Hilbert spaces.

Introduction. Euclidian geometry teaches that if A is some invertible $n \times n$-matrix whose elements are real numbers and A is regarded as a linear map from $\mathbf{R}^n$ into itself, then the image of the euclidian unit sphere S^{n-1} is an ellipsoid $\mathcal{E}_A$. Conversely if $\mathcal{E}$ is an ellipsoid there exists an invertible matrix A such that $\mathcal{E} = \mathcal{E}_A$.

The case $n = 2$. Let (x, y) be the coordinates in $\mathbf{R}^2$ and A the linear map

$$(x, y) \mapsto (x + y, y) \tag{0.1}$$

To get the image of the unit circle $x^2 + y^2 = 1$ we use polar coordinates and write $x = \cos\phi$ and $y = \sin\phi$. This gives the closed image curve

$$\phi \mapsto (\cos\phi + \sin\phi\,;\, \sin\phi) \quad :| \quad 0 \leq \phi \leq 2\pi \tag{i}$$

It is not obvious how to determine the principal axes of this ellipse. The gateway is to consider the *symmetric* 2×2-matrix $B = A^* A$. If u, v is a pair of vectors in $\mathbf{R}^2$ we have

$$\langle Bu, v\rangle = \langle Au, Av\rangle \tag{ii}$$

It follows that $\langle Bu, u\rangle > 0$ for all $u \neq 0$. By a wellknown result in elementary geometry it means that the symmetric matrix B is positive, i.e. the eigenvalues arising from zeros of the characteristic polynomial $\det(\lambda E_2 - B)$ are both positive. Moreover, the *spectral theorem* for symmetric matrices shows that there exists an orthonormal basis in $\mathbf{R}^2$ given by a pair of eigenvectors for B denoted by u_* and v_*. So here

$$B(u_*) = \lambda_1 \cdot u_* \quad : \quad B(v_*) = \lambda_2 \cdot v_*$$

Next, since (u_*, v_*) is an orthonormal basis in $\mathbf{R}^2$ points on the unit circle are of the form

$$\xi = \cos\phi \cdot u_* + \sin\phi \cdot v_*$$

Then we get

$$|A(\xi)|^2 = \langle A(\xi).A(\xi)\rangle = \langle B(\xi), \xi\rangle = \cos^2\phi \cdot \lambda_1 + \sin^2\phi \cdot \lambda_2$$

From this we see that the ellipse $\mathcal{E}_A$ has u_* and v_* as principal axes. It is a circle if and only if $\lambda_1 = \lambda_2$. If $\lambda_1 > \lambda_2$ the largest principal axis has length $2\sqrt{\lambda_1}$ and the smallest has length $2\sqrt{\lambda_2}$. The reader should now compute the specific example (*) and find $\mathcal{E}_A$.

0.2 A Historic Remark. The fact that $\mathcal{E}_A$ is an ellipsoid was wellknown in the Ancient Greek mathematics when $n = 2$ and $n = 3$. After general matrices and their determinants were introduced, the spectral theorem for symmetric matrices was established by A. Cauchy in 1810 under the assumption that the eigenvalues are different. Later Weierstrass gave the proof in the general case where multiple eigenvalues appear, and independently Gram and Weierstrass found a method to produce an orthonormal basis of eigenvectors for a symmetric $n \times n$-matrix B. To find an eigenvector with largest eigenvalue one studies the extremal problem

$$\max_x \langle Bx, x\rangle \quad : \quad ||x|| = 1 \tag{1}$$

If a unit vector x_* maximises (1) then it is an eigenvector, i.e.

$$Bx_* = a_1 x_*$$

holds for a real number a. In the next stage one takes the orthogonal complement $x_*^\perp$ and proceed to the restricted extremal problem where x say in this orthogonal complement which gives an eigenvector whose eigenvalue $a_2 \leq a_1$. After n steps we obtain an n-tuple of pairwise orthogonal eigenvectors to B. In the orthonormal basis given by this n-tuple the linear operator of B is represented by a diagonal matrix. One often refers to this as the Gram-Schmidt construction of an orthonormal basis with respect to B.

Singular values. *Mathematica* has implemented programs which for every invertible $n \times n$-matrix A determines the ellipsoid $\mathcal{E}_A$ numerically. This is presented under the headline *singular values for matrices.* In general the A-matrix is not symmetric but the spectral theorem is applied to the symmetric matrix A^*A which determines the ellipsoid $\mathcal{E}_A$ and whose principal axis are pairwise disjoint.

0.3 Rotating bodies. The spectral theorem in dimension $n = 3$ is best illustrated by regarding a rotating body. Consider a bounded 3-dimensional body K in which some distribution of mass is given. The body is placed i $\mathbf{R}^3$ where (x_1, x_2, x_3) are the coordinates and the distribution of mass is expressed by a positive function $\rho(x, y, z)$ defined in K. The *center of gravity* in K is the point $(\bar{x}_1, \bar{x}_2, \bar{x}_3)$ where

(i) $$\bar{x}_\nu = \iiint_K x_\nu \cdot \rho(x_1, x_2, x_3) \cdot dx_1 dx_2 dx_3 \quad : 1 \leq \nu \leq 3$$

After a translation we may assume that the center of mass is the origin. Now we imagine that a rigid bar which stays on a line ℓ is attached to K with its two endpoints p and q, i.e. if γ is the unit vector in $\mathbf{R}^3$ which determines the line then

$$p = A \cdot \gamma \quad : \quad q = -A \cdot \gamma$$

where A is so large that p and q are outside K. The mechanical experiment is to rotate around ℓ with some constant angular velocity ω while the two points p and q are kept fixed. The question arises if such an imposed rotation of K around ℓ implies that external forces at p and q are needed to prevent these to points from moving. It turns out that there exist so called free axes where no such forces are needed, i.e. for certain directions of ℓ the body rotates nicely around the axis with constant angular velocity. The free axes are found from the spectral theorem. More precisely, one introduces the symmetric 3×3-matrix A whose elements are

(i) $$a_{pq} = \bar{x}_\nu = \iiint_K x_p \cdot x_q \cdot \rho(x_1, x_2, x_3) \cdot dx_1 dx_2 dx_3$$

Using the expression for the centrifugal force by C. Huyghen's one has the *Law of Momentum* which in the present case shows that the body has a free rotation along the lines which correspond to eigenvectors of the symmetric matrix A above. In view of the historic importance of this example we present the proof of this in a separate section even though some readers may refer to this as a subject in classical mechanics rather than linear algebra. Hence the spectral theorem becomes evident by this mechanical experiment, i.e. just as Stokes Theorem the spectral theorem for symmetric matrices is rather a Law of Nature than a mathematical discovery.

0.4 Inner product norms

Let A be an invertible $n \times n$-matrix. The ellipsoid $\mathcal{E}_A$ defines a norm on $\mathbf{R}^n$ by the general construction in XX. This norm has a special property. For if $B = A^*A$ and x, y is a pair of n-vectors, then

(i) $$||x + y||^2 = \langle B(x + y), B(x + y) \rangle = ||x||^2 + ||y||^2 + 2 \cdot B(x, y)$$

It means that the map

(ii) $$(x, y) \mapsto ||x + y||^2 - ||x||^2 - ||y||^2$$

is linear both with respect to x and to y, i.e. it is a bilinear map given by

(iii) $$(x, y) \mapsto 2 \cdot B(x, y)$$

We leave as an exercise for the reader to prove that if K is a symmetric convex set in $\mathbf{R}^n$ defining the ρ_K-norm as in xx, then this norm satisfies the bi-linearity (ii) if and only if K is an ellipsoid and therefore equal to $\mathcal{E}_A$ for an invertible $n \times n$-matrix A. Following Hilbert we refer to a norm defined by some bilinear form $B(x, y)$ as an *inner product norm.* The spectral theorem asserts that there exists an orthonormal basis in $\mathbf{R}^n$ with respect to this norm.

0.5 The complex case. Consider a Hermitian matrix A, i.e an $n \times n$-matrix with complex elements satisfying

(*) $$a_{qp} = \bar{a}_{pq} \quad : \quad 1 \leq p, q \leq n$$

Consider the n-dimensional complex vector space $\mathbf{C}^n$ with the basis $e_1, \ldots, e_n$. An inner product is defined by

$$(**) \qquad \langle x, y \rangle = x_1 \bar{y}_1 + \ldots + x_n \bar{y}_n$$

where $x_\bullet = \sum x_\nu \cdot e_\nu$ and $y_\bullet = \sum y_\nu \cdot e_\nu$ is a pair of complex n-vectors. If A as above is a Hermitian matrix we obtain

$$(***) \qquad \langle Ax, y \rangle = \sum \sum a_{pq} x_q \cdot \bar{y}_p \sum \sum x_p \cdot \bar{a}_{qp} \bar{y}_q = \langle x, Ay \rangle$$

Let us consider the characteristic polynomial $\det(\lambda \cdot E_n - A)$. If λ is a root there exists a non-zero eigenvector x such that $Ax = \lambda \cdot x$. Now (***) entails that

$$\lambda \cdot ||x||^2 = \langle Ax, x \rangle = \langle x, Ax \rangle = \bar{\lambda} \cdot ||x||^2$$

It follows that λ is *real*, i.e. the roots of the characteristic polynomial of a Hermitian matrix are always real numbers. If all roots are > 0 one say that the Hermitiain matrix is *positive.*

0.6 Unitary matrices. An $n \times n$-matrix U is called unitary if

$$\langle Ux, Ux \rangle = \langle x, x \rangle$$

hold for all $x \in \mathbf{C}^n$. The spectral theorem for Hermitian matrices asserts that if A is Hermitian then there exists a unitary matrix U such that

$$UAU^* = \Lambda$$

where Λ is a diagonal matrix whose elements are real.

