(A)=\|P_{T_{\mathcal{M}_1}}P_{T_{\mathcal{M}_2}}-P_{T_{\mathcal{M}_1\cap\mathcal{M}_2}}\|=\|P_{T_{\mathcal{M}_1}}P_{T_{\mathcal{M}_2}}-P_{T_{\mathcal{M}_1}\cap T_{\mathcal{M}_2}}\|=\sigma_{clas}(A,T_{\mathcal{M}_1},T_{\mathcal{M}_2}),$$

where we used (4.4) in the crucial step. □

Proposition 4.7. *The function σ in Definition 4.2 is C^{p-1} (on the set of regular points in $\mathcal{M}_1\cap\mathcal{M}_2$). In particular, non-tangentiality is a local property, i.e. if A is non-tangential, then the same holds for all $B\in\mathcal{M}_1\cap\mathcal{M}_2$ in a neighborhood of A.*

Proof. Let $A\in\mathcal{M}_1\cap\mathcal{M}_2$ be a regular point and let ϕ be the usual chart. We need to show that $\sigma\circ\phi$ is C^{p-1} around 0. By standard differential geometry (see e.g. [6]) there exists C^{p-1}-functions $f_1,\dots,f_{n-m}$ defined on the domain of ϕ such that

$$\left(T_{\mathcal{M}_1\cap\mathcal{M}_2}(\phi(x))\right)^{\perp}=\mathsf{Span}\ \{f_1(x),\dots,f_{n-m}(x)\}.$$

It is easy to see that (with $j=1$ or $j=2$)

$$T_{\mathcal{M}_j}(\phi(x))\ominus T_{\mathcal{M}_1\cap\mathcal{M}_2}(\phi(x))=\mathsf{Span}\ \{P_{T_{\mathcal{M}_j}(\phi(x))}f_1(x),\dots P_{T_{\mathcal{M}_j}(\phi(x))}f_{n-m}(x)\}. \tag{4.5}$$

Since the dimension of $T_{\mathcal{M}_j}(\phi(x))$ is constant m_j, it is easy to see that the functions on the right are C^{p-1} as well. Moreover, as the dimension of the space on the left in (4.5) is constant m_j-m, we can pick m_j-m of the functions on the right and put them as columns of an injective matrix $M_j(x)$ such that

$$T_{\mathcal{M}_j}(\phi(x))\ominus T_{\mathcal{M}_1\cap\mathcal{M}_2}(\phi(x))=\mathsf{Ran}\ M_j(x)$$

for x in a neighborhood of 0. By (4.1) we have for x around 0 that

$$\sigma(\phi(x))=\|P_{\mathsf{Ran}\ M_1(x)}P_{\mathsf{Ran}\ M_2(x)}\|. \tag{4.6}$$

(Note that this is not true with the classical definition when the two ranges have a non-trivial intersection, but with the definition here it works.) Moreover, it is easy to see that

$$P_{\mathsf{Ran}\ M_j}=M_j(M_j^*M_j)^{-1}M_j^*,$$

where x has been omitted for readability. Combining this with (4.6) and the fact that singular values are C^∞-functions of the matrix entries, it is clear that σ is C^{p-1} near 0, as desired. □

5. Properties of the projection operators

Let A $\mathcal{M}_1$, $\mathcal{M}_2$, ϕ_1 etc. be as before, i.e., C^p-manifolds with a non-tangential intersection point A. We assume that $p \geq 1$ and continue to use the convention of denoting objects related to the C^p-manifold $\mathcal{M} = \mathcal{M}_1 \cap \mathcal{M}_2$ without subindex. In Propositions 3.2 and 3.3 the quantities s_2 and s_3 appears. These are not necessarily the same and they depend on an auxiliary constant ϵ_2 originating from Proposition 3.2. In this section we let ϵ_2 be a fixed number (which we will determine later), and we let s_4 denote the minimum of all possible s's from Section 3 related to the 3 manifolds. The above will not be repeated in the statements of the results below. Thus, we can apply any result from Section 3 to either of the manifolds considered here. Moreover, letting j denote either $1, 2$ or nothing, we let $r_4 > 0$ be such that

$$\mathcal{M}_j \cap \mathcal{B}_{\mathcal{K}}(A, s_4) = \phi_j(\mathcal{B}(0, r_4)) \cap \mathcal{B}_{\mathcal{K}}(A, s_4),$$

and such that the results of Section 3 applies to each $\phi_j(x)$ with $\|x\| < r_4$. We also assume that $\epsilon_2 < 1$.

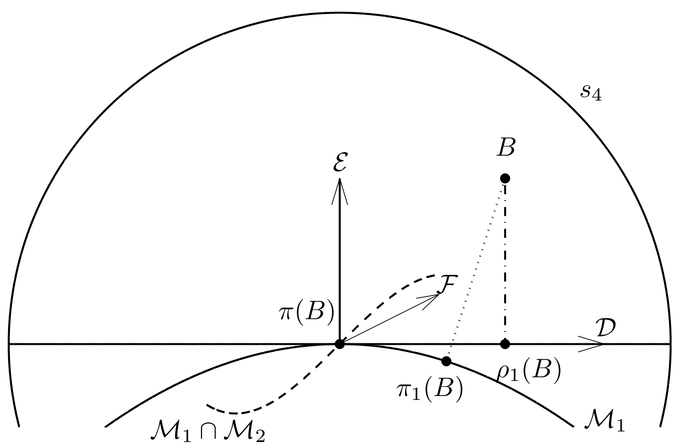


Figure 6. Illustration of the difference between ρ_1 and π_1. $\mathcal{D}$ and $\mathcal{E}$ appear in the proof of Proposition 5.2.

