12.8 Thm ("Douglas lemma") Let $A \in \mathcal{L}(G;H)$ and $B \in \mathcal{L}(G;K)$ (Hilb. spaces). Then A can be factored into $A = ZB$ for some contraction $Z \in \mathcal{L}(K;G)$ if and only if

$$(*) \quad A^*A \le B^*B.$$

Proof: If $A = ZB$, then for all $x \in G$,

$$(A^*Ax, x) = (Ax, Ax) = \|Ax\|^2 = \|ZBx\|^2$$
$$\le \|Z\|^2 \|Bx\|^2 \le \|Bx\|^2 = (Bx, Bx) = (B^*Bx, x),$$

So $A^*A \le B^*B$.

Conversely, assume $(*)$. (Following proof "cleaner" than in the book.)

Step 1. Assume first that B is one-to-one (so that $\mathcal{N}(B) = \{0\}$). Then, to each $x \in G$ there is exactly one $y \in K$ so that $y = Bx$. We can therefore define an inverse B^{-1}, whose domain is the range of B: $x = B^{-1}y$ if $y \in \mathcal{R}(B)$ and $y = Bx$. Define (for $y \in \mathcal{R}(B)$)

$$Zy = Ax \quad (= AB^{-1}y).$$

Then Z is linear (easy), and

$$\|Zy\|^2 = \|Ax\|^2 = (Ax, Ax) = (A^*Ax, x)$$
$$\le (B^*Bx, x) = (Bx, Bx) = \|y\|^2,$$

So Z is a contraction defined on $\mathcal{R}(B)$. Moreover, $\boxed{Ax = ZBx}$ for all $x \in G$.

Step 2: Extend the domain of Z to $\overline{\mathcal{R}(B)}$ as follows:

To each $y \in \overline{\mathcal{R}(B)}$ we can find a sequence $y_n \in \mathcal{R}(B)$ such that $y_n \to y$ in K as $n \to \infty$. Since Z is a contraction, if we define $z_n = Zy_n$, then

$$\|z_n - z_m\| = \|Z(y_n - y_m)\| \le \|y_n - y_m\|,$$

So z_n is a Cauchy seq in H. Therefore it converges to some $z \in H$. Define $Zy = z$. This way Z becomes defined on the closure of the range of B. Still true that Z is a linear contraction (not difficult to show). This has no effect on the relation $Ax = ZBx$ (since $Bx \in \mathcal{R}(B)$ always).

Step 3. Extend Z to all of K as follows: Split K into $K = K_0 \oplus K_0^\perp$, where $K_0 = \overline{\mathcal{R}(B)}$. Every $x \in K$ can be written as $x = x_0 + x_1$, where $x_0 \in K_0$ and $x_1 \perp K_0$. Put

$$Zx = Zx_0 \quad (\text{i.e., } Zx_1 = 0).$$

This is still a linear operator on K, and it is a contraction since

$$\|Zx\|^2 = \|Zx_0\|^2 \le \|x_0\|^2 \le \|x_0\|^2 + \|x_1\|^2 = \|x\|^2.$$

Step 4: If B is not one-to-one, then we split G into $G = G_0 \oplus G_0^\perp$, where $G_0 = \mathcal{N}(B)$. Write each $x \in G$ as $x = x_0 + x_1$, where $x_0 \in G_0$ and $x_1 \in G_0^\perp$. Then $Bx_0 = 0$, so

$$0 = \|Bx_0\|^2 = (Bx_0, Bx_0) = (B^*Bx_0, x_0)$$
$$\ge (A^*Ax_0, x_0) = (Ax_0, Ax_0) = \|Ax_0\|^2,$$

so $Ax_0 = 0$. Thus,

$$\boxed{Bx = Bx_1, \quad \text{and} \quad Ax = Ax_1}.$$

The operator B is one-to-one on $G_0^\perp$. We construct Z as in Steps 1–3 with G replaced by $G_0^\perp$. Then for all $x_1 \in G_0^\perp$, and

$$Ax_1 = ZBx_1, \text{ so}$$

for all $x \in G$, $Ax = Ax_1 = ZBx_1 = ZBx$. □

12.9 Cor. Let $A \in \mathcal{L}(G;H)$ and $B \in \mathcal{L}(K;H)$ (Hilbert spaces). Then $A = BZ$ for some contraction $Z \in \mathcal{L}(G;K)$ if and only if

$$AA^* \le BB^*.$$

Proof. Apply Thm 12.8 to A^* and B^*. □

XII.4 Operator Matrices (Block Matrices)

12.10 Defn. Let H_1 and H_2 be Hilbert spaces. Then the cross product (or orthogonal direct sum) H of H_1 and H_2 consists of all pairs of vectors $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ where $x_1 \in H_1$ and $x_2 \in H_2$.

