An algorithm for the Cartan-Dieudonné theorem on generalized scalar product spaces

M.A. Rodríguez-Andrade [a], G. Aragón-González [b], J.L. Aragón [c], and Luis Verde-Star [d]

[a] *Departamento de Matemáticas, Escuela Superior de Física y Matemáticas, IPN, México D.F. 07300, México and Departamento de Matemática Educativa Centro de Investigación y Estudios Avanzados, IPN, México D.F. 07360, México.*

[b] *Programa de Desarrollo Profesional en Automatización, Universidad Autónoma Metropolitana, Azcapotzalco, San Pablo 180, Colonia Reynosa-Tamaulipas, México D.F. 02200, México.*

[c] *Centro de Física Aplicada y Tecnología Avanzada, Universidad Nacional Autónoma de México, Apartado 1-1010, Querétaro 76000, México.*

[d] *Departamento de Matemáticas, Universidad Autónoma Metropolitana, Iztapalapa, Apartado 55-534, México D.F. 09340, México.*

Abstract

We present an algorithmic proof of the Cartan-Dieudonné theorem on generalized real scalar product spaces with arbitrary signature. We use Clifford algebras to compute the factorization of a given orthogonal transformation as a product of reflections with respect to hyperplanes. The relationship with the Cartan-Dieudonné-Scherk theorem is also discussed in relation to the minimum number of reflections required to decompose a given orthogonal transformation.

Key words: Cartan-Dieudonné, orthogonal matrices, Householder transformations, orthogonal group, Clifford algebras
AMS classification: 15A23, 15A30, 15A66

Email addresses: marco@esfm.ipn.mx (M.A. Rodríguez-Andrade), gag@correo.azc.uam.mx (G. Aragón-González), aragon@fata.unam.mx (J.L. Aragón), verde@xanum.uam.mx (Luis Verde-Star).

1 Introduction

The Cartan-Dieudonné theorem is a fundamental result in the theory of metric vector spaces. It states that every orthogonal transformation is the composition of reflections with respect to hyperplanes. The classical proofs of the Cartan-Dieudonné theorem use induction on the dimension of the vector space and are not constructive. See [1] and [2]. Recently, Uhlig [3] presented a constructive proof of the Cartan-Dieudonné theorem for the case of vector spaces with a positive definite inner product, and also a constructive proof of a weaker version of the theorem for generalized scalar product spaces of signature (p, q) [3, Thm.3].

The matrix representation of a reflection with respect to a hyperplane is called a Householder matrix [3]. The analogues of Householder transformations on spaces with a non-degenerate bilinear or sesquilinear form are studied in [4]. Householder matrices are also used in Gallier's book [5, Ch. 7], which discusses the Cartan-Dieudonné theorem for linear and affine isometries, including applications to QR decomposition.

In the present paper we present an alternative proof of the Cartan-Dieudonné theorem for generalized real scalar product spaces of arbitrary signature. The proof yields an algorithm for the factorization of a given orthogonal transformation as a product of reflections with respect to hyperplanes. This work is a generalization to spaces of arbitrary signature of a previous one [6], where we provide an algorithmic proof of the Cartan-Dieudonné theorem in $\mathbb{R}^n$ valid over the fields $\mathbb{Q}$, $\mathbb{R}$ or $\mathbb{C}$.

In the theory of Clifford algebras there is an alternative way to find the image of a vector under a reflection with respect to a hyperplane. This is done using vector multiplications under the rules of the Clifford algebras. This is the method that we propose for the computations involved in the factorization algorithm. Aditionally, our approach produces an alternative way for calculating the Householder matrices with respect to orthogonal bases, but we don't use these matrices in the numerical examples.

In section 2 we present some results about vector spaces with non-degenerate bilinear forms. We include some results about Artinian spaces because they are important for the development of our proof of the Cartan-Dieudonné theorem. In section 3 we present the proof for the case of spaces with a symmetric non-degenerate bilinear form of arbitrary signature. In section 4 we propose the use of Clifford algebras as a computational tool to obtain the reflections that give the factorization of a given orthogonal transformation. In section 5 we present some examples of the factorization of orthogonal matrices. Finally, in section 6 some conclusions are given, including a comment about the relationship with the Cartan-Dieudonné-Scherk theorem, and the minimum number of reflections required to decompose a given orthogonal transformation.

2 Generalized scalar product spaces

In this section we present some basic results concerning real vector spaces equipped with a non-degenerate symmetric bilinear form. We call such spaces (generalized) scalar product spaces. They are also known as metric vector spaces or real orthogonal spaces. The proofs of the results presented here can be found in [2].

Definition 2.1 *Let $\mathcal{X}$ be a real vector space and let $\mathcal{B} : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ be a map that satisfies the conditions*

(B1) (Bilinearity) For all $v, v', w, w' \in \mathcal{X}$ and for all $\lambda \in \mathbb{R}$

$$\mathcal{B}\left(\lambda v + v', w\right) = \lambda \mathcal{B}\left(v, w\right) + \mathcal{B}\left(v', w\right),$$

and

$$\mathcal{B}(v, \lambda w + w') = \lambda \mathcal{B}\left(v, w\right) + \mathcal{B}\left(v, w'\right).$$

(B2) (Symmetry) For all $v, w \in \mathcal{X}$, $\quad \mathcal{B}\left(v, w\right) = \mathcal{B}\left(w, v\right)$.

(B3) (Non-degeneracy) For each non-zero v in $\mathcal{X}$ there exists w in $\mathcal{X}$ such that $\mathcal{B}\left(v, w\right) \neq 0$.

Then the pair $(\mathcal{X}, \mathcal{B})$ is said to be a (generalized) scalar product space*.*

Definition 2.2 *Let $\mathcal{X}$ be a real vector space and $\mathcal{B}$ a bilinear form on $\mathcal{X}$.*

(1) The vectors u, v in $\mathcal{X}$ are orthogonal if $\mathcal{B}\left(u, v\right) = 0$.
(2) A vector u in $\mathcal{X}$ is called isotropic if $\mathcal{B}\left(u, u\right) = 0$.
(3) We say that u is invertible if u is not isotropic, that is, if $\mathcal{B}\left(u, u\right) \neq 0$. (This terminology will be justified in section 4.)
(4) Let $\mathcal{W}$ and $\mathcal{V}$ be vector subspaces of $\mathcal{X}$. We say that $\mathcal{W}$ and $\mathcal{V}$ are orthogonal if $\mathcal{B}\left(v, u\right) = 0$, for all $u \in \mathcal{V}$ and $v \in \mathcal{W}$.
(5) A subspace $\mathcal{V}$ of $\mathcal{X}$ is called null subspace if $\mathcal{B}\left(v, u\right) = 0$, for all $u, v \in \mathcal{V}$.
(6) Let $\mathcal{W}$ be a subspace of $\mathcal{X}$. The orthogonal complement of $\mathcal{W}$ is the subspace $\mathcal{W}^{\perp} = \{u \in \mathcal{X} \mid \mathcal{B}(v, u) = 0, \text{ for all } v \in \mathcal{W}\}$.
(7) Let $\mathcal{W}$ be a subspace of $\mathcal{X}$. We say that $\mathcal{W}$ is a non-degenerate subspace, relative to $\mathcal{B}$, if the restriction of $\mathcal{B}$ to $\mathcal{W} \times \mathcal{W}$ is non degenerate.

Note that the vector subspace generated by an isotropic vector u is a null subspace of $\mathcal{X}$.

The next proposition states when a subspace and its orthogonal complement decompose the space $\mathcal{X}$ as a direct sum.

Proposition 2.3 *Let $(\mathcal{X}, \mathcal{B})$ be a generalized scalar product space of dimension n and let $\mathcal{W}$ be a subspace of $\mathcal{X}$. Then $\mathcal{X} = \mathcal{W} \oplus \mathcal{W}^{\perp}$ if and only if $\mathcal{W}$ is non-degenerate. That is, the space $(\mathcal{W}, \mathcal{B}\mid_{\mathcal{W}})$ is non-degenerate, where $\mathcal{B}\mid_{\mathcal{W}}$ is the restriction of $\mathcal{B}$ to the subspace $\mathcal{W}$. In particular, if $a \in \mathcal{X}$ we have $\mathcal{X} = \mathbb{R}a \oplus (\mathbb{R}a)^{\perp}$ if and only if $\mathcal{B}(a, a) \neq 0$, where $\mathbb{R}a$ denotes the subspace generated by a.*

For a proof see [2, Prop. 149.1].

In the remaider of this article $(\mathcal{X}, \mathcal{B})$ will denote a (generalized) scalar product space of dimension n over the field of real numbers.

Let $e = \{e_1, e_2, \ldots, e_n\}$ be an ordered basis of the vector space $\mathcal{X}$. For each pair of indices i, j in $\{1, 2, \ldots, n\}$ let $a_{i,j} := \mathcal{B}(e_i, e_j)$. The matrix $A = [a_{i,j}]$ is called the matrix of $\mathcal{B}$ relative to the basis e. It describes the bilinear form $\mathcal{B}$ in the following way. Let v, w in $\mathcal{X}$ and let $x = [v]_e$, $y = [w]_e$ be the coordinate (column) vectors of v and w with respect to the basis e. Then $\mathcal{B}(v, w) = x^t A y$. Since $\mathcal{B}$ is non-degenerate and symmetric, we see that A is a non-singular symmetric matrix.

