EL3370 Mathematical Methods in Signals, Systems and Control

Topic 8: Linear Operators

Cristian R. Rojas

Division of Decision and Control Systems
KTH Royal Institute of Technology

Outline

Outline

Solving Linear Equations

Many problems in physics and engineering involve solving linear equations $Lf = g$, where L is, *e.g.*, a differential operator. Some questions are:

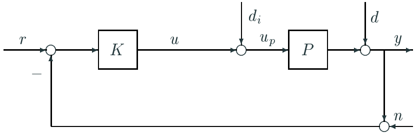
(1) Is there a solution of $Lf = g$?

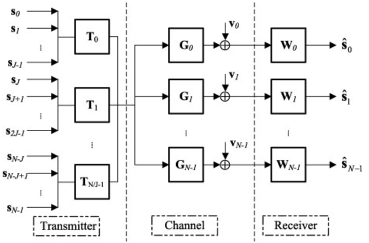
(2) Is it unique?

(3) How does it change if g is slightly perturbed?

Transfer functions

In systems theory, signals are represented by elements of normed spaces ($\ell_2, \ell_\infty, L_2, L_\infty, \ldots$), and systems are described by *operators* between these spaces.





Definitions

If E, F are vector spaces, a *linear operator* from E to F is a mapping $T: E \to F$ s.t.

$$T(\lambda x + \mu y) = \lambda Tx + \mu Ty \quad \text{for all } x, y \in E \text{ and scalars } \lambda, \mu.$$

If E, F are normed, T is *bounded* if there is an $M > 0$ s.t. $\|Tx\| \leqslant M\|x\|$ for all $x \in E$. If so, the *norm* of T is the smallest such M, *i.e.*,

$$\|T\| := \sup\{\|Tx\| : x \in E, \|x\| \leqslant 1\}.$$

The *kernel*, Ker T, of $T: E \to F$ is the subspace $\{x \in E : Tx = 0\} \subseteq E$, and the *range* of T, $\mathscr{R}(T)$, is the subspace $\{Tx : x \in E\} \subseteq F$.

The operator $I_E : E \to E$, given by $I_E(x) = x$ for all $x \in E$, is the *identity operator* on E. When there is no ambiguity, it will be written simply as I.

Examples

1. *Matrices*
   A matrix $A \in \mathbb{R}^{n\times m}$ corresponds to a linear operator $\mathbf{A}\colon \mathbb{R}^m \to \mathbb{R}^n$ where $\mathbf{A}x = Ax$ for every $x \in \mathbb{R}^m$. If we consider the 2-norm on $\mathbb{R}^m$ and $\mathbb{R}^n$, then $\|\mathbf{A}\| = \sup_{\|x\|_2=1} \|Ax\|_2$ is the largest singular value of A.
2. *Multiplication*
   Define M_f on $L_2[a,b]$ by: $(M_f x)(t) = f(t)x(t)$, where $f \in C[a,b]$. M_f is linear, and

$$\|M_f x\|^2 = \int_a^b |f(t)|^2 |x(t)|^2 dt \leqslant \max_{\tau\in[a,b]} |f(\tau)|^2 \int_a^b |x(t)|^2 dt = \|f\|^2 \|x\|^2,$$

   so $\|M_f\| \leqslant \|f\|$. In fact, $\|M_f\| = \|f\|$ (by choosing an appropriate (x_n)).

3. *Integral operator*
   Let $a,b,c,d \in \mathbb{R}$, and $k : [c,d]\times[a,b] \to \mathbb{R}$ continuous. Then, define $K : L_2[a,b] \to L_2[c,d]$ as

$$(Kx)(t) = \int_a^b k(t,s)x(s)ds, \quad c \leqslant t \leqslant d.$$

   K is linear, and, by Cauchy-Schwarz, $\|Kx\|^2 \leqslant \left(\int_c^d \int_a^b |k(t,s)|^2 ds dt\right) \|x\|^2$, so K is bounded.

Examples (cont.)

3. *Differential operator*

   Let $\mathscr{D} \subseteq C(\mathbb{R})$ be the space of differentiable functions $f \in C(\mathbb{R})$ s.t. $f' \in C(\mathbb{R})$. Then,

$$\frac{d}{dx} : \mathscr{D} \to C(\mathbb{R})$$

   is a linear operator, but it is not bounded (*why?*).

4. *Shift operator*

   Define S on ℓ_2 by:

$$S(x_1, x_2, x_3, \dots) = (0, x_1, x_2, \dots).$$

   S is an *isometry* (*i.e.*, $\|Sx\| = \|x\|$ for all $x \in \ell_2$), so it is bounded and $\|S\| = 1$. We can also define the backward shift operator S^* on ℓ_2 by $S^*(x_1, x_2, x_3, \dots) = (x_2, x_3, x_4, \dots)$, which is bounded and s.t. $\|S^*\| = 1$, but it is not an isometry (*why not?*).

Theorem

Let E, F be normed spaces, and $T: E \to F$ be a linear operator. The following are equivalent:

(1) T is continuous,

(2) T is continuous at 0,

(3) T is bounded.

Proof. Similar to the case for linear functionals. □

Outline

The Banach Space $\mathscr{L}(E,F)$

Definition
Let E,F be normed spaces. $\mathscr{L}(E,F)$ is the space of bounded linear operators from E to F, and $\mathscr{L}(E) = \mathscr{L}(E,E)$.

If F is a Banach space, so is $\mathscr{L}(E,F)$ (*similar to the proof that V^* is Banach*, in Topic 7).

The *composition* of operators $A : E \to F$ and $B : F \to G$, BA, is $BA(x) = B(Ax)$ for all $x \in E$.

Theorem. If $A \in \mathscr{L}(E,F)$ and $B \in \mathscr{L}(F,G)$, then $BA \in \mathscr{L}(E,G)$, and $\|BA\| \leqslant \|B\|\|A\|$.
Proof. BA is linear, and, since A,B are continuous, so is BA. Also,

$$\|BAx\|_G = \|B(Ax)\|_G \leqslant \|B\|\|Ax\|_F \leqslant \|B\|\|A\|\|x\|_E, \qquad x \in E,$$

so $\|BA\| \leqslant \|B\|\|A\|$. □

Observation. This last result shows that $\mathscr{L}(E)$ is not only a normed space, but also a *normed algebra* (since we have defined a product). If $\mathscr{L}(E)$ is complete, we say that it is a *Banach algebra*.

Outline

Inverses of Operators

Solving an equation $Ax = y$ involves computing "$x = A^{-1}y$".

Definition. Let E, F be normed spaces. $A \in \mathscr{L}(E,F)$ is *invertible* if there is a $B \in \mathscr{L}(F,E)$ s.t. $AB = I_F$ and $BA = I_E$. In this case, B is unique (*why?*) and is called the *inverse* of A, A^{-1}.

If E, F are Banach spaces, and $A \in \mathscr{L}(E,F)$ is bijective, its inverse is necessarily bounded (*Banach-Schauder* / *Open mapping theorem*) and linear (*why?*).

Examples

1. The shift operators S and S^* on ℓ_2 satisfy $S^*S = I$, but $SS^* \neq I$ (*why?*), so S, S^* are not invertible.

2. The multiplication operator M_t on $L_2[0,1]$ given by $(M_t x)(t) = tx(t)$ $(0 \leqslant t \leqslant 1)$ is injective but not surjective:

   $M_t x = 0$ implies $tx(t) = 0$, so $x(t) = 0$ (for almost all t).

   However, there is no $x \in L_2[0,1]$ s.t. $(M_t x)(t) = 1$, since $t \mapsto 1/t \notin L_2[0,1]$.

One way to produce inverses is as follows:

Theorem. Let E be a Banach space, and $A \in \mathscr{L}(E)$ s.t. $\|A\| < 1$. Then $I - A$ is invertible (in the normed space $\mathscr{L}(E)$), and

$$(I-A)^{-1} = \sum_{n=0}^{\infty} A^n = \lim_{N\to\infty}\left(I + A + A^2 + \cdots + A^N\right).$$

Proof. Let $x \in E$. Then $((I + A + A^2 + \cdots + A^n)x)$ is Cauchy: If $m > n$,

$$\left\|\sum_{k=0}^{m} A^k x - \sum_{k=0}^{n} A^k x\right\| = \left\|\sum_{k=n+1}^{m} A^k x\right\| \leqslant \sum_{k=n+1}^{m} \|A\|^k \|x\| \leqslant \frac{\|A\|^{n+1}}{1-\|A\|}\|x\| \to 0 \quad \text{as } n, m \to \infty \quad (m > n), \qquad (*)$$

so $\sum_{k=0}^{n} A^k x \to Tx$. T is linear, and letting $m \to \infty$ in $(*)$ gives $\left\|Tx - \sum_{k=0}^{n} A^k x\right\| \leqslant \dfrac{\|A\|^{n+1}}{1-\|A\|}\|x\|$, hence $Tx - \sum_{k=0}^{n} A^k x$ is bounded, and so is T.

Proof (cont.)

Also, $\left\| T - \sum_{k=0}^{n} A^k \right\| \leqslant \dfrac{\|A\|^{n+1}}{1 - \|A\|}$, so $\sum_{k=0}^{\infty} A^k = T$.

Finally, since $\|A^n x\| \leqslant \|A\|^n \|x\| \to 0$ as $n \to \infty$ (so $\lim A^n x = 0$),

$$(I - A)Tx = (I - A)\lim \sum_{k=0}^{n} A^k x = \lim \sum_{k=0}^{n} (A^k - A^{k+1})x = x - \lim(A^{n+1}x) = x,$$

and similarly $T(I - A) = I$. Therefore $T = (I - A)^{-1}$. □

Corollary. If E is a Banach space, the set of invertible operators on E is open in $\mathscr{L}(E)$.

Proof. Let $A \in \mathscr{L}(E)$ be invertible. Then for every $B \in \mathscr{L}(E)$ s.t. $\|B\| \leqslant 1/\|A^{-1}\|$, we have that $I + A^{-1}B$ is invertible, since $\|A^{-1}B\| \leqslant \|A^{-1}\| \|B\| < 1$, and $[(I + A^{-1}B)^{-1}A^{-1}](A + B) = (I + A^{-1}B)^{-1}(I + A^{-1}B) = I$, while $(A + B)[(I + A^{-1}B)^{-1}A^{-1}] = A(I + A^{-1}B)[(I + A^{-1}B)^{-1}A^{-1}] = AA^{-1} = I$, so $A + B$ is invertible and it has inverse $(A + B)^{-1} = (I + A^{-1}B)^{-1}A^{-1}$. This means that every invertible element of $\mathscr{L}(E)$ has a nbd of invertible elements, hence the set of invertible operators on E is open in $\mathscr{L}(E)$. □

Inverses of Operators (cont.)

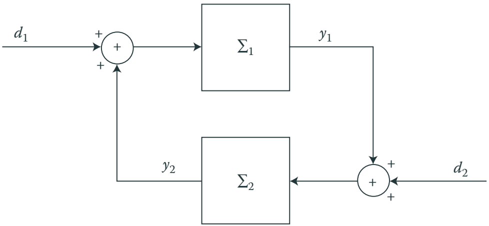
Application to small gain theorem in control, and to structured SVD

A linear discrete-time system $G\colon \ell_2 \to \ell_2$ is *stable w.r.t. the* ℓ_2 *norm* if $G \in \mathscr{L}(\ell_2)$.

