# A simple proof of polar decomposition in pseudo-Euclidean geometry

by

## Maciej P. Wojtkowski (Olsztyn)

**Abstract.** We give a simple direct proof of the polar decomposition for separated linear maps in pseudo-Euclidean geometry.

**1. Introduction.** Ricardo Mañé [M] introduced the concept of a dominated splitting for a diffeomorphism of a compact manifold, which is a weak form of hyperbolicity of dynamical systems. This property is the subject of several recent papers, in particular [P-S], which contains an extensive bibliography. It was observed in [W1] that the dominated splitting is associated with a pseudo-Riemannian structure on the manifold and the derivative having a special property with respect to that structure, namely that it takes positive tangent vectors to positive tangent vectors. Such linear maps were called *separated* in [W2] for the reason that this property leads to the gap in the spectrum, for instance the separation of Lyapunov exponents. Other terms can be found in the literature, e.g., they were called *plus operators* by Krein and Shmul'jan [K-S1].

Non-degenerate separated linear maps have the polar decomposition, and the corresponding singular values, similarly to the Euclidean case. It was shown in [W2] how to estimate Lyapunov exponents in terms of these singular values.

The polar decomposition of separated linear maps was first obtained by Potapov [P]. We give a simple direct proof of this fact.

**2. Separated linear maps.** Let us consider an $n$-dimensional real linear vector space $V$ with a chosen *pseudo-Euclidean structure*, i.e., a non-degenerate symmetric bilinear form $[\cdot,\cdot]$ with the positive index of inertia

equal to $p$ and the negative index of inertia equal to $q$, $p + q = n$, $p \geq 1$, $q \geq 1$. We will refer to it as the *J-form*.

Depending on the sign of the J-form, we have positive and negative vectors. We denote by

$$C_{\pm} = \{v \in V \mid \pm[v,v] > 0\} \cup \{0\}$$

the open cones of, respectively, positive and negative vectors (with the zero vector included), and by $C_0$ their common boundary,

$$C_0 = \{v \in V \mid [v,v] = 0\}.$$

Without loss of generality we can assume that $V = \mathbb{R}^p \oplus \mathbb{R}^q$ and for $v = (v_1, v_2), v_1 \in \mathbb{R}^p, v_2 \in \mathbb{R}^q$,

$$[v,v] = \langle Jv, v \rangle = v_1^2 - v_2^2, \quad \text{where} \quad J = \begin{bmatrix} I_p & 0 \\ 0 & -I_q \end{bmatrix},$$

where $I_p, I_q$ are the identity matrices in $\mathbb{R}^p$ and $\mathbb{R}^q$ respectively. By $\langle \cdot, \cdot \rangle$ we denote here the arithmetic scalar product, but we will avoid using it in our paper. The terms orthogonality, isometry, etc., will refer exclusively to the J-form.

A linear subspace $E$ is called non-degenerate if the restriction of the J-form to this subspace is non-degenerate. We know that $E$ is non-degenerate if and only if its orthogonal complement $E^{\perp}$ is complementary, i.e., $E \cap E^{\perp} = \{0\}$, or equivalently $E \oplus E^{\perp} = V$.

DEFINITION 1. A linear non-singular operator $A : V \to V$ is

(i) *separated* if $AC_+ \subset C_+$,
(ii) *strictly separated* if $A(C_0 \cup C_+) \subset C_+$,
(iii) *monotone* if $[Av, Av] \geq [v,v]$ for every $v \in V$.

DEFINITION 2. For a linear operator $B : V \to V$ the *adjoint operator* $B^{\dagger} : V \to V$ is defined by the condition that for every $v, w \in V$,

$$[B^{\dagger}v, w] = [v, Bw].$$

An operator $B$ is *symmetric* (or *self-adjoint*) if $B^{\dagger} = B$.

In coordinates we have

$$B^{\dagger} = JB^T J.$$

Clearly for any operator $A$ the operators $A^{\dagger}A$ and $AA^{\dagger}$ are symmetric. An operator $U$ is an isometry if and only if $U^{\dagger}U = I = UU^{\dagger}$.

Potapov [P] obtained the following fundamental result:

THEOREM 1. *Every non-degenerate separated operator $A$ has a unique representation $A = UR$, where $U$ is an isometry, and $R$ is symmetric and has only real positive eigenvalues.*

The operator $R$ is called the *modulus* of $A$. Strictly speaking, Potapov considered only monotone operators, but he was well aware that for every separated operator $A$ there is a positive real number $\alpha > 0$ such that $\alpha A$ is monotone, so there is no actual generalization (the range of such numbers $\alpha$ will be described in Theorem 3). The proof of Potapov is elegant and intricate. We propose a simple direct proof of the Potapov modulus theorem.