0.7 The passage to infinite dimension.

Around 1900 the need for a spectral theorem in infinite dimensions became urgent. In his article *Sur une nouvelle méthode pour la resolution du problème de Dirichlet* from 1900, Ivar Fredholm extended earlier construction by Volterra and studied systems of linear equations in an infinite number of variables with certain bounds. In Fredholm's investiagations one starts with a sequence of matrices $A_1, A_2, \ldots$ where A_n is an $n \times n$-matrix and an infinite dimensional vector space

$$V = \mathbf{R}e_1 + \mathbf{R}e_2 + \ldots$$

To each $N \geq 1$ we get the finite dimensional subspace $V_N = \mathbf{R}e_1 + \ldots \mathbf{R}e_N$. Now A_N is regarded as a linear operator on V_N and we assume that the A-sequence is matching , i.e. if $M > N$ then the restriction of A_M to V_N is equal to A_N. This means that we take any infinite matrix A_∞ with elements $\{a_{ik}\}$ and here A_N is the $N \times N$-matrix which appears as an upper block with N^2-elements $a_{ik} : 1 \leq i, k \leq N$. To each N the ellipsoid $\mathcal{E}_N = \mathcal{E}_{A_N}$ on V_N where defines a norm. As N increases the norms are matching and hence V is equipped with a norm which for every $N \geq 1$ restricts to the norm defined by $\mathcal{E}_N$ on the finite dimensional subspace V_N. Notice that the norm of any vector $\xi \in V$ is finite since ξ belongs to V_N for some N, i.e. by definition any vector in V is a finite $\mathbf{R}$-linear combination of the basis vectors $\{e_\nu\}$. Moreover, the norm on V satisfies the bilinear rule from (0.3), i.e. on $V \times V$ there exists a bilinear form B such that

$$(*) \qquad ||x + y||^2 - ||x||^2 - ||y||^2 = 2B(x, y) \quad : \quad x, y \in V$$

Remark. Inequalities for determinants due to Hadamard play an important role in Fredholm's work and since the Hadamard inequalities are used in many other situations we announce some of his results, leaving proofs as an exercise or consult the literature where an excellent source is the introduction to integral equations by the former professor at Harvard University Maxime Bochner [Cambridge University Press: 1914):

0.8 Two inequalities. Let $n \geq 2$ and $A = \{a_{ij}\}$ some $n \times n$-matrix whose elements are real numbers. Show that if

$$a_{i1}^2 + \ldots + a_{in}^2 = 1 \quad : \quad 1 \leq i \leq n$$

then the determinant of A has absolute value ≤ 1. Next, assume that there is a constant M such that the absolute values $|a_{ij|}| \leq M$ hold for all pairs i, j. Show that this gives

$$\big|\det(A)\big| \leq \sqrt{n^n} \cdot M^n$$

0.9 The Hilbert space $\mathcal{H}_V$. This is the completition of the normed space V. That is, exactly as when the field of rational numbers is completed to the real number system one regards Cauchy sequences for the norm of vectors in V and in this way we get a normed vector space denoted by $\mathcal{H}_V$ where the norm topology is complete. Under this process the bi-linearity is preserved, i.e. on $\mathcal{H}_V$ there exists a bilinear form $B_{\mathcal{H}}$ such that (*) above holds for pairs $x, y \in \mathcal{H}_V$. Following Hilbert we refer to $B_{\mathcal{H}}$ as the *inner product* attached to the norm. Having performed this construction starting from any infinite matrix A_∞ it is tempting to make a further abstraction. This is precisely what Hilbert did, i.e. he ignored the "source" of a matrix A_∞ and defined a complete normed vector space over $\mathbf{R}$ to be a real Hilbert space if the there exists a bilinear form B on $V \times V$ such that (*) holds.

Remark. If V is a "abstract" Hilbert space the restriction of the norm to any finite dimensional subspace W is determined by an ellipsoid and exactly as in linear algebra one constructs an orthonormal basis on W. By the Gram-Schmidt construction there exists an orthonormal sequence $\{e_n\}$ in V. However, in order to be sure that it suffices to take a *denumerable* orthonormal basis it is necessary and sufficient that the normed space V is *separable.* Assuming this every $v \in V$ has a unique representation

(i) $$v = \sum c_n \cdot e_n \quad : \quad \sum |c_n|^2 = ||v||^2$$

The existence of this orthonormal family means that every separable Hilbert space is isomorphic to the standard space ℓ^2 whose vectors are infinite sequences $\{c_n\}$ where the square sum $\sum c_n^2 < \infty$. So in order to prove general results about separable Hilbert spaces it is sufficient to regard ℓ^2. However, the abstract notion of a Hilbert space is useful since inner products on specific linear spaces appear in many different situations. For example, in complex analysis one regards the space of analytic functions which are square integrable in a bounded domain or whose boundary values are square integrable. Here the inner product is given in advance but it can be a highly non-trivial affair to exhibit an orthonormal basis.

0.10 Linear operators on ℓ^2. A bounded linear operator T from the complex Hilbert space ℓ^2 into itself is described by an infinite matrix $\{a_{p,q}\}$ whose elements are complex numbers. Namely, for each $p \geq 1$ we set

(i) $$T(e_p) = \sum_{q=1}^{\infty} a_{pq} \cdot e_q$$

For each fixed p we get

(ii) $$||T(e_p)||^2 = \sum_{q=1}^{\infty} |a_{pq}[^2$$

Next, let $x = \sum \alpha_\nu \cdot e_\nu$ and $y = \sum \beta_\nu \cdot e_\nu$ be two vectors in ℓ^2. Then we get

$$||x+y||^2 = \sum |\alpha_\nu + \beta_\nu|^2 \cdot e_\nu$$

For each ν we have the pair of complex numbers α_ν, β_ν and here we have the inequality

$$|\alpha_\nu + \beta_\nu|^2 \leq 2 \cdot |\alpha_\nu|^2 + 2 \cdot |\beta_\nu|^2$$

It follows that

(iii) $$||x+y||^2 \leq 2 \cdot ||x||^2 + 2 \cdot ||y||^2$$

In (iii) equality holds if and only if the two vectors x and y are linearly dependent, i.e. if there exists some complex number λ such that $y = \lambda \cdot x$. Let us now return to the linear operator T. In

(ii) we get an expression for the norm of the T-images of the orthonormal basis vectors. So when T is bounded with operator norm M then the sum of the squared absolute values in each row of the matrix $A = \{a_{p,q}\}$ is $\leq M^2$. However, this condition along is not sufficient to guarantee that T is a bounded linear operator. For example, suppose that the row vectors in T are all equal to a given vector in ℓ^2, i.e. $a_{p,q} = \alpha_q$ hold for all pairs where $\sum, |\alpha_q|^2 = 1$. Then

$$T(e_1 + \ldots + e_N) = N \cdot v \quad : \quad v = \sum \alpha_q \cdot e_q$$

The norm in the right hand side is N while the norm of $e - 1 + \ldots + e_n$ is $\sqrt{N}$. Since $N >> \sqrt{N}$ when N increases this shows that T cannot be bounded. So the condition on the matrix A in order that T is bounded is more subtle. In fact, given a vector $x = \sum \alpha_\nu \cdot e_\nu$ as above with $||x|| = 1$ we have

$$(*) \qquad ||T(x)||^2 = \sum_{p=1}^{\infty} \sum_q \sum_k a_{p,q} \cdot \alpha_q \cdot \bar{a}_{pk} \cdot \bar{\alpha}_k$$

So we encounter an involved triple sum. Notice also that for each fixed p we get a *non-negative* term

$$\rho_p = \sum_q \sum_k a_{p,q} \cdot \alpha_q \cdot \bar{a}_{pk} \cdot \bar{\alpha}_k = |\sum_{q=1}^{\infty} a_{pq} \cdot \alpha_q \,|^2$$

Final remark. Thus, the description of the Banach space $L(\ell^2, \ell^2)$ of all bounded linear operators on ℓ^2 is not easy to grasp. In fact, no "comprehensible" description exists of this space.