The previous theorem allows us to derive a simple sufficient criterion for stability (w.r.t. the ℓ_2 norm) of feedback systems:

Theorem (Small Gain)

Consider two stable (w.r.t. the ℓ_2 norm), causal and linear systems Σ_1, Σ_2 in a feedback interconnection as shown below. The closed loop system, with d_1, d_2 as inputs and y_1, y_2 as outputs, is ℓ_2-stable if $\|\Sigma_1\|\|\Sigma_2\| < 1$.



Application to small gain theorem in control, and to structured SVD (cont.)

Proof. The feedback interconnection yields, $y_2 = \Sigma_2(d_2 + y_1) = \Sigma_2 d_2 + \Sigma_2\Sigma_1 d_1 + \Sigma_2\Sigma_1 y_2$. This means that the closed loop system is stable iff $I - \Sigma_2\Sigma_1$ is invertible, since in that case

$$y_2 = [I - \Sigma_2\Sigma_1]^{-1}(\Sigma_2 d_2 + \Sigma_2\Sigma_1 d_1).$$

The previous theorem tells us that a sufficient condition for $I - \Sigma_2\Sigma_1$ to be invertible is that $\|\Sigma_2\Sigma_1\| < 1$, and this condition is fulfilled if $\|\Sigma_1\|\|\Sigma_2\| < 1$, since $\|\Sigma_2\Sigma_1\| \leqslant \|\Sigma_1\|\|\Sigma_2\|$. □

In multivariable control, Σ_1 may correspond to a feedback loop, while Σ_2 represents a source of uncertainty in the plant being controlled. If only the norm of Σ_2 were known, the small gain theorem states that Σ_1 should satisfy $\|\Sigma_1\|\|\Sigma_2\| < 1$ to ensure stability.

If Σ_2 had a known structure, *e.g.*, $\Sigma_2 = \text{diag}(\delta_1, \dots, \delta_n)$, one can define the *structured singular value* $\mu(\Sigma_1) = \sup\left\{\|\Sigma_2\|^{-1} : \Sigma_2 = \text{diag}(\delta_1, \dots, \delta_n), \|\Sigma_1\Sigma_2\| \geqslant 1\right\}$, so the condition $\mu(\Sigma_1) < 1$ implies that $\|\Sigma_1\Sigma_2\| < 1$ for all $\Sigma_2 = \text{diag}(\delta_1, \dots, \delta_n)$ with $\|\Sigma_2\| < 1$, and thus, by the small gain theorem, (Σ_1, Σ_2) is stable for those Σ_2.

Outline

Adjoint Operators

The transpose of a matrix $A \in \mathbb{R}^{n\times n}$ satisfies $(Ax, y) = y^T Ax = (A^T y)^T x = (x, A^T y)$ for $x, y \in \mathbb{R}^n$.

We can generalize the transpose to general normed spaces:

Theorem. Let $A \in \mathscr{L}(E,F)$, where E,F are normed spaces. Then there is a unique $A^* \in \mathscr{L}(F^*,E^*)$ s.t. $\langle Ax, y^*\rangle_F = \langle x, A^* y^*\rangle_E$ for all $x \in E$, $y^* \in F^*$, and $\|A\| = \|A^*\|$.

Proof. Fix $y^* \in F^*$. $x \mapsto \langle Ax, y^*\rangle_F$ is a linear functional on E. Also, $|\langle Ax, y^*\rangle| \leqslant \|y^*\|\|Ax\| \leqslant \|y^*\|\|A\|\|x\|$, so $x \mapsto \langle Ax, y^*\rangle_F$ is a bounded linear functional, say, $x^* \in E^*$. Define $A^* y^* = x^*$. A^* is unique and linear (*why?*). Furthermore, $|\langle x, A^* y^*\rangle_E| = |\langle Ax, y^*\rangle_F| \leqslant \|y^*\|\|Ax\| \leqslant \|y^*\|\|A\|\|x\|$, so $\|A^* y^*\| \leqslant \|A\|\|y^*\|$, *i.e.*, $\|A^*\| \leqslant \|A\|$, and if $x_0 \in E$ is non-zero, by Corollary 2 of Hahn-Banach, there is a $y_0^* \in F^*$, $\|y_0^*\| = 1$, s.t. $\langle Ax_0, y_0^*\rangle_F = \|Ax_0\|$, so $\|Ax_0\| = |\langle x_0, A^* y_0^*\rangle_E| \leqslant \|A^* y_0^*\|\|x_0\| \leqslant \|A^*\|\|x_0\|$, thus $\|A\| \leqslant \|A^*\|$. Thus, $\|A\| = \|A^*\|$. □

A^* is the *adjoint* of A. It can be shown that, when E,F are reflexive, $A^{**} = A$.

Note. If E,F are inner product spaces, one can also define the *inner product adjoint* of $A \in \mathscr{L}(E,F)$ via $(Ax, y) = (x, A^* y)$ for all $x \in E$, $y \in F$; this differs from the normed adjoint in that $(\alpha A)^* = \overline{\alpha} A^*$ for the inner product adjoint, while $(\alpha A)^* = \alpha A^*$ for the normed adjoint.

Properties of the Adjoint

(1) $I^* = I$.

(2) If $A_1, A_2 \in \mathscr{L}(E,F)$, then $(A_1 + A_2)^* = A_1^* + A_2^*$.

(3) If $A \in \mathscr{L}(E,F)$ and $\alpha \in \mathbb{C}$, then $(\alpha A)^* = \alpha A^*$. For inner product adjoints, $(\alpha A)^* = \overline{\alpha} A^*$.

(4) If $A \in \mathscr{L}(E,F)$, $B \in \mathscr{L}(F,G)$, then $(A_2 A_1)^* = A_1^* A_2^*$.

(5) If $A \in \mathscr{L}(E,F)$ and A has a bounded inverse, then $(A^{-1})^* = (A^*)^{-1}$.

Proof

Properties (1)-(4) are straightforward. Regarding (5), assume $A \in \mathscr{L}(E,F)$ has a bounded inverse A^{-1}. To show that A^* has an inverse, we will establish that A^* is injective and surjective. If $y_1^*, y_2^* \in F^*$, $y_1^* \neq y_2^*$, then $\langle x, A^* y_1^* \rangle - \langle x, A^* y_2^* \rangle = \langle Ax, (y_1^* - y_2^*) \rangle \neq 0$ for some $x \in E$, so $A^* y_1^* \neq A^* y_2^*$ and A^* is injective. Now, given some $x^* \in E^*$, and $x \in E$, $Ax = y$, we have $\langle x, x^* \rangle = \langle A^{-1} y, x^* \rangle = \langle y, (A^{-1})^* x^* \rangle = \langle Ax, (A^{-1})^* x^* \rangle = \langle x, A^* (A^{-1})^* x^* \rangle$, so $x^* \in \mathscr{R}(A^*)$, and also $(A^*)^{-1} = (A^{-1})^*$. $\square$

Adjoint Operators (cont.)

Examples of inner product adjoints

1. Consider the multiplication operator on $L_2[a,b]$, $(M_f x)(t) = f(t)x(t)$:

$$(x, M_f^* y) = (M_f x, y) \quad \Leftrightarrow \quad \int_a^b x(t)\overline{[M_f^* y](t)}dt = \int_a^b f(t)x(t)\overline{y(t)}dt \quad \Leftrightarrow \quad [M_f^* y](t) = \overline{f(t)}y(t).$$

2. Consider the integral operator $K : L_2[a,b] \to L_2[c,d]$ with kernel k. Then

$$\begin{aligned}
(x, K^* y) = (Kx, y) \quad \Leftrightarrow \quad \int_a^b x(t)\overline{[K^* y](t)}dt &= \int_c^d [Kx](t)\overline{y(t)}dt \\
&= \int_c^d \int_a^b k(t,s)x(s)\overline{y(t)}dsdt \\
&= \int_a^b x(s) \int_c^d k(t,s)\overline{y(t)}dtds \\
\Leftrightarrow \quad (K^* y)(t) = \int_c^d \overline{k(s,t)}y(s)ds.&
\end{aligned}$$

3. The adjoint of the shift operator S on ℓ_2 is the backward shift operator S^* (*exercise!*).

Outline

Self-Adjoint and Non-Negative Operators

Definition

Let H be a Hilbert space. $A \in \mathscr{L}(H)$ is *self-adjoint* (or *Hermitian*) if $A = A^*$.
An operator $A \in \mathscr{L}(H)$ is *non-negative* ($A \geqslant 0$) if $(Ax, x) \geqslant 0$ for all $x \in H$, and it is *positive* if, in addition, $(Ax, x) = 0$ implies that $x = 0$. $A \leqslant B$ means that $(Ax, x) \leqslant (Bx, x)$ for all $x \in H$.

Examples

1. The multiplication operator in $L_2[a, b]$ where f is real valued is self-adjoint, and non-negative if $f(x) \geqslant 0$ for all $x \in [a, b]$.
2. The integral operator in $L_2[a, b]$ with kernel k is self-adjoint iff $k(t, s) = \overline{k(s, t)}$, $t, s \in [a, b]$.

Theorem. If $A \in \mathscr{L}(H)$ is self-adjoint, then $\|A\| = \sup_{\|x\|=1} |(Ax, x)|$.

Proof (for real H**).** For every $x \in H$, $\|x\| = 1$, $|(Ax, x)| \leqslant \|Ax\| \|x\| \leqslant \|A\|$, hence $m := \sup_{\|x\|=1} |(Ax, x)| \leqslant \|A\|$. On the other hand, $(A(x \pm y), x \pm y) = (Ax, x) \pm 2(Ax, y) + (y, y)$, so

$$|(Ax, y)| = \frac{1}{4} |(A(x+y), x+y) - (A(x-y), x-y)| \leqslant \frac{m}{4} \left(\|x+y\|^2 + \|x-y\|^2 \right) \leqslant \frac{m}{2} \left(\|x\|^2 + \|y\|^2 \right).$$

Taking $y = (\|x\| / \|Ax\|) Ax$ gives $\|x\| \|Ax\| \leqslant m \|x\|^2$, or $\|Ax\| \leqslant m$ whenever $\|x\| = 1$, so $\|A\| \leqslant m$. □

Theorem. If $A \in \mathscr{L}(H)$, where H is a *complex* Hilbert space, and $(Ax, x) = 0$ for all $x \in H$, then $A = 0$.

Proof. Since $(A(x+y), x+y) = 0$, we have that $(Ay, x) + (Ax, y) = 0$ for all $x, y \in H$. Replacing y by iy yields $i(Ay, x) - i(Ax, y) = 0$, *i.e.*, $(Ay, x) - (Ax, y) = 0$. Adding these expressions gives $(Ay, x) = 0$, which holds for every $x, y \in H$; therefore, $Ay = 0$ for all $y \in H$, *i.e.*, $A = 0$. □

Corollary. If $A \in \mathscr{L}(H)$ is non-negative, where H is a complex Hilbert space, then it is also self-adjoint.

Proof. If $A \in \mathscr{L}(H)$ is non-negative, (Ax, x) is real, so $(x, A^* x) = (Ax, x) = (x, Ax)$, *i.e.*, $(x, [A - A^*]x) = 0$ for every $x \in H$, so by the theorem above, $A = A^*$. □

Outline

Goal: Extend the concept of eigenvalues to linear operators on a Banach space E.

