do with perturbation theory (1912). Autonne came up with the polar decomposition (1902, 1915). Eckart and Young extended SVD to rectangular matrices (1936, 1939).

The next three theorems deal with a linear map $f\colon E \to E$ over a Euclidean space E. We will show later on how to generalize these results to linear maps $f\colon E \to F$ between two Euclidean spaces E and F.

Theorem 12.1.3 *Given a Euclidean space E of dimension n, for any linear map $f\colon E \to E$ there are two positive self-adjoint linear maps $h_1\colon E \to E$ and $h_2\colon E \to E$ and an orthogonal linear map $g\colon E \to E$ 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 f is invertible, and $h_1 = h_2$ if f is normal.

Proof. By Lemma 12.1.2 there are two (unique) positive self-adjoint linear maps $h_1\colon E \to E$ and $h_2\colon E \to E$ such that $f^* \circ f = h_1^2$ and $f \circ f^* = h_2^2$. Note that

$$\langle f(u), f(v)\rangle = \langle h_1(u), h_1(v)\rangle$$

for all $u, v \in E$, since

$$\langle f(u), f(v)\rangle = \langle u, (f^* \circ f)(v)\rangle = \langle u, (h_1 \circ h_1)(v)\rangle = \langle h_1(u), h_1(v)\rangle,$$

because $f^* \circ f = h_1^2$ and $h_1 = h_1^*$ (h_1 is self-adjoint). From Lemma 12.1.1, $\mathrm{Ker}\, f = \mathrm{Ker}\,(f^* \circ f)$, and from Lemma 12.1.2, $\mathrm{Ker}\,(f^* \circ f) = \mathrm{Ker}\, h_1$. Thus,

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

If r is the rank of f, then since h_1 is self-adjoint, by Theorem 11.3.1 there is an orthonormal basis $(u_1, \ldots, u_n)$ of eigenvectors of h_1, and by reordering these vectors if necessary, we can assume 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), and that $\mu_{r+1} = \cdots = \mu_n = 0$. Observe that $(u_{r+1}, \ldots, u_n)$ is an orthonormal basis of $\mathrm{Ker}\, f = \mathrm{Ker}\, h_1$, and that $(u_1, \ldots, u_r)$ is an orthonormal basis of $(\mathrm{Ker}\, f)^\perp = \mathrm{Im}\, f^*$. Note that

$$\langle f(u_i), f(u_j)\rangle = \langle h_1(u_i), h_1(u_j)\rangle = \mu_i \mu_j \langle u_i, u_j\rangle = \mu_i^2 \delta_{i\,j}$$

when $1 \le i, j \le n$ (recall that $\delta_{i\,j} = 1$ if $i = j$, and $\delta_{i\,j} = 0$ if $i \ne j$). Letting

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

when $1 \le i \le r$, observe that

$$\langle v_i, v_j\rangle = \delta_{i\,j}$$