Given an affine linear manifold $\mathcal{N} \subset \mathcal{K}$, we denote by $P_{\mathcal{N}}$ the orthogonal projection onto $\mathcal{N}$. We introduce maps $\rho_j : \mathcal{B}(A, s_4) \to \mathcal{K}$, ($j = 1$ or $j = 2$), via

$$\rho_j(B) = P_{\tilde{T}_{\mathcal{M}_j}(\pi(B))}(B).$$

Thus, ρ_j resemble π_j but is slightly different. π_j projects onto $\mathcal{M}_j$ whereas ρ_j projects onto the tangent plane of $\mathcal{M}_j$ taken at the closest point to B in $\mathcal{M}_1 \cap \mathcal{M}_2$, i.e. $\pi(B)$, (see Fig 5). A proper estimate of the difference is given in Proposition 5.2.

Lemma 5.1. *The operators ρ_1 and ρ_2 are C^{p-1}-maps in $\mathcal{B}_{\mathcal{K}}(A, s_4)$. Moreover, we can select a number $s_5 < s_4$ such that the image of $\mathcal{B}(A, s_5)$ under $\rho_1, \rho_2, \pi, \pi_1, \pi_2$, as well as any composition of two of those maps, is contained in $\mathcal{B}(A, s_4)$.*

Proof. The second part is an immediate consequence of the continuity of the maps. Let j denote 1 or 2 and set $M_j(B) = d\phi_j(\pi(B))$. Note that M_j is a C^{p-1}-map in

$\mathcal{B}_{\mathcal{K}}(A, s_4)$ by Proposition 3.3 and the choice of s_4. Moreover,

$$\rho_j(B) = P_{\tilde{T}_{\mathcal{M}_j}(\pi(B))}(B) = \pi(B) + P_{T_{\mathcal{M}_j}(\pi(B))}(B - \pi(B)) =$$
$$= \pi(B) + M_j(B)\big(M_j^*(B)M_j(B)\big)^{-1}M_j^*(B)\big(B - \pi(B)\big),$$

from which the result follows. □

Proposition 5.2. *Given any $B \in \mathcal{B}(A, s_5)$ and $j = 1$ or $j = 2$, we have*

$$\|\pi_j(B) - \rho_j(B)\| < 5\sqrt{\epsilon_2}\|B - \pi(B)\|.$$

Proof. By Lemma 5.1 we have that Proposition 3.2 applies to the point $C = \pi(B)$. It is no restriction to assume that $\pi(B) = 0$, which we now do. Denote $\mathcal{D} = T_{\mathcal{M}_j}(0)$, $\mathcal{E} = \mathsf{Span}\,\{B - \rho_j(B)\}$ and $\mathcal{F} = \mathcal{K} \ominus (\mathcal{D} \oplus \mathcal{E})$, (see Fig. 5). Let D_B and E_B be elements of $\mathcal{D}$ and $\mathcal{E}$ such that $B = D_B + E_B$, and note that

$$\rho_j(B) = D_B.$$

We thus have to show that

$$\|\pi_j(B) - D_B\| < 5\sqrt{\epsilon_2}\|B\|. \tag{5.1}$$

First note that by Proposition 3.2 (ii) there exists a point in $\mathcal{M}_1 \cap \mathcal{M}_2 \cap \mathcal{B}(D_B, \epsilon_2\|D_B\|)$. Thus

$$\|B - \pi_j(B)\| \leq \|B - D_B\| + \epsilon_2\|D_B\| = \|E_B\| + \epsilon_2\|D_B\|$$

and hence $\pi_j(B)$ is a member of the set

$$\Big\{(D, E, F) : \|D - D_B\|^2 + \|E - E_B\|^2 + \|F\|^2 < (\|E_B\| + \epsilon_2\|D_B\|)^2\Big\}. \tag{5.2}$$

However, by Proposition 3.2 (i) we have that $\pi_j(B)$ also is a member of

$$\Big\{(D, E, F) : \|E\|^2 + \|F\|^2 < \epsilon_2^2\|D\|^2\Big\}. \tag{5.3}$$

The left hand side of (5.1) is thus dominated by the supremum of the function

$$f(D, E, F) = \|D - D_B\|^2 + \|E\|^2 + \|F\|^2,$$

subject to the conditions in (5.2) and (5.3). Either by geometrical considerations or the method of Lagrange multipliers, it is not hard to deduce that this supremum is attained for $F = 0$ and D, E of the form $D = dD_B$ and $E = eE_B$ where $d, e \in \mathbb{R}$. We now have a two-dimensional problem of circles and cones, and in the remainder of the proof we treat D and E as elements of $\mathbb{R}^2$. Given $D + E$ in the intersection of (5.2) and (5.3), it is easily seen (see Fig. 5) that

$$\|E\| \leq \epsilon_2(\|D_B\| + \|E_B\| + \epsilon_2\|D_B\|),$$

since this is the height of the cone at the outer edge of the circle. The problem becomes simpler if we replace the cone in (5.3) by the following strip:

$$\Big\{E : \|E\| < \epsilon_2((1 + \epsilon_2)\|D_B\| + \|E_B\|)\Big\}. \tag{5.4}$$

