April 18, 2016

Lemma 2 shows that S itself is also an orthogonal matrix. It is well known that this is equivalent to s being an orthogonal transformation, that is,

$$(s(\lambda), s(\mu)) = (\lambda, \mu) \quad (\lambda, \mu \in \mathbf{R}^n). \tag{12}$$

This can be directly verified as follows. First, let $s = s_\alpha$ with $\alpha \neq 0$ and set

$$\pi(\lambda) = \lambda - \frac{(\lambda, \alpha)}{(\alpha, \alpha)}\alpha.$$

Then $(\pi(\lambda), \alpha) = 0$, so

$$\lambda = \frac{(\lambda, \alpha)}{(\alpha, \alpha)}\alpha + \pi(\lambda)$$

is the representation of λ as an element of $\mathbf{R}\alpha \oplus (\mathbf{R}\alpha)^\perp$. By the definition of a reflection, we obtain

$$\begin{aligned} s_\alpha(\lambda) &= -\frac{(\lambda, \alpha)}{(\alpha, \alpha)}\alpha + \pi(\lambda) \\ &= \lambda - \frac{2(\lambda, \alpha)}{(\alpha, \alpha)}\alpha. \end{aligned}$$

Note that this is a direct generalization of our formula (2) originally established in $\mathbf{R}^2$ only. Now

$$\begin{aligned} (s_\alpha(\lambda), s_\alpha(\mu)) &= (\lambda - \frac{2(\lambda, \alpha)}{(\alpha, \alpha)}\alpha, \mu - \frac{2(\mu, \alpha)}{(\alpha, \alpha)}\alpha) \\ &= (\lambda, \mu) - (\lambda, \frac{2(\mu, \alpha)}{(\alpha, \alpha)}\alpha) - (\mu, \frac{2(\lambda, \alpha)}{(\alpha, \alpha)}\alpha) + (\frac{2(\lambda, \alpha)}{(\alpha, \alpha)}\alpha, \frac{2(\mu, \alpha)}{(\alpha, \alpha)}\alpha) \\ &= (\lambda, \mu) - \frac{2(\mu, \alpha)}{(\alpha, \alpha)}(\lambda, \alpha) - \frac{2(\lambda, \alpha)}{(\alpha, \alpha)}(\mu, \alpha) + \frac{2(\lambda, \alpha)}{(\alpha, \alpha)}\frac{2(\mu, \alpha)}{(\alpha, \alpha)}(\alpha, \alpha) \\ &= (\lambda, \mu) - \frac{2(\lambda, \alpha)(\mu, \alpha)}{(\alpha, \alpha)} - \frac{2(\lambda, \alpha)(\mu, \alpha)}{(\alpha, \alpha)} + \frac{4(\lambda, \alpha)(\mu, \alpha)}{(\alpha, \alpha)} \\ &= (\lambda, \mu). \end{aligned}$$

Therefore, s_α is an orthogonal transformation.

For a real vector space V with an inner product, the set of orthogonal transformation is denoted by $O(V)$. Thus, every reflection in V is an element of $O(V)$. It is necessary to consider a more general vector space V than just $\mathbf{R}^n$, since we sometimes need to consider linear transformation defined on a subspace of $\mathbf{R}^n$.

Let us recall how the transformation rule (10) was used to derive every word in $\langle s, t\rangle$ is one of the $2m$ possible forms. We now formalize this by ignoring the fact that s, t are reflections. Instead we only assume $s^2 = t^2 = 1$. In order to facilitate this, we consider a set of formal symbols X and consider the set of all words of length n. This is the set of sequence of length n, so it can be regarded as the cartesian product

$$X^n = \underbrace{X \times X \times \cdots \times X}_{n}.$$

Then we can form a disjoint union

$$X^* = \bigcup_{n=0}^{\infty} X^n,$$

where X^0 consists of a single element called the empty word, denoted by 1.

A word $x = (x_1, x_2, \ldots, x_n) \in X^n$ is said to be *reduced* if $x_i \neq x_{i+1}$ for $1 \leq i < n$. By definition, the word 1 of length 0 is reduced, and every word of length 1 is reduced. For brevity, we write $x = x_1 x_2 \cdots x_n \in X^n$ instead of $x = (x_1, x_2, \ldots, x_n) \in X^n$. We denote the set of all reduced words by $F(X)$.