Notation: $H = \begin{bmatrix} H_1 \\ H_2 \end{bmatrix}$ or $H = H_1 \oplus H_2$.

12.10a Lemma $H_1 \oplus H_2$ is a Hilbert space with inner product

$$\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \right)_{\left[\begin{smallmatrix} H_1 \\ H_2 \end{smallmatrix}\right]} = (x_1, y_1)_{H_1} + (x_2, y_2)_{H_2}.$$

Proof: See Analysis II.

Note that $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ x_2 \end{pmatrix}$, where $x_1 \in H_1$ and $x_2 \in H_2$. If we "identify" H_1 with the subset $\begin{bmatrix} H_1 \\ 0 \end{bmatrix}$ of H and H_2 with the subspace $\begin{bmatrix} 0 \\ H_2 \end{bmatrix}$ of H, then both H_1 and H_2 become subspaces of H and we are back in the situation described in Defn. 4.26 on p. 39.

12.11 Defn. Let $A_{ij} \in \mathcal{L}(H_j; K_i)$, $i,j = 1,2$. We define the "block matrix" or "operator matrix" $A \in \mathcal{L}\left(\begin{bmatrix} H_1 \\ H_2 \end{bmatrix}; \begin{bmatrix} K_1 \\ K_2 \end{bmatrix}\right)$ by

$$A \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{bmatrix} A_{11}x_1 + A_{12}x_2 \\ A_{21}x_1 + A_{22}x_2 \end{bmatrix}$$

(Same rule as for matrix multiplication, but A_{ij} are operators and x_j are vectors) (think of A_{ij} as "submatrices")

12.13 Thm Let $A_1 \in \mathcal{L}(H; K_1)$ and $A_2 \in \mathcal{L}(H; K_2)$. Then $\begin{bmatrix} A_1 \\ A_2 \end{bmatrix}$ is a contraction from H to $\begin{bmatrix} K_1 \\ K_2 \end{bmatrix}$ if and only if

$$A_1 = Z_1 (I - A_2^* A_2)^{1/2} \quad (*)$$

for some contraction $Z_1 : H \to K_1$, or equivalently, if and only if

$$A_2 = Z_2 (I - A_1^* A_1)^{1/2}$$

for some contraction $Z_2 : H \to K_2$.

Proof: By Thm 12.6, $\begin{bmatrix} A_1 \\ A_2 \end{bmatrix}$ is a contraction

$$\iff I - \begin{bmatrix} A_1^* & A_2^* \end{bmatrix} \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} \ge 0$$

$$\iff I - A_1^* A_1 - A_2^* A_2 \ge 0$$

$$\iff A_1^* A_1 \le I - A_2^* A_2 \iff A_2^* A_2 \le I - A_1^* A_1.$$

Now use Thm 12.8. □

12.14 Coroll. $[A_1\ A_2]$ is a contraction
$\iff A_1 = (I - A_2 A_2^*)^{1/2} Z_1$ for some contr. Z_1
$\iff A_2 = (I - A_1 A_1^*)^{1/2} Z_2$ — " — Z_2.

Proof. Apply Thm 12.13 to $[A_1, A_2]^*$. $\square$

XVI Singular values

(See pp. 204–207 in the book).

When an operator maps one space into another it cannot have eigenvalues. We then need the concept of a singular value. (Much used in H^∞ control theory).

XVI. 1. Singular values

Idea: A compact self-adjoint operator K has an expansion

$$Kx = \sum_{k=1}^{\infty} \lambda_k (x, \varphi_k) \varphi_k .$$

We can order the eigenvalues so that $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \dots$. Then the "best approximation of K with rank n" is given by

$$K_n x = \sum_{k=1}^{n} \lambda_k (x, \varphi_k) \varphi_k .$$

The norm of the error is

$$\| K - K_n \| = \sup_{k \geq n+1} |\lambda_k| .$$

Thus:

$$\lambda_1 = \|K\|$$
$$\lambda_2 = \|K - K_1\|$$
$$\lambda_3 = \|K - K_2\|, \text{ etc.}$$

16.1 Defn. Let $T \in \mathcal{L}(H;K)$ (Hilbert spaces). For $k = 0, 1, 2, \dots$, we define

$$s_k(T) = \inf \{ \|T - R\| \mid R \in \mathcal{L}(H;K) \text{ and } \operatorname{rank} R \leq k \}.$$

We call $s_k(T)$ the k-th singular value of T. (Obviously $s_0(T) \geq s_1(T) \geq s_2(T) \geq \dots$).

Note: $s_k(T)$ = "the distance of T to the set of all operators with rank k". Taking $k=0$ we get the distance of T to 0 (rank $=0 \iff R=0$).

16.1a Defn. Let $E \subset H$ (Hilbert spaces) be a closed subspace of E. Then the codimension of E in H is the dimension of $E^\perp$.