1. General results about Hilbert spaces.

Let $\mathcal{H}$ be a real Hilbert space. The construction of the inner product norm entails that

(*) $$||x+y||^2 + ||x-y||^2 = 2\cdot ||x||^2 + 2\cdot ||y||^2$$

for every pair x, y in $\mathcal{H}$. Using this one solves an extremal problem. For every closed convex subset K of $\mathcal{H}$ and if $\xi \in \mathcal{H}\setminus K$ there exists a unique $k_* \in K$ such that

(**) $$\min_{k\in K} ||\xi - k|| = ||\xi - k_*||$$

To prove (**) we let ρ denote the minimal distance. We find a sequence $\{k_n\}$ in K such that $||\xi - k_n|| \to \rho$. Now we show that $\{k_n\}$ is a Cauchy sequence. For let $\epsilon > 0$ which gives some integer N_* such that

(i) $$||\xi - k_n|| < \rho + \epsilon \quad : \quad n \geq N_*$$

The convexity of K implies that if $n, m \geq N_*$ then $\frac{k_n + k_m}{2} \in K$. Hence

(ii) $$\rho^2 \leq ||\xi - \frac{k_n + k_m}{2}||^2 \implies 4\rho^2 \leq ||(\xi - k_n) + (\xi - k_m]||^2$$

By the identity (*) the right hand side is

(iii) $$2||\xi - k_n||^2 + 2||\xi - k_m||^2 - ||k_n - k_m||^2$$

It follows from (i-iii) that

$$||k_n - k_m||^2 \leq 4(\rho+\epsilon)^2 - 4\rho^2 = 8\rho\cdot\epsilon + 4\epsilon^2$$

Since ϵ can be made arbitrary small $\{k_n\}$ is a Cauchy sequence and hence there exists a limit $k_n \to k_*$ where $k_* \in K$ since K is closed. Finally, the uniqueness of k_* follows from the equality

$$||\xi - k_1||^2 + ||\xi - k_2||^2 = 2\cdot ||\xi - \frac{k_1 + k_2}{2}||^2 + \frac{1}{2}\cdot ||k_1 - k_2||^2$$

for every pair k_1, k_2 in K. In fact, this equality entails that if $\epsilon > 0$ and k_1, k_2 is a pair such that

$$||\xi - k_\nu|| \leq \rho^2 + \epsilon \quad : \nu = 1, 2$$

then we have

$$||k_1 - k_2||^2 \leq 4\epsilon$$

from which the uniquness of k_* follows.

1.1 The decomposition theorem. Let V be a closed subspace of H. Its orthogonal complement is defined by

(i) $$V^\perp = \{x \in H \quad : \langle x, V\rangle = 0\}$$

It is obvious that $V^\perp$ is a closed subspace of H and that $V \cap V^\perp = 0$. There remains to prove the equality

(ii) $$H = V \oplus V^\perp$$

To see this we take some $\xi \in H \setminus V$. Now V is a closed convex set so we find v_* such that

(iii) $$\rho = ||\xi - v_*|| = \min_{v\in V} ||\xi - v||$$

If we prove that $\xi - v_* \in V^\perp$ we get (ii). To show this we consider some $\eta \in V$. If $\epsilon > 0$ we have

$$\rho^2 \leq ||\xi - v_* + \epsilon\cdot\eta||^2 = ||\xi - v_*||^2 + \epsilon^2\cdot ||\eta||^2 + \epsilon\langle \xi - v_*, \eta\rangle$$

Since $||\xi - v_*||^2 = \rho^2$ and $\epsilon > 0$ it follows that

$$\langle \xi - v_*, \eta\rangle + \epsilon\cdot ||\eta||^2 \geq 0$$

here ϵ can be arbitrary small and we conclude that $\langle \xi - v_*, \eta\rangle \geq 0$. Using $-\eta$ instead we get the opposed inequality and hence $\langle \xi - v_*, \eta\rangle = 0$ as required.

1.2 Complex Hilbert spaces. On a complex vector space similar results as above hold provided that we regard convex sets which are **C**-invariant. Here the inner product is hermitian. It means that

$$\langle x, y\rangle = \overline{\langle y, x\rangle}$$

hold for every pair of vectors. When y is fixed

$$x \mapsto \langle x, y\rangle$$

is **C**-linear. By the construction of the norm and the Cacuhy-Schweartz inequality one gets

$$|\langle x, y\rangle| \leq ||x|| \cdot ||y||$$

In particular (xx) gives a linear funtional on $\mathcal{H}$ of norm $\leq ||y||$ and by taking $x = y$ we see that the norm is equal to $||y||$.

1.4 Hilbert spsces are self-dual. A fundamental fact is that (x) exhibit vectors in the dual space $\mathcal{H}^*$. To see this we consider some $\phi \in \mathcal{H}^*$ whose null-space is a complex hyperplane $\ker(\phi$. The complex verion of (1.1) whose verification is left as an exercise gives

$$\mathcal{H} = \ker(\phi) \oplus \mathbf{C} \cdot \xi$$

for some vector ξ which belongs to $\ker(\phi)^\perp$. Now the reader can check that there exists a complex number c such that

$$\phi(x) = c \cdot \langle x, \xi\rangle$$

for every $x \in \mathcal{H}$. The conclusion is that one has a bijective map from $\mathcal{H}$ to $\mathcal{H}^*$ where every $\in \in \mathcal{H}$ gives the linear functional

$$x \mapsto \langle x, \xi\rangle$$

In particular Hilbert spaces are reflexive Banach spaces.

1.6 Self-adjoint prjections. They consist of linear operator Π on the complex Hlbert space which satisfy

$$\Pi^2 = \Pi \quad \& \quad \Pi = \Pi^*$$

For every such operator we get a decompsition

$$\mathcal{H} = \Pi(\mathcal{H}) \oplus \mathrm{Ker}(\Pi)$$

So the range and the kenrle of Π are orthogonal subspace of $\mathcal{H}$ snd thier direct sum is the whole space. More genwerally one can consider a finite srt ot self-adjoint projections $\{\Pi_\nu\}$ where $\Pi_\nu \circ \Pi_k = 0$ when $k \neq \nu$ while $\sum Pi_\nu(\mathcal{H}) = \mathcal{H}$. This gives a decompdiion

$$\mathcal{H} = \Pi_1(\mathcal{H}) \oplus \ldots \oplus \Pi_k(\mathcal{H})$$

1.x Hilbert-Schmidt operators. We restrict the study to separable Hilbert spaces and assume that $\mathcal{H} = \ell^2$. A linear operator S on ℓ^2 gives a doubly indexed sequence of complex numbers $\{a_{pq}\}$ where

$$\text{(i)} \qquad S(e_p) = \sum_{q=1}^{\infty} a_{pq} \cdot e_q$$

Impose the condition

$$(*) \qquad \sum \sum |a_{pq}|^2 < \infty$$

If (*) holds one says that S is an operator of the Hilbert Schmidt type. Every such operator is bounded. For let $x = \sum x_p \cdot e_p$ be a vector of unit norm in ℓ^2, i.e. $\sum |x_p|^2 = 1$. For each integer $q \geq 1$ we set

$$\rho_q = \sum a_{qp} \cdot x_p$$

Then (i) gives

(ii)
$$Sx = \sum \rho_q \cdot e_q \implies ||Sx||^2 = \sum_q |\rho_q|^2$$

Apply the Cauchy-Schwarz inequality and sibnce x is a unit vector the last term in (ii) is majorised by (*). In partocuoar the operator norm of S satisfies

(iii)
$$||S|| \leq \sqrt{\sum\sum |a_{pq}|^2}$$

One refers to the right hand side as the Hilbert-Schmidt norm of S. Examples show that the inequality above in generla is strict.

Exercise. Let N be a positive integer and consider the opertor S_N respresented by the truncated matrix where we keep a_{pq} for pairs $1 \leq p, q \leq N$ and otherwise all matrix elements are zero. Sow that (*) entails that

$$\{lim_{N\to\infty} ||S - S_N|| = 0$$

Hence the Hilbert-Scmhidt operator S can be approximated in the operator norm by linear operators with a finite-dimensional range. in particular S is a compact operator. The converse is not true, i.e. there exist compact operators on ℓ^2 which are not of the Hilbert-Schmidt type. An example is when we take a sequence of complex numbers $\{c_n\}$ whixh convefges to zero while $\sum |c_n|^2 = \infty$. Let S be the operator defined via the diagonsl matrix where $a_{p,p} = c_p$ for every p. Then S is not a Hilbert-Schmidt operator while the reader should verify that it is a compact operator on ℓ^2.

Carleman's inequality.

We announce a result for matrices and remark that Theorem xx below can be used to derive important results about spectea for Hilbert-Schmidt operators on infinite dimensional Hilbert spaces. See § for an account.

BLABLA ....

Hilbert's spectral theorem.

In his book *Integralgleichungen* from 1904 Hilbert established a fundamental result which extends the spectral theorem for hermitian matrices. A bounded linear operator A on a complex Hilbert space is self-adjoint if

$$\langle Ax, y\rangle = \langle x, Ay\rangle$$

hold for every pair of vectors. It can be expressed by the equality $A = A^*$. Let us now regard a complex number $\lambda = a + ib$ where $b \neq 0$. For a vector x the reader can check that

(i) $$||\lambda \cdot x + Ax||^2 = (a^2 + b^2)||x||^2 + ||Ax||^2 + (a + ib)\langle x, Ax\rangle + (a - ib)\langle Ax, x\rangle$$

Since A is self-adjoint two terms cancel and (i) becomes

$$(a^2 + b^2)||x||^2 + ||Ax||^2 + 2 \cdot \langle ax, Ax\rangle$$

Next, the Cauchy-Schwarz inequality gives

$$2 \cdot \big|\langle ax, Ax\rangle\big| \leq 2 \cdot ||ax|| \cdot ||Ax||$$

Since $r^2 + s^2 - 2rs = (r - s)^2 \geq 0$ for every pair ofm non-negative numbers we conclude that

$$||(a + ib) \cdot x + Ax||^2 \geq b^2 \cdot ||x||^2$$

It follows that the linear operator $S = \lambda \cdot E + A$ is injective and has closed range. Moreover it is surjective for if $y \perp S(\mathcal{H})$ we have

$$0 = \langle \lambda \cdot x + Ax, y\rangle = \langle x, \bar{\lambda} \cdot y + Ay\rangle$$

for every x and then

$$\bar{\lambda} \cdot y + Ay = 0$$

Since the complex conjugate $\bar{\lambda}$ also has a non-zero imaginary part it follows that $y = 0$ and the requested surjectivity follows. Hence S in invertible anbd since λ was an arbitrary non-real complex number the spectrum $\sigma(A)$ consists of real numbers.