We can define a binary operation $\mu : F(X) \times F(X) \to F(X)$ as follows.

$$\mu(1, x) = \mu(x, 1) = 1 \quad (x \in F(X)), \tag{13}$$

and for $x = x_1 \cdots x_m \in X^m \cap F(X)$ and $y = y_1 \cdots y_n \in X^n \cap F(X)$ with $m, n \geq 1$, we define

$$\mu(x, y) = \begin{cases} x_1 \cdots x_m y_1 \cdots y_n \in X^{m+n} & \text{if } x_m \neq y_1, \\ \mu(x_1 \cdots x_{m-1}, y_2 \cdots y_n) & \text{otherwise.} \end{cases} \tag{14}$$

This is a recursive definition. Note that if $x_m \neq y_1$, then $x_1 \cdots x_m y_1 \cdots y_n$ is a reduced word. Note also that there is no guarantee that $x_1 \cdots x_{m-1} y_2 \cdots y_n$ is a reduced word. If it is not, then $x_{m-1} = y_2$, so we define this to be $\mu(x_1 \cdots x_{m-2}, y_3 \cdots y_n)$. Since the length is finite, we eventually reach the case where the last symbol of x is different from the first symbol of y, or one of x, y is 1.

Definition 3. A set G with binary operation $\mu : G \times G \to G$ is said to be a *group* if

(i) μ is associative, that is, $\mu(\mu(a, b), c) = \mu(a, \mu(b, c))$ for all $a, b, c \in G$,

(ii) there exists an element $1 \in G$ such that $\mu(1, a) = \mu(a, 1) = a$ for all $a \in G$,

(iii) for each $a \in G$, there exists an element $a' \in G$ such that $\mu(a, a') = \mu(a', a) = 1$.

The element 1 is called the *identity* of G, and a' is called the *inverse* of a.

Theorem 4. *The set of reduced words $F(X)$ forms a group under the binary operation μ defined by* (13)–(14).

Proof. Clearly, the empty word 1 is the identity in $F(X)$, i.e.,

$$\mu(1, a) = \mu(a, 1) = a \quad (a \in F(X)). \tag{15}$$

Next we prove associativity (i), by a series of steps.

Step 1.

$$\mu(\mu(a,x),\mu(x,b)) = \mu(a,b) \quad (a,b \in F(X),\ x \in X). \tag{16}$$

Indeed, denote by a_{-1} the last entry of a, and by b_1 the first entry of b. Write

$$\begin{aligned} a &= a'x && \text{if } a_{-1} = x,\\ b &= xb' && \text{if } b_1 = x. \end{aligned}$$

Since

$$\begin{aligned} ax &\in F(X) && \text{if } a_{-1} \neq x,\\ xb &\in F(X) && \text{if } b_1 \neq x, \end{aligned}$$

we have

$$\begin{aligned} \mu(\mu(a,x),\mu(x,b)) &= \begin{cases} \mu(a',b') & \text{if } a_{-1} = x, b_1 = x,\\ \mu(a',xb) & \text{if } a_{-1} = x, b_1 \neq x,\\ \mu(ax,b') & \text{if } a_{-1} \neq x, b_1 = x,\\ \mu(ax,xb) & \text{if } a_{-1} \neq x, b_1 \neq x \end{cases}\\ &= \mu(a,b). \end{aligned}$$

Step 2.

$$\mu(x,\mu(x,c)) = c \quad (c \in F(X),\ x \in X). \tag{17}$$

Indeed,

$$\begin{aligned} \mu(x,\mu(x,c)) &= \mu(\mu(1,x),\mu(x,c)) && \text{(by (13))}\\ &= \mu(1,c) && \text{(by (16))}\\ &= c && \text{(by (13))}. \end{aligned}$$

Step 3.

$$\mu(x,\mu(b,c)) = \mu(\mu(x,b),c) \quad (b,c \in F(X),\ x \in X). \tag{18}$$

Assume $b \in X^m$. We prove (18) by induction on m. If $m = 0$, then $b = 1$, so

$$\begin{aligned} \mu(x,\mu(b,c)) &= \mu(x,\mu(1,c))\\ &= \mu(x,c) && \text{(by (15))}\\ &= \mu(\mu(x,1),c) && \text{(by (15))}\\ &= \mu(\mu(x,b),c). \end{aligned}$$

Next assume $m > 0$. If $b = xb'$, then

$$\begin{aligned} \mu(x,\mu(b,c)) &= \mu(x,\mu(\mu(x,b'),c))\\ &= \mu(x,\mu(x,\mu(b',c))) && \text{(by induction)} \end{aligned}$$

$$
\begin{aligned}
&= \mu(b', c) && \text{(by (17))} \\
&= \mu(\mu(x, b), c).
\end{aligned}
$$

If $b = b'y$ and $c = yc'$ for some $b', c' \in F(X)$ and $y \in X$, then

$$
\begin{aligned}
\mu(x, \mu(b, c)) &= \mu(x, \mu(b', c')) && \text{(by (14))} \\
&= \mu(\mu(x, b'), c') && \text{(by induction)} \\
&= \mu(\mu(\mu(x, b'), y), \mu(y, c')) && \text{(by (16))} \\
&= \mu(\mu(\mu(x, b'), y), c) \\
&= \mu(\mu(x, \mu(b', y)), c) && \text{(by induction)} \\
&= \mu(\mu(x, b), c).
\end{aligned}
$$

Finally, if $b_1 \neq x$ and $b_{-1} \neq c_1$, then $\mu(x, b) = xb$ and $\mu(b, c) = bc$, and $xbc \in F(X)$. Thus

$$
\begin{aligned}
\mu(x, \mu(b, c)) &= \mu(x, bc) \\
&= xbc \\
&= \mu(xb, c) \\
&= \mu(\mu(x, b), c).
\end{aligned}
$$

This completes the proof of (18).