From Figure 5, and some freshman formulas, it is readily verified that the sought supremum is dominated by $\alpha(\|E_B\| + \epsilon_2\|D_B\|)$, where α is the angle given by

$$\alpha = \cos^{-1}\left(\frac{\|E_B\| - \epsilon_2((1 + \epsilon_2)\|D_B\| + \|E_B\|)}{(\|E_B\| + \epsilon_2\|D_B\|)}\right).$$

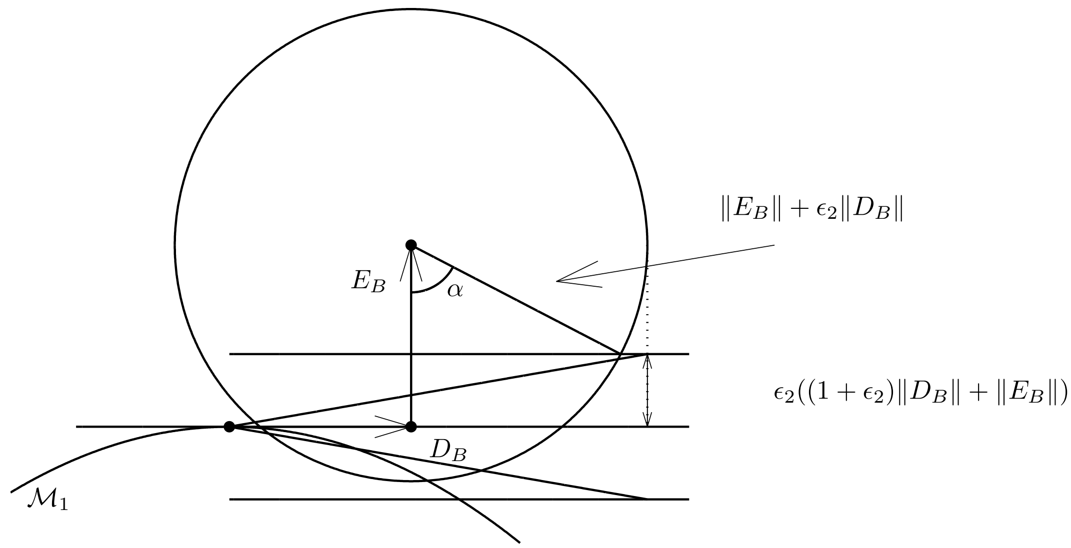


FIGURE 7. Proof illustration.

Since we have assumed that $\epsilon_2 < 1$, the proposition thus follows if we establish that

$$\cos^{-1}\left(\frac{\|E_B\| - \epsilon_2(2\|D_B\| + \|E_B\|)}{(\|E_B\| + \epsilon_2\|D_B\|)}\right)\frac{\|E_B\| + \epsilon_2\|D_B\|}{\sqrt{\|E_B\|^2 + \|D_B\|^2}} < 5\sqrt{\epsilon_2}$$

By dividing with $\|E_B\|$ at suitable places, one sees that this is equivalent to

$$\sup_{t>0}\left\{\cos^{-1}\left(\frac{1 - 2\epsilon_2 t - \epsilon_2}{1 + \epsilon_2 t}\right)\frac{1 + \epsilon_2 t}{\sqrt{1 + t^2}}\right\} < 5\sqrt{\epsilon_2}$$

Using $\cos^{-1}(1 - x) \leq \pi\sqrt{x/2}$ and $\epsilon_2 < 1$ we obtain

$$\cos^{-1}\left(\frac{1 - 2\epsilon_2 t - \epsilon_2}{1 + \epsilon_2 t}\right)\frac{1 + \epsilon_2 t}{\sqrt{1 + t^2}} = \cos^{-1}\left(1 - \frac{3\epsilon_2 t + \epsilon_2}{1 + \epsilon_2 t}\right)\frac{1 + \epsilon_2 t}{\sqrt{1 + t^2}} \leq$$
$$\leq \frac{\pi}{\sqrt{2}}\sqrt{\frac{3\epsilon_2 t + \epsilon_2}{1 + \epsilon_2 t}}\frac{1 + \epsilon_2 t}{\sqrt{1 + t^2}} \leq \sqrt{\epsilon_2}\frac{\pi}{\sqrt{2}}\frac{\sqrt{(3t+1)(1+t)}}{\sqrt{1 + t^2}} < 5\sqrt{\epsilon_2}$$

since the function on the right clearly is bounded, and a few calculations show that the maximum is attained at $t = (1 + \sqrt{5})/2$ which gives $5\sqrt{\epsilon_2}$ as an upper bound. □