Proposition 2.4 *Let $(\mathcal{X}, \mathcal{B})$ be a generalized scalar product space. Then there exist an ordered basis $e^* = (e_1^*, \ldots, e_n^*)$ of $\mathcal{X}$ and nonnegative integers p and q, with $p + q = n$, such that*

(1) $\mathcal{B}\left(e_j^*, e_j^*\right) = 1$ *for* $j = 1, 2, \ldots, p$.

(2) $\mathcal{B}\left(e_j^*, e_j^*\right) = -1$ *for* $j = p+1, p+2, \ldots, p+q$.

(3) $\mathcal{B}\left(e_i^*, e_j^*\right) = 0$ *for* $i \neq j$.

This means that the basis e^* diagonalizes the matrix associated with the bilinear form $\mathcal{B}$.

(4) The number of elements that satisfy $\mathcal{B}\left(e_j^*, e_j^*\right) = 1$ is independent of the basis that diagonalizes the bilinear form $\mathcal{B}$.

For a proof see [2, Prop. 159.1].

The basis e^* is called an orthonormal basis with respect to $\mathcal{B}$. If $q = 0$ then the space $(\mathcal{X}, \mathcal{B})$ is called positive definite.

Let p and q be nonnegative integers such that $n = p + q$. The bilinear form $\mathcal{B}^*$ on the space $\mathbb{R}^n$ defined by

$$\mathcal{B}^*(x, y) = \sum_{i=1}^{p} x_i y_i - \sum_{i=p+1}^{p+q} x_i y_i,$$

where $x = (x_1, x_2, \ldots, x_{p+q})$ and $y = (y_1, y_2, \ldots, y_{p+q})$, is symmetric and non-degenerate. The generalized scalar product space $(\mathbb{R}^n, \mathcal{B}^*)$ is denoted by $\mathbb{R}^{p,q}$.

Any generalized scalar product space $(\mathcal{X}, \mathcal{B})$ that satisfies conditions 1,2 and 3 of the previous proposition is isomorphic to the space $\mathbb{R}^{p,q}$. Any such space $(\mathcal{X}, \mathcal{B})$ is said to have characteristic or signature (p, q), and it is usually identified with $\mathbb{R}^{p,q}$. See [2, Thm. 177.1].

2.1 The orthogonal group

Among the linear operators on $(\mathcal{X}, \mathcal{B})$ the most interesting are clearly those that preserve the bilinear form.

Definition 2.5 *Let $T : \mathcal{X} \to \mathcal{X}$ be a linear operator. Then T is an orthogonal transformation if and only if*

$$\mathcal{B}(Tv, Tw) = \mathcal{B}(v, w), \qquad v, w \in \mathcal{X}.$$

The set of all the orthogonal transformations is a group, called the orthogonal group of $(\mathcal{X}, \mathcal{B})$, and denoted by $\mathcal{O}(\mathcal{X})$.

The group $\mathcal{O}(\mathbb{R}^{p,q})$ can be considered as the set of invertible $n \times n$ matrices Q that satisfy $Q^t A Q = A$, where A is the matrix associated with the bilinear form $\mathcal{B}$ with respect to the canonical basis of $\mathbb{R}^{p,q}$. We denote $\mathcal{O}(\mathbb{R}^{p,q})$ by $\mathcal{O}(p, q)$.

Definition 2.6 $\mathcal{SO}(p, q) := \{Q \in \mathcal{O}(p, q) \mid \det(Q) = 1\}$ *is the group of special orthogonal transformations or rotations of $\mathbb{R}^{p,q}$.*

If a is an invertible vector ($\mathcal{B}(a, a) \neq 0$) then the subspace $(\mathbb{R}a)^\perp$ has dimension $n-1$ and it is called the hyperplane associated with a. In this case, every v in $\mathcal{X}$ has a unique representation of the form $v = \lambda a + b$, with $b \in (\mathbb{R}a)^\perp$ and $\lambda \in \mathbb{R}$. The next proposition shows that the linear transformation $\varphi_a : \mathcal{X} \to \mathcal{X}$, defined by $\varphi_a(v) = -\lambda a + b$ is orthogonal. It is called the reflection with respect to the hyperplane $(\mathbb{R}a)^\perp$. For the sake of convenience we denote $(\mathbb{R}a)^\perp$ by H_a.

Proposition 2.7 *Let $\mathcal{W}$ be a non-degenerate subspace of $\mathcal{X}$. Define the linear map $T : \mathcal{X} = \mathcal{W} \oplus \mathcal{W}^\perp \to \mathcal{X}$ by $T(v) = x - y$, where $v = x + y$,, with x in $\mathcal{W}$ and y in $\mathcal{W}^\perp$. Then T is an orthogonal transformation.*

Lemma 2.8 *Let a, b be invertible vectors such that $\mathcal{B}(a, a) = \mathcal{B}(b, b)$. Then there exists a linear map φ such that $\varphi(a) = b$ and φ is either the reflection with respect to a hyperplane or the composition of two reflections with respect to hyperplanes.*

Proof: If $\mathcal{B}(a,a) = \mathcal{B}(b,b)$ and a, b are invertible then $a+b$ and $a-b$ are orthogonal. This follows from $\mathcal{B}(a+b, a-b) = \mathcal{B}(a,a) - \mathcal{B}(b,b)$.

We have $\mathcal{B}(a+b, a+b) + \mathcal{B}(a-b, a-b) = 4\mathcal{B}(a,a) \neq 0$. We deduce that either $a+b$ is invertible, or $a-b$ is invertible, because at least one of the summands in the above equation must be non-zero.

If $a-b$ is invertible then $\varphi_{a-b} : \mathcal{X} \to \mathcal{X}$ is a reflection and

$$\varphi_{a-b}(a) = \varphi_{a-b}\left(\frac{1}{2}(a-b) + \frac{1}{2}(a+b)\right) = -\frac{1}{2}(a-b) + \frac{1}{2}(a+b) = b.$$

If $a+b$ is invertible, then $\varphi_{a+b} : \mathcal{X} \to \mathcal{X}$ is a reflection and

$$\begin{aligned} \varphi_b \varphi_{a+b}(a) &= \varphi_b \varphi_{a+b}\left(\frac{a-b}{2} + \frac{a+b}{2}\right) \\ &= \varphi_b\left(\frac{a-b}{2} - \frac{a+b}{2}\right) \\ &= \varphi_b(-b) = b. \end{aligned}$$

■

The Cartan-Dieudonné theorem states that every orthogonal transformation T on an n-dimensional generalized scalar product space is the composition of at most n reflections with respect to hyperplanes. In the following section we will present our proof.

The main difficulty to obtain the proof of the Cartan-Dieudonné theorem appears in the case when $T(x) - x$ is a nonzero isotropic vector for every nonisotropic vector x. This case leads us to consider Artinian spaces. We present next some basic properties of Artinian spaces that we will use in the next section.

2.2 Artinian spaces

The simplest example of an Artinian space is the Lorentz plane $\mathbb{R}^{1,1}$. In this plane the subspace generated by $u = (1,1)$ is a null space of dimension 1.

An Artinian space is a generalized scalar product space of the form $\mathbb{R}^{p,p}$ for some positive integer p. Every Artinian space $\mathcal{X}$ has the following properties.

(1) $\dim(\mathcal{X}) = 2p$ is even.
(2) $\mathcal{X}$ contains a null subspace of dimension p.

(3) If U is a maximal null subspace of $\mathcal{X}$ then
   (a) $\dim(U) = p$.
   (b) If T is an element of $\mathcal{O}(\mathcal{X})$ such that $T(U) = U$ then T is a rotation, that is, $\det(T) = 1$.

For our purposes, the main property of Artinian spaces is the following lemma [2, Prop. 247.1, Lemma 249.2].

Lemma 2.9 *Let T be an element of $\mathcal{O}(p,q)$ such that $T(x) - x$ is a nonzero isotropic vector for every nonisotropic vector x. Then*

(1) $p = q$ and $2p$ is a multiple of 4.
(2) T is a rotation with fixed space U, where U is a maximal null subspace and $\quad Im(T - I) = U = Ker\,(T - I)$.

This result, combined with induction over the dimension of the space, is used in the proof of the Cartan-Dieudonné theorem presented in [2, Thm. 254.1]. In the next section we present an alternative proof based on an algorithm to decompose a given orthogonal transformation as product of reflections. In section 4.2 we present an explict formula for calculating the matrix representations of the reflections.

3 An alternative proof of the Cartan-Dieudonné theorem

In order to simplify the notation, from here on we write $uv = \mathcal{B}(u, v)$, whenever $\mathcal{B}$ is the symmetric non-degenerate bilinear form on the space $\mathbb{R}^{p,q}$, and u and v are in $\mathbb{R}^{p,q}$. The subspace generated by the vectors $u_1, u_2, \ldots, u_k$ is denoted by $\langle u_1, u_2, \ldots, u_k \rangle$. Recall that, if a is a vector such that $a^2 \neq 0$ then φ_a denotes the reflection with respect to the hyperplane H_a of all vectors that are orthogonal to a.

The next lemma may be considered as a weak version of the Cartan-Dieudonné theorem. The proof is very similar to Uhlig's proof of the analogous result [3, Thm. 3].