The equality $||A||^2 = ||A^2//$**.** Let x be a unit vector. Now we get

$$||Ax||^2 = \langle Ax, Ax\rangle = \langle A^2x, x\rangle$$

By the Cauchy-Schwarz inequslity the absolute vaue of the right hand side above is b ounded by $||A^2x||$ and hence

$$||Ax||^2|| \leq ||A^2x||$$

hold for every x. We can take x so that $||Ax||$ approxiumates the operator norm of A and conclude that

$$||A||^2|| \leq ||A^2||$$

Equality holds since the submultiplicative inequality for operator norms of composed linear operators entail that the right hand side is majorised by $||A||^2||$. We cabn repat ths using the self-adjoint operator A^2 and so on. It follows that

$$||A||^{2m}|| \leq ||A^{2m}||$$

hold for every positive integer m. Taking m:th roots and passing to the limit we get the equality

$$||A|| = \rho(A) = \limsup\, ||A^n||^{\frac{1}{n}}$$

where the right hand side is the spectral radius of A. Now we can apply reults from the general theoru about commutative Banach algebras, applied to the closed subalgebra of $L(\mathcal{H})$ generted by A. The spectral radius formula in Banach algebras together with (xx) entails that $\sigma(A)$ is a compact interval on the real line which contains at least one of the points $||A||$ or $-||A||$.

At this stage Hilbert adopted Carl Neumann's calculus together with measure theory which at this time had been sufficiently devloped by Stieltjes, Borel and Lebesgue. The great merit was of course how Hilbert applied this to a situation where one regards operators on infinite dimensional spaces. The crucial point is the following:

1. Proposition. *For every polynomial $p(t)$ with real coefficients one has the equality*

$$||p(A)|| = \max_{t\in\sigma(A)} |p(t)| \tag{1.1}$$

Proof. By the general formula for spectra in § xx one has

$$\sigma(p(A)) = p(\sigma(A))$$

next, since p has real coefficients it is clear that $p(A)$ is self-adjoint and from the above $\sigma(p(A))$ contains at least on of the points $||p(A)||$ or $-||p(A)||$ and (1.1) follows.

2. Spectral measures. Let A be self-adjoint. Following Hilbert we consider a pair of vectors x, y in $\mathcal{H}$. Then

$$p(t) \mapsto \langle p(A)x, y\rangle \tag{2.1}$$

is a linear functional on the space of real-valued polynomials on the t-line. The right hand side in (2.1) has absolute value bounded by

$$||x|| \cdot ||y|| \cdot ||p(A)|| \tag{2.2}$$

From (1.1) this is equal to $||x|| \cdot ||y||$ times the maximum norm of $P(t)$ on $\sigma(A)$. Now we use a wellkonown result due to Weierstrass which assetts that the space iof real-valued polynomials is dense in the normed space $C^0(\sigma(A)$ of real-valued and continuous functions on the compact set $\sigma(A)$. Together with the Riesz' represention theoren for the dual space we find a unique meausre $\mu_{x,y}$ oin $\sigma(A)$ whose total variation is at most $||x|| \cdot |y||$ and

$$\langle p(A)x, y\rangle = \int_{\sigma(A)} p(t) \cdot d\mu_{x,y}(t) \tag{2.3}$$

hold for every real-valued polynomial. Next, let g be some real-valued continuous function on $\sigma(A)$. Weierstrass theorem gives a sequence of polynomials $\{p_\nu\}$ whixh converge uniformly to g on $\sigma(A)$. In particular the maxium norms

$$\lim |p_n - p_m|_{\sigma(A)} = 0$$

when n and m tend to ∞. From (1.1) it follows that

$$\lim |p_n(A) - p_m(A)|| = 0$$

Now we use that the space of bounded liunear operators on $\mathcal{H}$ is a Banach space, a fact which of course was wellknown already in 1904. We conckude that the given continuous function g yields a unique bounded linear operator $g(A)$ where

$$\lim |p_n(A) - g(A)|| = 0$$

and

$$\langle g(A)x, y\rangle = \int_{\sigma(A)} g(t) \cdot d\mu_{x,y}(t)$$

hold for every pair x, y in $\mathcal{H}$.

3. The spectral resolution. Consider the algebra $\mathcal{B}^\infty(\sigma(A)$ of real-valued bounded Borel functions on the spectrum of A. As explained in [Meausre) one can integrate evety Borel function with resopect to the Riesz measures $\{\mu_{x,y}\}$ when x, y vary in $\mathcal{H}$. Let us fix a bounded Borel function ϕ. Keeping y fixed in $\mathcal{H}$ we get a linear functional on the Hilbert space defined by

$$x \mapsto \int_{\sigma(A)} \phi(t) \cdot d\mu_{x,y}(t)$$

whose norm is majorised by $||y||$ times the sup-norm of ϕ. Since the Hilbert space is self-dual we find a unique vector $\Phi(y)$ such that

$$\langle x, \Phi(y)\rangle = \int_{\sigma(A)} \phi(t) \cdot d\mu_{x,y}(t) \tag{3.1}$$

hold for every x.

Exericse. Verify that the map

$$y \to \Phi(y)$$

is linear and conclude that ϕ yields a bounded linear oeprator Φ for which (3.1) hold for all pairs x, y. Show also that the operator norm

$$||\Phi|| = |\phi|_{\sigma(A)}$$

and that

$$\phi \mapsto \Phi$$

is an algebra homorphism from the Borel algebra to a commutative subalgebra of bounded linear operators on the Hilbert space. Ths, if ϕ and ψ is a pair of bounded Borel functions then $\Phi \circ \Psi$ is the operator assigned to $\phi \cdot \psi$.

4. The operators $E(\delta)$. For every Borel set δ in $\sigma(A)$ its characterstic function is a Borel function which gives an operator denoted by $E(\delta)$. The exercise shows that

$$E(\delta_1) \circ E(\delta_2) = E(\delta_1 \cap \delta_2) \tag{4.1}$$

hold for every pair of Borel sets. In particular

$$E(\delta) \circ E(\delta) = E(\delta)$$

for every single Borel set. These E-operators commute with A and are self-adjoint which means that every $E(\delta)$ is an orthogonal projection from $\mathcal{H}$ onto its range. Moreover, if $\delta_1, \ldots, \delta_N$ is a finite family of disjoint Brel sets whose union is $\sigma(A)$, then $\{E(\delta_\nu)\}$ are pairwise orthogonal and their sum is the identy operator in $\mathcal{H}$. It can also be expressed by the direct sum decomposition

$$\mathcal{H} = \oplus E(\delta_\nu)(\mathcal{H})$$

5. A more refined study. For each vector x and every Borel set δ one has

$$||E(\delta)(x)||^2 = \langle E(\delta)(x), x \rangle = \int_\delta d\mu_{x,x}(t) \tag{5.1}$$

Since the left hand side is non-negative and the Borel set was arbitrary, we conclude that the measure $d\mu_{x,x}$ is non-negative. For every real number s we put

$$\sigma(A)(s) = \sigma(A) \cap (-\infty, s]$$

Keeping x fixed we obtain a non-derasing function

$$s \mapsto \int_{\sigma(A)(s)} d\mu_{x,x}(t) \tag{5.2}$$

This function can have jumps at some ponts, which by general facts is at most denumerable. or it is a continuous function, but not necessarily absolutely contonous in the sense of Lebesgue. For example, absolute continuity fails for every x if $\sigma(A)$ is a null set on the real line. The behaviour of (5.2) depends on the specific vector x. So the overall picture is quite involved.

6. Decomposition of A. To every Borel set δ contained in $\sigma(A)$ we get the operator $A(\delta) = A \cdot E(\delta)$ and fior its spectrum one has the inclusion

$$\sigma(A(\delta)) \subset \overline{\delta}$$

In this way Hilbert decomposed A into a sum of self-adjoint operators whose spectra are confined to small Borel sets in $\sigma(A)$. This constitutes Hilbert's extension of the spectral theorem for matrices where the spectrum is finite, To fully grasp - and also to appreciate - Hilbert's theorem one should of course regard specific cases of self-adjoint operators whose spectra are not reduced to a finite set, and in general contains whole intervals.

The spectral theorem for normal operators.