16.2 Thm. Notation as in Defn. 16.1. Then

$$s_k(T) = \inf_E \|T_{|E}\|,$$

where $T_{|E}$ is the restriction of T to E, and E is a subspace of H of codimension k.

Proof:

Let $E \subset H$, with codimension $\le k$. Define

$$Rx = \begin{cases} Tx, & x \in E^\perp, \\ 0, & x \in E. \end{cases}$$

Then $\operatorname{rank}(R) \le k$ (the dimension of the range is always $\le$ dimension of domain). Clearly,

$$(T-R)x = \begin{cases} 0, & x \in E^\perp, \\ Tx, & x \in E, \end{cases}$$

so, denoting $x = x_1 + x_2$ where $x_1 \in E$, $x_2 \in E^\perp$,

$$\|T-R\| = \sup_{\substack{x \in H \\ \|x\| \le 1}} \|(T-R)x\| = \sup_{\substack{x \in H \\ \|x\| \le 1}} \|Tx_1\|$$

$$= \sup_{\substack{x_1 \in E \\ \|x_1\| \le 1}} \|Tx_1\| = \|T_{|E}\|.$$

Therefore $\inf_{\operatorname{codim}(E) \le k} \|T_{|E}\| \ge \inf \{ \|T-R\| \mid \operatorname{Rank} R \le k \}$
$= s_k(T)$.

For the converse, pick $\varepsilon > 0$ and some R with rank $R \le k$ so that

$$\|T-R\| \le s_k(T) + \varepsilon.$$

Then $\operatorname{codim}(\ker(R)) \le k$ (the dimension of $\ker(R)^\perp$ is $\le$ dimension of the range). Now

$$\|T_{|\ker R}\| \qquad (R=0 \text{ on } \ker(R))$$

$$= \|(T-R)_{|\ker R}\| \le \|T-R\| \le s_k(T)+\varepsilon,$$

so $\inf_{\operatorname{codim}(E) \le k} \|T_{|E}\| \le s_k(T)$. ◻

16.3 Corollary: If $\|Tx\| \ge t\|x\|$ for all x in a subspace of dimension $k+1$, then $s_k(T) \ge t$ (here $k = 0, 1, 2, \dots$)

Proof: Call this subspace F. If E is any subspace with codimension k, then $F \not\subset E^\perp$, hence $F \cap E \ne \{0\}$. Therefore E contains a vector x so that $Tx \ge t\|x\|$, and $\|T_{|E}\| \ge t$. ◻

16.4 Thm (Important) Let T be compact, $T \in \mathcal{L}(H;K)$ (Hilbert spaces). Expand T^*T into

$$T^*T = \sum_{k=0}^{\infty} \lambda_k (x, \varphi_k) \varphi_k,$$

where $\lambda_0 \ge \lambda_1 \ge \lambda_2 \dots$ and φ_k is a complete orthonormal sequence. Then

$$\boxed{s_k(T) = +\sqrt{\lambda_k}}.$$

(In particular, $\|T\| = \sqrt{\lambda_0}$).
(Note: $s_k(T) = 0 \iff T$ is of rank $\le k$).

<u>Proof.</u> Let E = closed linear span of $\{\varphi_k, \varphi_{k+1}, \varphi_{k+2} \dots\}$.
Then $E^\perp$ = span of $\{\varphi_0, \varphi_1, \dots, \varphi_{k-1}\}$, with dimension k. Thus, codim$(E) = k$. Now, for any $x \in E$,

$$\|Tx\|^2 = \langle Tx, Tx\rangle = \langle T^*Tx, x\rangle.$$

Take the sup over all $\underline{\underline{x \in E}}$ with $\|x\| \leq 1$ we get

$$\|T|_E\|^2 = \sup_{\substack{x\in E\\ \|x\|\leq 1}} \langle T^*Tx, x\rangle = \quad (\text{Thm. } 7.18)$$

$$= \|T^*T|_E\| \quad (\text{see argument on p. } 126)$$

$$= \lambda_k.$$

Thus, by Thm. 16.2, $s_k(T) \leq \sqrt{\lambda_k}$.
Conversely, take any closed $E \subset H$ with codimension k. Let F = span of $\{\varphi_0, \varphi_1, \dots, \varphi_k\}$. Then F has dimension $k+1$, and $F \not\subset E^\perp$, i.e., $F \cap E \neq \{0\}$. Take any $x \neq 0$, $x \in F \cap E$. Then

$$\|Tx\|^2 = (Tx, Tx) = (T^*Tx, x).$$

$$= \Big(\sum_{j=0}^{k} \lambda_j (x, \varphi_j)\varphi_j, \; x\Big)$$

$$= \sum_{j=0}^{k} \lambda_j |(x,\varphi_j)|^2$$

$$\geq \lambda_k \sum_{j=0}^{k} |(x,\varphi_j)|^2 \quad (\text{Pythagoras})$$

$$= \lambda_k \|x\|^2. \quad (\text{take } \|x\| = 1)$$

Thus, $\sup\limits_{\substack{x\in E\\ \|x\|\leq 1}} \|Tx\|^2 = \|T|_E\|^2 \geq \lambda_k$, and