Motivating example: Separation of variables in PDEs
To solve the differential equation $\dot{x}(t) = Ax(t)$, with $x(t) \in \mathbb{R}^n$, one can decompose the matrix A as $A = TDT^{-1}$, where $D = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$ has the eigenvalues of A (assumed distinct) and $T = [v_1 \ \cdots \ v_n]$ the corresponding eigenvectors as columns, which satisfy $Av_k = \lambda_k v_k$ for $k = 1, \dots, n$. Then, re-defining $x(t) = Ty(t)$, one obtains $\dot{y}(t) = Dy(t)$, so $y_k(t) = c_k \exp(\lambda_k t)$ and the general solution is

$$x(t) = c_1 v_1 \exp(\lambda_1 t) + \cdots + c_n v_n \exp(\lambda_n t).$$

Consider now a partial differential equation (PDE) such as

$$\frac{\partial y}{\partial t} = k \frac{\partial^2 y}{\partial x^2} \qquad \text{heat equation in } y(x,t); \quad x, t \in \mathbb{R}$$

subject to an initial condition $y(x,0)$ s.t. $\lim_{x \to \pm\infty} y(x,0) = 0$.

Motivating example: Separation of variables in PDEs (cont.)

This equation can be solved in a similar manner if one consider $\underline{y}(t) = y(\cdot, t)$ as an "infinite-dimensional vector" or function for each fixed t. Then, the PDE can be written as $\dot{\underline{y}} = A\underline{y}$, where A is a linear operator satisfying

$$(A\underline{y}(t))(x) = k\frac{\partial^2 y(x,t)}{\partial x^2}.$$

One can then diagonalize A by solving the equation $Av_\lambda = \lambda v_\lambda$ for $v_\lambda : x \mapsto v_\lambda(x)$, or $kv''_\lambda = \lambda v_\lambda$, which gives $v_\lambda(x) = a_\lambda \exp(\sqrt{\lambda/k}x) + b_\lambda \exp(-\sqrt{\lambda/k}x)$. Under the given initial condition, $\lambda < 0$, so the general solution of the PDE is, informally,

$$y(x,t) = \int_0^\infty \left\{\tilde{a}(\lambda)\exp\left(i\sqrt{-\frac{\lambda}{k}}x\right) + \tilde{b}(\lambda)\exp\left(-i\sqrt{-\frac{\lambda}{k}}x\right)\right\}\exp(-\lambda t)d\lambda,$$

where the functions $\tilde{a}, \tilde{b}$ are determined from the initial condition $y(\cdot, 0)$.

This is the standard method of *separation of variables for solving PDEs*! To formalize it, one needs to extend the notion of eigenvalues and eigenvectors to infinite dimensional spaces.

Spectrum (cont.)

Some operators do not have eigenvalues! (λ's for which $(\lambda I - A)x = 0$ for some $x \neq 0$).
Recall the shift operator S on ℓ_2: $S(x_1, x_2, \dots) = (0, x_1, x_2, \dots)$ If $Sx = \lambda x$, then $x = 0$!

Definition
The *spectrum* of $A \in \mathscr{L}(E)$ is $\sigma(A) := \{\lambda \in \mathbb{C}\colon \lambda I - A$ does not have an inverse in $\mathscr{L}(E)\}$.

$\sigma(A) \neq \emptyset$, and may have not only *eigenvalues*.

Example
Consider the multiplication operator $M_f \in \mathscr{L}(L_2[a,b])$ for an $f \in C[a,b]$. Then $\sigma(M_f) = \mathscr{R}(f)$:

If $\lambda \notin f([a,b])$, then $\lambda I - M_f$ has a bounded inverse $M_{(\lambda - f)^{-1}}$, so $\lambda \notin \sigma(M_f)$. Conversely, if $\lambda = f(t)$ for some $t_0 \in [a,b]$, and $\lambda I - M_f$ had an inverse $T \in L_2[a,b]$, then consider a sequence (x_n) in $L_2[a,b]$, $x_n(t) \geqslant 0$ s.t. $x_n(t) \to 0$ for $t \neq t_0$ and $\int_a^b |x_n(t)|^2 dt = 1$: $(\lambda I - M_f)x_n \to 0$ but $T(\lambda I - M_f)x_n = x_n$, even though $\|x_n\| = 1$! This means that $\lambda = \sigma(M_f)$.

Hence, $\sigma(M_f) = \mathscr{R}(f)$. However, for many f's, M_f does not have eigenvalues (*e.g.*, $f(t) = t$).

Theorem. $\sigma(A)$ is compact, and it is contained in $\overline{B(0, \|A\|)}$.

Proof. Define $F : \mathbb{C} \to \mathscr{L}(E)$ as $F(\lambda) = \lambda I - A$. Since $\|F(\lambda) - F(\mu)\| = |\lambda - \mu|$, F is continuous. Therefore, since $\sigma(A) = F^{-1}(G^c)$, where G is the set of invertible operators in $\mathscr{L}(E)$, which is open, we have that $F^{-1}(G^c)$ is closed.
Let $|\lambda| > \|A\|$. Then, $\|\lambda^{-1}A\| < 1$, so $I - \lambda^{-1}A$ is invertible, and hence $\lambda I - A$ is invertible. Therefore, $\lambda \notin \sigma(A)$. In other words, $\sigma(A) \subseteq \overline{B(0, \|A\|)}$.

Since $\sigma(A)$ is closed and bounded in $\mathbb{C}$, it is compact (by Heine-Borel). $\square$

It can also be shown that $\sigma(A) \neq \emptyset$ using complex analysis: if $\sigma(A) = \emptyset$, pick an $f \in \mathscr{L}(E)^*$ s.t. $f(A^{-1}) \neq 0$. It can be shown that $g(\lambda) = f([\lambda I - A]^{-1})$ is analytic in $\lambda \in \mathbb{C}$. Since $g(\lambda) \to 0$ as $|\lambda| \to \infty$, g is bounded and analytic, so by Liouville's theorem (from complex analysis), $g = 0$, which contradicts the fact that $g(0) = f(A^{-1}) \neq 0$, thus $\sigma(A) \neq \emptyset$.

Self-adjoint and non-negative operators have similar spectral properties to Hermitian and positive semi-definite matrices, which can be deduced using the following lemma:

Lemma. If for a self-adjoint operator $A \in \mathscr{L}(H)$, where H is a Hilbert space, there is a $\delta > 0$ s.t. $\|Ax\| \geqslant \delta \|x\|$ for all $x \in H$, then A is invertible.

Proof. The inequality implies that T is injective (*why?*). Now, $x \in \text{Ker } A$ iff $0 = (Ax, y) = (x, Ay)$ for all $y \in H$, *i.e.*, iff $x \in \mathscr{R}(A)^{\perp}$, so $\mathscr{R}(A)^{\perp} = \{0\}$, that is, $\mathscr{R}(A)$ is dense in H. On the other hand, $\mathscr{R}(A)$ is closed, since if (y_n), $y_n = Ax_n$, is a sequence in $\mathscr{R}(A)$ that converges to, say, $y \in H$, then (y_n) is Cauchy, and so is (x_n) (by the stated inequality), so $x_n \to x \in H$, say, and by continuity $y = Ax \in \mathscr{R}(A)$. Therefore, A is bijective, and its inverse is bounded due to the inequality, so A is invertible. □

Theorem. If $A \in \mathscr{L}(H)$ is self-adjoint, then $\sigma(A) \subseteq \mathbb{R}$. Furthermore, if $A \geqslant 0$, $\sigma(A) \subseteq [0, \infty)$.
Proof. Since $A = A^*$, if $\lambda = a + bi \in \sigma(A)$, then $\|(A - \lambda I)x\|^2 = \|Ax - ax\|^2 + b^2 \|x\|^2$ for every $x \in H$, so $\|(A - \lambda I)x\| \geqslant |b| \|x\|$. If $b \neq 0$, then $A - \lambda I$ by the lemma above, so $\lambda \notin \sigma(A)$ If $A \geqslant 0$, then for every $\lambda < 0$ one has that $|\lambda| \|x\|^2 = (-\lambda x, x) \leqslant ([A - \lambda I]x, x) \leqslant \|(A - \lambda I)x\| \|x\|$ for every $x \in H$, so $|\lambda| \|x\| \leqslant \|(A - \lambda I)x\|$, and by the lemma above $A - \lambda I$ is invertible, hence $\lambda \notin \sigma(A)$. □

The previous result can be strengthened to

Theorem. If $A \in \mathscr{L}(H)$ is self-adjoint, $m := \inf_{\|x\|=1}(Ax, x)$, $M := \sup_{\|x\|=1}(Ax, x)$, then $\sigma(A) \subseteq [m, M]$, and $m, M \in \sigma(A)$.

Proof. Let $\lambda > M$. Since $(Ax, x) \leqslant M(x, x)$ for all $x \in H$, we have that $\|(\lambda I - A)x\| \|x\| \geqslant (\lambda x - Ax, x) \geqslant (\lambda - M)\|x\|^2$, where $\lambda - M > 0$, or $\|(\lambda I - A)x\| \geqslant (\lambda - M)\|x\|$, so $\lambda I - A$ is invertible, *i.e.*, $\lambda \notin \sigma(A)$. Similarly, if $\lambda < m$ then $\lambda \notin \sigma(A)$, so $\sigma(A) \subseteq [m, M]$.

To prove that $M \in \sigma(A)$, consider the bilinear form $a(x, y) := (Mx - Ax, y)$, which is symmetric (because A is self-adjoint) and s.t. $a(x, x) = (Mx, x) - (Ax, x) \geqslant 0$ for all $x \in H$. Cauchy-Schwarz applied to a yields $|a(x, y)| \leqslant \sqrt{a(x, x)}\sqrt{a(y, y)}$, or $|(Mx - Ax, y)| \leqslant \sqrt{(Mx - Ax, x)}\sqrt{(My - Ay, y)}$. Taking sup over $\|y\| = 1$, we obtain

$$\|Mx - Ax\| \leqslant C\sqrt{(Mx - Ax, x)} \text{ for all } x \in H, \qquad (*)$$

where $C = \sup_{\|y\|=1}\sqrt{(My - Ay, y)}$. By definition of M, there is a sequence (x_n) s.t. $\|x_n\| = 1$ and $(Ax_n, x_n) \to M$. From $(*)$, $\|Mx_n - Ax_n\| \to 0$, so $M \in \sigma(A)$, since otherwise $MI - A$ would be invertible, so $x_n = (MI - A)^{-1}(Mx_n - Ax_n) \to 0$, a contradiction. Similarly, $m \in \sigma(A)$. □

Corollary. If $A \in \mathscr{L}(H)$ is self-adjoint and $\sigma(A) \subseteq [0, \infty)$, then A is non-negative.

Outline

Infinite Matrices

Linear operators in infinite dimensions can be represented by infinite matrices, resembling linear algebra.

Definition. Let E,F be separable Hilbert spaces, and $A \in \mathscr{L}(E,F)$. The *matrix* of A with respect to orthonormal bases (e_n) and (f_n) of E,F, respectively, is the array $[a_{jk}]_{j,k=1}^{\infty}$ of complex numbers given by $a_{jk} = (Ae_k, f_j)$.