Now we prove

$$
\mu(a, \mu(b, c)) = \mu(\mu(a, b), c) \quad (a, b, c \in F(X)). \tag{19}
$$

by induction on n, where $a \in X^n$. The cases $n = 0$ is trivial because of (15). Assume $a = a'x$, where $a' \in F(X)$ and $x \in X$. Then

$$
\begin{aligned}
\mu(a, \mu(b, c)) &= \mu(\mu(a', x), \mu(b, c)) \\
&= \mu(a', \mu(x, \mu(b, c))) && \text{(by induction)} \\
&= \mu(a', \mu(\mu(x, b), c)) && \text{(by (18))} \\
&= \mu(\mu(a', \mu(x, b)), c) && \text{(by induction)} \\
&= \mu(\mu(\mu(a', x), b), c) && \text{(by induction)} \\
&= \mu(\mu(a, b), c).
\end{aligned}
$$

Therefore, we have proved associativity.

If $a = x_1 \cdots x_n \in F(X) \cap X^n$, then the reversed word $a' = x_n \cdots x_1 \in F(X) \cap X^n$ is the inverse of a. $\square$

We call $F(X)$ the *free group generated by the set of involutions X*. From now on, we omit μ to denote the binary operation in $F(X)$ by juxtaposition. So we write ab instead of $\mu(a, b)$ for $a, b \in F(X)$. Also, for $a = x_1 \cdots x_n \in F(X) \cap X^n$, its inverse $x_n \cdots x_1$ will be denoted by a^{-1}.

Let s and t be the linear transformation of $\mathbf{R}^2$ represented by the matrices

$$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \text{ and } \begin{bmatrix} \cos\frac{2\pi}{m} & \sin\frac{2\pi}{m} \\ \sin\frac{2\pi}{m} & -\cos\frac{2\pi}{m} \end{bmatrix},$$

respectively. Let $G = \langle s, t \rangle$ be the set of all linear transformation expressible as a product of s and t. We know

$$G = \{(st)^j \mid 0 \le j < m\} \cup \{(st)^j s \mid 0 \le j < m\}.$$

and $|G| = 2m$. The product of linear transformations defines a binary operation on G, and G forms a group under this operation. This group is called the *dihedral group* of order $2m$. In order to connect the dihedral group with a free group, we make a definition.

Definition 5. Let G_1 and G_2 be groups. A mapping $f : G_1 \to G_2$ is called a *homomorphism* if

$$f(ab) = f(a)f(b) \quad (\forall a, b \in G_1), \tag{20}$$

where the product ab is computed under the binary operation in G_1, the product $f(a)f(b)$ is computed under the binary operation in G_2. A bijective homomorphism is called an *isomorphism*. The groups G_1 and G_2 are said to be *isomorphic* if there exists an isomorphism from G_1 to G_2.

Let $X = \{x, y\}$ be a set of two distinct formal symbols. Clearly, there is a homomorphism $f : F(X) \to G$ with $f(x) = s$ and $f(y) = t$, where $G = \langle s, t \rangle$ is the dihedral group of order $2m$ defined above. Note that $f((xy)^m) = (st)^m = 1$, but $(xy)^m \in F(X)$ is not the identity. This suggests introducing another transformation rule $(xy)^m = 1$, in addition to $x^2 = y^2 = 1$ as we adopted when constructing the group $F(X)$. We do this by introducing an equivalence relation on $F(X)$. Let $a, b \in F(X)$. If there exists $c \in F(X)$ such that $a = bc^{-1}(xy)^m c$, then $f(a) = f(b)$ holds. So we write $a \sim b$ if there is a finite sequence $a = a_0, a_1, \ldots, a_n = b \in F(X)$ such that for each $i \in \{1, 2, \ldots, n\}$, a_i is obtained by multiplying a_{i-1} by an element of the form $c^{-1}(xy)^m c$ for some $c \in F(X)$. Then $\sim$ is an equivalence relation, since $a = bc^{-1}(xy)^m c$ implies $b = a(xc)^{-1}(xy)^m(xc)$. Clearly, $a \sim b$ implies $f(a) = f(b)$. In other words, f induces a mapping from the set of equivalence classes to G. In fact, the set of equivalence classes forms a group under the binary operation inherited from $F(X)$. We can now make this more precise.