The discussion of the general case of polar decomposition can be found in [B-R], and for the infinite-dimensional case, see [K-S2], [B]. The paper [M-R-R] addresses the general case of plus matrices (in our terminology: separated linear maps without the assumption of non-singularity).

Suppose we have $S = UR$ with $R$ symmetric and $U$ an isometry. Then $S^\dagger S = RU^\dagger UR = R^2$. Hence our task is to obtain the square root of the operator $S^\dagger S$. We will take advantage of the well known fact that if a linear map $L : V \to V$ has only real positive eigenvalues than it has a unique square root with positive eigenvalues. It is easy to obtain the square root using the Jordan normal form of $L$.

The Potapov modulus theorem follows from

THEOREM 2. *For every non-degenerate separated operator $S$ the operator $S^\dagger S$ has only real positive eigenvalues.*

Indeed, once we have $R = (S^\dagger S)^{1/2}$ we can define $U = SR^{-1}$. We see immediately that $U$ is an isometry:

$$U^\dagger U = R^{-1} S^\dagger S R^{-1} = I.$$

**3. Three lemmas.** Let us formulate three lemmas which apply to any pseudo-Euclidean structure. The first two are standard and we give them without proof.

LEMMA 1. *If $B$ is symmetric and $E$ is an invariant subspace of $B$, $B(E) \subset E$, then $E^\perp$, the orthogonal complement of $E$, is also invariant, $B(E^\perp) \subset E^\perp$. Also, any two eigenvectors of $B$ with different eigenvalues are orthogonal.*

LEMMA 2. *If $B$ is symmetric then the function $[Bv, v]$, $v \in \mathbb{V}$, restricted to $[v, v] = $ const has a critical point at $w \neq 0$ if and only if $w$ is an eigenvector of $B$.*

COROLLARY 1. *If for a symmetric $B$ there is a real number $r$ such that $[Bv, v] \geq r[v, v]$ for all $v$, and there is $w \neq 0$ such that $[Bw, w] = r[w, w]$, then $w$ is an eigenvector of $B$ and $r$ is the corresponding eigenvalue.*

*Proof.* The form $[Bv, v]$ has a conditional minimum at $w$ under the condition $[v, v] = [w, w]$. By Lemma 2, $w$ is an eigenvector with an eigenvalue $b$. If $[w, w] \neq 0$ then clearly $b = r$. If $[w, w] = 0$ then let us consider a vector

$v_0 \in C_0$ such that $[w, v_0] = 1$. We now get, for every $\epsilon$,

$$[B(w + \epsilon v_0), w + \epsilon v_0] = 2\epsilon b + \epsilon^2 [Bv_0, v_0] \geq r[w + \epsilon v_0, w + \epsilon v_0] = 2\epsilon r.$$

It follows immediately that $b = r$. ∎

Let us consider the case of a plane with the Lorentzian form

$$J = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

We find that $B$ is symmetric if and only if $a = d$. Clearly this condition does not guarantee that $B$ has real eigenvalues, nor that it has a basis of eigenvectors. However, we have the following

LEMMA 3. *If for a symmetric $B$, $[Bv, v] \geq 0$ for $v$ in the positive cone $C_+$, and the form $[B\cdot, \cdot]$ assumes also negative values, then $B$ has only positive eigenvalues. If the eigenvalues are different then the eigenvector of the larger eigenvalue is in $C_+$, and the eigenvector of the smaller one is in $C_-$. If the eigenvalue is double then either $B$ is a multiple of identity, or $B$ is a Jordan block with the unique eigenvector in $C_0$.*

*Proof.* Putting $v = (x, y)$ we have $[v, v] = 2xy$ and $[Bv, v] = bx^2 + cy^2 + 2axy$. For this quadratic form to be non-negative in $C_+$ it is necessary that $b, c \geq 0$. We have further

$$[Bv, v] = (\sqrt{b}x - \sqrt{c}y)^2 + 2(a + \sqrt{bc})xy = (\sqrt{b}x + \sqrt{c}y)^2 + 2(a - \sqrt{bc})xy.$$

Since the quadratic form is assumed to have negative values somewhere in $C_-$, the second formula leads to $a - \sqrt{bc} > 0$. We have

$$[Bv, v] \geq (a + \sqrt{bc})[v, v] \quad \text{and} \quad [Bv, v] \geq (a - \sqrt{bc})[v, v].$$

Assuming that $b, c > 0$ we have equalities occurring at the non-zero vectors $v$ in $C_+$ and $C_-$, respectively. By Corollary 1 they must be eigenvectors of $B$ with eigenvalues $a \pm \sqrt{bc}$.