It is difficult to determine from a matrix representation if an operator is bounded.

Example (Linear system)
Let $k \in C[-\pi,\pi]$ be 2π-periodic, and consider the integral operator K on $L_2[-\pi,\pi]$ given by

$$(Kx)(t) = \int_{-\pi}^{\pi} k(t-s)x(s)ds.$$

If $(e_n)_{n\in\mathbb{Z}}$ denotes the Fourier basis of $L_2[-\pi,\pi]$, then

$$(Ke_n)(t) = \int_{-\pi}^{\pi} k(t-s)e_n(s)ds = \frac{1}{\sqrt{2\pi}}\int_{-\pi}^{\pi} k(s-t)e^{ins}ds = \frac{e^{int}}{\sqrt{2\pi}}\int_{-\pi}^{\pi} k(\tau)e^{-in\tau}d\tau = c_n e_n(t),$$

where c_n is the n-th Fourier coefficient of k. Therefore, the matrix of K with respect to (e_n) is $[a_{jk}]$ with $a_{jk} = (Ae_k, e_j) = c_k\delta_{j-k}$:

$$[A] = \begin{bmatrix} \ddots & & & \\ & c_{-1} & & 0 \\ & & c_0 & \\ & 0 & & c_1 \\ & & & & \ddots \end{bmatrix}. \qquad \text{(diagonal matrix)}$$

Next Topic

Optimization of Functionals

Outline

Bonus: Applications of the Adjoint

Let $A \in \mathscr{L}(E,F)$, where E,F are Hilbert spaces.

Theorem. Let $y \in F$. Then the vector $x \in E$ minimizes $\|y - Ax\|$ iff $A^* Ax = A^* y$.
Proof. By the projection theorem, $x \in E$ minimizes $\|y - Ax\|$ iff $(y - Ax, A\tilde{x}) = 0$ for all $\tilde{x} \in E$. However, $(y - Ax, A\tilde{x}) = (A^*[y - Ax], \tilde{x})$, so the latter holds iff $A^*[y - Ax] = 0$. □

Theorem (Fredholm Alternative). $[\mathscr{R}(A)]^\perp = \text{Ker } A^*$.
Proof. $x \in \text{Ker } A^*$ iff $A^* x = 0$, *i.e.*, iff $(x, Ay) = (A^* x, y) = 0$ for all y, that is, iff $x \in [\mathscr{R}(A)]^\perp$. □

Corollary. Assume that $\mathscr{R}(A^*)$ is closed and $y \in \mathscr{R}(A)$. The vector $x \in E$ of minimum norm s.t. $Ax = y$ is given by $x = A^* z$, where $z \in E$ is any solution of $AA^* z = y$.
Proof. Every $x \in E$ satisfying $Ax = y$ is of the form $x = x_0 + m$, where $Ax_0 = y$ and $m \in \text{Ker } A$. By Fredholm's Alternative, $\text{Ker } A = [\mathscr{R}(A^*)]^\perp$, and by the minimum norm theorem, the sought $x \in E$ satisfies $x \perp [\mathscr{R}(A^*)]^\perp$, or $x \in [\mathscr{R}(A^*)]^{\perp\perp} = \mathscr{R}(A^*)$ (since $\mathscr{R}(A^*)$ is closed), so $x = A^* z$ for some $z \in E$, and plugging this expression into $Ax = y$ gives $AA^* z = y$. □

Example (control)
Consider a linear system of the form $\dot{x}(t) = Ax(t) + Bu(t)$. We want to drive $x(0) = 0$ to $x(T) = x_0$ by designing a control input $u(t)$ of minimum energy $\displaystyle\int_0^T u^2(t)dt$.

Let $u \in L_2[0,T]$. We know that $x(T) = \int_0^T e^{A(T-t)}Bu(t)dt$, so let us define an operator $\Phi: L_2[0,T] \to \mathbb{R}^n$ as

$$\Phi u = \int_0^T e^{A(T-t)}Bu(t)dt.$$

The problem is to find a $u \in L_2[0,T]$ of minimum norm s.t. $\Phi u = x_0$. Since $\mathscr{D}(\Phi^*) = \mathbb{R}^n$, the range of Φ^* is finite dimensional, and hence it is closed, so by the last corollary we have that the optimal solution is $u^{\text{opt}} = \Phi^* z$, where $\Phi\Phi^* z = x_0$

...so we need expressions for Φ^* and $\Phi\Phi^*$.

Example (control) (cont.)

For every $u \in L_2[0,T]$ and $y \in \mathbb{R}^n$,

$$(\Phi u, y) = y^T \int_0^T e^{A(T-t)} B u(t) dt = \int_0^T y^T e^{A(T-t)} B u(t) dt = (u, \Phi^* y),$$

so $(\Phi^* y)(t) = B^T e^{A^T(T-t)} y$, and

$$\Phi\Phi^* y = \int_0^T e^{A(T-t)} B B^T e^{A^T(T-t)} y dt = \underbrace{\int_0^T e^{A(T-t)} B B^T e^{A^T(T-t)} dt}_{\in \mathbb{R}^{n\times n} \text{ (Controllability Gramian)}} y.$$

The optimal control is given by

$$u^{\text{opt}}(t) = (\Phi^*[\Phi\Phi^*]^{-1} x_0)(t) = B^T e^{A^T(T-t)} \left[\int_0^T e^{A(T-\tau)} B B^T e^{A^T(T-\tau)} d\tau \right]^{-1} x_0,$$

assuming that the inverse exists. Notice that $\mathscr{R}(\Phi\Phi^*)$ corresponds to the states reachable from the origin in T seconds/minutes/..., and that $\mathscr{R}(\Phi\Phi^*) = \mathscr{R}(\Phi)$ (*why?*).

Together with the Hahn-Banach theorem, the Uniform Boundedness principle, the Closed-Graph theorem and the Open Mapping theorem are considered to be the cornerstones of Banach space theory.

Theorem (Uniform Boundedness Principle / Banach-Steinhaus)
Let $\mathscr{F}$ be a family of bounded linear operators from a Banach space X to a normed space Y. If $\sup_{A \in \mathscr{F}} \|Ax\| < \infty$ for every $x \in X$, then $\sup_{A \in \mathscr{F}} \|A\| < \infty$.

Proof. Assume that $\sup_{A \in \mathscr{F}} \|A\| = \infty$, and choose a sequence (A_n) in $\mathscr{F}$ s.t. $\|A_n\| \geqslant 4^n$. Set $x_0 = 0 \in X$ and, for $n \in \mathbb{N}$, choose $x_n \in X$ as follows: note that for every $\|\xi\| \leqslant 3^{-n}$,

$$\max\{\|A_n(x_{n-1}+\xi)\|, \|A_n(x_{n-1}-\xi)\|\} \geqslant \frac{1}{2}\|A_n(x_{n-1}+\xi)\| + \frac{1}{2}\|A_n(x_{n-1}-\xi)\| \geqslant \|A_n\xi\|,$$

so taking sup over $\|\xi\| \leqslant 3^{-n}$ shows that there is a $\|\xi_n\| \leqslant 3^{-n}$ s.t., say, $\|A_n(x_{n-1}+\xi_n)\| \geqslant (2/3)3^{-n}\|A_n\|$; choose $x_n = x_{n-1} + \xi_n$. On the other hand, (x_n) is a Cauchy sequence (*why?*), which converges to, say, $x \in X$, and in addition, $\|x - x_n\| \leqslant (1/2)3^{-n}$, hence

$$\|A_n x\| = \|A_n(x - x_n) + A_n x_n\| \geqslant |\|A_n x_n\| - \|A_n(x - x_n)\|| \geqslant \left|\frac{2}{3}3^{-n}\|A_n\| - \frac{1}{2}3^{-n}\|A_n\|\right| \geqslant \frac{1}{6}(4/3)^n,$$

which tends to ∞ as $n \to \infty$. □

Application to divergence of Fourier series

From Topic 5, the Fourier series of an $f \in C[-\pi,\pi]$, truncated to N terms, is

$$f_N(x) = \sum_{n=-N}^{N} (f, e_n) e_n(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x+y) D_N(y) dy, \qquad D_N(y) := \frac{\sin([N+1/2]y)}{\sin(y/2)}.$$

Define $T_N : C[-\pi,\pi] \to \mathbb{R}$ by $T_N f = f_N(0) = (2\pi)^{-1} \int_{-\pi}^{\pi} f(y) D_N(y) dy$, whose norm is $\|T_N\| = (2\pi)^{-1} \int_{-\pi}^{\pi} |D_N(y)| dy$. However,

$$\int_{-\pi}^{\pi} |D_N(y)| dy = \int_{-\pi}^{\pi} \left| \frac{\sin([N+1/2]y)}{\sin(y/2)} \right| dy \geqslant 4 \int_0^{\pi} \left| \frac{\sin([N+1/2]y)}{y} \right| dy = 4 \int_0^{(N+1/2)\pi} |\sin(y)| \frac{dy}{y}$$
$$> 4 \sum_{k=1}^{N} \frac{1}{k\pi} \int_{(k-1)\pi}^{k\pi} |\sin(y)| dy = \frac{4}{\pi} \sum_{k=1}^{N} \frac{1}{k} \to \infty \quad \text{as } N \to \infty,$$

so by the uniform boundedness principle: *there is an* $f \in C[-\pi,\pi]$ *s.t.* $f_N(0)$ *diverges*.

Bonus: Closed Graph Theorem

Definitions

- The *graph* of a function $T: \mathscr{D}(T) \subseteq X \to Y$ is $\mathscr{G}(T) = \{(x, T(x)) \in X \times Y : x \in \mathscr{D}(T)\}$. If X, Y are vector spaces and T is linear, then $\mathscr{G}(T)$ is a linear subspace of $X \times Y$.
- If X, Y are normed spaces, a norm can be introduced in $X \times Y$, *e.g.*, $\|(x, y)\| = \|x\| + \|y\|$. An operator $T: \mathscr{D}(T) \subseteq X \to Y$ is *closed* if $\mathscr{G}(T)$ is closed in $X \times Y$; equivalently, T is closed iff whenever (x_n) is a sequence in $\mathscr{D}(T)$ s.t. $x_n \to x \in \mathscr{D}(T)$ and $y_n := T(x_n) \to y \in Y$, then $y = T(x)$.
- An *adjoint* of a linear (but not necessarily bounded) operator $T: \mathscr{D}(T) \subseteq X \to Y$ is an operator $T^*: \mathscr{D}(T^*) \subseteq Y^* \to X^*$ s.t. $\langle Tx, y^* \rangle = \langle x, T^* y^* \rangle$ for all $x \in \mathscr{D}(T)$, $y^* \in \mathscr{D}(T^*)$. Adjoints in general are non-unique, unless $\mathscr{D}(T)$ is dense in X, and $\mathscr{D}(T^*)$ consists of those $y^* \in Y^*$ for which $x \mapsto \langle Tx, y^* \rangle$ is bounded on $\mathscr{D}(T)$.

If $T: \mathscr{D}(T) \to Y$ is linear and closed, where X, Y are Banach spaces, $\mathscr{D}(T)$ is itself a Banach space under the *graph norm* $\|x\|_g := \|x\| + \|T(x)\|$, since $x \mapsto (x, T(x))$ is an isometry from $\mathscr{D}(T)$ to $\mathscr{G}(T)$, which is complete (*why?*). Also, T is bounded under this norm.