Lemma 3.1 *Every orthogonal transformation on the space $\mathbb{R}^{p,q}$ can be expressed as the composition of at most $2n$ reflections with respect to hyperplanes, where $n = p + q$.*

Proof: Let $T \in \mathcal{O}(p,q)$. Let $\{w_1, w_2, \ldots, w_n\}$ be an orthogonal basis for $\mathbb{R}^{p,q}$ such that $w_i^2 \neq 0$ for $i = 1, 2, \ldots, n$. Define $V_j = \langle w_j, w_{j+1}, \ldots, w_n \rangle$ for $j = 1, 2, \ldots, n$.

Consider the vectors $T(w_1)$ and w_1. Define the linear function $\varphi_1 : \mathbb{R}^{p,q} \to \mathbb{R}^{p,q}$

by

$$\varphi_1 = \begin{cases} I, & \text{if } T(w_1) = w_1, \\ \varphi_{c_1}, & \text{if } T(w_1) \neq w_1 \text{ and } (T(w_1) - w_1)^2 \neq 0, \\ \varphi_{w_1}\varphi_{d_1}, & \text{if } T(w_1) \neq w_1 \text{ and } (T(w_1) - w_1)^2 = 0, \end{cases}$$

where $c_1 = T(w_1) - w_1$ and $d_1 = T(w_1) + w_1$. By Lemma 2.8 it is easy to see that $\varphi_1 T(w_1) = w_1$, and $\varphi_1 T(V_2) \subseteq V_2$.

Consider now the vectors $\varphi_1 T(w_2)$ and w_2. Define the linear function $\varphi_2 : \mathbb{R}^{p,q} \to \mathbb{R}^{p,q}$ by

$$\varphi_2 = \begin{cases} I, & \text{if } \varphi_1 T(w_2) = w_2, \\ \varphi_{c_2}, & \text{if } \varphi_1 T(w_2) \neq w_2 \text{ and } (\varphi_1 T(w_2) - w_2)^2 \neq 0, \\ \varphi_{w_2}\varphi_{d_2}, & \text{if } \varphi_1 T(w_2) \neq w_2 \text{ and } (\varphi_1 T(w_2) - w_2)^2 = 0, \end{cases}$$

where $c_2 = \varphi_1 T(w_2) - w_2$ and $d_2 = \varphi_1 T(w_2) + w_2$. We know that $\varphi_2\varphi_1 T(w_2) = w_2$, and since $c_2, d_2 \in V_2$, we see that $\varphi_2\varphi_1 T(w_1) = w_1$. Therefore we have $\varphi_2\varphi_1 T(w_i) = w_i$ for $i = 1, 2$, and $\varphi_2\varphi_1 T(V_3) \subseteq V_3$.

Consider now $\varphi_2\varphi_1 T(w_3)$ and w_3. Define the linear function $\varphi_3 : \mathbb{R}^{p,q} \to \mathbb{R}^{p,q}$ by

$$\varphi_3 = \begin{cases} I, & \text{if } \varphi_2\varphi_1 T(w_3) = w_3, \\ \varphi_{c_3}, & \text{if } \varphi_2\varphi_1 T(w_3) \neq w_3 \text{ and } (\varphi_2\varphi_1 T(w_3) - w_3)^2 \neq 0, \\ \varphi_{w_3}\varphi_{d_3}, & \text{if } \varphi_2\varphi_1 T(w_3) \neq w_3 \text{ and } (\varphi_2\varphi_1 T(w_3) - w_3)^2 = 0, \end{cases}$$

where $c_3 = \varphi_2\varphi_1 T(w_3) - w_3$ and $d_3 = \varphi_2\varphi_1 T(w_3) + w_3$. We know that $\varphi_3\varphi_2\varphi_1 T(w_3) = w_3$, and since $c_3, d_3 \in V_3$, we can show that $\varphi_3\varphi_2\varphi_1 T(w_i) = w_i$ for $i = 1, 2, 3$. Therefore we have $\varphi_3\varphi_2\varphi_1 T(w_i) = w_i$ for $i = 1, 2, 3$, and $\varphi_3\varphi_2\varphi_1 T(V_4) \subseteq V_4$.

We introduce the notation $\Phi_k = \varphi_k\varphi_{k-1}\cdots\varphi_1$ for $k \geq 1$. Then, following the procedure used above, we can get orthogonal transformations $\varphi_1, \varphi_2, \ldots, \varphi_n$ such that $\Phi_k T(w_i) = w_i$ for $i = 1, 2, \ldots, k$, and $\Phi_k T(V_{k+1}) \subseteq V_{k+1}$.

The orthogonal transformations φ_j are defined by

$$\varphi_{k+1} = \begin{cases} I, & \text{if } \Phi_k T(w_{k+1}) = w_{k+1}. \\ \varphi_{c_{k+1}}, & \text{if } \Phi_k T(w_{k+1}) \neq w_{k+1} \text{ and } (\Phi_k T(w_{k+1}) - w_{k+1})^2 \neq 0, \\ \varphi_{w_{k+1}}\varphi_{d_{k+1}}, & \text{if } \Phi_k T(w_{k+1}) \neq w_{k+1} \text{ and } (\Phi_k T(w_{k+1}) - w_{k+1})^2 = 0, \end{cases}$$

where $c_{k+1} = \Phi_k T(w_{k+1}) - w_{k+1}$, and $d_{k+1} = \Phi_k T(w_{k+1}) + w_{k+1}$.

Therefore we have $\Phi_n T(w_i) = w_i$ for $i = 1, 2, \ldots, n$, and thus $\Phi_n T = I$, and then

$$T = \Phi_n^{-1} = \varphi_1^{-1}\varphi_2^{-1}\cdots\varphi_n^{-1}.$$

Since each φ_k^{-1} is either the identity, a reflection with respect to a hyperplane, or the composition of two reflections with respect to hyperplanes, we see that T is the composition of at most $2n$ reflections with respect to hyperplanes and this completes the proof. ■

Reviewing the ideas used in the proof of the previous lemma, we see that one way to reduce the number of reflections needed to factor T is the following. Suppose we have found $\varphi_1, \varphi_2, \ldots, \varphi_k$, reflections with respect to hyperplanes, such that $\Phi_k T(w_i) = w_i$ for $i = 1, 2, \ldots, \ell$, and $\Phi_k T(V_{\ell+1}) \subseteq V_{\ell+1}$. If there exists $w_j \in \{w_{\ell+1}, w_{\ell+2}, \ldots, w_n\}$ such that $\Phi_k T(w_j) = w_j$, or $\Phi_k T(w_j) \neq w_j$ and $(\Phi_k T(w_j) - w_j)^2 \neq 0$, then we can reorder the elements to force $j = \ell+1$, and then, using the construction of the φ_i of the previous lemma we would have that φ_{k+1} is either the identity or a reflection with respect to a hyperplane.

Lemma 2.9 tells us under what conditions it can happen that no element of $\{w_{\ell+1}, w_{\ell+2}, \ldots, w_n\}$ satisfies the conditions described above, and in such situation we can not assure that φ_{k+1} is the identity or a reflection with respect to a hyperplane.

In order to get a proof of the Cartan-Dieudonné theorem, we must find an algorithm for the construction of the orthogonal transformations φ_j that avoids in some way reaching a situation where the hypothesis of Lemma 2.9 is satisfied. This can be done by introducing an additional reflection in the way we describe next.

Let $T \in \mathcal{O}(p, q)$ and suppose that $T(x) - x$ is a nonzero isotropic vector for every nonisotropic x. Then, by Lemma 2.9 we must have $p = q$, $n = p + q$ is a multiple of 4, and T is a rotation, that is, $\det(T) = 1$.

Let y be an invertible element of $\mathbb{R}^{p,q}$ and let φ_y be the reflection with respect to the hyperplane H_y. Define $S = \varphi_y T$. Since $\det(S) = -1$ we see that $S \neq I$ and S is not a rotation. Therefore S does not satisfy the hypothesis of Lemma 2.9.

Let $\{w_1, w_2, \ldots, w_n\}$ be an orthogonal basis for $\mathbb{R}^{p,q}$. We can reorder the elements of the basis so that either $S(w_1) = w_1$ or $S(w_1) - w_1$ is an invertible vector. Then we can find φ_1 that is either the identity or a reflection with respect to a hyperplane, and satisfies $\varphi_1 S(w_1) = w_1$ and $\varphi_1 S(V_2) \subseteq V_2$.

Since the dimension of V_2 is not a multiple of 4, the orthogonal transformation $\varphi_1 S$ restricted to V_2 does not satisfy the hypothesis of Lemma 2.9. Therefore there exists j such that $2 \leq j \leq n$ and either $\varphi_1 S(w_j) = w_j$ or $\varphi_1 S(w_j) - w_j$ is an invertible vector. Reordering the basis of V_2 if necessary, we can suppose that $j = 2$. Then we can find φ_2 that is either the identity or a reflection with respect to a hyperplane and satisfies $\varphi_2 \varphi_1 S(w_i) = w_i$ for $i = 1, 2$ and $\varphi_2 \varphi_1 S(V_3) \subseteq V_3$.