Clearly one can get this conclusion and the rest of the lemma by direct calculation. We chose to apply Corollary 1 to illustrate its efficacy. ∎

Another consequence of the proof is that for any number $r$ with $a - \sqrt{bc} \leq r \leq a + \sqrt{bc}$ we have $[Bv, v] \geq r[v, v]$ for all $v \in \mathbb{R}^2$. By Corollary 1 the equality occurs only for an eigenvector $v$ (or $v = 0$), and $r$ equal to the eigenvalue.

## 4. Proof of the main theorem

THEOREM 3. *If a non-degenerate linear map $A$ is strictly separated then $B = A^\dagger A$ has only positive eigenvalues $\lambda_p \geq \cdots \geq \lambda_1 > \mu_1 \geq \cdots \geq \mu_q > 0$ and a basis of respective mutually orthogonal eigenvectors $v_p, \ldots, v_1 \in C_+$ and $w_1, \ldots, w_q \in C_-$. Moreover, for any $r \in [\mu_1, \lambda_1]$, we have $[Bv, v] =$*

$[Av, Av] > r[v, v]$ *for every non-zero* $v \in V$, *unless* $v$ *is an eigenvector with eigenvalue* $r$.

*Proof.* Note that the quadratic form $[Av, Av] = [Bv, v]$ has the same type as the form $[v, v]$. In particular, it is indefinite. Moreover, it is positive for $v \notin C_-$.

Consider the Euclidean sphere $\mathbb{S} = \{v \in \mathbb{R}^n \mid \langle v, v \rangle = 1\}$ and the function $f(v) = [Bv, v]/[v, v]$, well defined on $\mathbb{S} \setminus C_0$. The numerator $[Bv, v]$ is positive on the compact subset $\mathbb{S} \setminus C_-$, hence it is bounded away from zero there. The denominator is zero at the boundary of this set. It follows that $f(v)$ goes to $+\infty$ at the boundary of $\mathbb{S} \cap C_+$, and finally $f(v)$ attains a minimum value $\lambda_1 > 0$ at $v_1 \in \mathbb{S} \cap C_+$. We have shown that the quadratic form $[Bv, v]$ has a minimum at $v_1$ under the condition $[v, v] = 1$. By Lemma 2, $v_1$ is an eigenvector with eigenvalue $\lambda_1$. Moreover, $[Bv, v] \geq \lambda_1[v, v]$ for $v \notin C_-$.

Since the quadratic form $[Bv, v]$ assumes negative values at some $v \in \mathbb{S} \cap C_-$, the function $f(v)$ assumes positive values at the same $v$. Near the boundary of $\mathbb{S} \cap C_-$ the form $[Bv, v]$ has positive values and is bounded away from zero on a neighborhood of the boundary. Again it follows that $f(v)$ goes to $-\infty$ as $v \in \mathbb{S} \cap C_-$ tends to the boundary. We conclude that $f(v)$ attains a maximum value $\mu_1 > 0$ at $w_1 \in C_-$, and by Lemma 2, $w_1$ is an eigenvector with eigenvalue $\mu_1$. Moreover, $[Bv, v] \geq \mu_1[v, v]$ for $v \notin C_+$.

Let $E$ be the plane spanned by $v_1$ and $w_1$. Then $B(E) = E$. Hence we can apply Lemma 3 to the restriction of the form $[B\cdot, \cdot]$ to the plane $E$. It follows that $\mu_1 < \lambda_1$ and that the eigenvectors are orthogonal.

Applying now Corollary 1 we get the last part of our theorem.

Let $F$ be the orthogonal complement of $E$. Since the subspace $E$ is non-degenerate (i.e., the restriction of the J-form to $E$ is non-degenerate), the subspace $F$ is complementary to $E$, that is, $E \cap F = \{0\}$ and $E \oplus F = V$. We restrict the quadratic forms $[\cdot, \cdot]$ and $[B\cdot, \cdot]$ to $F$. They must have the same type. If they are both positive (or negative) definite we get the conclusion of our theorem by observing that now $B$ is symmetric with respect to the Euclidean scalar product $\pm[\cdot, \cdot]$. If the forms are indefinite we can repeat the argument above and obtain another pair of eigenvalues

$$\lambda_2 = \inf_{v \in C_+ \cap F} \frac{[Bv, v]}{[v, v]} \geq \inf_{v \in C_+} \frac{[Bv, v]}{[v, v]} = \lambda_1,$$

$$0 < \mu_2 = \sup_{v \in C_- \cap F} \frac{[Bv, v]}{[v, v]} \leq \sup_{v \in C_-} \frac{[Bv, v]}{[v, v]} = \mu_1$$

with orthogonal eigenvectors $v_2 \in C_+ \cap F$ and $w_2 \in C_- \cap F$.