As $\langle (x, -Tx), (T^* y^*, y^*) \rangle = \langle x, T^* y^* \rangle - \langle Tx, y^* \rangle = 0$, $\mathscr{G}'(T^*) = \mathscr{G}(-T)^\perp$ if $\mathscr{D}(T) \subseteq X$ is dense, where $\mathscr{G}'(T^*) := \{(T^* y^*, y^*) : y^* \in \mathscr{D}(T^*)\}$ is the *reversed graph* of T^*, so T^* is always closed.

Lemma. Let $T: X \to Y$ be linear and closed, where X, Y are Banach spaces. Then, $\mathscr{D}(T^*) = Y^*$.

Proof. First we will show that $\mathscr{D}(T^*)$ is weak*-dense in Y^*. If not, there is a $y \in Y \setminus \{0\}$ s.t. $\langle y, y^* \rangle = 0$ for all $y^* \in \mathscr{D}(T^*)$. But then $(0, y) \in {}^{\perp}\mathscr{G}'(-T^*) = \mathscr{G}(T)$ (since $\mathscr{G}(T)$ is closed), *i.e.*, $T(0) = y \neq 0$, which is impossible because T is linear.
Next we will show that $\mathscr{D}(T^*)$ is weak*-closed, which implies that $\mathscr{D}(T^*) = Y^*$. By Krein-Smulian, it suffices to show that $V = \mathscr{D}(T^*) \cap \{y^* \in Y^* : \|y^*\| \leqslant 1\}$ is weak*-closed. Now, $\sup_{y^* \in V} |\langle x, T^* y^* \rangle| = \sup_{y^* \in V} |\langle Tx, y^* \rangle| \leqslant \|Tx\|$, hence $\sup_{y^* \in V} \|T^* y^*\| =: K < \infty$ by uniform boundedness. Thus, $|\langle Tx, y^* \rangle| = |\langle x, T^* y^* \rangle| \leqslant K\|x\|$ for all $x \in X$, $y^* \in V$; since $y^* \mapsto \langle Tx, y^* \rangle$ is weak*-continuous, $|\langle Tx, y^* \rangle| \leqslant K\|x\|$ for all y^* in the weak*-closure of V, $\overline{V}$, *i.e.*, $x \mapsto \langle Tx, y^* \rangle$ is bounded on $\overline{V}$, so V is weak*-closed. □

Theorem (Closed graph theorem)

Let $T: X \to Y$ be linear and closed, where X, Y are Banach spaces. Then, T is bounded.

Proof. Assume T is unbounded. Then, there is a (x_n) in X, $\|x_n\| = 1$, s.t. $\|Tx_n\| \to \infty$, but $\sup_n |\langle Tx_n, y^* \rangle| = \sup_n |\langle x_n, T^* y^* \rangle| \leqslant \|T^* y^*\|$. Thus, (Tx_n) is a point-wise bounded but norm-unbounded family in X^{**}, which contradicts uniform boundedness. Thus, T is bounded. □

Corollary (Hellinger-Toeplitz theorem)

Let $T: H \to H$ be a linear self-adjoint operator in a Hilbert space H. Then, T is bounded.

Proof. Let (x_n) is in H, s.t. $x_n \to x \in H$ and $Tx_n \to y \in H$. For every $z \in H$, $(Tx, z) = (x, Tz) = \lim (x_n, Tz)$ $= \lim (Tx_n, z) = (y, z)$, so $Tx = y$ and T is closed. Then, by the closed graph theorem, T is bounded. □

Theorem (Banach inverse theorem)
Let $T \in \mathscr{L}(X,Y)$, where X,Y are Banach spaces. If T is bijective, then T^{-1} is continuous.

Proof. Since $T\colon X \to Y$ is bounded, its graph $\mathscr{G}(T)$ is closed in $X \times Y$: indeed, if (x_n) is a sequence in X converging to, say, $x \in X$, and (y_n), where $y_n = Tx_n$, converges to, say, $y \in Y$, then by continuity $y = Tx$, so $\mathscr{G}(T)$ is closed. Then, $\mathscr{G}(T^{-1}) = \mathscr{G}'(T)$ is closed in $Y \times X$, and by the closed graph theorem, T^{-1} is continuous. □

Corollary (Open mapping / Banach-Schauder)
Let $T \in \mathscr{L}(X,Y)$ be surjective, where X,Y are Banach spaces. Then, T is an *open mapping*, *i.e.*, $T(U)$ is open in Y whenever U is open in X.

Proof. Define an equivalence relation on X, where $x \sim y$ iff $x - y \in \text{Ker } T$. Since T is bounded, $\text{Ker } T \subseteq X$ is closed, so the set of equivalence classes, $X/\text{Ker } T$, is a Banach space with norm $\|[x]\| := \inf_{k \in \text{Ker } T} \|x + k\|$ (*exercise!*). T induces a bijective bounded linear operator $\bar{T}\colon X/\text{Ker } T \to Y$ by $\bar{T}([x]) = T(x)$, so by the Banach inverse theorem, $\bar{T}^{-1}$ is continuous, *i.e.*, $\bar{T}$ maps open sets onto open sets. Also, $T = \bar{T} \circ \pi$, where $\pi\colon X \to X/\text{Ker } T$, given by $\pi(x) = [x]$, is linear, surjective and open (because if $\|[x-y]\| < \varepsilon$, then $\varepsilon > \inf_{m \in \text{Ker } T} \|x - y - m\|$, so there is an $m^* \in \text{Ker } T$ such that $\|x - y - m^*\| < \varepsilon$, thus $B([x], \varepsilon) \subseteq \pi(B(x, \varepsilon))$), and the composition of open maps is open, hence T is open. □

Spectral theorems correspond to a class of results that allow one to "diagonalize" a linear operator (thus resembling the eigenvalue decomposition result from linear algebra). Here we will establish one version for self-adjoint operators, based on the following facts:

(1) *Bounded monotone sequences of self-adjoint operators converge to a self-adjoint operator.*

Assume $0 \leqslant A_1 \leqslant A_2 \leqslant \cdots \leqslant I$, and let $B = A_{n+k} - A_n$ for some $n, k \in \mathbb{N}$. Note that $0 \leqslant B \leqslant I$, so Cauchy-Schwarz applies to the bilinear form (Bx, y); in particular, $(Bx, Bx)^2 \leqslant (Bx, x)(B^2x, Bx) \leqslant (Bx, x)(Bx, Bx)$, so $\|Bx\|^2 = (Bx, Bx) \leqslant (Bx, x)$. Thus, $\|A_{n+k}x - A_n x\|^2 \leqslant (A_{n+k}x, x) - (A_n x, x)$ for every $x \in H$. Now, since $((A_n x, x))_{n\in\mathbb{N}}$ is a bounded monotone sequence in $\mathbb{R}$, it converges, so $(A_n x)$ is Cauchy in H, and $\lim_{n\to\infty} A_n x = Ax$ exists. A is linear, and by uniform boundedness, it is bounded. Furthermore, letting $n \to \infty$ in $(A_n x, y) = (x, A_n y)$ shows that A is self-adjoint. □

Let $\mathbb{R}[t]$ ($\mathbb{C}[t]$) be the set of polynomials in t with real (complex) coefficients. If $p \in \mathbb{C}[t]$, where $p(t) = p_n t^n + p_{n-1} t^{n-1} + \cdots + p_1 t + p_0$, one can define, for every $A \in \mathscr{L}(H)$,

$$\tilde{p}(A) = p_n A^n + p_{n-1} A^{n-1} + \cdots + p_1 A + p_0 I.$$

(2) *Every operator $A \geqslant 0$ has a unique non-negative square root $A^{1/2}$: $(A^{1/2})^2 = A$.*
Firstly, we can assume w.l.o.g., by scaling A, that $0 \leqslant A \leqslant I$. Consider the sequence of operators $(T_n)_{n\in\mathbb{N}}$ given by $T_1 = 0$ and $T_{n+1} = T_n + (1/2)[A - T_n^2]$ for $n \in \mathbb{N}$. Note that $0 = T_1 \leqslant I$, $T_2 - T_1 = (1/2)A \geqslant 0$, and that if $0 \leqslant T_n \leqslant I$ and $T_n \leqslant T_{n+1}$, then $I - T_n \geqslant 0$, so $0 \leqslant (1/2)(I - T_n)^2 + (1/2)(I - A) = I - T_n - (1/2)(A - T_n^2) = I - T_{n+1}$, *i.e.*, $T_{n+1} \leqslant I$, and $T_{n+2} - T_{n+1} = T_{n+1} + (1/2)[A - T_{n+1}^2] - T_n - (1/2)[A - T_n^2] = (1/2)(T_{n+1} - T_n)(I - T_{n+1} + I - T_n) \geqslant 0$, so $T_{n+1} \leqslant T_{n+2}$. Hence, from (1), $T_n \to T$, where $T = T + (1/2)[A - T^2]$, or $T^2 = A$. Let $A^{1/2} := T$.
Consider another operator $B \geqslant 0$ s.t. $B^2 = A$. Then, $BA = B^3 = AB$, so $BA^n = A^n B$ for every $n \in \mathbb{N}$, thus $BT_n = T_n B$, and taking $n \to \infty$, $BA^{1/2} = A^{1/2}B$. Let $M = (A^{1/2})^{1/2}$ and $N = B^{1/2}$. Then, given $x \in H$, let $y = (A^{1/2} - B)x$. We have that $\|My\|^2 + \|Ny\|^2 = (M^2 y, y) + (N^2 y, y) = ([A^{1/2} + B]y, y) = ([A - B^2]x, y) = 0$, so $My = Ny = 0$ and $M^2 y = N^2 y = 0$, *i.e.*, $A^{1/2}y = By = 0$, so $\|(A^{1/2} - B)x\|^2 = ([A^{1/2} - B]^2 x, x) = ([A^{1/2} - B]y, x) = 0$, that is, $A^{1/2} = B$. □

(3) *Let A,B be commuting non-negative, linear, bounded operators. Then, $AB \geqslant 0$.*
From the proof of (2), since $AB = BA$, also $AB^{1/2} = B^{1/2}A$ holds. Thus, for all $x \in H$, $(ABx, x) = (AB^{1/2}B^{1/2}x, x) = (B^{1/2}AB^{1/2}x, x) = (AB^{1/2}x, B^{1/2}x) \geqslant 0$. □

The map $\phi\colon \mathbb{C}[t] \to \mathscr{L}(H)$ given by $\phi(p) = \tilde{p}(A)$ is linear, *multiplicative* (*i.e.*, $\phi(pq) = \phi(p)\phi(q)$) and *unital* (*i.e.*, $\phi(1) = I$). ϕ is also *order-preserving*:

(4) *If $p \in \mathbb{R}[t]$ satisfies $p(t) \geqslant 0$ for all $t \in [m, M]$, and the self-adjoint operator A satisfies $mI \leqslant A \leqslant MI$, then $\tilde{p}(A) \geqslant 0$.*
p can be factorized as $p(t) = c\prod_j (t - \alpha_j)\prod_k(\beta_k - t)\prod_l \left[(t-\gamma_l)^2 + \delta_l^2\right]$, where $c > 0$, $\alpha_j \leqslant m \leqslant M \leqslant \beta_k$ and $\gamma_l, \delta_l \in \mathbb{R}$. By (3), we have that $\tilde{p}(A) \geqslant 0$. □

Corollary. The map ϕ can be extended to $C[m, M]$. Moreover, if $f \in C[m, M]$, $\|\tilde{f}(A)\| \leqslant \|f\|$.
Proof. Since $\mathbb{C}[t]$ is dense in $C[m, M]$, ϕ can be extended uniquely by continuity. The inequality follows because, for every $p \in \mathbb{C}[t]$, $\|p\| \pm p$ is a non-negative polynomial in $[m, M]$, so $\|p\|I \geqslant \pm\tilde{p}(A)$, *i.e.*, $\|p\| \geqslant \|\tilde{p}(A)\|$; this inequality extends by continuity to $C[m, M]$. □

The extension of ϕ to $C[m, M]$ defines a *functional calculus* for operators, *i.e.*, given a self-adjoint $A \in \mathscr{L}(H)$, and $f \in C[m, M]$, $\tilde{f}(A)$ is another self-adjoint operator in H.