Proceeding in the same way we can get φ_3 and φ_4, that are either the identity or reflections, that satisfy $\varphi_4\varphi_3\varphi_2\varphi_1 S\,(w_i) = w_i$ for $i = 1, 2, 3, 4$, and $\varphi_4\varphi_3\varphi_2\varphi_1 S\,(V_5) \subseteq V_5$.

Consider now the composition $\Phi_4 S$, where $\Phi_4 = \varphi_4\varphi_3\varphi_2\varphi_1$. (Recall that we defined $\Phi_k = \varphi_k\varphi_{k-1}\cdots\varphi_2\varphi_1$ for $k \geq 1$). There are two possible cases:

(1) $\varphi_i \neq I$ for $i = 1, 2, 3, 4$.

In this case we have $\det(\Phi_4 S) = -1$ and hence $\Phi_4 S$ restricted to V_5 does not satisfy the hypothesis of Lemma 2.9. Therefore we can find orthogonal transformations φ_j, for $j = 5, 6, 7, 8$ such that $\Phi_8 S\,(w_i) = w_i$ for $i = 1, 2, \ldots, 8$ and $\Phi_8 S\,(V_9) \subseteq V_9$. Then, for the composition $\Phi_8 S$ we have again the same two possible cases that we had for $\Phi_4 S$, but now considering the maps φ_i for $i = 5, 6, 7, 8$.

(2) $\varphi_i = I$ for at least one i, with $1 \leq i \leq 4$.

In this case $\dim(V_5)$ is a multiple of 4 and it is possible that $\det(\Phi_4 S) = 1$.

If $\Phi_4 S$ restricted to the space V_5 does not satisfy the hypothesis of Lemma 2.9 then we can find φ_5, which is either the identity or a reflection such that $\Phi_5 S\,(w_i) = w_i$ for $i = 1, 2, \ldots, 5$ and $\Phi_5 S\,(V_6) \subseteq V_6$. Notice that, since at least one of the φ_i is the identity, for $1 \leq i \leq 4$, the number of reflections in the composition Φ_5 is at most equal to 4.

If $\Phi_4 S$ restricted to the space V_5 satisfies the hypothesis of Lemma 2.9 then we take an invertible vector z in V_5 and form the composition $\varphi_z\Phi_4 S$. This map restricted to V_5 can not satisfy the hypothesis of Lemma 2.9, and consequently, we can find φ_5, which is either the identity or a reflection and satisfies $\varphi_5\varphi_z\Phi_4 S\,(w_i) = w_i$ for $i = 1, 2, \ldots, 5$ and $\varphi_5\varphi_z\Phi_4 S\,(V_6) \subseteq V_6$. Since at least one of the φ_i is the identity, for $1 \leq i \leq 4$, the number of reflections in the composition $\varphi_5\varphi_z\Phi_4$ is at most equal to 5.

Applying the procedure described above we see that for each ℓ such that $1 \leq \ell \leq n$, we can find reflections with respect to hyperplanes $\varphi_1, \varphi_2, \ldots, \varphi_s$ such that $\Phi_s S\,(v_i) = v_i$ for $1 \leq i \leq \ell$, where $\{v_1, \ldots, v_n\}$ is a reordering of the orthogonal basis $\{w_1, \ldots, w_n\}$, and $s \leq \ell$. This last inequality is very important.

In particular, for $\ell = n$ we get $\Phi_s S = I$, with $s \leq n$. We claim that the case $s = n$ is not possible. If $s = n$ then $\det(\Phi_s) = (-1)^n = 1$, because n is a multiple of 4. On the other hand, $\det(\Phi_s)\det(S) = \det(I) = 1$. But we know that $\det(S) = -1$. Therefore $s = n$ is not possible and we conclude that $s < n$.

Since $S = \varphi_y T$, we have $\Phi_s\varphi_y T = I$ and therefore T is the composition of at most n reflections with respect to hyperplanes.

We have proved the following result.

Lemma 3.2 *Let T be an element of $\mathcal{O}(p,q)$. If $T(x)-x$ is a nonzero isotropic vector for every nonisotropic vector x then T is the composition of at most $p+q$ reflections with respect to hyperplanes.*

Lemma 3.3 *Let T be an element of $\mathcal{O}(p,q)$. If there exists a basis $\{w_1,\ldots,w_{p+q}\}$, where all the elements are nonisotropic, such that*

$$T(w_i)-w_i$$

is a nonzero isotropic vector for $i=1,\ldots,p+q$, then T is the composition of at most $p+q$ reflections with respect to hyperplanes.

Proof: We can proceed as follows.

Step 1 In each step we have to deal with three possible cases. In particular, here we have:

(1) There exists a nonzero and nonisotropic element $v_1 \in \mathbb{R}^{p,q}$ such that $T(v_1)-v_1=0$.

(2) There exists a nonzero and nonisotropic element v_1 of the basis such that $T(v_1)-v_1 \neq 0$ and $(T(v_1)-v_1)^2 \neq 0$ (*i.e.* $T(v_1)-v_1$ is nonisotropic).

(3) For every nonzero and nonisotropic $x \in \mathbb{R}^{p,q}$ we have that $T(x)-x \neq 0$ is isotropic.

In case 3), from Lemma 3.2 we obtain that T is the composition of at most $p+q$ reflections with respect to hyperplanes.

In cases 1) or 2), we can find φ_1, which is either the identity or a reflection, that satisfies

$$\varphi_1 T(v_1)=v_1$$

and also $\varphi_1 T(W_1)=W_1$, where $W_1=\langle w_1\rangle^{\perp}$.

Step 2 Now, for the orthogonal transformation $\varphi_1 T$ restricted to the space W_1, we have

(1) There exists a nonzero and nonisotropic element $v_2 \in W_1$ such that $\varphi_1 T(v_2)-v_2=0$.

(2) There exists a nonzero and nonisotropic element $v_2 \in W_1$ such that $\varphi_1 T(v_2)-v_2 \neq 0$ is nonisotropic.

(3) For each nonzero and nonisotropic $x \in \mathbb{R}^{p,q}$ we have that $\varphi_1 T(x)-x \neq 0$ is isotropic.

In case 3), we have that $\varphi_1 T=S$ and thus, from Lemma 3.2, S is the composition of at most $p+q-1$ reflections with respect to hyperplanes.

In cases 1) or 2), we can find φ_2, which is either the identity or a reflection, that satisfies

$$\varphi_2\varphi_1 T(v_i)=v_i,$$

where i can be chosen as 1 or 2. Also, it is fulfilled that $\varphi_2\varphi_1 T(W_2)=W_2$, where $W_2=\langle w_1,w_2\rangle^{\perp}$.

By following these steps, we can end up with one of the two following situations:

A We can find an orthogonal set of nonisotropic elements $\{v_1, v_2, \ldots, v_{p+q}\}$ and a finite sequence of linear transformations $\varphi_1, \varphi_2, \ldots, \varphi_{p+q}$, such that:

- φ_i is either a reflection or the identity, for $i = 1, 2, \ldots, p+q$.
- $\varphi_l \varphi_{l-1} \cdots \varphi_2 \varphi_1 T(v_i) = v_i$, for $i = 1, 2, \ldots, l$ and $l = 1, 2, \ldots p+q$.
- $\varphi_l \varphi_{l-1} \cdots \varphi_2 \varphi_1 T(W_i) = W_i$, for $l = 1, 2, \ldots p+q$, where $W_i = \langle v_1, v_2, \ldots, v_i \rangle^\perp$ for $i = 1, 2, \ldots p+q$.

B We can find an orthogonal set of nonisotropic elements $\{v_1, v_2, \ldots, v_k\}$ and a finite sequence of linear transformations $\varphi_1, \varphi_2, \ldots, \varphi_k$, where $k < p+q$, such that:

- φ_i is either a reflection or the identity, for $i = 1, 2, \ldots, k$.
- $\varphi_l \varphi_{l-1} \cdots \varphi_2 \varphi_1 T(v_i) = v_i$, for $i = 1, 2, \ldots, l$ and $l = 1, 2, \ldots k$.
- $\varphi_l \varphi_{l-1} \cdots \varphi_2 \varphi_1 T(W_i) = W_i$, for $l = 1, 2, \ldots k$, where $W_i = \langle v_1, v_2, \ldots, v_i \rangle^\perp$ for $i = 1, 2, \ldots k$.
- Consider the orthogonal transformation $\varphi_k \varphi_{l-1} \cdots \varphi_2 \varphi_1 T$, restricted to W_k. Then we have that for every nonzero and nonisotropic $x \in W_k$ it is fulfilled that $\varphi_k \varphi_{l-1} \cdots \varphi_2 \varphi_1 T(x) - x$ is nonzero and isotropic.

In case **A**, we have that

$$\varphi_{p+q} \varphi_{p+q-1} \cdots \varphi_2 \varphi_1 T = I.$$

Thus T is the composition of at most $p+q$ reflections.

In case **B**, we have that

$$\varphi_k \varphi_{l-1} \cdots \varphi_2 \varphi_1 T = S,$$

restricted to W_k (whose dimension is $p+q-k$). From Lemma 3.2 we conclude that S is the composition of at most $p+q-k$ reflections.

In summary, T is the composition of at most $p+q$ reflections. ■

Now we are ready to prove the Cartan-Dieudonné theorem.