A bounded linear operator R on the complex Hilbert space is normal if it commutes with its adjont R^*. When this holds and x is a unit vector one gets

$$||R^2x||^2 = \langle R^2x, R^2x\rangle = \langle Rx, R^*R^2x\rangle = \langle Rx, RR^*R2x\rangle = \langle R^*Rx, R^*Rx\rangle$$

The last term is equal to $||R^*Rx||^2$. We have also

$$||Rx||^2 = \langle Rx, Rx\rangle = \langle R^*Rx, x\rangle$$

By the Cauchy-Schwarz inequality the right hand side is majorised by $||R^*Rx||$ when x is a unit vector. We conclude that

$$||R^2x||^2 \geq ||Rx||^4$$

Exactly as in § xx it follows that

$$||R^2|| = ||R||^2$$

and taking higher powers one gets the equality

$$||R|| = \rho(R)$$

Exercise. Show that if $p(z) = c_0 + c_1z + \ldots + c_mz^m$ is a polynomials with complex coefficents, then

$$||p(R)|| = \max_{z\in\sigma(R)} |p(z)|$$

Above $\sigma(R)$ is a compact subset of the complex z-plane which in general contains non-real points. More geenrally, consider a polynomial in z and $\bar{z}$:

$$q(z, \bar{z}) = \sum\sum\ c_{jk} \cdot z^j \cdot \bar{z}^k$$

where a doubly-indexed family of complex coefficients appear. To q one associates the operator

$$q(R, R^*) = \sum\sum\ c_{jk} \cdot R^j \cdot (R^*)^k$$

It turns out that one has the equality

$$||q(R, R^*)|| = \max_{z\in\sigma(R)} |q(z, \bar{z})|$$

To prove this we consider the conmutative subalgebra $L_*(R, R^*)$ of $L(\mathcal{H})$ generated by R and R^*. We can take its closure in thr normed space $L(\mathcal{H})$ which gives a commutative Banach algebra denoted by B. Now we apply Gelfand's general theory from § xx and find the maximal ideal space of this Banach algebra which we denote by $\mathcal{M}(B)$. Every operator S in B is normal and hence the equality in xxx holds. It follows that B is a sup-norm algebra, i.e. the Gelfand transform is norm-preserving. In oteher words, for every operator S in B one has the equality

$$||S|| = \max_{p\in\mathcal{M}(B)} |\widehat{S}(p)|$$

where $\widehat{S}$ is the Gelfand transform. Among the B-elements we have R and R^* and hence also R^*R. The biduality equation $R = R^{**}$ entails that R^*R is self-adjoint and therefore has a real spectrum. This impes that the funvtion

$$p \mapsto \widehat{R^*}(p) \cdot \widehat{R}(p)$$

is real-valued on $\mathcal{M}(B)$. From this the reader can ckeck that the Gelfand transform of R^* is the complex conjugate of that of R. Next, tghe spectrum

$$\sigma(R) = \widehat{R}(\mathcal{M}(B))$$

It turnd out that

$$p \mapsto \widehat{R}(p)$$

is bijective and therefore identifies $\mathcal{M}(B)$ with the spectrum of R. To see this we notice that since R and R^* generate B, it follows that when$p$ and q are teo distinc points in (B) then we cannot have the tewo equalities

$$\widehat{R}(p) = \widehat{R}(q) \quad \& \quad \widehat{R^*}(p) = \widehat{R}(^*q)$$

Since $\widehat{R^*}$ is the compex cojnugte of $\widehat{R}$ this means that

$$\widehat{R}(p) \neq \widehat{R}(q)$$

whoch proves that the map in (xx) is bijective.

Exercise. Confimrm from the above that the eaquality in (xx) holds for every q-polynomial of z and $\bar{z}$.

Spectral resolutions. Exactly as for self-adjoint operatos we find spectral measures starting from a normal operator R. More precisely, one employs Weiertrass's theorem for self-adjoint algebra of continuous funtions, i.e. that every $g \in C^0(\sigma(R)$ can be uniformly approximated by q-polynomiasl as above. This leads to an spectral calcluus where every bounded Borel function ϕ on $\sigma(R)$ gives a bounded linear operator Φ and for every pair x, y, in the Hilbert space one has a unique measure $\mu_{x,y}$ on $\sigma(R)$ and

$$\langle \Phi(x), y \rangle = \int_{\sigma(R)} \phi(z) \cdot d\mu_{x,y}(z)$$

In partocular we can take charactersitic functions of Borel sets in $\sigma(R)$ and obtain operators $E(\delta)$, and exaclty as for self-adjoint oeprstors they are orthogonal projections. We also get the normal operators $R(\delta = R \cdot E(\delta)$ where the spectrum is confined to the clisure of δ for every Borel set.

From the above one has in particular

$$\langle R(x), y \rangle = \int_{\sigma(R)} z \cdot d\mu_{x,y}(z) \quad \& \quad \langle R^*(x), y \rangle = \int_{\sigma(R)} \bar{z} \cdot d\mu_{x,y}(z)$$

It follows that the normal operator R is self-adjoint if and only if the spectrum is confined to the real line. If ϕ is a real-vaöued and bounded Borel function on $\sigma(R)$, it follows that Φ is self-adjoint, i.e. starting form a norma operator R one can produce a quite extensive family of self-adjoint operators with the property that they commute with each other, as well as with R.

4:B. Eigenvalues of matrices.

Using the Hermitian inner product on $\mathbf{C}^n$ we study eigenvalues of an $n \times n$-matrices A with complex elements. The spectrum $\sigma(A)$ is the n-tuple of roots $\lambda_1, \ldots, \lambda_n$ of the characteristic polynomial $P_A(\lambda) = \det(\lambda \cdot E_n - A)$, where eventual multiple eigenvalues are repeated.

4:B.1 Polarisation. Let A be an arbitrary $n \times n$-matrix. Then there exists a unitary matrix U such that the matrix U^*AU is upper triangular. To prove this we first use the wellknown fact that there exists a basis $\xi_1, \ldots, \xi_n$ in $\mathbf{C}^n$ in which A is upper triangular, i.e.

$$A(\xi_k) = a_{1k}\xi_1 + \ldots a_{kk}\xi_k \quad : \ , 1 \leq k \leq n$$

The *Gram-Schmidt orthogonalisation* gives an orthonormal basis $e_1, \ldots, e_n$ where

$$\xi_k = c_{1k} \cdot e_1 + \ldots c_{kk} \cdot e_k \quad \text{for each} \quad 1 \leq k \leq n$$

Let U be the unitary matrix which sends the standard basis in $\mathbf{C}^n$ to the ξ-basis. Now the reader can verify that the linear operator U^*AU is represented by an upper triangular matrix in the ξ-basis.

A theorem by H. Weyl. Let $\{\lambda_k\}$ be the spectrum of A where the λ-sequence is chosen with non-increasing absolute values, i.e. $|\lambda_1| \geq \ldots \geq |\lambda_n|$. We have also the Hermitian matrix A^*A which is non-negative so that $\sigma(A^*A)$ consists of non-negative real numbers $\mu_1 \geq \mu_2 \geq \ldots \geq \mu_n$. In particular one has

$$\mu_1 = \max_{|x|=1} \langle Ax, Ax \rangle \tag{1}$$

4:B.2 Theorem. *For every $1 \leq p \leq n$ one has the inequality*

$$|\lambda_1 \cdots \lambda_p| \leq \sqrt{\mu_1 \cdots \mu_p}$$

where $\{\mu_k\}$ are the eigenvalues of A^*A

First we consider the case $p = 1$ and prove the inequality

$$|\lambda_1| \leq \sqrt{\mu_1} \tag{i}$$

Since λ_1 is an eigenvalue there exists a vector x_* with $|x_*| = 1$ so that $A(x_*) = \lambda_1 \cdot x_*$. It follows from (1) above that

$$\mu_1 \geq \langle A(x_*), A(x_*) \rangle = |\lambda_1|^2$$

Remark. The inequality is in general strict. Consider the 2×2-matrix

$$A = \begin{pmatrix} 1 & a \\ 0 & b \end{pmatrix}$$

where $0 < b < 1$ and $a \neq 0$ some complex number which gives

$$A^*A = \begin{pmatrix} 1 & a \\ a & a^2 + b^2 \end{pmatrix}$$

Here $\lambda_1 = 1$ and the eigenvector $x_* = e_1$ and we see that $\langle A(x_*), A(x_*) \rangle = 1 + |a|^2$.

Proof when $p \geq 2$ We employ a construction of independent interest. Let $e_1, \ldots, e_n$ be some orthonormal basis in $\mathbf{C}^n$. For every $p \geq 2$ we get the inner product space V^p whose vectors are

$$v = \sum c_{i_1, \ldots, i_p} \cdot e_{i_1} \wedge \ldots \wedge e_{i_p}$$

where the sum extends over p-tuples $1 \leq i_1 < \ldots < i_p$. This is an inner product space of dimension $\binom{n}{p}$ where $\{e_{i_1} \wedge \ldots \wedge e_{i_p}\}$ is an orthonormal basis. Consider a linear operator A on $\mathbf{C}^n$ which in the e-basis is represented by a matrix with elements

$$a_{ik} = \langle Ae_i, e_k \rangle$$

If $p \geq 1$ we define the linear operaror $A^{(p)}$ on $V^{(p)}$ by

$$A^{(p)}(e_{i_1} \wedge \ldots \wedge e_{i_p}) = A(e_{i_1}) \wedge \ldots \wedge A(e_{i_p}) = \sum a_{j_1 i_1} \cdots a_{j_p i_p} \cdot e_{j_1} \wedge \ldots \wedge e_{j_p}$$

with the sum extended over all $1 \leq j_1 < \ldots < j_p$.

Sublemma. The eigenvalues of $A^{(p)}$ consists of the $\binom{n}{p}$-tuple given by the products

$$(*) \qquad \lambda_{i_1} \cdots \lambda_{i_m} \quad : \quad 1 \leq i_1 < \ldots < i_p \leq n$$

Proof. The eigenvalues above are independent of the chosen orthonormal basis $e_1, \ldots, e_n$ since a change of this basis gives another orthonormal basis in $V^{(p)}$ which does not affect the eigenvalues of $A^{(p)}$. Using a polarisation from 4:B.1 we may assume from the start that A is an upper triangular matrix and then reader can verify (*) in the sublemma.