Given a self-adjoint operator $A \in \mathscr{L}(H)$, where H is a separable Hilbert space, a *cyclic vector* of A is an element $\xi \in H$ s.t. $\text{lin}\{A^k\xi : k \in \mathbb{N}_0\} = \text{lin}\{\tilde{p}(A)\xi : p \in \mathbb{C}[t]\}$ is dense in H.

Next we present a version of the Spectral Theorem for self-adjoint operators in a separable Hilbert space:

Spectral Theorem

If the self-adjoint operator $A \in \mathscr{L}(H)$, where H is a separable Hilbert space, has a cyclic vector ξ, then there is a unitary operator $U: H \to L_2(l)$ identifying H with $L_2(l)$ for some $l \in C[m,M]^*$, s.t. $UAU^* = M_t$, where $M_t: L_2(l) \to L_2(l)$ is the multiplication operator $(M_t x)(t) = t\,x(t)$ for $t \in [m,M]$, and $m, M \in \mathbb{R}$ are s.t. $m\|x\|^2 \leqslant (Ax,x) \leqslant M\|x\|^2$ for all $x \in H$.

$L_2(l)$ is the completion of $C[m,M]$, with inner product $(f,g) = l(f\overline{g})$, where $l \in C[m,M]^*$ is *positive* (*i.e.*, $l(f) \geqslant 0$ if $f(t) \geqslant 0$ for all $t \in [m,M]$). To ensure that $(f,f) > 0$ if $f \neq 0$, one actually considers $C[m,M]/N$ instead of $C[m,M]$, where $N = \{f \in C[m,M] : l(\tilde{f}^2) = 0\}$.

An operator $A \in \mathscr{L}(E,F)$ is *unitary* if $AA^* = A^*A = I$; thus, $(Ax,Ay)_F = (x,y)_E$ for all $x, y \in E$.

Proof. Define the linear functional $l \in C[m,M]^*$ by $l(f) := (\tilde{f}(A)\xi,\xi)$ for all $f \in C[m,M]$. Note that $l \geqslant 0$, since $f(A) \geqslant 0$ if $f(x) \geqslant 0$ on $[m,M]$, and that $(f,g) := l(f\overline{g}) = (\tilde{f}(A)\xi, \tilde{g}(A)\xi)$ defines an inner product in $C[m,M]/N$, where $N = \{f \in C[m,M] \colon l(\tilde{f}^2) = 0\}$. Denote by $L_2(l)$ the completion of $C[m,M]/N$.
Define the operator $U \colon H \to L_2(l)$ by $U\tilde{p}(A)\xi = p$ for all $p \in \mathbb{C}[t]$, which specifies it on a dense set of H (since ξ is cyclic). This operator is well defined, since $\tilde{p}_1(A)\xi = \tilde{p}_2(A)\xi$ iff $0 = \|\tilde{p}_1(A)\xi - \tilde{p}_2(A)\xi\|^2 = l([p_1 - p_2]^2)$, *i.e.*, $p_1 - p_2 \in N$. Also, U has the following properties:

(1) U is *isometric*: $(U\tilde{p}_1(A)\xi, U\tilde{p}_2(A)\xi)_H = (p_1, p_2)$ for every $p_1, p_2 \in \mathbb{C}[t]$.

(2) $\mathscr{R}(U)$ is dense in $L_2(l)$, since is contains all polynomials in $[m,M]$ modulo N. This property, together with (1), show that the extension of U to H by continuity is a unitary operator.

(3) $(UA\tilde{p}(A)\xi)(t) = tp(t) = t(U\tilde{p}(A)\xi)(t)$, so, by the density of the polynomials and the cyclic nature of ξ, $UAv = M_t Uv$ for all $v \in H$, *i.e.*, $UAU^* = M_t$. Note in particular that $U\xi = 1$. □

Note. Assuming that A has a cyclic vector is not very restrictive, since otherwise one can pick a ξ_1 from a complete orthonormal sequence (e_n) in H, and define $H_1 = \text{clin}\{A^n\xi \colon n \in \mathbb{N}\}$; if $H_1 \neq H$, apply iteratively this procedure to $(H_1 \oplus \cdots \oplus H_{k-1})^\perp$, so H can be written as a countable direct sum, $H = H_1 \oplus H_2 \oplus \cdots$. The spectral theorem can then be applied to each of these subspaces individually. By transfinite induction, it can be further extended to general (non-separable) Hilbert spaces.

Bonus: Generalized Functions

In many applications, the concept of a function needs to be extended in order to define solutions of some functional (*e.g.*, ordinary/partial differential) equations.

Example. The original motivation for P. Dirac to define his *delta function* was to find eigenvectors of the linear operator $T\colon C([a,b]) \to C([a,b])$ given by $(Tf)(x) = xf(x)$ ($f \in C([a,b])$). Such an eigenvector $f_\lambda \in C([a,b])$ should be the solution of $(x-\lambda)f_\lambda(x) = 0$ for all $x \in [a,b]$, which implies that $f(x) = 0$ for all $x \neq \lambda$, *i.e.*, $f = 0$. Thus, T has no eigenvalues (in the ordinary sense).

Alternatively, we can notice that the eigenvalue equation is equivalent to

$$\int_a^b f_\lambda(x)(x-\lambda)\phi(x)dx = 0, \quad \text{for all } \phi \in C^\infty([a,b]),$$

where can be interpreted as finding a linear functional $\delta \in C^\infty([a,b])^*$ s.t. $\delta[(T-\lambda\mathbb{1})\phi] = 0$ for all $\phi \in C^\infty([a,b])$. Now, every ϕ can be written as $\phi(x) = (x-\lambda)\tilde{\phi}(x) + \phi(\lambda)$, where $\tilde{\phi} \in C^\infty([a,b])$, so

$$\delta(\phi) = \delta[(T-\lambda\mathbb{1})\tilde{\phi} + \phi(\lambda)\mathbb{1}] = \phi(\lambda)\delta(\mathbb{1}).$$

If we define $\delta(\mathbb{1}) = 1$, then δ should satisfy $\delta(\phi) = \phi(\lambda)$. This is the definition of a Dirac delta at λ!

Note that δ cannot be related to a normal function f, as $\delta(\phi) = \int_a^b f(x)\phi(x)dx$ is not possible; it is a *generalized function*. Thus, it is not possible to “evaluate” f (or δ) at a point in $[a,b]$, but it only makes sense as an extension of expressions like $\int_a^b f(x)\phi(x)dx$.

Formalization

Let K be the vector space of all infinitely differentiable functions $\phi\colon \mathbb{R} \to \mathbb{R}$ of compact support (*i.e.*, $\phi(x) = 0$ for all $x \in \mathbb{R}$ outside some compact set). K is not a normed space, but it becomes a topological vector space when endowed with this notion of convergence:

Definition. A sequence (ϕ_n) in K *converges* to $\phi \in K$ if there is a compact set $C \subset \mathbb{R}$ s.t.

- $\phi_n(x) = 0$ for all $x \in C^c$ and $n \in \mathbb{N}$, and
- the sequences $(\phi_n^{(k)})_n$ converge uniformly on C to $\phi^{(k)}$ for all $k \in \mathbb{N}_0$.

K, with this notion of convergence, is a *test space*, and its elements are *test functions*.

Definition. A continuous linear functional ℓ on the test space K is a *generalized function* (or *distribution*) on $\mathbb{R}$, where "continuity" means that if $\phi_n \to \phi$ in K then $\ell(\phi_n) \to \ell(\phi)$.

Example. The *Dirac delta function* δ_λ ($\lambda \in \mathbb{R}$) is a generalized function on K, defined by $\delta_\lambda(\phi) = \phi(\lambda)$ for all $\phi \in K$ (*Exercise:* prove that it is a continuous linear functional).

If a generalized function ℓ can be written as $\ell(\phi) = T_f(\phi) := \int_{-\infty}^{\infty} f(x)\phi(x)dx$ ($\phi \in K$), where $f: \mathbb{R} \to \mathbb{R}$ is *locally integrable* (*i.e.*, integrable on every bounded interval), then ℓ is a *regular generalized function*. Otherwise, ℓ is a *singular generalized function*; *e.g.*, δ_λ is singular.

Convergence of generalized functions. A sequence (ℓ_n) of generalized functions is said to converge to ℓ if $\ell_n(\phi) \to \ell(\phi)$ for every $\phi \in K$.
The space of generalized functions, with this notion of convergence, is the topological vector space K^*.

Operations on generalized functions. Since generalized functions are linear functionals, their addition and scalar multiplication can be directly defined as

$$[\ell_1 + \ell_2](\phi) := \ell_1(\phi) + \ell_2(\phi)$$
$$[\lambda\ell](\phi) := \lambda\ell(\phi), \qquad \phi \in K,$$

for $\ell_1, \ell_2, \ell \in K^*$, and $\lambda \in \mathbb{R}$.

If $f : \mathbb{R} \to \mathbb{R}$ is locally integrable and $\alpha \in C^\infty(\mathbb{R})$, then $T_{\alpha f}(\phi) = \int_{-\infty}^{\infty} \alpha(x) f(x)\phi(x)dx = \int_{-\infty}^{\infty} f(x)\alpha(x)\phi(x)dx = T_f(\alpha\phi)$ for $\phi \in K$, which motivates the definition of $\alpha\ell \in K^*$ as

$$[\alpha\ell](\phi) := \ell(\alpha\phi), \quad \alpha \in C^\infty(\mathbb{R}), \phi \in K.$$

Note. The multiplication of generalized functions is in general **not** well defined!

Bonus: Generalized Functions (cont.)

Differentiation. If $f : \mathbb{R} \to \mathbb{R}$ is locally integrable and differentiable, and $\phi \in K$, then by integration by parts

$$T_{f'}(\phi) = \int_{-\infty}^{\infty} f'(x)\phi(x)dx = f(x)\phi(x)\Big|_{x\to-\infty}^{x\to\infty} - \int_{-\infty}^{\infty} f(x)\phi'(x)dx = -\int_{-\infty}^{\infty} f(x)\phi'(x)dx = -T_f(\phi'),$$

which motivates the definition of the derivative $\ell' \in K^*$ of $\ell \in K^*$ as

$$\ell'(\phi) := -\ell(\phi'), \quad \phi \in K.$$

Since ϕ is infinitely differentiable, every generalized function is infinitely differentiable. Also, if $\ell_n \to \ell$, then we also have that $\ell_n^{(k)} \to \ell^{(k)}$ for every $k \in \mathbb{N}$ (*why?*).