Repeating this step we will exhaust the dimension and obtain an orthogonal basis of eigenvectors with all eigenvalues positive, which is the content of our theorem. ∎

**5. The limit case.** Let us finally discuss the limit case when $A$ is separated but not strictly separated. There are strictly separated operators $D_\epsilon$ which are $\epsilon$-close to the identity. Then $A_\epsilon = D_\epsilon A$ is strictly separated and $\epsilon$-close to $A$. By continuity of the spectrum, Theorem 3 immediately implies

THEOREM 4. *If $A$ is separated then $B = A^\dagger A$ has only positive eigenvalues $\lambda_p \geq \cdots \geq \lambda_1 \geq \mu_1 \geq \cdots \geq \mu_q > 0$, and $A$ is strictly separated if and only if $\mu_1 < \lambda_1$. Moreover, for any $r \in [\mu_1, \lambda_1]$ we have $[Bv, v] = [Av, Av] > r[v, v]$ for every non-zero $v$, unless $v$ is an eigenvector with eigenvalue $r$.*

*Proof.* This follows from Theorem 3 by taking the limit over strictly separated operators as explained above. In particular, for any $r$ with $\mu_1 \leq r \leq \lambda_1$ we get $[Bv, v] \geq r[v, v]$ for all $v \in V$. To prove the criterion for strict separatedness, assume that $A$ is non-strictly separated. Then there is a nonzero $w \in C_0$ such that $Aw \in C_0$. It follows that $[Bw, w] = 0 = r[w, w]$ and by Corollary 1 such a $w$ must be an eigenvector with eigenvalue $r$. It follows that $r = \mu_1 = \lambda_1$. ∎

It remains to inspect what can happen to the orthogonal bases in the limit of non-strictly separated $A$, when $\mu_1 = \lambda_1$. As opposed to the Euclidean case, the orthogonal bases can collapse and give rise to Jordan blocks.

If two or more orthogonal eigenvectors collapse to one then the limit eigenvector must belong to $C_0$, and its eigenvalue is equal to $r = \mu_1 = \lambda_1$ (cf. the proof of Theorem 4 above). It is interesting that the resulting Jordan block must be of dimension 2, which was known to Frobenius [P]. Indeed, we have

PROPOSITION 1. *If $A$ is non-strictly separated then any Jordan block of $B = A^\dagger A$ with eigenvalue $r = \mu_1 = \lambda_1$ is of dimension 2.*

*Proof.* We have $[Bv, v] > r[v, v]$ for any non-zero $v$, unless $v$ is an eigenvector with eigenvalue $r$. Consider the operator $H = B - rI$ which has the eigenvalue zero. $B$ has a Jordan block with eigenvalue $r$ of dimension 3 or more if and only if there is a non-zero vector $z$ such that $H^3 z = 0$ and $H^2 z \neq 0$. The inequality above yields $[Hv, v] > 0$ for non-zero $v$, unless $v$ is an eigenvector with eigenvalue zero, i.e. $Hv = 0$. Since $[H(Hz), Hz] = [H^3 z, z] = 0$, we conclude that $Hz$ is an eigenvector with eigenvalue zero, and hence $H^2 z = 0$. The contradiction proves our claim. ∎

Let us finally describe the complete structure of $B = A^\dagger A$ for a non-strictly separated $A$, with $\lambda_1 = \mu_1 = r > 0$ and $[Bv, v] = [Av, Av] \geq r[v, v]$.

PROPOSITION 2. *There is a splitting $V = E \oplus F \oplus G$ invariant under $B = A^\dagger A$ with the following properties:*

(i) *The subspaces $E, F, G$ are mutually orthogonal, and hence the restriction of $[\cdot, \cdot]$ to all three subspaces is non-degenerate (unless a subspace is trivial).*

(ii) *$[Bv, v] = [Av, Av] > r[v, v]$ for all non-zero $v \in G$, and there is an orthogonal basis of eigenvectors of $B$ in $G \setminus C_0$; the eigenvalues of the eigenvectors in $G \setminus C_-$ and in $G \setminus C_+$ are, respectively, smaller and larger than $r$.*

(iii) *$Bv = rv$ for $v \in F$.*

(iv) *The subspace $E$ has dimension $2k$, $E = E_1 \oplus \cdots \oplus E_k$ with mutually orthogonal planes $E_1, \ldots, E_k$, invariant under $B$. Each of these planes gives rise to a Jordan block for $B$, i.e., there is only one eigenvector with eigenvalue $r$ in each of the planes. The restriction of the J-form to any of the planes is indefinite, and hence the restriction of $B$ to such a plane is covered by Lemma 3, and the unique eigenvector belongs to $C_0$.*

*Proof.* If the spectrum of $B$ contains only one eigenvalue $r$ then the subspace $G$ is trivial. Otherwise it is spanned by a maximal orthogonal family of eigenvectors with eigenvalues different from $r$. We construct it in the same way as in the Euclidean case, taking advantage of the fact that none of these eigenvectors can lie in $C_0$, and hence its orthogonal complement is indeed complementary.