Final part of the proof. If $p \geq 2$ it is clear that one has the equality

$$\text{(i)} \qquad (A^{(p)})^* \cdot A^{(p)} = (A^* \cdot A)^{(p)}$$

If $\lambda_1, \ldots \lambda_p$ is the large p-tuple in Weyl's Theorem the product appears as an eigenvalue of $A^{(p)}$ and using the case $p = 1$ one gets Weyl's inequality since the product $\mu_1 \cdots \mu_p$ appears as an eigenvalue of $(A^* \cdot A)^{(p)}$.

4.B.3 An inequality by Pick.

Let $C = \{c_{ik}\}$ be a skew-symmetric $n \times n$-matrix, i.e. $c_{ik} = -c_{ki}$ hold for all pairs i, k. Denote by g the maximum of the absolute values of the matrix elements of C.

4:B.4 Theorem. *One has the inequality*

$$(*) \qquad \max_{|x|=1} \big|\langle Cx, x\rangle\big| \leq g \cdot \cot(\frac{\pi}{2n}) \cdot \sqrt{n(n-1)/2}$$

Proof. Since g is unchanged if we permute the columns of the given C-matrix it suffices to prove (*) for a vector x of unit length such that

$$(1) \qquad \mathfrak{Im}(x_k\bar{x}_i - x_i\bar{x}_k) \geq 0 \quad : \quad 1 \leq i < k \leq n$$

Now one has

$$(3) \qquad \langle Cx, x\rangle = \sum\sum c_{ik}x_k\bar{x}_i = \sum_{i<k} c_{ik}x_k\bar{x}_i + \sum_{i>k} c_{ik}x_k\bar{x}_i = \sum_{i<k} c_{ik}(x_k\bar{x}_i - \bar{x}_k x_i)$$

where the last equality follows since C is skew-symmetric. Put

$$\gamma_{ik} = \mathfrak{Im}\,(x_k\bar{x}_i - \bar{x}_k x_i)$$

Then (1) and the triangle inequality give

$$|\langle Cx, x\rangle| \leq \sum_{i<k} |c_{ik}| \cdot \gamma_{ik} \leq g \cdot \sum_{i<k} \gamma_{ik}$$

Hence there only remains to show that

$$(4) \qquad \sum_{i<k} \gamma_{ik} \leq \cot(\frac{\pi}{2n}) \cdot \sqrt{n(n-1)/2}$$

To prove this we write $x_k = \alpha_k + i\beta_k$ and the reader can verify that (4) follows from the inequality

$$(5) \qquad \sum_{i \neq k} a_k b_i \leq \cot(\frac{\pi}{2n}) \cdot \sqrt{n(n-1)/2}$$

whenever $\{a_k\}$ and $\{b_i\}$ are n-tuples of non-negative real numbers for which $\sum_{k=1}^{k=n} a_k^2 + b_k^2 = 1$. Finally, (4) follows when one applies Lagrange's multiplier for extremals of a quadratic form.

4.B.4 Results by A. Brauer.

Let A be an $n \times n$-matrix. To each $1 \le k \le n$ we set

$$r_k = \min\,[\sum_{j \neq k} |a_{jk}| : \sum_{j \neq k} |a_{kj}|\,]$$

4:B.5 Theorem. *Denote by C_k the closed disc of of radius r_k centered at the diagonal element a_{kk}. Then one has the inclusion:*

(*) $$\sigma(A) \subset C_1 \cup \ldots \cup C_n$$

Proof. Consider some eigenvalue λ so that $Ax = \lambda \cdot x$ for a non-zero eigenvector. It means that

$$\sum_{j=1}^{j=n} a_{j\nu} \cdot x_\nu = \lambda \cdot x_j \quad : \quad 1 \le j \le n$$

Choose k so that $|x_k| \ge |x_j|$ for all j. Now we have

(1) $$(\lambda - a_{kk}) \cdot x_k = \sum_{j \neq k} a_{j\nu} \cdot x_\nu \implies |\lambda - a_{kk}| \le \sum_{j \neq k} |a_{kj}|$$

At the same time the adjoint A^* satisfies $A^*(x) = \bar{\lambda} \cdot x$ which gives

$$\sum_{j=1}^{j=n} \bar{a}_{\nu,j} \cdot x_\nu = \bar{\lambda} \cdot x_j \quad : \quad 1 \le j \le n$$

Exactly as above we get

(2) $$|\lambda - a_{kk}| = |\bar{\lambda} - \bar{a}_{kk}| \le \sum_{j \neq k} |a_{jk}|$$

Hence (1-2) give the inclusion $\lambda \in C_k$.

4:B.6 Theorem. *Assume that the closed discs $C_1, \ldots, C_n$ are disjoint. Then the eigenvalues of A are simple and for every k there is a unique $\lambda_k \in C_k$.*

Proof. Let D be the diagonal matrix where $d_{kk} = a_{kk}$. For ever $0 < s < 1$ we consider the matrix

$$B_s = sA + (1-s)D$$

Here $b_{kk} = a_{kk}$ for every k and the associated discs of the B-matrix are $C_1(s), \ldots, C_b(s)$ where $C_k(s)$ is again centered at a_{kk} while the radius is $s \cdot r_k$. When $s \simeq 0$ the matrix $B \simeq D$ and then it is clear that the previous theorem implies that B_s has simple eigenvalues $\{\lambda_k(s)\}$ where $\lambda_k(s) \in C_k(s)$ for every k. Next, since the "large discs" $C_1, \ldots, C_n$ are disjoint, it follows by continuity that these inclusions holds for every s and with $s = 1$ we get the theorem.

Exercise. Assume that the elements of A are all real and the discs above are disjoint. Show that the eigenvalues of A are all real.

Results by Perron and Frobenius

Let $A = \{a_{pq}\}$ be a matrix where all elements are real and positive. Denote by Δ_+^n the standard simplex of n-tuples $(x_1, \ldots, x_n)$ where $x_1 + \ldots + x_n = 1$ and every $x_k \ge 0$. The following result was established by Perron in [xx]:

4:B.7 Theorem. *There exists a unique $\mathbf{x}^* \in \Delta_+^n$ which is an eigenvector for A with an eigenvalue s^*. Moreover. $|\lambda| < s^*$ holds for every other eigenvalue.*

We leave the proof as an exercise to the reader. In [Frob] the following addendum to Theorem 4:B.7 is proved.

4:B.8 Theorem. *Let A as above be a positive matrix which gives the eigenvalue s^*. For every complex $n \times n$-matrix $B = \{b_{pq}\}$ such that $|b_{pq}| \le a_{pq}$ hold for all pairs p, q, it follows that every root of $P_B(\lambda)$ has absolute value $\le s^*$ and equality holds if and only if $B = A$.*

4:B.9 The case of probability matrices. Let A have positive elements and assume that the sum in every column is one. In this case $s^* = 1$ for with $\mathbf{x}^* = (x_1^*, \ldots, x_n^*)$ we have

$$s^* = s^* \cdot \sum x_p^* = \sum \sum a_{pq} \cdot x_q^* = \sum x_q^* = 1$$

The components of the Perron vector $\mathbf{x}^*$ yields the probabilities to arrive at a station q after many independent motions in an associated stationary Markov chain where the A-matrix defines the transition probabilities.

Example. Let $n = 2$ and take $a_{11} = 3/4$ and $a_{21} = 1/4$, while $a_{12} = a_{22} = 1/2$. A computation gives $s^* = 2/3$ which in probabilistic terms means that the asymptotic probability to arrive at station 1 after many steps is 2/3 while that of station 2 is 1/3. Here we notice that the second eigenvalue is $s_* = 1/4$ and an associated eigenvector is $(1, -2)$.

4.B.10 Extension to infinite dimensions. The Perron-Forbenius result was extended to positive operstors on Hilbert spaces by Pietsch in the article [1912]. We refer to § xx the proof.

Unitary operators.

An operator U on the complex Hilbert space $\mathcal{H}$ is unitary if U is invertible and its inverse is equal to the adjoint U^*. It follows that

$$\langle Ux, Ux\rangle = \langle x, U^*Ux\rangle = ||x||^2$$

Hence $||Ux|| = ||x|$— for every x, i.e. the operator is norm-preserving. Moreover, since every invertible operator commutes with its invrse, it follows that U and U^* commute, i.r. unitary operators are normal. From (xx) the reader can chekc that $\sigma(U)$ is confined to the unit circle. We can apply Hilbert's spectral theorem for bounded normal operators. So to every Brel set δ in T we get a self-adjoint projection $E(\delta)$. A class of unitary operatoies aries when we take a finite family of pairwise disjoint Borel sets $\{\delta_k\}$ whosr union is the unit circle T and define

$$U_* = \sum e^{i\theta_k} \cdot E(\delta_k)$$

The given unitry operator U can be approximated in the operator norm by unitary ooerators in (xx). Notice that U_* has the finite spectrum $\{e^{i\theta_k}\}$.

The approximatins of U lead to a useful inequality. Namely, let P be some self-adjoint prpjection from $\mathcal{H}$ onto a subspace $P(\mathcal{H})$. Then one has

$$||P \circ U(x)|| \leq ||x||$$

for every vector $x \in \mathcal{H}$. In other words, $P \circ U$ is a contraction. To prove this it suffices to regard a unitary operator with a finte spectrum . given by U_* in (1). Iif $x \in \mathcal{H}$ we have

$$x = \sum E_U(\delta_k)(x)$$

where the vectors $\{x_k = E_U(\delta_k)(x)$ are orthogonal and

$$||x||^2 = \sum ||x_k||^2$$

Now

$$P \circ U(x) = \sum e^{i\theta_k} \cdot P(x_k)$$

where the vectors $\{P(x_k)\}$ are orthogonal. It follows that

$$||P \circ U(x)||^2 = \sum ||P(x_k)||^2 \leq ||x||^2$$

where the last inequality holds since $||Py|| \leq ||y||$ for every $y \in \mathcal{H}$. Hence, for every self-adjoint projection P the composed operator $P \circ U$ is a contraction.