Example. Consider the step function $\mu : \mathbb{R} \to \mathbb{R}$,

$$\mu(x) = \begin{cases} 0, & x < 0 \\ 1, & x \geqslant 0. \end{cases}$$

Then, for each $\phi \in K$, $T_\mu(\phi) = \int_0^\infty \phi(x)dx$, so $T'_\mu(\phi) = -\int_0^\infty \phi'(x)dx = \phi(0) = \delta_0(\phi)$. Thus, the derivative of the step function is a Dirac delta at 0!

Generalized functions are very useful in the theory of ordinary and partial differential equations, as well as in Fourier analysis (where Dirac deltas abound). Also, they can be defined on other sets like $\mathbb{R}^n$, $\mathbb{C}$ and $\mathbb{T}$!

Bonus: Application to SOS Optimization

Motivation: Minimization of (*non-convex*) polynomials subject to polynomial constraints:

$$\begin{array}{cl} \min\limits_{x=(x_1,\ldots,x_n)} & p_0(x) \\ \text{s.t.} & p_k(x) \geqslant 0, \quad k=1,\ldots,m \end{array} \quad \Leftrightarrow \quad \begin{array}{cl} \min\limits_{t\in\mathbb{R}} & t \\ \text{s.t.} & t - p_0(x) \geqslant 0 \text{ for all } x \text{ s.t. } p_k(x) \geqslant 0,\ k=1,\ldots,m. \end{array}$$

We need to characterize which polynomials $p \in \mathbb{R}[x]$ are positive, *i.e.*, $p(x) \geqslant 0$, either in $\mathbb{R}^n$ or in a set defined by other polynomials, *e.g.*, $\{x \in \mathbb{R}^n : p_k(x) \geqslant 0 \text{ for all } k=1,\ldots,m\}$.

Definitions

- $p \in \mathbb{R}[x]$ ($x \in \mathbb{R}^n$) is a *sum-of-squares* (SOS) polynomial if $p(x) = (q(x))^2$ for some $q \in \mathbb{R}[x]$.
- The set of SOS polynomials in $\mathbb{R}[x]$ is denoted $\Sigma^2\mathbb{R}[x]$.
- The set of polynomials $p \in \mathbb{R}[x]$ which are non-negative in $\mathbb{R}^n$ is denoted $\mathscr{P}_+(\mathbb{R}^n)$.
- The *quadratic module generated by a finite set of polynomials* $F = \{f_1,\ldots,f_N\} \subseteq \mathbb{R}[x]$ is

$$\mathrm{QM}(F) = \sum_{f \in F\cup\{1\}} f\,\Sigma^2\mathbb{R}[x] = \left\{ q_0^2(x) + f_1(x)q_1^2(x) + \cdots + f_N(x)q_N^2(x) \colon q_k \in \mathbb{R}[x] \right\}.$$

- A quadratic module is *Archimedean* if there is a $C > 0$ s.t. $C - x_1^2 - \cdots - x_n^2 \in \mathrm{QM}(F)$.

Bonus: Application to SOS Optimization (cont.)

In general $\Sigma^2\mathbb{R}[x] \subseteq \mathscr{P}_+(\mathbb{R}^n)$, and both sets are typically strictly different (Hilbert, 1888).

While $\mathscr{P}_+(\mathbb{R}^n)$ may be difficult to characterize, the coefficients of SOS polynomials have a simple, convex characterization (Parrilo, 2000): Since $p \in \Sigma^2\mathbb{R}[x]$ iff $p(x) = q^2(x)$, and a polynomial $q \in \mathbb{R}[x]$ can be written as a linear combination of *monomials* (*e.g.*, $q(x) = x_1^2 + 3x_1x_2 + 4x_2^2 = [1\ 3\ 4][x_1^2\ \ x_1x_2\ \ x_2^2]^T =: \alpha^T m(x)$), one has that

$$p(x) = m(x)^T \underbrace{\alpha\alpha^T}_{A} m(x).$$

The coefficients of p appear in $A \succeq 0$. Conversely, if $p(x) = m(x)^T A m(x)$ for some matrix $A \succeq 0$, decomposing A as $v_1 v_1^T + \cdots + v_m v_m^T$ yields $p(x) = [v_1^T m(x)]^2 + \cdots + [v_m^T m(x)]^2$, so $p \in \Sigma^2\mathbb{R}[x]$.

Note. The decomposition $p(x) = m(x)^T A m(x)$ is not unique: $x_1^2 + 2x_1x_2 + x_2^2$ can be written as $[x_1\ \ x_2]\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}[x_1\ \ x_2]^T$ or $[x_1\ \ x_2]\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}[x_1\ \ x_2]^T$; however, the set of all A that yield p is a linear subspace (*e.g.*, $\{A \in \mathbb{R}^{2\times 2} : a_{11} = a_{22} = 1, a_{12} + a_{21} = 2\}$), so the characterization of an SOS polynomial in terms of A is convex.

An impressive result, due to M. Putinar (1993), shows that, under mild conditions, the set of polynomials which are strictly positive on a set $\mathscr{D}_F := \{x \in \mathbb{R}^n : f(x) \geqslant 0 \text{ for all } f \in F\}$ defined by a finite set $F \subseteq \mathbb{R}[x]$ can be characterized in terms of SOS polynomials:

Theorem (Putinar's *Positivstellensatz*)
Consider a finite set $F \subseteq \mathbb{R}[x]$, $x \in \mathbb{R}^n$, s.t. QM(F) is Archimedean. Then, every polynomial strictly positive on $\mathscr{D}_F$ is in QM(F).

In other words, every p which is strictly positive on $\mathscr{D}_F$ can be written as

$$p(x) = p_0(x) + f_1(x)p_1(x) + \cdots + f_N(x)p_N(x), \qquad F = \{f_1, \ldots, f_N\},$$

where $p_0, \ldots, p_N$ are SOS polynomials, so if one fixes the degrees of these polynomials, it is possible to characterize p in a convex manner!

The assumption of QM(F) being Archimedean implies that $\mathscr{D}_F$ should be compact, and is easy to fulfill by adding to F the polynomial $C - x_1^2 - \cdots - x_n^2$, with $C \geqslant 1$ sufficiently large.

Putinar's Positivstellensatz is a purely algebraic result from real semi-algebraic geometry, but we will provide a functional analytical proof, based on Hahn-Banach and some spectral properties. However, first we need to generalize the notion of spectrum to a set of operators, and establish the *spectral mapping theorem* :

Definition. Let $A_1, \ldots, A_n \in \mathscr{A} \subseteq \mathscr{L}(H)$, where $\mathscr{A}$ is a *commutative algebra* of operators on a Hilbert space H, *i.e.*, a subset of $\mathscr{L}(H)$ s.t. if $A, B \in \mathscr{A}$ and $\alpha \in \mathbb{C}$, then $AB = BA$ and $A + B, \alpha A, AB \in \mathscr{A}$. The *joint spectrum* of $A = (A_1, \ldots, A_n)$ in $\mathscr{A}$, denoted $\sigma(A)$, is the set of $\lambda \in \mathbb{C}^n$ for which there exist no $B_1, \ldots, B_n \in \mathscr{A}$ s.t. $B_1(A_1 - \lambda_1 I) + \cdots + B_n(A_n - \lambda_1 I) = I$. Note that $\sigma(A) \subseteq \sigma(A_1) \times \cdots \times \sigma(A_n)$.

If $f : \mathbb{C}^n \to \mathbb{C}$ is a polynomial of the form $f(x) = \sum_{i_1, \ldots, i_n \in \mathbb{N}_0} \alpha_{i_1 \cdots i_n} x_1^{i_1} \cdots x_n^{i_n}$, and $A_1, \ldots, A_n \in \mathscr{L}(H)$ are commuting operators, let $\tilde{f} : \mathscr{L}(H)^n \to \mathscr{L}(H)$ be given by $\tilde{f}(A) = \sum_{i_1, \ldots, i_n \in \mathbb{N}_0} \alpha_{i_1 \cdots i_n} A_1^{i_1} \cdots A_n^{i_n}$, where $A = (A_1, \ldots, A_n) \in \mathscr{L}(H)^n$. This definition extends to systems of polynomials $f : \mathbb{C}^n \to \mathbb{C}^m$.

Theorem (Spectral Mapping)
Let $A = \{A_1, \ldots, A_n\}$ be a subset of a commutative algebra of operators $\mathscr{A}$ on a Hilbert space H, and $f : \mathbb{C}^n \to \mathbb{C}^m$ a system of polynomials. Then, $f(\sigma(A)) = \sigma(\tilde{f}(A))$.

Lemma. If $A \in \mathscr{L}(H)$, and $\lambda \in \partial\sigma(A)$, then there is a sequence (T_n) in $\mathscr{L}(H)$ s.t. T_n is invertible and $\|T_n\| = 1$ for all $n \in \mathbb{N}$, and $(A - \lambda I)T_n \to 0$.

Proof. Since $\lambda \in \partial\sigma(A)$, pick a sequence (λ_n) in $\sigma(A)^c$ s.t. $\lambda_n \to \lambda$, and let $R_n := (A - \lambda_n I)^{-1}$. Then, $R_n(A - \lambda I) - I = R_n(A - \lambda_n I + (\lambda_n - \lambda)I) - I = (\lambda_n - \lambda)R_n$. Then, $(\|R_n\|)$ is unbounded; otherwise there is an $M > 0$ s.t. $\|R_n\| \leqslant M$ for all n, and $\|R_n(A - \lambda I) - I\| = |\lambda_n - \lambda|\|R_n\| \to 0$, so $\|R_{n^*}(A - \lambda I) - I\| < 1$ for some n^*, thus $R_{n^*}(A - \lambda I)$ is invertible, and so is $A - \lambda I = (A - \lambda_n I)R_{n^*}(A - \lambda I)$, a contradiction. Thus, assume that $\|R_n\| \to \infty$, and let $T_n := R_n/\|R_n\|$, so $\|T_n\| = 1$. Then, $\|(A - \lambda I)T_n\| = \|(A - \lambda I)R_n\|/\|R_n\| = \|I/\|R_n\| + (\lambda_n - \lambda)T_n\| \leqslant 1/\|R_n\| + |\lambda_n - \lambda|\|T_n\| \to 0$. $\square$

Proof of Spectral Mapping Theorem (Harte, 1972). If $f_k : \mathbb{C}^n \to \mathbb{C}$ is a polynomial, then by the remainder theorem, for every $\lambda \in \mathbb{C}^n$, $\tilde{f}_k(A) - f_k(\lambda)I = \sum_j B_j(A_j - \lambda_j I)$ for some $B_1, \ldots, B_n \subseteq \mathscr{A}$, so if $f(\lambda) \notin f(\sigma(A))$, then $\lambda \notin \sigma(A)$, *i.e.*, $f(\sigma(A)) \subseteq \sigma(\tilde{f}(A))$.