Lemma 5.3. *Given $B \in \mathcal{B}(A, s_5)$ we have*

$$\pi(B) = \pi(\rho_1(B)) = \pi(\rho_2(\rho_1(B))).$$

Proof. If we prove the first equality the second follows by reversing the roles of 1 and 2 and applying the first equality to $\rho_1(B)$. To see the first equality, note that $B - \rho_1(B) \perp T_{\mathcal{M}_1}(\pi(B))$ and obviously $B - \pi(B) \perp T_{\mathcal{M}_1 \cap M_2}(\pi(B))$ so

$$\rho_1(B) - \pi(B) = (\rho_1(B) - B) + (B - \pi(B)) \perp T_{\mathcal{M}_1 \cap M_2}(\pi(B)),$$

which by Lemma 5.1 and Proposition 3.3 implies $\pi(B) = \pi(\rho_1(B))$, as desired. □

Lemma 5.4. *Let $c_5 > 1$ and $\epsilon_5 > 0$ be given. If $E, F \in \mathcal{K}$ satisfies $\|E\| > c_5\|F\|$ and $\|E - F\| < \epsilon_5$, then*

$$\|E\| < \epsilon_5 \frac{c_5}{c_5 - 1}.$$

Proof. If $\|E\| < \epsilon_5$ we are done, otherwise

$$c_5 < \frac{\|E\|}{\|F\|} < \frac{\|E\|}{\|E\| - \epsilon_5}$$

which easily gives the desired estimate. □

The next result will be the main tool for proving convergence of the alternating projections.

Theorem 5.5. *Let $A \in \mathcal{M}_1 \cap \mathcal{M}_2$ be a non-tangential point and assume that $\mathcal{M}_1$ and $\mathcal{M}_2$ are C^2-manifolds. Then for each $c_6 > \sigma(A)$ and each $\epsilon_6 > 0$ there exists a positive $s_6 < s_5$ such that for all $B \in \mathcal{M}_2 \cap \mathcal{B}(A, s_6)$ we have*

(i) $\|\pi_1(B) - \pi(B)\| < c_6\|B - \pi(B)\|$

(ii) $\|\pi(\pi_1(B)) - \pi(B)\| < \epsilon_6\|B - \pi(B)\|$

Moreover the same holds true with the roles of $\mathcal{M}_1$ and $\mathcal{M}_2$ reversed.

Proof. Fix c_1 such that $\sigma(A) < c_1 < c_6$ and pick an $s_7 < s_5$ such that

$$\sup\{\sigma(C) : C \in \mathcal{M}_1 \cap \mathcal{M}_2 \cap \mathcal{B}(A, s_7)\} < c_1, \tag{5.5}$$

which we can do since σ is continuous by Proposition 4.7. Let $c_2 > 1$ be such that $c_2c_1 < c_6$. By Lemma 5.1 and Proposition 3.3, ρ_1 and π are C^1-functions, and hence we can pick $c_3 > 0$ such that

$$\|\rho_j(B) - \rho_j(B')\| \leq c_3\|B - B'\| \text{ and } \|\pi(B) - \pi(B')\| \leq c_3\|B - B'\| \tag{5.6}$$

for all $B, B' \in \mathcal{B}(A, s_4)$, (recall that s_4 was chosen in the beginning of this section). Finally, we fix ϵ_2 such that

$$5\sqrt{\epsilon_2}(c_3 + c_3^2) < \epsilon_6 \text{ and } 5\sqrt{\epsilon_2}(1 + c_3)\frac{c_2}{c_2 - 1} < c_6 \text{ and } (1 + 5\sqrt{\epsilon_2})c_2c_1 < c_6. \tag{5.7}$$

This may seem like a circle argument, because c_1, c_2 and c_3 depends on c_6 which depends on ϵ_2 via s_5, and the c_j's appears in (5.7). However, this is easily circumvented by first choosing s_5 with $\epsilon_2 = 1$, say, and pick values of c_1, c_2, c_3. Then, once the real ϵ_2 has been chosen via (5.7) we can redefine s_4, s_5 and s_7 accordingly without violating (5.5) or (5.6).

Now, pick an s_6 such that

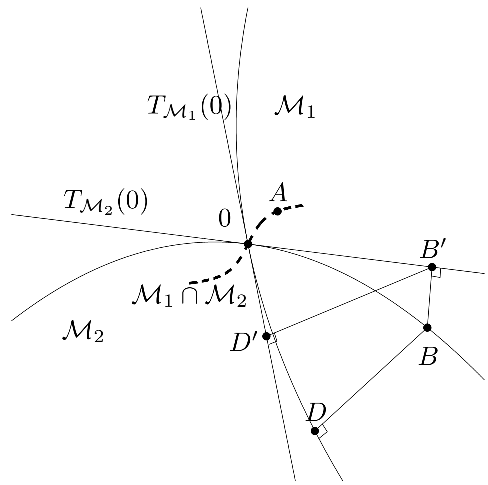
$$\pi(\mathcal{B}(A, s_6)) \subset \mathcal{B}(A, s_7),$$

and let $B \in \mathcal{M}_2 \cap \mathcal{B}(A, s_6)$. We begin with proving (ii). Denote $C = \pi(B)$ and $D = \pi_1(B)$ and note that $C \in \mathcal{M}_1 \cap \mathcal{M}_2 \cap \mathcal{B}(A, s_7)$ so $\sigma(C) < c_6$. There is no restriction to assume that $C = 0$, which we do from now on. We thus need to show that $\|\pi(D)\| < \epsilon_6\|B\|$. Put $B' = \rho_2(B)$ and $D' = \rho_1(B')$. (See Figure 5 and recall the $\pi(B') = \pi(B) = 0$ by Lemma 5.3). First, note that by Proposition 5.2