Theorem 3.4 *Let T be an element of $\mathcal{O}(p,q)$. Then T is the composition of at most $p+q$ reflections with respect to hyperplanes.*

Proof: Let $n = p+q$ and let $\{w_1, w_2, \ldots, w_n\}$ be an orthogonal basis for $\mathbb{R}^{p,q}$ with $w_i^2 \neq 0$ for $1 \leq i \leq n$.

If there exists an element w_j of the basis such that $T(w_j) = w_j$ or $T(w_j) \neq w_j$ and $(T(w_j) - w_j)^2 \neq 0$, then, reordering the basis (keeping the notation w_i for the basic elements), we can get $j = 1$. Thus we can find φ_1, which is either the identity or a reflection, that satisfies $\varphi_1 T(w_1) = w_1$, and $\varphi_1 T(V_2) \subseteq V_2$.

If there exists an element $w_j \in \{w_2, w_3, \ldots, w_n\}$ such that $\varphi_1 T(w_j) = w_j$ or $\varphi_1 T(w_j) \neq w_j$ and $(\varphi_1 T(w_j) - w_j)^2 \neq 0$, reordering the basis $\{w_2, w_3, \ldots, w_n\}$

we can assume that $j = 2$, and then we can find φ_2, that is either the identity or a reflection with respect to a hyperplane, that satisfies $\varphi_2\varphi_1 T(w_i) = w_i$ for $i = 1, 2$, and $\varphi_2\varphi_1 T(V_3) \subseteq V_3$.

This process can be repeated until we reach either of two possible cases. In the first case we can find φ_i, for $1 \leq i \leq n$, such that $\Phi_n T = I$, and therefore T is the composition of at most n reflections with respect to hyperplanes.

In the second case we find φ_i, for $1 \leq i \leq k$, where $1 \leq k < n$, that satisfy $\Phi_k T(w_i) = w_i$ for $1 \leq i \leq k$, $\Phi_k T(V_{k+1}) \subseteq V_{k+1}$, and $\Phi_k T(w_j) - w_j$ is a nonzero isotropic vector for every para $k < j \leq n$. Applying Lemma 3.3 to the orthogonal transformation $\Phi_k T$, on the space V_{k+1}, we obtain reflections with respect to hyperplanes τ_i, for $1 \leq i \leq s$, with $s \leq n - k$, such that $\tau_s \tau_{s-1} \cdots \tau_1 \Phi_k T$ is the identity on $\mathbb{R}^{p,q}$. Since $s + k \leq n$, we conclude that, also in this case, T is the composition of at most n reflections with respect to hyperplanes. ■

4 Clifford algebras

In this section we present some definitions and basic results about Clifford algebras associated with a generalized scalar product space $(\mathcal{X}, \mathcal{B})$ of signature (p, q). We use the notation of [7, Ch. I]. We also describe how the Clifford algebra structure can be used to deal with orthogonal transformations, and in particular, with reflections with respect to hyperplanes.

Definition 4.1 *Let $(\mathcal{X}, \mathcal{B})$ be a generalized scalar product space of dimension n and let $\mathcal{A}$ be a real associative algebra with identity 1 such that*

(C1) $\mathcal{A}$ contains copies of $\mathbb{R}$ and of $\mathcal{X}$ as linear subspaces.

(C2) For all $v \in \mathcal{X}$ we have $v^2 = \mathcal{B}(v, v)$.

(C3) $\mathcal{A}$ is generated as a ring by the copies of $\mathbb{R}$ and of $\mathcal{X}$, or equivalently, as a real algebra by $\{1\}$ and $\mathcal{X}$.

Then $\mathcal{A}$ is called a real Clifford algebra for $(\mathcal{X}, \mathcal{B})$ and it is denoted by $\mathcal{A} = \mathcal{C}(\mathcal{X})$.

Note that (C2) links the multiplication in the algebra with the bilinear form on $\mathcal{X}$.

Here we describe a basis for $\mathcal{C}(\mathcal{X})$ in terms of an orthonormal basis for $\mathcal{X}$. Let $(\mathcal{X},\mathcal{B})$ be a generalized scalar product space of signature (p,q) and let $e=\{e_1,e_2,\ldots,e_n\}$ be an orthonormal basis for $\mathcal{X}$. Then, by (C2) we have

$$e_i^2=\mathcal{B}(e_i,e_i)=\begin{cases}1, & \text{for } i=1,2,\ldots,p\\ -1, & \text{for } i=p+1,p+2,\ldots,p+q,\end{cases}$$

and it is easy to show that

$$e_ie_j+e_je_i=0, \qquad i\neq j,$$

and

$$\frac{1}{2}(vw+wv)=\mathcal{B}(v,w), \qquad u,v\in\mathcal{X}.$$

Define $N=\{1,2,\ldots,n\}$. Let $\beta_1,\beta_2,\ldots,\beta_s$ be distinct elements of N. Then

$$e_{\beta_1}e_{\beta_2}\cdots e_{\beta_s}=(-1)^{\sigma}e_{\alpha_1}e_{\alpha_2}\cdots e_{\alpha_s},$$

where $(\alpha_1,\alpha_2,\ldots,\alpha_s)$ is the permutation of $(\beta_1,\beta_2,\ldots,\beta_s)$ that satisfies $\alpha_1<\alpha_2<\cdots<\alpha_n$, and σ is the number of transpositions of the permutation that sends $(\alpha_1,\alpha_2,\ldots,\alpha_s)$ to $(\beta_1,\beta_2,\ldots,\beta_s)$.

Since $e_i^2=\pm1$, it is easy to see that, up to a change of sign, every product of basic elements, possibly with repeated factors, can be reduced to a product of at most n factors with indices in increasing order. This implies that every element of $\mathcal{C}(\mathcal{X})$ can be written in the form

$$\sum_A \lambda_A\, e_{\alpha_1}e_{\alpha_2}\cdots e_{\alpha_s},$$

where the sum runs over all the subsets $A=\{\alpha_1,\alpha_2,\ldots,\alpha_n\}$ of N (with $\alpha_1<\alpha_2<\cdots<\alpha_n$) and λ_A is a real coefficient. In order to simplify the notation we define $e_A=e_{\alpha_1}e_{\alpha_2}\cdots e_{\alpha_s}$, and we put $e_\emptyset=1$. Therefore the collection of elements of the form e_A, where A is a subset of N, is a generating set for $\mathcal{C}(\mathcal{X})$ and consequently, $\dim(\mathcal{C}(\mathcal{X}))\leq 2^n$.

Recall that a vector $s\in\mathcal{X}$ is called invertible if $\mathcal{B}(s,s)\neq 0$. This condition is equivalent to $s^2=\mathcal{B}(s,s)\neq 0$. Therefore, the element $(1/\mathcal{B}(s,s))\,s$ is the inverse of s with respect to the multiplication in the algebra $\mathcal{C}(\mathcal{X})$, and we can write $s^{-1}=s/s^2=s/\mathcal{B}(s,s)$.

Theorem 4.2 *Let $\mathcal{C}(\mathcal{X})$ be a Clifford algebra for an n-dimensional generalized scalar product space $(\mathcal{X},\mathcal{B})$ with signature (p,q) and let*

$e=\{e_1,\ldots,e_p,e_{p+1},\ldots,e_{p+q}\}$ *be an orthonormal basis for $(\mathcal{X},\mathcal{B})$. Then* *(i) If n is even then* $\dim \mathcal{C}(\mathcal{X}) = 2^n$ *and* $\{e_A : A \subseteq N\}$ *is a basis for* $\mathcal{C}(\mathcal{X})$*.*

(ii) If n is odd and $e_N \notin \mathbb{R}$, then $\dim \mathcal{C}(\mathcal{X}) = 2^n$ *and* $\{e_A : A \subseteq N\}$ *is a basis for* $\mathcal{C}(\mathcal{X})$*.*

(iii) If n is odd and $e_N \in \mathbb{R}$, then $e_N = \pm 1$ and $p - q \equiv 1 \mod (4)$. In this case $\dim \mathcal{C}(\mathcal{X}) = 2^{n-1}$ *and* $\{e_A : A \subseteq N,\ \#A \text{ even }\}$ *is a basis for* $\mathcal{C}(\mathcal{X})$*.*

Theorem 4.3 *Let $\mathcal{C}(\mathcal{X})$ be a Clifford algebra for the generalized scalar product space $(\mathcal{X}, \mathcal{B})$ and let W be a non-degenerate subspace of $\mathcal{X}$. Then the subalgebra of $\mathcal{C}(\mathcal{X})$ generated by W is a Clifford algebra for W.*

The proof of these results can be found in [7], [8], and [9].

A result of Chevalley states that every n-dimensional generalized scalar product space $(\mathcal{X}, \mathcal{B})$ has a Clifford algebra of dimension 2^n (see [7]). If $(\mathcal{X}, \mathcal{B})$ has signature (p, q) then it is isomorphic to $\mathbb{R}^{p,q}$. The Clifford algebra of dimension 2^n of $\mathbb{R}^{p,q}$ is denoted by $\mathbb{R}_{p,q}$.