A remarkable inequality. Let $p(z) = c_0 + c - 1z + \ldots + c_m z^m$ be a polynomial with complex coefficents. Given a unitsry operator U and a self-adjoint prjection P we put

$$S = \sum c_\nu \cdot (P \circ U)^\nu$$

Let x be a unit vector in $\mathcal{H}$. Approximating U by a unitsry operator with s finite spectrum wehave with rthe notstions form xxxx:

$$S(x) = \sum \sum c_\nu \cdot e^{i\nu\theta_k} \cdot P(x_k)$$

it follows that

$$||S(x)||^2 = \sum \sum_{\nu=1}^{\nu=m} |c_\nu \cdot e^{i\nu\theta_k}| \cdot ||P(x_k)||^2$$

Put

$$|p|_T = \max_\theta |p(e^{i\theta}|$$

Then the right hsnd side in (xx) is majorised by

$$|p[_T \cdot \sum ||P(x_k)||^2 \leq |p|_T \cdot \sum ||x_k||^2 = |p|_T \cdot ||x||^2 = |p|_T$$

where the last equality follows since x has unit norm. Since x was arbitary we get the inequality below for the operator norm:

$$||S|| \leq |p|_T$$

Notice that we can take $P = E$ aobve and hence one has in particular

$$||p(U)|| \leq |p|_T$$

for every polynomial $p(z9$. Let us remakr that this inequality is an immediate consequence of the result in § xx which shows thst when x is a vecotr of unit length and $p(z)$ aa polynomial, then

$$||p(U)x||^2 = \int_T |p(e^{i\theta})|^2 \cdot d\mu_{x,x}(\theta)$$

where $\mu_{x,x}$ is a probability measure. Then (xx follows since the L^2integral in the right hand side above is majorised by $|p|_T^2$.

§ 11. Contractions and the Nagy-Szegö theorem

A linear operator A on the Hilbert space $\mathcal{H}$ is a contraction if its operator norm is ≤ 1, i.e.

(1) $$||Ax|| \leq ||x|| \quad : \quad x \in \mathcal{H}$$

Let E be the identity operator on $\mathcal{H}$. Now $E - A^*A$ is a bounded self-adjoint operator and (1) gives:

$$\langle x - A^*Ax, x \rangle = ||x||^2 - ||Ax||^2 \geq 0$$

The result in § 8.xx shows that this non-negative self-adjoint operator has a square root:

$$B_1 = \sqrt{E - A^*A}$$

Next, the operator norms of A and A^* are equal so A^* is also a contraction and the equation $A^{**} = A$ gives exactöy as above the self-adjoint operator

$$B_2 = \sqrt{E - AA^*}$$

Since $AA^* = A^*A$ is not assumed the self-adjoint operators B_1, B_2 need not be equal. However, the following hold:

11.3.1 Propostion. *One has the equations*

$$AB_1 = B_2A \quad \text{and} \quad A^*B_2 = B_1A^*$$

Proof. If n is a positive integer we notice that

(i) $$A(A^*A)^n = (AA^*)^nA$$

Now A^*A is a self-adjoint operator whose compact spectrum is confined to the closed unit interval $[0, 1]$. If $f \in C^0[0, 1]$ is a real-valued continuous function it can be approximated uniformly by a sequence of polynomials $\{p_n\}$ and the operational calculus from § XX yields an operator $f(A^*A)$ where

$$\lim_{n\to\infty} ||p_n(A^*A) - f(A^*A)|| = 0$$

Since the spectrum of AA^* also is confined to $[0, 1]$, the same polynomial sequence $\{p_n\}$ gives an operator $f(AA^*)$ where

$$\lim ||p_n(AA^*) - f(AA^*)|| = 0$$

Now (i) and the two limit formulas above give:

(ii) $$A \circ f(A^*A) = f(AA^*) \circ A$$

In particular we can take $f(t) = \sqrt{1-t}$ which gives the left hand side in Proposition 11.3.1. By a similar reasoning one proves the equality in the right hand side.

11.2 The unitary operator U_A. On the Hilbert space $\mathcal{H} \times \mathcal{H}$ we define a linear operator U_A represented by the block matrix

(*) $$U_A = \begin{pmatrix} A & B_2 \\ B_1 & -A^* \end{pmatrix}$$

11.3 Proposition. *U_A is a unitary operator on $\mathcal{H} \times \mathcal{H}$.*

Proof. For a pair of vectors x, y in $\mathcal{H}$ we must prove the equality

(i) $$||U_A(x \oplus y)||^2 = ||x||^2 + ||y||^2$$

To get (i) we notice that for every vector $h \in \mathcal{H}$ the self-adjointness of B_1 gives

(ii) $$||B_1h||^2 = \langle B_1h, B_1h \rangle = \langle B_1^2h, h \rangle = \langle h - A^*Ah, h \rangle = ||h||^2 - ||Ah||^2$$

where the last equality holds since we have $\langle A^*Ah, h \rangle = \langle Ah, A^{**}h \rangle = ||Ah||^2$ and the biduality formula $A = A^{**}$. In the same way one has:

(iii) $$||B_2h||^2 = ||h||^2 - ||A^*h||^2$$

Next, by the construction of U_A the left hand side in (i) becomes

(iv) $$||Ax + B_2y||^2 + ||B_1x - A^*y||^2$$

Using (iii) we have

$$||Ax + B_2y||^2 = ||Ax||^2 + ||y||^2 - ||A^*y||^2 + \langle Ax, B_2y\rangle + \langle B_2y, Ax\rangle$$

Similarly, (ii) gives

$$||B_1x - A^*y||^2 = ||x||^2 - ||Ax||^2 + ||A^*y||^2 - \langle B_1x, A^*y\rangle - \langle A^*y, B_x\rangle$$

Adding these two equations we conclude that (i) follows from the equality

(v) $$\langle Ax, B_2y\rangle + \langle B_2y, Ax\rangle = \langle B_1x, A^*y\rangle + \langle A^*y, B_x\rangle$$

To get (v) we use Proposition 11.5.1 which gives

$$\langle Ax, B_2y\rangle = \langle x, A^*B_2y\rangle = \langle x, B_1A^*y\rangle = \langle B_1x, A^*y\rangle$$

where the last equality used that B_1 is self-adjoint. In the same way one verifies that

$$\langle B_2y, Ax\rangle = \langle A^*y, B_1x\rangle$$

and (v) follows.

11.4 The Nagy-Szegö theorem.

The constructions above were applied by Nagy and Szegö to give:

11.4.1 Theorem *For every bounded linear operator A on a Hilbert space $\mathcal{H}$ there exists a Hilbert space $\mathcal{H}^*$ which contains $\mathcal{H}$ and a unitary operator U_A on $\mathcal{H}^*$ such that*

$$A^n = (P \cdot U_A)^n \quad : \quad n = 1, 2, \ldots$$

where $\mathcal{P}\colon \mathcal{H}^* \to \mathcal{H}$ is the orthogonal projection.

Proof. On the product $\mathcal{H}_1 = \mathcal{H} \times \mathcal{H}$ we have the unitary operator U_A from (*) in 11.3.2 and notice that

$$U(x \oplus 0) = Ax$$

This gives the equations in (xx) above.

Application. The Nagy-Szegö result has an interesting consequence. Let A be a contraction. If $p(z) = c_0 + c_1 < + \ldots + c_n z^n$ is a polynomial with complex coefficients we get the operator $p(A) = \sum c_\nu A^\nu$ and with these notations one has:

11.4.2 Theorem *For every pair $A, p(z)$ as above one has*

$$||p(A)|| \leq \max_{z \in D} |p(z)|$$

where the the maximum in the right hand side is taken on the unit disc.

Proof. Theorem 11.4.1 gives $p(A) = p(P \circ U_A)$. Since the orthogonal $\mathcal{P}$-projection is norm decreasing we get

$$||p(A)(\xi)||^2 \leq ||p(U_A)(\xi, 0)||^2$$

Let ξ be a unit vector such that $||p(A)(\xi)|| = ||p(A)||$. The operational calculus in § 7 XX applied to the unitary operator U_A yields a probablity measure μ_ξ on the unit circle such that

$$||p(U_A)(\xi, 0)||^2 = \int_0^{2\pi} |p(e^{i\theta})|^2 \cdot d\mu_\xi(\theta)$$

The right hand side is majorized by $|p|_D^2$ and Theorem 11.4.2 follows.

11.4.3 An application. Let $A(D)$ be the disc algebra. Since each $f \in A(D)$ can be uniformly approximated by analytic polynomials, Theorem 11.4.2 entails that if a linear operator A on the Hilbert space $\mathcal{H}$ is a contraction then each $f \in A(D)$ gives a bounded linear operator $f(A)$, i.e. we have norm-preserving map from the supnorm algebra $A(D)$ into the space of bounded linear operators on $\mathcal{H}$.