$$\|B - B'\| = \|\pi_2(B) - \rho_2(B)\| < 5\sqrt{\epsilon_2}\|B\| \tag{5.8}$$

and moreover by (5.6) and Proposition 5.2 we have that

$$\begin{aligned} \|D - D'\| = \|\rho_1(B') - \pi_1(B)\| &\leq \|\rho_1(B') - \rho_1(B)\| + \|\rho_1(B) - \pi_1(B)\| \leq \\ &\leq c_3\|B' - B\| + 5\sqrt{\epsilon_2}\|B\| < 5\sqrt{\epsilon_2}(1 + c_3)\|B\| \end{aligned} \tag{5.9}$$



By Lemma 5.3 we have $0 = \pi(B) = \pi(D')$ so part (ii) follows by (5.6), (5.7), (5.9) and the calculation

$$\|\pi(D)\| = \|\pi(D) - \pi(D')\| < c_3(5\sqrt{\epsilon_2}(1 + c_3)\|B\|) < \epsilon_6\|B\|.$$

We turn to part (i). Clearly $B' \in T_{\mathcal{M}_2}(0)$ and by Lemma 5.3 we also have $D' \in T_{\mathcal{M}_1}(0)$. Thus

$$\frac{\|D'\|}{\|B'\|} \leq \sigma(0) < c_1 \tag{5.10}$$

whereas (i) amounts to showing that $\|D\|/\|B\| < c_6$. Recall (5.9) and apply Lemma 5.4 with $E = D$, $F = D'$, $\epsilon_5 = 5\sqrt{\epsilon_2}(1 + c_3)\|B\|$ and $c_5 = c_2$. We see that either

$$\frac{\|D\|}{\|D'\|} \leq c_2 \tag{5.11}$$

or $\|D\| < 5\sqrt{\epsilon_2}(1 + c_3)\frac{c_2}{c_2 - 1}\|B\|$, in which case we are done since the constant is less than c_6 by (5.7). We thus assume that (5.11) holds. Note that

$$\frac{\|B'\|}{\|B\|} \leq 1 + 5\sqrt{\epsilon_2} \tag{5.12}$$

by (5.8). Combining (5.10), (5.11) and (5.12) we get

$$\frac{\|D\|}{\|B\|} = \frac{\|D\|}{\|D'\|}\frac{\|B'\|}{\|B\|}\frac{\|D'\|}{\|B'\|} < c_2(1 + 5\sqrt{\epsilon_2})c_1 < c_6$$

by (5.7). □

6. Alternating projections

We are finally ready for the main theorem. The third conclusion below was not mentioned in the introduction. It basically says that if the angle between $\mathcal{M}_1$ and $\mathcal{M}_2$ is not too close to 0, the sequence of alternating projections $B_0,\ B_1,\ B_2, \ldots$ will converge within machine precision within a fairly low number of iterations. In the terminology of [8], $(B_k)_{k=0}^{\infty}$ converges "R-linearly" with rate less than c. For example, if $\sigma(A) = 1/2$, then we will hit single precision ($\approx 10^{-7.5}$) at B_{24}, since single precision has 24 bits of significant precision.

Theorem 6.1. *Let $\mathcal{M}_1 \cap \mathcal{M}_2$ be locally non-tangential C^2-manifolds around $A \in \mathcal{M}_1 \cap \mathcal{M}_2$, and let $\epsilon > 0$ and $1 > c > \sigma(A)$ be given. Then there exists an $s > 0$ such that the sequence of alternating projections*

$$B_0 = \pi_1(B),\ B_1 = \pi_2(\pi_1(B)),\ B_2 = \pi_1(\pi_2(\pi_1(B))),\ B_3 = \pi_2(\pi_1(\pi_2(\pi_1(B)))), \dots$$

(*i*) *converges to a point $B_\infty \in \mathcal{M}_1 \cap \mathcal{M}_2$*

(*ii*) $\|B_\infty - \pi(B)\| < \epsilon \|B - \pi(B)\|$

(*iii*) $\|B_\infty - B_k\| < c^k \|B - \pi(B)\|$

Proof. Let c_6 and ϵ_6 in Theorem 5.5 be given by $c_6 = c$ and

$$\epsilon_6 = (1-c)\epsilon/2. \tag{6.1}$$

Let s_6 be given by Theorem 5.5 and pick

$$s < \frac{s_6(1-\epsilon)}{4(2+\epsilon)} \tag{6.2}$$

such that $\pi(\mathcal{B}(A,s)) \subset \mathcal{B}(A, s_6/4)$, (recall that $\epsilon \le \epsilon_2 < 1$, by assumption). The latter condition ensures that

$$\|\pi(B) - A\| < s_6/4. \tag{6.3}$$

Let $l = \|B - \pi(B)\|$ and note that

$$l \le \|B - A\| + \|A - \pi(B)\| \le s + s_6/4. \tag{6.4}$$

First note that $\|B_0 - B\| = \|\pi_1(B) - B\| \le \|\pi(B) - B\| = l$ and that $\pi(B) = \pi(B_0)$ by Lemma 5.3, so