The algebra $\mathbb{R}_{p,q}$ is a graded vector space. Let s be an integer such that $0 \leq s \leq n$. Let $\mathbb{R}^s_{p,q}$ denote the subspace generated by $\{e_A : A \subseteq N,\ \#A = s\}$. It has dimension $\binom{n}{s}$. The elements of $\mathbb{R}^s_{p,q}$ are called s-vectors. Note that

$$\mathbb{R}_{p,q} = \bigoplus_{s=1}^{n} \mathbb{R}^s_{p,q}.$$

The space of 0-vectors is generated by $\{1\}$, and its elements are also called scalars. The space $\mathbb{R}^1_{p,q}$ can be identified with $\mathbb{R}^{p,q}$. The space $\mathbb{R}^n_{p,q}$ is generated by $\{e_N\}$. The element e_N is called pseudoscalar.

Every element $a \in \mathbb{R}_{p,q}$ can be expressed in the form

$$a = \sum_{A \subseteq N} \lambda_A e_A, \qquad \lambda_A \in \mathbb{R}.$$

and also as

$$a = \sum_{r=0}^{n} a_r, \qquad a_r \in \mathbb{R}^r_{p,q}.$$

Definition 4.4 *Let $a = \sum_{r=0}^{n} a_r$, where $a_r \in \mathbb{R}^r_{p,q}$. If $a_r = 0$ for $r > t$ and $a_t \neq 0$, then we say that the grade of the multivector a is t, and write $gr(a) = t$.*

Notice that if $\lambda \in \mathbb{R}$ then $\mathrm{gr}(\lambda a) = \mathrm{gr}(a)$. If $s_1, \ldots, s_k$ are non-zero vectors in $\mathbb{R}^{p,q}$, then $\mathrm{gr}(s_1 \cdots s_k) = l \leq k$ and $\dim\langle s_1, \ldots, s_k \rangle = l$.

The multiplication in the algebra $\mathbb{R}_{p,q}$ can be used to represent reflections with respect to a hyperplane in the space $\mathbb{R}^{p,q}$ as follows. Let s be an invertible vector in $\mathbb{R}^{p,q}$. Define the map $\varphi_s : \mathbb{R}^{p,q} \to \mathbb{R}^{p,q}$ by

$$\varphi_s(x) = -sxs^{-1}, \qquad x \in \mathbb{R}^{p,q}.$$

In order to show that $\varphi_s(x)$ is a vector in $\mathbb{R}^{p,q}$ for all $x \in \mathbb{R}^{p,q}$ we first compute sxs and get

$$\begin{aligned} sxs &= \frac{1}{2}(sx + xs + (sx - xs))\, s \\ &= \frac{1}{2}(2\mathcal{B}(x,s) + sx - xs)\, s \\ &= \mathcal{B}(x,s)\, s + \frac{1}{2} sxs - \frac{1}{2} xs^2. \end{aligned}$$

The last equation yields $sxs = 2\mathcal{B}(x,s)\, s - xs^2$ and hence

$$-sxs^{-1} = -\frac{2\mathcal{B}(x,s)}{s^2} s + x, \tag{1}$$

which is clearly a vector in $\mathbb{R}^{p,q}$.

Lemma 4.5 *Let s be an invertible vector in $\mathbb{R}^{p,q}$. Then the linear map $\varphi_s : \mathbb{R}^{p,q} \to \mathbb{R}^{p,q}$ is an orthogonal transformation. Furthermore, it is the reflection with respect to the hyperplane $H_s = \{x \in \mathbb{R}^{p,q} : \mathcal{B}(x,s) = 0\}$.*

Proof: For $x, y \in \mathbb{R}^{p,q}$ we have

$$\begin{aligned} \mathcal{B}(\varphi_s(x), \varphi_s(y)) &= \mathcal{B}\left(-sxs^{-1}, -sys^{-1}\right) \\ &= \frac{1}{2}\left(\left(-sxs^{-1}\right)\left(-sys^{-1}\right) + \left(-sys^{-1}\right)\left(-sxs^{-1}\right)\right) \\ &= \frac{1}{2} s(xy + yx)\, s^{-1} \\ &= \mathcal{B}(x,y). \end{aligned}$$

If $x \in H_s$ we have $\varphi_s(x) = -sxs^{-1} = -\left(2\mathcal{B}(x,s)/s^2\right) s + x = x$.

If $x = \lambda s$ where λ is a real number then $\varphi_s(x) = -s(\lambda s)\, s^{-1} = -\lambda s = -x$. Therefore φ_s is the reflection with respect to H_s. ■

It is easy to verify that φ_s satisfies

(1) $\varphi_s = \varphi_{\lambda s}$ for every nonzero real λ.
(2) $\varphi_s^{-1} = \varphi_s$.
(3) If $s_1, s_2, \ldots, s_k$ are invertible elements of $\mathbb{R}^{p,q}$ then

$$\varphi_{s_1}\varphi_{s_2}\cdots\varphi_{s_k}(x) = (-1)^k s_1 s_2 \cdots s_{k-1} s_k x s_k^{-1} s_{k-1}^{-1} \cdots s_2^{-1} s_1^{-1}.$$

Using the Clifford algebra multiplication the Cartan-Dieudonné theorem reads as follows.

Theorem 4.6 *Let T be an orthogonal transformation on $\mathbb{R}^{p,q}$. Then there exist invertible elements $s_1, s_2, \ldots, s_k$ in $\mathbb{R}^{p,q}$, with $k \leq p+q$, such that*

$$T(x) = (-1)^k s_1 s_2 \cdots s_{k-1} s_k x s_k^{-1} s_{k-1}^{-1} \cdots s_2^{-1} s_1^{-1}, \qquad x \in \mathbb{R}^{p,q}.$$

If T is an orthogonal transformation on $\mathbb{R}^{p,q}$ and there exist invertible multivectors A, B in the Clifford algebra $\mathbb{R}_{p,q}$ such that

$$T(x) = \pm A x A^{-1} = \pm B x B^{-1}, \qquad x \in \mathbb{R}^{p,q},$$

then $A = \lambda B$ for some real λ. The proof of this result can be found in [9]. We will illustrate this fact in the examples of the next section.

4.2 A matrix representation of φ_s

Having in mind that, in many applications, matrix representations of reflections are useful, we introduce another algebraic expression for φ_s.

Define $\lambda = \frac{\mathcal{B}(x,s)}{x^2}$, then

$$s = \lambda x + y,$$

where y is orthogonal to x, that is

$$xy = -yx.$$

Thus

$$\begin{aligned} -sxs^{-1} &= -\frac{1}{s^2}(\lambda x + y)\, x\, (\lambda x + y), \\ &= -\frac{1}{s^2}\left(\left(\lambda^2 x^2 - y^2\right) x + 2\lambda x^2 y\right), \\ &= -\frac{1}{s^2}\left(\left(\frac{(\mathcal{B}(s,x))^2}{x^2} - y^2\right) x + 2\mathcal{B}(s,x)\, y\right). \end{aligned} \tag{2}$$

This last result (2) turns out to be useful to find a matrix representaction of φ_s with respect to a given orthogonal basis $B = \{w_1, w_2, \ldots, w_{p+q}\}$, where $w_i^2 \neq 0$ for $i = 1, 2, \ldots, p+q$. Indeed, consider $s = \sum_{i=1}^{p+q} \alpha_i w_i$, then

$$
\begin{aligned}
s &= \alpha_k w_k + \sum_{\substack{i=1 \\ i\neq k}}^{p+q} \alpha_i w_i, \\
&= \frac{\mathcal{B}(s, w_k)}{w_k^2} w_k + \sum_{\substack{i=1 \\ i\neq k}}^{p+q} \frac{\mathcal{B}(s, w_i)}{w_i^2} w_i .
\end{aligned}
$$

Define

$$
y_k = \sum_{\substack{i=1 \\ i\neq k}}^{p+q} \alpha_i w_i = \sum_{\substack{i=1 \\ i\neq k}}^{p+q} \frac{\mathcal{B}(s, w_i)}{w_i^2} w_i .
$$

Using (2) to get

$$
\begin{aligned}
\varphi_s(w_k) &= -\frac{1}{s^2}\left(\left(\frac{(\mathcal{B}(s, w_k))^2}{w_k^2} - y_k^2\right) w_k + (2\mathcal{B}(s, w_k))\, y_k\right), \\
&= -\frac{1}{s^2}\left(\left(\frac{(\mathcal{B}(s, w_k))^2}{w_k^2} - \left(\sum_{\substack{i=1 \\ i\neq k}}^{p+q} \frac{(\mathcal{B}(s, w_i))^2}{w_i^2}\right)\right) w_k + 2\mathcal{B}(s, w_k)\, y_k\right).
\end{aligned}
$$

From this result we obtain that the k-th column of the matrix $A_s = [\varphi_s]_B$ is given by

$$
[\varphi_s(w_k)]_B = -\frac{1}{s^2}\begin{pmatrix}
2\mathcal{B}(s, w_k)\frac{\mathcal{B}(s,w_1)}{w_1^2} \\
2\mathcal{B}(s, w_k)\frac{\mathcal{B}(s,w_2)}{w_2^2} \\
\vdots \\
\frac{(\mathcal{B}(s,w_k))^2}{w_k^2} - \left(\sum_{\substack{i=1 \\ i\neq k}}^{p+q} \frac{(\mathcal{B}(s,w_i))^2}{w_i^2}\right) \leftarrow \quad (k\text{-th row}) \\
2\mathcal{B}(s, w_k)\frac{\mathcal{B}(s,w_{k+1})}{w_{k+1}^2} \\
\vdots \\
2\mathcal{B}(s, w_k)\frac{\mathcal{B}(s,w_{p+q})}{w_{p+q}^2}
\end{pmatrix}.
$$

