are two positive self-adjoint linear maps $h_1 : E \to E$ *and* $h_2 : F \to F$ *and a weakly orthogonal linear map* $g : E \to F$ *such that*

$$f = g \circ h_1 = h_2 \circ g.$$

Furthermore, if f *has rank* r*, the maps* h_1 *and* h_2 *have the same positive eigenvalues* $\mu_1, \ldots, \mu_r$*, which are the singular values of* f*, i.e., the positive square roots of the nonnull eigenvalues of both* $f^* \circ f$ *and* $f \circ f^*$*. Finally,* h_1, h_2 *are unique,* g *is unique if* $\mathrm{rank}(f) = \min(m, n)$ *and* $h_1 = h_2$ *if* f *is normal.*

Proof. By Lemma 12.1.2 there are two (unique) positive self-adjoint linear maps $h_1 : E \to E$ and $h_2 : F \to F$ such that $f^* \circ f = h_1^2$ and $f \circ f^* = h_2^2$. As in the proof of Theorem 12.1.3,

$$\mathrm{Ker}\, f = \mathrm{Ker}\, h_1,$$

and letting r be the rank of f, there is an orthonormal basis $(u_1, \ldots, u_n)$ of eigenvectors of h_1 such that $(u_1, \ldots, u_r)$ are associated with the strictly positive eigenvalues $\mu_1, \ldots, \mu_r$ of h_1 (the singular values of f). The vectors $(u_{r+1}, \ldots, u_n)$ form an orthonormal basis of $\mathrm{Ker}\, f = \mathrm{Ker}\, h_1$, and the vectors $(u_1, \ldots, u_r)$ form an orthonormal basis of $(\mathrm{Ker}\, f)^\perp = \mathrm{Im}\, f^*$. Furthermore, letting

$$v_i = \frac{f(u_i)}{\mu_i}$$

when $1 \leq i \leq r$, using the Gram–Schmidt orthonormalization procedure, we can extend $(v_1, \ldots, v_r)$ to an orthonormal basis $(v_1, \ldots, v_m)$ of F (even when $r = 0$). Also note that $(v_1, \ldots, v_r)$ is an orthonormal basis of $\mathrm{Im}\, f$, and $(v_{r+1}, \ldots, v_m)$ is an orthonormal basis of $\mathrm{Im}\, f^\perp = \mathrm{Ker}\, f^*$.

Letting $p = \min(m, n)$, we define the linear map $g : E \to F$ by its action on the basis $(u_1, \ldots, u_n)$ as follows:

$$g(u_i) = v_i$$

for all i, $1 \leq i \leq p$, and

$$g(u_i) = 0$$

for all i, $p + 1 \leq i \leq n$. Note that $r \leq p$. Just as in the proof of Theorem 12.1.3, we have

$$(g \circ h_1)(u_i) = f(u_i)$$

when $1 \leq i \leq r$, and

$$(g \circ h_1)(u_i) = g(h_1(u_i)) = g(0) = 0$$

when $r + 1 \leq i \leq n$ (since $(u_{r+1}, \ldots, u_n)$ is a basis for $\mathrm{Ker}\, f = \mathrm{Ker}\, h_1$), which shows that $f = g \circ h_1$. The fact that g is weakly orthogonal follows easily from the fact that it maps the orthonormal vectors $(u_1, \ldots, u_p)$ to the orthonormal vectors $(v_1, \ldots, v_p)$.

We can show that $f = h_2 \circ g$ as follows. Just as in the proof of Theorem 12.1.3,

$$h_2^2(v_i) = \mu_i^2 v_i$$

when $1 \leq i \leq r$, and

$$h_2^2(v_i) = (f \circ f^*)(v_i) = f(f^*(v_i)) = 0$$

when $r + 1 \leq i \leq m$, since $(v_{r+1}, \ldots, v_m)$ is a basis for $\operatorname{Ker} f^* = (\operatorname{Im} f)^\perp$. Since h_2 is positive self-adjoint, so is h_2^2, and by Lemma 12.1.2, we must have

$$h_2(v_i) = \mu_i v_i$$

when $1 \leq i \leq r$, and

$$h_2(v_i) = 0$$

when $r + 1 \leq i \leq m$. This shows that $(v_1, \ldots, v_m)$ are eigenvectors of h_2 for $\mu_1, \ldots, \mu_m$ (letting $\mu_{r+1} = \cdots = \mu_m = 0$), and thus h_1 and h_2 have the same nonnull eigenvalues $\mu_1, \ldots, \mu_r$. As a consequence,

$$(h_2 \circ g)(u_i) = h_2(g(u_i)) = h_2(v_i) = \mu_i v_i = f(u_i)$$

when $1 \leq i \leq m$. Since h_1, h_2, $f^* \circ f$, and $f \circ f^*$ are positive self-adjoint, $f^* \circ f = h_1^2$, $f \circ f^* = h_2^2$, and $\mu_1, \ldots, \mu_r$ are the eigenvalues of both h_1 and h_2, it follows that $\mu_1, \ldots, \mu_r$ are the singular values of f, i.e., the positive square roots of the nonnull eigenvalues of both $f^* \circ f$ and $f \circ f^*$.

Finally, if $m \geq n$ and $\operatorname{rank}(f) = n$, then $\operatorname{Ker} h_1 = \operatorname{Ker} f = (0)$ and h_1 is invertible and if $n \geq m$ and $\operatorname{rank}(f) = m$, then $\operatorname{Ker} h_2 = \operatorname{Ker} f^* = (0)$ and h_2 is invertible. By Lemma 12.1.2 h_1 and h_2 are unique and since

$$f = g \circ h_1 \quad \text{and} \quad f = h_2 \circ g,$$

if h_1 is invertible then $g = f \circ h_1^{-1}$ and if h_2 is invertible then $g = h_2^{-1} \circ f$, and thus g is also unique. If h is normal, then $f^* \circ f = f \circ f^*$ and $h_1 = h_2$. $\square$

In matrix form, Theorem 12.2.3 can be stated as follows. For every real $m \times n$ matrix A, there is some weakly orthogonal $m \times n$ matrix R and some positive symmetric $n \times n$ matrix S such that

$$A = RS.$$

The proof also shows that if $n > m$, the last $n - m$ columns of R are zero vectors. A pair (R, S) such that $A = RS$ is called a *polar decomposition of* A.

Remark: If E is a Hermitian space, Theorem 12.2.3 also holds, but the weakly orthogonal linear map g becomes a weakly unitary map. In terms of matrices, the polar decomposition states that for every complex $m \times n$