$$\|B_0 - \pi(B_0)\| \le \|B_0 - B\| + \|B - \pi(B)\| \le 2l.$$

Applying Theorem 5.5 we get

$$\|B_{k+1} - \pi(B_{k+1})\| \le \|B_{k+1} - \pi(B_k)\| \le c\|B_k - \pi(B_k)\|,$$

as long as

$$B_k \in \mathcal{B}(A, s_6). \tag{6.5}$$

Assuming this for the moment we get

$$\|B_k - \pi(B_k)\| \le 2lc^k \tag{6.6}$$

and (Theorem 5.5 and Lemma 5.3)

$$\|\pi(B_{k+1}) - \pi(B_k)\| \le \epsilon_6(2lc^k). \tag{6.7}$$

The sequence $(\pi(B_k))_{k=1}^\infty$ is thus a Cauchy sequence, and hence converges to some point B_∞. By (6.6) the sequence $(B_k)_{k=1}^\infty$ must also converge, and the limit point is again B_∞, which thus satisfies $B_\infty = \pi(B_\infty)$ since π is continuous. By the triangle inequality, the fact $\pi(B) = \pi(B_0)$, (6.1) and (6.7) we have

$$\|\pi(B_k) - \pi(B)\| < \frac{2\epsilon_6 l}{1-c} < \epsilon l,$$

and combining this with (6.3), (6.4), (6.6) we also have

$$\|A - B_k\| \leq \|A - \pi(B)\| + \|\pi(B) - \pi(B_k)\| + \|\pi(B_k) - B_k\| < s_6/4 + \epsilon l + 2l \leq$$
$$\leq s_6/4 + \epsilon(s + s_6/4) + 2(s + s_6/4) < s_6,$$

where the last inequality follows by (6.2). With these estimates at hand, it is easy to turn the above argument into a proper induction proof in which (6.5) is verified at each step. We omit the details. □

7. Rank k matrices versus Hankel matrices; numerical examples

We now continue the example in Section 2 and the application to approximation of a signal by sums of k exponential functions, as outlined towards the end. By Theorem 4.4, an arbitrary intersection point of $\mathcal{R}_k$ and $\mathcal{H}$ is almost surely non-tangential. Given the signal f we let g be the closest "sum of k exponentials" to f in the $\|\cdot\|_w$-norm, see (2.6). Theorem 6.1 thus says that if $\|f-g\|_w$ is not too large, the sequence of alternating projections will converge to an $H(g_\infty)$, where $g_\infty \approx g$ is a sum of k exponentials. More precisely, let

$$\mathcal{K} = \left\{ \sum_{j=1}^{k} c_j (\alpha_j^l)_{l=0}^{2n-2} : \ c_j \in \mathbb{C}, \ \alpha_j \in \mathbb{C} \right\} \subset \mathbb{C}^{2n-1}.$$

Then for every $\epsilon > 0$ there exists an $s > 0$ such that if $\mathsf{dist}(f, \mathcal{K}) < s$, then

$$\|g_\infty - g\|_w < \epsilon \ \mathsf{dist}(f, \mathcal{K}).$$

We now discuss what happens if f is not close enough to g that Theorem 6.1 can be applied for any ϵ. Since both $\pi_{\mathcal{H}}$ and $\pi_{\mathcal{R}_k}$ are contractions, the sequence of alternating projections $(B_j)_{j=0}^{\infty}$ (with $B_0 = H(f)$) will be bounded. Thus it has a convergent subsequence, and the limit point is easily seen to be in $\mathcal{H}_k$. It is not hard to deduce that $\mathsf{dist}(B_j, \mathcal{H}_k) \to 0$ as $j \to \infty$. The only way the whole sequence could avoid converging is thus if it switches endlessly along the valleys of the thin set of tangential points, thereby avoiding the surrounding hills made by the open sets $\cup_{A \in \mathcal{H}_k} \mathcal{B}(A, s_A)$, where $s_A > 0$ is given by Theorem 6.1 (with $\epsilon = 1$ say) for all non-tangential A's, and $s_A = 0$ for the thin set of tangential A's. That this could happen seems highly unlikely to us, and we have certainly never encountered it in practice. However, we leave it as an open problem to prove that this can not occur.

For practical purposes, it is of course of interest that s be a relatively large number, (given some fixed ϵ and c). Inspecting the proof Theorem 6.1, one gets the impression that $B(A, s)$ will not be distinguishable even with binoculars. To test the actual relationship between s and ϵ in Theorem 6.1, (with $c = 1$), we conducted a few experiments which we now present. A function[2] g is written as a sum of 10 exponential functions, and then f is generated by adding a smaller function n to g. g is normalized such that $\|g\|_w = 1$, (recall (2.6)), n is chosen such that $H(n)$ is orthogonal to $\mathcal{H}_{10}^n$ at Hg, (with respect to the Hilbert-Schmidt norm). By Proposition 3.3, $H(g)$ is likely very close to $\pi(H(f))$, and so we will estimate $\mathsf{dist}(H(f), \mathcal{H}_{10}) = \|H(f) - \pi(H(f))\|$ by $\|H(f) - H(g)\| = \|n\|_w$ and $\|H(f_\infty) - \pi(H(f))\|$ by $\|H(f_\infty) - H(g)\|$. Note that there is no explicit way to