Equivalently

$$
(A_s)_{lj} = \begin{cases}
-2\mathcal{B}(s, w_j)\frac{\mathcal{B}(s,w_l)}{s^2 w_l^2} \text{ if } l \neq j, \\
-\frac{1}{s^2}\left(\frac{(\mathcal{B}(s,w_j))^2}{w_j^2} - \left(\sum_{\substack{i=1 \\ i\neq j}}^{p+q} \frac{(\mathcal{B}(s,w_i))^2}{w_i^2}\right)\right) \text{ for } l = j.
\end{cases} \tag{3}
$$

5 Examples

Let T be the orthogonal transformation on the space $\mathbb{R}^{2,3}$ represented by the matrix

$$T_E = \begin{bmatrix} 1 & 5 & 4 & 3 & 0 \\ -5 & 1 & 3 & -4 & 0 \\ 4 & 3 & 1 & 5 & 0 \\ 3 & -4 & -5 & 1 & 0 \\ 0 & 0 & 0 & 0 & -1 \end{bmatrix}$$

with respect to the canonical basis $E = \{e_1, e_2, e_3, e_4, e_5\}$ of $\mathbb{R}^{2,3}$. Then we have $T(e_i) - e_i \neq 0$, $(T(e_i) - e_i)^2 = 0$ for $i = 1, 2, 3, 4$, and $T(e_5) - e_5 = -2e_5$. We can take $c_1 = e_5$.

It is easy to see that $\varphi_{c_1} T$ restricted to $\langle e_1, e_2, e_3, e_4 \rangle$ satisfies the conditions of Lemma 2.9. By Lemma 2.8 and the proof of Lemma 3.1 we obtain $c_2 = \varphi_{c_1} T(e_1) + e_1 = 2e_1 - 5e_2 + 4e_3 + 3e_4$ and $c_3 = e_1$. Then we have $\varphi_{c_3} \varphi_{c_2} \varphi_{c_1} T(e_i) = e_i$, for $i = 1, 5$. Notice that $\varphi_{c_2} \varphi_{c_1} T$ does not satisfy the hypothesis of Lemma 2.9 on $\langle e_1, e_2, e_3, e_4 \rangle$.

Now we have $(\varphi_{c_3} \varphi_{c_2} \varphi_{c_1} T(e_i) - e_i)^2 \neq 0$ for $i = 2, 3, 4$. We can take

$$c_4 = \varphi_{c3} \varphi_{c_2} \varphi_{c_1} T(e_2) - e_2 = \frac{25}{2} e_2 - 7e_3 - \frac{23}{2} e_4.$$

Then we have $\varphi_{c_4} \varphi_{c_3} \varphi_{c_2} \varphi_{c_1} T(e_i) = e_i$ for $i = 1, 2, 5$, and

$(\varphi_{c_4} \varphi_{c_3} \varphi_{c_2} \varphi_{c_1} T(e_i) - e_i)^2 \neq 0$ for $i = 3, 4$.

We can take $c_5 = \varphi_{c_4} \varphi_{c3} \varphi_{c_2} \varphi_{c_1} T(e_3) - e_3 = \frac{-18}{25} e_3 + \frac{24}{25} e_4$.

It is easy to verify that $\varphi_{c_5} \varphi_{c_4} \varphi_{c_3} \varphi_{c_2} \varphi_{c_1} T(e_i) = e_i$ for $i = 1, 2, 3, 4, 5$, that is $T = \varphi_{c_1} \varphi_{c_2} \varphi_{c_3} \varphi_{c_4} \varphi_{c_5}$.

Let us denote by A_j the matrix representation with respect to the canonical basis E of φ_{c_j}, for $1 \leq j \leq 5$. Using formula (3) we obtain:

$$A_5 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & \frac{7}{25} & \frac{24}{25} & 0 \\ 0 & 0 & \frac{24}{25} & -\frac{7}{25} & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \quad A_4 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & \frac{27}{2} & 7 & \frac{23}{2} & 0 \\ 0 & -7 & -\frac{73}{25} & -\frac{161}{25} & 0 \\ 0 & -\frac{23}{2} & -\frac{161}{25} & -\frac{479}{50} & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \quad A_3 = \begin{bmatrix} -1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix},$$

$$A_2 = \begin{bmatrix} -1 & 5 & 4 & 3 & 0 \\ 5 & -\frac{23}{2} & -10 & -\frac{15}{2} & 0 \\ -4 & 10 & 9 & 6 & 0 \\ -3 & \frac{15}{2} & 6 & \frac{11}{2} & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \quad A_1 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & -1 \end{bmatrix}.$$

A direct computation yields $T_E = A_1 A_2 A_3 A_4 A_5$.

Computing the Clifford product $c_1 c_2 c_3 c_4 c_5$ we obtain a linear combination of s-vectors where $s \leq 3$. This suggests that $c_1 c_2 c_3 c_4 c_5$ may be equal to the Clifford product of 3 vectors and that T could be factored as the product of 3 reflections. We present next another way to factor T that confirms this conjecture.

We apply the factorization algorithm to T, but now using the orthogonal basis $W = \{w_1, w_2, w_3, w_4, w_5\}$, where

$w_1 = e_3 + e_4 - e_5$, $w_2 = e_1 + e_2$, $w_3 = e_1 + e_4 + 2e_5$, $w_4 = e_3 - e_4$ and $w_5 = e_1 - e_2$.

We obtain $(T(w_i) - w_i)^2 \neq 0$, for $1 \leq i \leq 5$. Therefore we can take

$$\begin{aligned} d_1 &= T(w_1) - w_1, \\ \varphi_{d_1} T(w_i) - w_i &\neq 0 \text{ and } (\varphi_{d_1} T(w_i) - w_i)^2 \neq 0, \text{ for } i = 2,3,4,5, \\ d_2 &= \varphi_{d_1} T(w_2) - w_2, \\ \varphi_{d_2}\varphi_{d_1} T(w_i) - w_i &\neq 0 \text{ and } (\varphi_{d_2}\varphi_{d_1} T(w_i) - w_i)^2 \neq 0, \text{ for } i = 3,4,5, \\ d_3 &= \varphi_{d_2}\varphi_{d_1} T(w_3) - w_3, \\ 0 &= \varphi_{d_3}\varphi_{d_2}\varphi_{d_1} T(w_i) - w_i, \text{ for } i = 4,5. \end{aligned}$$

and therefore $T = \varphi_{d_1}\varphi_{d_2}\varphi_{d_3}$.

Let B_i denote the matrix representation of φ_{d_i} with respect to the canonical basis E, for $i = 1, 2, 3$. We have

$$B_1 = \begin{bmatrix} \frac{51}{2} & -\frac{7}{2} & -\frac{35}{2} & \frac{35}{2} & -7 \\ -\frac{7}{2} & \frac{3}{2} & \frac{5}{2} & -\frac{5}{2} & 1 \\ \frac{35}{2} & -\frac{5}{2} & -\frac{23}{2} & \frac{25}{2} & -5 \\ -\frac{35}{2} & \frac{5}{2} & \frac{25}{2} & -\frac{23}{2} & 5 \\ 7 & -1 & -5 & 5 & -1 \end{bmatrix}, \qquad d_1 = 7e_1 - e_2 + 5e_3 - 5e_4 + 2e_5,$$

$$B_2 = \begin{bmatrix} \frac{347}{9} & -\frac{104}{9} & -\frac{286}{9} & \frac{208}{9} & -\frac{26}{3} \\ -\frac{104}{9} & \frac{41}{9} & \frac{88}{9} & -\frac{64}{9} & \frac{8}{3} \\ \frac{286}{9} & -\frac{88}{9} & -\frac{233}{9} & \frac{176}{9} & -\frac{22}{3} \\ -\frac{208}{9} & \frac{64}{9} & \frac{176}{9} & -\frac{119}{9} & \frac{16}{3} \\ \frac{26}{3} & -\frac{8}{3} & -\frac{22}{3} & \frac{16}{3} & -1 \end{bmatrix}, \qquad d_2 = 26e_1 - 8e_2 + 22e_3 - 16e_4 + 6e_5,$$

$$B_3 = \begin{bmatrix} \frac{43}{18} & -\frac{25}{18} & -\frac{35}{18} & \frac{5}{18} & -\frac{5}{3} \\ -\frac{25}{18} & \frac{43}{18} & \frac{35}{18} & -\frac{5}{18} & \frac{5}{3} \\ \frac{35}{18} & -\frac{35}{18} & -\frac{31}{18} & \frac{7}{18} & -\frac{7}{3} \\ -\frac{5}{18} & \frac{5}{18} & \frac{7}{18} & \frac{17}{18} & \frac{1}{3} \\ \frac{5}{3} & -\frac{5}{3} & -\frac{7}{3} & \frac{1}{3} & -1 \end{bmatrix}, \qquad d_3 = -5e_1 + 5e_2 - 7e_3 + e_4 - 6e_5.$$

It is easy to verify that $T_E = B_1 B_2 B_3$, and notice that $d_1 d_2 d_3 = 6c_1c_2c_3c_4c_5$.