To prove the converse, we will show that if $C = (C_1, \ldots, C_m) \in \mathscr{A}^m$, and $\mu \in \sigma(C) \subseteq \mathbb{C}^m$, then there exists a $\lambda \in \mathbb{C}^n$ s.t. $(\lambda, \mu) \in \sigma(A, C)$. This is done by induction on n, so we will only consider $n = 1$:

Let $\mathscr{N} := \overline{\left\{\sum_j B_j(C_j - \mu_j I) : B_1, \ldots, B_m \in \mathscr{A}\right\}}$. Note that $A\mathscr{N} \subseteq \mathscr{N}$ for every $A \in \mathscr{A}$ and that $I \neq \mathscr{N}$ (since $\mu \in \sigma(C)$), so $\mathscr{A}/\mathscr{N} \neq \{[0]\}$. Define $L_{A_1} : \mathscr{A}/\mathscr{N} \to \mathscr{A}/\mathscr{N}$ as $L_{A_1}([B]) = [A_1 B]$. $\sigma(L_{A_1}) \neq \emptyset$ is compact, so pick a $\lambda_1 \in \partial\sigma(L_{A_1})$. Then, by the lemma above, there is a sequence (T_n) of invertible operators in $\mathscr{A}/\mathscr{N}$ s.t. $\|[T_n]\|_{\mathscr{A}/\mathscr{N}} = 1$ for all n and $\|[(A_1 - \lambda_1 I)T_n]\|_{\mathscr{A}/\mathscr{N}} = \inf_{N \in \mathscr{N}} \|(A_1 - \lambda_1 I)T_n + N\| \to 0$.

Proof (cont.)
Based on this result, we claim that $(\lambda_1, \mu) \in \sigma(A_1, C)$, since otherwise there would be $A_1', C_1', \dots, C_n' \in \mathscr{A}$ s.t. $A_1'(A_1 - \lambda_1 I) + C_1'(C_1 - \lambda_1 I) + \cdots + C_n'(C_n - \lambda_n I) = I$, hence for an arbitrary $D \in \mathscr{A}$ we have that $D = A_1'(A_1 - \lambda_1 I)D + C_1'(C_1 - \lambda_1 I)D + \cdots + C_n'(C_n - \lambda_n I)D \in A_1'(A_1 - \lambda_1 I)D + \mathscr{N}$, but then $\|[D]\|_{\mathscr{A}/\mathscr{N}} = \inf_{N \in \mathscr{N}} \|A_1'(A_1 - \lambda_1 I)D + N\| \leqslant \inf_{N \in \mathscr{N}} \|A_1'(A_1 - \lambda_1 I)D + A_1' N\| = \inf_{N \in \mathscr{N}} \|A_1'[(A_1 - \lambda_1 I)D + N]\| \leqslant \|A_1'\| \|[(A_1 - \lambda_1 I)D]\|_{\mathscr{A}/\mathscr{N}}$, which contradicts the properties of (T_n). Thus, $(\lambda_1, \mu) \in \sigma(A_1, C)$.

Therefore, in general, for every $\mu \in \sigma(\tilde{f}(A))$ there is a $\lambda \in \mathbb{C}^n$ s.t. $(\lambda, \mu) \in \sigma(A, \tilde{f}(A))$. Since $\sigma(A, \tilde{f}(A)) \subseteq \sigma(A) \times \sigma(\tilde{f}(A))$, $\lambda \in \sigma(A)$. We just need to show that $\mu \in f(\lambda)$. Consider the system of polynomials $g\colon \mathbb{C}^{n+m} \to \mathbb{C}^m$ given by $g(\lambda, \mu) = \mu - f(\lambda)$. Then, by our first result, $\mu - f(\lambda) = g(\lambda, \mu) \in g(\sigma(A, \tilde{f}(A))) \subseteq \sigma(\tilde{g}(A, \tilde{f}(A))) = \sigma(0) = \{0\}$, *i.e.*, $\mu = f(\lambda)$, so $\sigma(\tilde{f}(A)) \subseteq f(\sigma(A))$.
In conclusion, $f(\sigma(A)) \subseteq \sigma(\tilde{f}(A))$ and $\sigma(\tilde{f}(A)) \subseteq f(\sigma(A))$, thus $\sigma(\tilde{f}(A)) = f(\sigma(A))$. □

Definition. Let K be a convex set in a vector space V. $x \in K$ is an *algebraic interior point of K relative to V* if for every $v \in V$ there is an $\varepsilon > 0$ s.t. $x + tv \in K$ for all $t \in [0, \varepsilon]$. The set of all algebraic interior points of K is called the *algebraic interior* of K, aint K.

To establish Putinar's Positivstellensatz, note that Eidelheit's separating hyperplane theorem can be modified to this "algebraic" version: If K_1 and K_2 are convex sets in a real vector space V s.t. aint $K_1 \neq \emptyset$ and $K_2 \cap \text{aint } K_1 = \emptyset$. Let $x_0 \in \text{aint } K_1$. Then there is a linear functional $l: V \to \mathbb{R}$ s.t. $l(x) \leqslant 0$ for all $x \in K_2$, $l(x) \geqslant 0$ for all $x \in K_1$, and $l(x_0) > 0$. (*Exercise!*)

Lemma. 1 is an algebraic interior point of an Archimedean QM(F).

Proof. Since $C - x_1^2 - \cdots - x_n^2 \in \text{QM}(F)$ for some $C \geqslant 1$, and QM(F) is a convex set,

- $C - x_i^2 = C - x_1^2 - \cdots - x_n^2 + \sum_{j \neq i} x_j^2 \in \text{QM}(F)$ for all $i = 1, \ldots, n$.
- $C \pm x_i = \frac{1}{2}[(C-1) + (C - x_i^2) + (x_i \pm 1)^2] \in \text{QM}(F)$ for all $i = 1, \ldots, n$.
- If $K \pm q \in \text{QM}(F)$ ($q \in \mathbb{R}[x]$, $K > 0$), then $K^2 - q^2 = \frac{1}{2K}[(K+q)^2(K-q) + (K-q)^2(K+q)] \in \text{QM}(F)$.
- If $K_1 \pm q_1, K_2 \pm q_2 \in \text{QM}(F)$, then $K_1 + K_2 - (q_1 \pm q_2) \in \text{QM}(F)$, and $\frac{(C_1+C_2)^2}{4} \pm q_1 q_2 = \frac{(C_1+C_2)^2}{4} \pm \frac{1}{4}(q_1+q_2)^2 \mp \frac{1}{4}(q_1 - q_2)^2 \in \text{QM}(F)$.
- From the previous properties, for every $p \in \mathbb{R}[x]$ there is a $K > 0$ s.t. $N \pm p \in \text{QM}(F)$ for all $N \geqslant K$, *i.e.*, $1 \pm \varepsilon p \in \text{QM}(F)$ for all $\varepsilon \in [0, 1/K]$. Thus, 1 is an algebraic interior point of QM(F). $\square$

Proof of Putinar's Positivstellensatz (Helton and Putinar, 2008)

Firstly notice that $\mathrm{QM}(F)$ is a convex set. Assume, to the contrary, that p is a strictly positive polynomial in $\mathscr{D}_F$, but $p \notin \mathrm{QM}(F)$. By the modified separating hyperplane theorem, there is a linear functional l on $\mathbb{R}[x]$ s.t. $l(1) > 0$, $l(q) \geqslant 0$ for all $q \in \mathrm{QM}(F)$, and $l(p) \leqslant 0$; extend l algebraically to $\mathbb{C}[x]$. Construct a Hilbert space $L_2(l)$ as the completion of $\mathbb{C}[x]/N$, where $N = \{q \in \mathbb{C}[x] \colon l(q) = 0\}$, and $(q,r) = l(q\bar{r})$. Consider the tuple of multiplication operators $M = (M_{x_1}, \ldots, M_{x_n})$ on $L_2(l)$ where $M_{x_k} q(x) = x_k q(x)$, which are self-adjoint and commute with each other. Furthermore, these operators are bounded, since $([C - x_1^2 - \cdots - x_n^2]q, q) = l([C - x_1^2 - \cdots - x_n^2]q^2) \geqslant 0$ by the Archimedean property (*i.e.*, $[C - x_1^2 - \cdots - x_n^2]q^2 \in \mathrm{QM}(F)$) and this implies that $(M_{x_k} q, q) \leqslant C(q,q)$ for every $q \in \mathbb{C}[x]$.

For every $f \in F$, since $(\tilde{f}(M)p, p) = (fp, p) \geqslant 0$ for every $p \in \mathbb{C}[x]$, thus $\tilde{f}(M)$ is non-negative, *i.e.*, $\sigma(\tilde{f}(M)) \subseteq [0,\infty)$, so the spectral mapping theorem implies that $f(\sigma(M)) = \sigma(\tilde{f}(M)) \subseteq [0,\infty)$ for all $f \in F$, that is, $\sigma(M) \subseteq \mathscr{D}_F$.

Therefore, for every $q \in \mathbb{C}[x]$ s.t. $q(x) \geqslant 0$ on $\mathscr{D}_F$, it holds by the spectral mapping theorem that $\sigma(\tilde{q}(M)) = q(\sigma(M)) \subseteq [0,\infty)$, so, by the Corollary in Slide 29, $\tilde{q}(M)$ is non-negative, thus $l(q) = (q,1) = (\tilde{q}(M)1, 1) \geqslant 0$, *i.e.*, l is a positive functional on $\mathbb{R}[x]$.

Since $\mathscr{D}_F$ is compact, there is an $\varepsilon > 0$ s.t. $p(x) \geqslant \varepsilon$ for all $x \in \mathscr{D}_F$, so $l(p) \geqslant \varepsilon l(1) > 0$, a contradiction. Therefore, all strictly positive polynomials in $\mathscr{D}_F$ belong to $\mathrm{QM}(F)$. □

Example (from slides by C. Scherer and S. Weiland)

Consider the problem of testing whether the following polynomials are Hurwitz (*i.e.*, have all their roots inside the unit disk):

$$\{s^3 + (3 - \delta_1^2 + \delta_2)s^2 + (3 + \delta_1)s + (0.9 + \delta_1\delta_2) : \delta_1 \in [-1, 1],\ \delta_2 \in [-1, 1]\}.$$

By the Routh-Hurwitz criterion, this amounts to checking

$$\left.\begin{array}{l} 3 - \delta_1^2 + \delta_2 \geqslant 0, \text{ and} \\ (3 + \delta_1 + \delta_2)(3 + \delta_1) - (0.9 + \delta_1\delta_2) \geqslant 0 \end{array}\right\} \quad \text{for all } \delta_1, \delta_2 \text{ s.t. } \delta_1^2 \leqslant 1 \text{ and } \delta_2^2 \leqslant 1.$$

By Putinar's Positivstellensatz, the positivity of the first condition is equivalent to

$$3 - \delta_1^2 + \delta_2 = p_0(\delta_1, \delta_2) + p_1(\delta_1, \delta_2)(1 - \delta_1^2) + p_2(\delta_1, \delta_2)(1 - \delta_2^2) \qquad (*)$$

for some SOS polynomials $p_0, p_1, p_2 \in \Sigma^2\mathbb{R}[\delta_1, \delta_2]$. By setting upper bounds on the degrees of these polynomials, $(*)$ corresponds to an LMI feasibility problem that can be solved using standard convex optimization tools (CVX/Yalmip via Sedumi, SDPT3, Mosek, ...).