[2] In this section we will sloppily say function when we talk of its sampling, i.e. a sequence.

compute $\pi(H(f))$. We set

$$\varepsilon(g,n) = \frac{\|H(f_\infty) - H(g)\|}{\|H(n)\|} = \frac{\|f_\infty - g\|_w}{\|n\|_w}.$$

Let s be a parameter taking values $0.1, 0.01,\ 0.001,\ 0.0001$ and 0.00001. For each value of s we randomly generate 50 normalized g's as above, and for each g we randomly generate a n with $\|n\|_w = s$. The supremum of $\varepsilon(g,n)$ should give us an idea of what value of ϵ allows for such an s as in Theorem 6.1. The results are demonstrated in Figure 8 and Figure 9. The difference between the two is that Figure 8 used white noise, and in Figure 9, the "noise" consisted of sums of randomly generated exponentials. Judging from these figures, $\epsilon \approx \sigma$ seems to be a good rule of thumb, although this will of course vary substantially from one application to another. The method seems to work slightly better for the case of white noise.

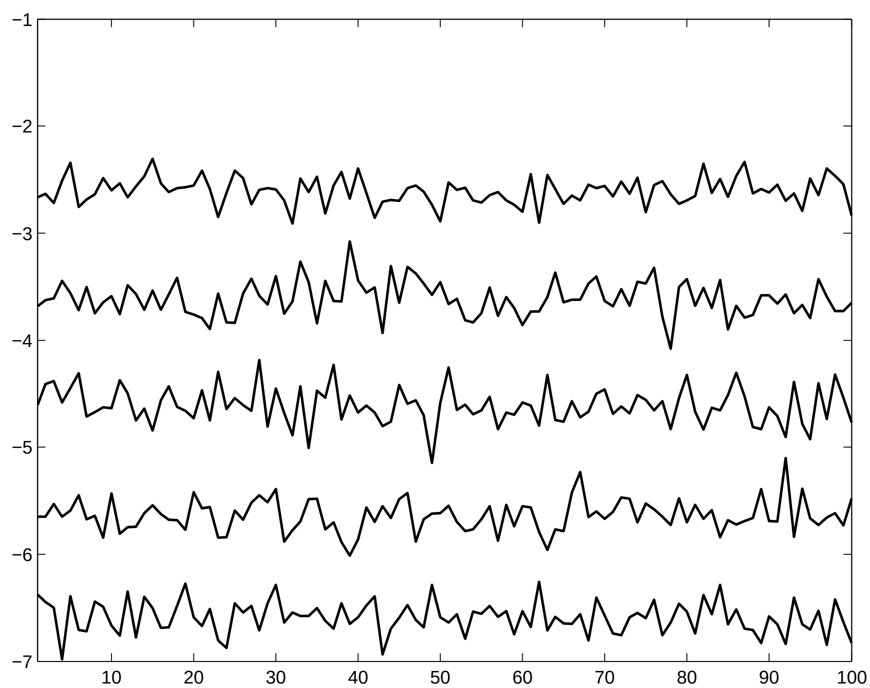


FIGURE 8. Signal with 10 exponentials and white noise.

8. Summary and open problems

We have developed a theoretical framework that enables us to understand convergence properties of alternating projection schemes. This provides substantially stronger results than previously available, e.g., for theorems relying solely on Zangwill's Global Convergence Theorem, which neither provide convergence rate estimates nor information about how far away the computed approximation is from the optimal value. Moreover, in contrast to the theory developed for the case of transversal manifolds, our framework provides more information under very non-restrictive requirements on the manifolds.

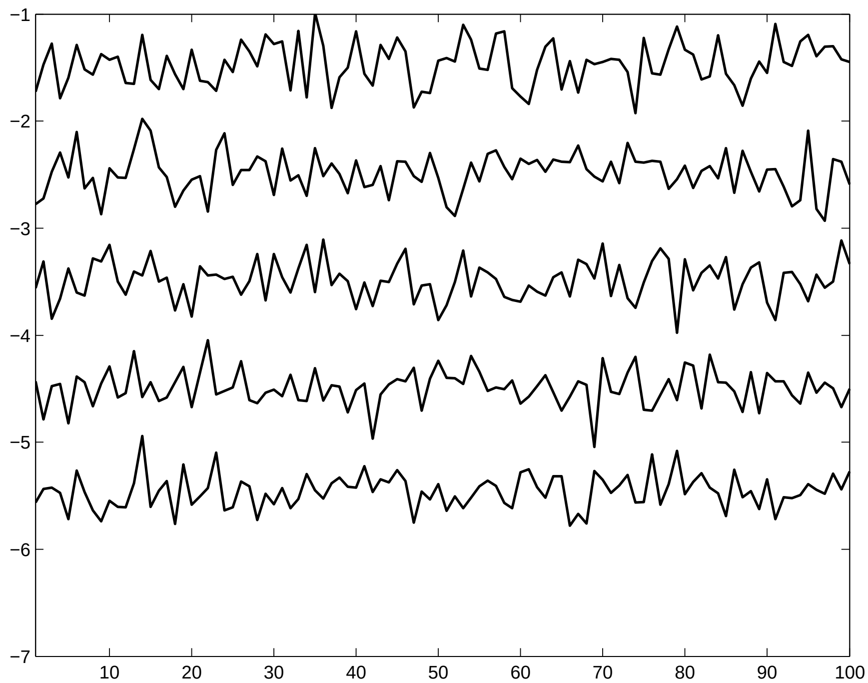


FIGURE 9. Signal with 10 exponentials and sums of other exponentials added.