Once we exhaust all the eigenvalues different from $r$ we obtain an invariant non-degenerate subspace $G$. The restriction of $B$ to the orthogonal complement of $G$ has only one eigenvalue, equal to $r$. If there are no eigenvectors with this eigenvalue which are not in $C_0$, then the subspace $F$ is trivial. Otherwise it is spanned by a maximal orthogonal family of eigenvectors with eigenvalue $r$, which are not in $C_0$. The construction is the same as before.

We are left with the invariant orthogonal complement $E$ of $F \oplus G$. The restriction of $B$ to $E$ has only one eigenvalue equal to $r$ and all the eigenvectors lie in $E \cap C_0$. Moreover, for non-zero $v \in E$ we have the inequality $[Bv, v] = [Av, Av] > r[v, v]$ unless $v$ is an eigenvector. The eigenvectors in $E$ form a subspace $E_0$ which is completely contained in $C_0$. If the restriction of the J-form to the subspace $E$ has a positive index of inertia $p'$, and a negative index of inertia $q'$, then the dimension of $E_0$ does not exceed the minimum of $p'$ and $q'$. Hence $B$ must have Jordan blocks in $E$. By Proposition 1 they must have dimension 2.

We found that there must be Jordan blocks in $E$ and all of them must be of dimension 2. Let $E_1$ be such an invariant plane. We want to establish that $E_1$ is non-degenerate. There is a basis $\{z, w\}$ in $E_1$ such that $Bz = rz + w, Bw = rw$. We get $[Bz, z] = r[z, z] + [w, z]$ and since $z$ is not an eigenvector we must have $[w, z] > 0$. This together with $[w, w] = 0$ guarantees

that the restriction of the J-form to $E_1$ is indefinite, and so $E_1$ is non-degenerate.

The subspace $E$ must be spanned by such mutually orthogonal planes. Indeed, on the complement of all the Jordan blocks in $E$ we would have $B = rI$, and the J-form would be non-degenerate. That would give us eigenvectors of eigenvalue $r$ outside $C_0$, and all of them were already included in $F$. ∎

## References

[B]       J. Bognar, *Indefinite Inner Product Spaces*, Springer, Berlin, 1974.

[B-R]     Y. Bolshakov and B. Reichstein, *Unitary equivalence in an indefinite scalar product: an analogue of singular-value decomposition*, Linear Algebra Appl. 222 (1995), 155–226.

[K-S1]    M. G. Krein and Ju. L. Shmul'jan, *On plus operators in a space with an indefinite metric*, Amer. Math. Soc. Transl. (2) 85 (1969), 93–113.

[K-S2]    —, —, *J-polar representation of plus operators*, ibid. (2) 85 (1969), 115–143.

[M]       R. Mañé, *An ergodic closing lemma*, Ann. of Math. 116 (1982), 503–540.

[M-R-R]   C. V. M. van der Mee, A. C. M. Ran and L. Rodman, *Classes of plus matrices in finite-dimensional indefinite scalar product spaces*, Integral Equations Operator Theory 30 (1998), 432–451.

[P]       V. P. Potapov, *A theorem on the modulus*, Amer. Math. Soc. Transl. (2) 138 (1988), 21–35.

[P-S]     E. R. Pujals and M. Sambarino, *On the dynamics of dominated splitting*, Ann. of Math. 69 (2009), 675–740.

[W1]      M. P. Wojtkowski, *W-flows on Weyl manifolds and Gaussian thermostats*, J. Math. Pures Appl. (9) 79 (2000), 953–974.

[W2]      —, *Monotonicity, J-algebra of Potapov and Lyapunov exponents*, in: Proc. Sympos. Pure Math. 69, Amer. Math. Soc., 2001, 499–521.

Maciej P. Wojtkowski
Department of Mathematics and Informatics
University of Warmia and Mazury
Żołnierska 14
10-561 Olsztyn, Poland
E-mail: wojtkowski@matman.uwm.edu.pl