$$s_k(T) \geq \sqrt{\lambda_k} \qquad \square$$

<u>XVI. 2 Schmidt pairs and Singular vectors</u>

Let T be compact, $T \in \mathcal{L}(H;K)$. If s is a singular value, then s^2 is an eigenvalue of T^*T (by Thm. 16.4). Therefore, there is an eigenvector $x \in H$ of T^*T. If $\boxed{s \neq 0}$, then we may take

$$y = s^{-1}Tx \in K.$$

Then $y \neq 0$, because $T^*y = \frac{1}{s} T^*Tx = x \neq 0$.
Moreover, y is an eigenvector of TT^*:

$$TT^*y = Tx = sy.$$

<u>16.5 Defn.</u> Let H, K be Hilbert spaces, $T \in \mathcal{L}(H;K)$, and let s be a <u>singular value</u> of T. A <u>Schmidt pair</u> of T corresponding to the singular value s is a pair of vectors (x, y), with $x \in H$, $y \in K$, satisfying

$$\boxed{Tx = sy, \quad T^*y = sx}.$$

A <u>singular vector</u> or <u>Schmidt vector</u> of T corresponding to s is an <u>eigenvector</u> of T^*T corresponding to the eigenvalue s^2. The <u>multiplicity</u> of the singular value s = the multiplicity of the eigenvalue s^2 of T^*T = dim$\{\ker\{s^2 I - T^*T\}\}$.

16.6 Representation Thm. Let $T \in \mathcal{L}(H;K)$ be <u>compact</u>, with H and K <u>separable</u> Hilbert spaces. Denote the singular values of T by s_k, $k = 0, 1, 2, \dots$, and let $\{e_k\}_{k=0}^{\infty}$ be a complete orthonormal sequence of eigenvectors of T^*T so that

$$(*) \quad \boxed{T^*Tx = \sum_{k=0}^{\infty} s_k^2 (x, e_k) e_k}.$$

For all those k for which $s_k \neq 0$, define $\phi_k = s_k^{-1} T e_k$. Then $\{\phi_k\}_{k=0}^{\infty}$ is an <u>orthonormal</u> sequence in K which spans $\overline{\text{Range}(T)}$ (it is complete iff the range of T is dense in K), and

$$(**) \quad \boxed{Tx = \sum_{k=0}^{\infty} s_k (x, e_k) \phi_k}.$$

(where we only need to sum over those k for which $s_k \neq 0$).

Proof: Without loss of generality, suppose that T is one-to-one (because every $x \in H$ can be split into $x = x_1 + x_2$, with $x_1 \in Ker(T)$ and $x_2 \in Ker(T)^\perp$, and $Tx = Tx_1$.). Then T^*T is also one-to-one (Homework 31c). This amounts to dropping all terms with $s_k = 0$ in the sum $(*)$.

Claim: $\{\phi_k\}_{k=0}^{\infty}$ is orthonormal.

Proof: $\|\phi_k\|^2 = \left(\frac{1}{s_k} Te_k, \frac{1}{s_k} Te_k\right)$

$= \frac{1}{s_k^2}(T^*Te_k, e_k) = \frac{1}{s_k^2}(s_k^2 e_k, e_k) = 1,$

and for $k \neq l$,

$$(\phi_k, \phi_l) = \frac{1}{s_k}\frac{1}{s_l}(Te_k, Te_l)$$

$$= \frac{1}{s_k s_l}(T^*Te_k, e_l)$$

$$= \frac{s_k}{s_l}(e_k, e_l) = 0.$$

Thus, $\{\phi_k\}_{k=0}^{\infty}$ is <u>orthonormal</u>.

Write $x \in H$ into a Fourier series:

$$x = \sum_{k=0}^{\infty} (x, e_k) e_k, \quad x_n = \sum_{k=0}^{n} (x, e_k) e_k.$$

Then $x_n \to x$, and

$$Tx_n = \sum_{k=0}^{n} (x, e_k) Te_k = \sum_{k=0}^{n} s_k (x, e_k) \phi_k.$$

We know that $Tx_n \to Tx$ (T is continuous), so the sum converges. This is an orthogonal sum ($\phi_k \perp \phi_l$), so by Thm. 4.11, $\sum_{k=0}^{\infty} |s_k(x, e_k)|^2 < \infty$, and these are the Fourier coefficients of the vector Tx. Thus, we get $(**)$. ☑

$\boxed{\text{The end}}$

Good luck in the exam!

Merry X-mas!

Happy New Year!