We end by noting a few open problems. We first discuss smoothness. By examples similar to Example 1.2, it is not hard to see that one needs the manifolds to be at least C^1 for the alternating projections to converge. By inspection of the proofs, one easily sees that for Theorem 6.1 to hold, it suffices with C^1_{Lip}, i.e. that $d\phi$ is Lipschitz continuous. It is thus an open question whether Theorem 6.1 is true assuming only C^1. Next we discuss non-tangentiality. Again, Example 1.2 can be tailored such that the alternating projections does not converge, despite assuming C^∞, if we allow the manifolds to be tangential. However, for all applications we are aware of, the manifolds are algebraic, and it seems quite unlikely that a similar thing could happen in this case. We conjecture that for algebraic manifolds, the sequence of alternating projections will always converge, without any additional assumptions. Recall Example 1.4, where the sequence clearly converges albeit *extremely slowly*. For practical purposes, non-tangentiality is thus still vital.

9. Acknowledgements

This work was supported by the Swedish Research Council and the Swedish Foundation for International Cooperation in Research and Higher Education.

Acknowledgement: This work was partially conducted at Lund University, supported by the Swedish Research Council; Grant 2008-23883-61232-34

References

1. Fredrik Andersson, Marcus Carlsson, and Per Anders Ivert, *A fast alternating projection method for complex frequency estimation.*, Submitted.

2. Catalin Badea, Sophie Grivaux, and Vladimir Müller, *A generalization of the Friedrichs angle and the method of alternating projections*, Comptes Rendus Mathematique **348** (2010), no. 1-2, 53 – 56.
3. 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.
4. Heinz H. Bauschke and Jonathan M. Borwein, *On projection algorithms for solving convex feasibility problems*, SIAM Rev. **38** (1996), 367–426.
5. Heinz H. Bauschke, Dominikus Noll, Anna Celler, and Jonathan M. Borwein, *An em algorithm for dynamic spect*, IEEE TRANS. MED. IMAG **18** (1999), no. 3.
6. M. Berger and B. Gostiaux, *Differential geometry: Manifolds, curves and surfaces.*, Springer Verlag, 1988.
7. J.A. Cadzow, *Signal enhancement-a composite property mapping algorithm*, Acoustics, Speech and Signal Processing, IEEE Transactions on **36** (1988), no. 1, 49 –62.
8. J. E. Dennis and R.B. Schnabel, *Numerical methods for unconstrained optimization and nonlinear equations.*, Prentice-Hall, New Jersey, 1983.
9. J. N. Franklin, *Matrix theory.*, Dover Publications in Mineola, N.Y., 2000.
10. K.M. Grigoriadis, A.E. Frazho, and R.E. Skelton, *Application of alternating convex projection methods for computation of positive toeplitz matrices*, Signal Processing, IEEE Transactions on **42** (1994), no. 7, 1873 –1875.
11. Roger A. Horn and Charles R. Johnson, *Topics in matrix analysis*, Cambridge University Press, Cambridge, 1994. MR 1288752 (95c:15001)
12. Aharon Levi and Henry Stark, *Signal restoration from phase by projections onto convex sets*, J. Opt. Soc. Am. **73** (1983), no. 6, 810–822.
13. Adrian S. Lewis and Jérôme Malick, *Alternating projections on manifolds*, Math. Oper. Res. **33** (2008).
14. Ye Li, K.J.R. Liu, and J. Razavilar, *A parameter estimation scheme for damped sinusoidal signals based on low-rank hankel approximation*, Signal Processing, IEEE Transactions on **45** (1997), no. 2, 481 –486.
15. Biao Lu, Dong Wei, B.L. Evans, and A.C. Bovik, *Improved matrix pencil methods*, Signals, Systems Computers, 1998. Conference Record of the Thirty-Second Asilomar Conference on, vol. 2, nov 1998, pp. 1433 –1437 vol.2.
16. Ivan Markovsky, *Structured low-rank approximation and its applications*, Automatica **44** (2008), 891–909.
17. Robert J. Marks, II, *Alternating projections onto convex sets*, pp. 476–501, Academic Press, Inc., Orlando, FL, USA, 1996.
18. John Von Neumann, *Functional operators, volume ii: The geometry of orthogonal spaces*, Princeton University Press, 1950.
19. V.U. Prabhu and D. Jalihal, *An improved ESPRIT based time-of-arrival estimation algorithm for vehicular OFDM systems*, Vehicular Technology Conference, 2009. VTC Spring 2009. IEEE 69th, april 2009, pp. 1 –4.
20. W. I. Zangwill, *Nonlinear programming*, Prentice Hall, Englewood Cliffs, N. J., 1969.

Centre for Mathematical Sciences, Lund University, Sweden
E-mail address: `fa@maths.lth.se`

Departamento de Matemáticas,Universidad de Santiago de Chile, Chile
E-mail address: `marcus.carlsson@usach.cl`