The matrix representation of T with respect to the basis W is

$$T_W = \begin{bmatrix} \frac{1}{3} & 2 & \frac{4}{3} & -\frac{10}{3} & \frac{8}{3} \\ 3 & 1 & 3 & 4 & -5 \\ \frac{2}{3} & 1 & -\frac{1}{3} & -\frac{5}{3} & \frac{4}{3} \\ 5 & 4 & 5 & 1 & -3 \\ 4 & 5 & 4 & -3 & 1 \end{bmatrix}.$$

Note that this second factorization avoids the Artinian case, that is, the situation where the hypothesis of Lemma 2.9 holds.

As another application of formula (3), we can find the matrix representation of each φ_{d_i}, for $i = 1, 2, 3$, but now with respect to the orthogonal basis W. Using the above values of d_1, d_2 and d_3, respectively, we obtain:

$$C_1 = \begin{bmatrix} \frac{1}{3} & -2 & \frac{4}{3} & \frac{10}{3} & -\frac{8}{3} \\ 3 & 10 & -6 & -15 & 12 \\ \frac{2}{3} & 2 & -\frac{1}{3} & -\frac{10}{3} & \frac{8}{3} \\ 5 & 15 & -10 & -24 & 20 \\ 4 & 12 & -8 & -20 & 17 \end{bmatrix},$$

$$C_2 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 10 & -9 & -19 & 17 \\ 0 & 3 & -2 & -\frac{19}{3} & \frac{17}{3} \\ 0 & 19 & -19 & -\frac{352}{9} & \frac{323}{9} \\ 0 & 17 & -17 & -\frac{323}{9} & \frac{298}{9} \end{bmatrix},$$

and

$$C_3 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & -2 & -\frac{4}{3} & \frac{5}{3} \\ 0 & 0 & -4 & -\frac{7}{9} & \frac{20}{9} \\ 0 & 0 & -5 & -\frac{20}{9} & \frac{34}{9} \end{bmatrix}.$$

A direct computation yields $T_W = C_1C_2C_3$. It should be noticed that the orthogonal basis W is not ordered according to the signature of $\mathbb{R}^{2,3}$.

6 Final remarks and conclusions

The factorization of an orthogonal transformation is not unique, as we have shown in our example. Actually, the number of reflections that factorizes an orthogonal transformation is not unique either. An interesting question related to this last point is to determine the minimum number of reflections by hyperplanes required to factorize a given orthogonal transformation. An answer to this question using matrices has been given in [10], and it was translated to the language of orthogonal transformations in [2, Pgs 260,261].

In our examples (Section 5) we use the grade of a multivector to find the minimum number of reflections by hyperplanes required to factorize T_E. We can formalize this procedure as follows:

Lemma 6.1 *Let T be an orthogonal transformation on $\mathbb{R}^{p,q}$. Assume that there exist invertible elements $s_1, \ldots, s_k$ in $\mathbb{R}^{p,q}$, with $k \leq p+q$, such that*

$$T(x) = (-1)^k s_1 s_2 \cdots s_{k-1} s_k x s_k^{-1} s_{k-1}^{-1} \cdots s_2^{-1} s_1^{-1}, \qquad x \in \mathbb{R}^{p,q}.$$

If $gr(s_1 \cdots s_{k-1} s_k) = t \leq k$, then T cannot be factored into less than t reflections with respect to hyperplanes.

Proof: Suppose that there exist $s_1', \ldots, s_l'$, with $l < t$, such that

$$T(x) = (-1)^k s_1' \cdots s_{l-1}' s_l' x (s_l')^{-1} (s_{l-1}')^{-1} \cdots (s_1')^{-1}, \qquad x \in \mathbb{R}^{p,q}.$$

We have that $\mathrm{gr}\left(s_1' \cdots s_{l-1}' s_l'\right) \leq l < t$. But $\left(s_1' \cdots s_{l-1}' s_l'\right) = \lambda (s_1 \cdots s_{k-1} s_k)$, where $\lambda \in \mathbb{R}$ and therefore $\mathrm{gr}\left(s_1' \cdots s_{l-1}' s_l'\right) = \mathrm{gr}(s_1 \cdots s_{k-1} s_k) = t$, which is a contradiction. ■

With this result, we can state the following theorem

Theorem 6.2 *Let T be an orthogonal transformation on the space $\mathbb{R}^{p,q}$. If there exist invertible elements $s_1, s_2, \ldots, s_k \in \mathbb{R}^{p,q}$, such that:*

- $T(x) = (-1)^k s_1 \cdots s_{k-1} s_k x s_k^{-1} s_{k-1}^{-1} \cdots s_2^{-1} s_1^{-1}, \qquad x \in \mathbb{R}^{p,q}$,
- *$gr(s_1, \ldots, s_{k-1}, s_k) = t$, and*
- *$Ker(T - I)$ is non-degenerate,*

then T can be factored into t reflections with respect to hyperplanes and, moreover, $dim(Ker(T-I))^{\perp} = t$.

Proof: By hypothesis $T = \varphi_{s_1} \cdots \varphi_{s_k}$ and $\mathrm{gr}(s_1, \ldots, s_{k-1}, s_k) = t$. By considering the space $V_1 = \langle s_1, \cdots, s_k \rangle$, where $\dim V_1 = t$, it is easy to show that $V_1^{\perp} \subset \mathrm{Ker}(T - I)$. If $\mathrm{Ker}(T-I)$ is non-degenerate. Then we can find an orthogonal basis $\mathcal{B} = \{w_1, \ldots, w_j, w_{j+1}, \ldots, w_{j+l}\}$, where $j + l = n$ and $\mathrm{Ker}(T - I) = \langle w_1, \ldots, w_{j-1}, w_j \rangle$ and applying the algorithm proposed in this paper we get $\varphi_i = I$, for $i = 1, 2, \ldots, j$. From the application of the Cartan-Dieudonné theorem to $T|_{V_{j+1}}$, where $V_{j+1} = \langle w_{j+1}, \ldots, w_{j+l} \rangle$, we obtain that T is the composition of at most l reflections by hyperplanes. From Lemma 6.1 we know that $t \leq l$.

Now, suppose that there exist $u_1, \ldots, u_m \in (\mathrm{Ker}(T - I))^{\perp}$, such that $T = \varphi_{u_1} \cdots \varphi_{u_m}$. If $\mathrm{gr}(u_1 \cdots u_m) < l$ then by considering $T|_{V_{j+1}}$ we can find a nonzero vector $u \in V_{j+1}$ orthogonal to the set $\{u_1, \ldots u_m\}$. But in such case $u \in \mathrm{Ker}(T - I) \cap (\mathrm{Ker}(T - I))^{\perp}$, which is a contradiction since $\mathrm{Ker}(T - I)$

is non-degenerate. Therefore T is the composition of l reflections through hyperplanes, that is, there exist $u_1, \ldots, u_l$ such that $T = \varphi_{u_1} \cdots \varphi_{u_l}$ and $\mathrm{gr}\,(u_1 \cdots u_l) = l$ and, moreover, $t = gr\,(s_1 s_2 \cdots s_k) = gr\,(u_1 u_2 \cdots u_l) = l = \dim\,(\mathrm{Ker}\,(T - I))^{\perp}$ ■

It should be remarked that the previous Theorem relates the grade of a multivector with the Cartan-Diedudonné-Scherk theorem. It remains however, to propose an algorithm capable of finding explicitly the minimum number of reflections required to decompose a given orthogonal transformation.

Acknowledgements

MAR would like to thank for financial support from COFAA-IPN. GAG was supported by the Program for the Professional Development in Automation, through the grant from the Universidad Autónoma Metropolitana and Parker Haniffin - México. JLA would like to thank DGAPA-UNAM (grant IN100310-3) and CONACyT (grant 50368 and 79641) for financial support.

References

[1] E. Artin, Geometric Algebra, Interscience, New York, 1957.

[2] E. Snapper, R.J. Troyer, Metric Affine Geometry, Dover, New York, 1971.

[3] F. Uhlig, Constructive ways for generating (generalized) real orthogonal matrices as products of (generalized) symmetries, Linear Algebra Appl. 332–334 (2001) 459–467.

[4] D.S. Mackey, N. Mackey, F. Tisseur, G-reflectors: analogues of Householder transformations in scalar product spaces, Linear Algebra Appl. 385 (2004) 187–213.

[5] J. Gallier, *Geometric methods and applications, For computer science and engineering*, Springer-Verlag, New York, 2001.

[6] G. Aragón-González, J.L. Aragón, M.A. Rodríguez-Andrade, The decomposition of an orthogonal transformation as a product of reflections, J. Math. Phys. 47 (2006) Art. No. 013509.

[7] R. Delanghe, F. Sommen, V. Souček, Clifford Algebra and Spinor-Valued Functions: A Function Theory for the Dirac Operator, Kluwer, Dordrecht, 1992.

[8] I.R. Porteous, Topological Geometry, Cambridge Univ. Press, Cambridge, 1981.

[9] M. Riesz, in: E.F. Bolinder, P. Lounesto (Eds.), Clifford Numbers and Spinors: with Riesz's Private Lectures to E. Folke Bolinder and a Historical Review by Pertti Lounesto, Kluwer, Dordrecht, 1993.

[10] P. Scherk, On the descomposition of Orthogonalities in symmetries, Proceedings of the American Mathematical Society, Vol. 1, No. 4 (Aug., 1950), pp. 481–491.