§ 12 Miscellanous results

Before Theorem 12.x is announced we recall that the product formula for matrices in § X asserts the following. Let $N \geq 2$ and T is some $N \times N$-matrix whose elements are complex numbers which as usual is regarded as a linear operator on the Hermitian space $\mathbf{C}^N$. Then there exists the self-adjoint matrix $\sqrt{T^*T}$ whose eigenvalues are non-negative. Notice that for every vector x one has

(i) $$||T^*T(x)||^{|}|Tx||^2 \implies ||\sqrt{T^*T}(x)|| = ||Tx||$$

and since $\sqrt{T^*T}$ is self-adjoint we have an orthogonal decomposition

(ii) $$\sqrt{T^*T}(\mathbf{C}^N) \oplus \mathrm{Ker}(\sqrt{T^*T}) = \mathbf{C}^N$$

where the self-adjointness gives the equality

(iii) $$\mathrm{Ker}(\sqrt{T^*T}) = \sqrt{T^*T}(\mathbf{C}^N)^\perp$$

The partial isometry operator. Show that there exists a unique linear operator P such that

(*) $$T = P \cdot \sqrt{T^*T}$$

where the P-kernel is the orthogonal complement of the range of $\sqrt{T^*T}$. Moreover, from (i) it follows that

$$||P(y)|| = ||y||$$

for each vector in the range of $\sqrt{T^*T}$. One refers to P as a partial isometry attached to T.

Extension to operators on Hilbert spaces.. Let T be a bounded opertor on the Hilbert space $\mathcal{H}$. The spectral theorem for bounded and self-adjoint operators gives a similar equation as in (*) above using the non-negative and self-adjoint operator $\sqrt{T^*T}$. More generally, let T be densely defined and closed. From § XX there exists the densely defined self-adjoint operator T^*T and we can also take its square root.

12.1 Theorem. *There exists a bounded partial isometry P such that*

$$T = P \cdot \sqrt{T^*T}$$

Proof. Since T has closed graph we have the Hilbert space $\Gamma(T)$. For each $x \in \mathcal{D}(T)$ we get the vector $x_* = (x, Tx)$ in $\Gamma(T)$. Now

$$(x_*.y_*) \mapsto \langle x, y \rangle$$

is a bounded Hermitiain bi-linear form on the Hilbert space $\Gamma(T)$. The self-duality of Hilbert spaces gives bounded and self-adjoint operator A on $\Gamma(T)$ such that

$$\langle x, y \rangle = \{Ax_*, y_*\}$$

where the right hand side is the inner product between vectors in $\Gamma(T)$. Let

$$j \colon (x, Tx) \mapsto x$$

be the projection from $\Gamma(T)$ onto $\mathcal{D}(T)$ and for each $x \in \mathcal{D}(T)$ we put

$$Bx = j(Ax_*)$$

Then B is a linear operator from $\mathcal{D}(T)$ into itself where

(i) $$\langle Bx, y \rangle = \{Ax_*, y_*\} = \{x_*, Ay_*\} = \langle x, By \rangle \quad : \quad x, y \in \mathcal{D}(T)$$

We have also

$$\langle Bx, x \rangle = \{A^2 x_*, x_*\} = \{Ax_*, Ax_*\} = \langle Bx, Bx \rangle + \langle TBx, TBx \rangle \implies$$

$$||Bx||^2 = \langle Bx, Bx \rangle \leq \langle Bx, x \rangle \leq ||Bx|| \cdot ||x||$$

where the Cauchy-Schwarz inequality was used in the last step. Hence

$$||Bx|| \leq ||x|| \quad : \quad x \in \mathcal{D}(T)$$

This entails that that the densey defined operator B extends uniquely to $\mathcal{H}$ as a bounded operator of norm ≤ 1. Moreover, since (i) hold for pairs x, y in the dense subspace $\mathcal{D}(T)$, it follows that B is self-adjoint. Next, consider a pair x, y in $\mathcal{D}(T)$ which gives

$$\langle x, y\rangle = \{Ax_*, y_*\} = \{x_*, Ay_*\} = \langle x, By\rangle + \langle Tx, TBy\rangle$$

Keeping y fixed the linear functiional

$$x \mapsto \langle Tx, TBy\rangle = \langle x, y\rangle - \langle x, By\rangle$$

is bounded on $\mathcal{D}(T)$. By the construction of T^* it follows that $TBy \in \mathcal{D}(T^*)$ and we also get the equality

(ii) $$\langle x, y\rangle = \langle x, By\rangle + \langle x, T^*TBy\rangle$$

Since (ii) holds for all x in the dense subspace $\mathcal{D}(T)$ we conclude that

(iii) $$y = By + T^*TBy = (E + T^*T)(By) \quad : \quad y \in \mathcal{D}(T)$$

Conclusion. From the above we have the inclusion

$$TB(\mathcal{D}(T)) \subset \mathcal{D}(T^*)$$

Hence $\mathcal{D}(T^*T)$ contains $B(\mathcal{D}(T))$ and (iii) means that B is a right inverse of $E + T^*T$ provided that the y-vectors are restricted to $\mathcal{D}(T)$.

FINISH ..

12.2 Positive operators on $C^0(S)$

Let S be a compact Hausdorff space and X the Banach space of continuous and complex-valued functions on S. A linear operator T on X is positive if it sends every non-negative and real-valued function f to another real-valued and non-negative function. Denote by $\mathcal{F}^+$ the family of positive operators T which satisfy the following: First

(1) $$\lim_{n\to\infty} \frac{1}{n} \cdot x^*(T^n x) = 0$$

hold for all pairs $x \in X$ and $x^* \in X^*$. The second condition is that $\sigma(T)$ is the union of a compact set in a disc $\{|\lambda| \leq r$ for some $r < 1$, and a finite set of points on the unit circle. The final condition is that $R_T(\lambda)$ is meromorphic in the exterior disc $\{|\lambda > r\}$, i.e. it has poles at the spectral points on the unit circle.

12.2.1. Theorem. *If $T \in \mathcal{F}^+$ then each spectral value $e^{i\theta} \in \sigma(T)$ is a root of unity.*

Proof. Frist we prove that $R_T(\lambda)$ has a simple pole at each $e^{i\theta} \in \sigma(T)$. Replacing T by $e^{-i\theta} \cdot T$ it suffices to prove this when $e^{i\theta} = 1$. If $R_T(\lambda)$ has a pole of order ≥ 2 at $\lambda = 1$ we know from § XX that there exists $x \in X$ such that

(i) $$Tx \neq x \quad \text{and} \quad (E - T)^2\, x = 0$$

This gives $T^2 + x = 2Tx$ and by an induction

(ii) $$\frac{1}{n} \cdot T^n x = \frac{1}{n} \cdot x + (E - T)x \quad : n = 1, 2, \ldots$$

Condition (1) and (ii) give for each $x^* \in X^*$:

$$0 = \lim_{n\to\infty} \frac{1}{n} \cdot x^*(T^n x) = \lim_{n\to\infty} x^*(\frac{1}{n} \cdot x + (E - T)x)$$

It follows that $x^*(E - T)(x) = 0$ and since x^* is arbitrary we get $Tx = x$ which contradicts (i). Hence the pole must be simple.

Next, with $e^{i\theta} \in \sigma(T)$ we have seen that R_T has a simple pole. By the general result in § xx there exists some $f \in C^0(S)$ which is not identically zero and

$$T(f) = e^{i\theta} \cdot f$$

Multiplying f with a complex scalar we may assume that its maximum norm on S is one and there exists a point $s_0 \in S$ such that

$$f(s_0) = 1$$

For each $n \geq 1$ we have a linear functional on X defined by $g \mapsto T^n(g)(s_0)$ whuch gives a Riesz measure μ_n such that

$$\int_S g \cdot d\mu_n = T^n g(s_0) \quad : g \in C^0(S)$$

Since T^n is positive the integrals in the left hand side are ≥ 0 when g are real-valued and non-negative which entails that the measures $\{\mu_n\}$ are real-valued and non-negative. For each $n \geq 1$ we put

$$A_n = \{x : \, e^{-in\theta} \cdot f(x) \neq 1\}$$

Since the sup-norm of f is one we notice that

(iii) $$A_n = \{x : \, \mathfrak{Re}(e^{-in\theta} f(x)) < 1\}$$

Now

(iv) $$0 = f(s_0) - e^{-in\theta} \cdot T^n f(s_0) = \int_S [1 - e^{-in\theta} f(s)] \cdot d\mu_n(s)$$

Taking real parts we get

(v) $$0 = \int_S [1 - \mathfrak{Re}(e^{-in\theta} f(s))] \cdot d\mu_n(s)$$

By (iii) the integrand in (v) is non-negative and since the whole integral is zero it follows that

(vi) $$\mu_n(A_n) = \mu_n(\{\mathfrak{Re}(e^{-in\theta} < 1\} = 0$$

Suppose now that there exists a pair $n \neq m$ such that

(vii) $$(S \setminus A_n) \cap (S_m \setminus A_m) \neq \emptyset$$

A point s_* in this non-empty intersection gives

$$1 = e^{in\theta} f(s_*) = e^{im\theta} \cdot f(s_*) \implies e^{in\theta} = e^{im\theta}$$

and hence $e^{i\theta}$ is a root of unity. $m - n \neq 0$. So the proof of Theorem 6.1 is finished if we have established the following

Sublemma. The sets $\{S \setminus A_n\}$ cannot be pairwise disjoint.

Proof. First, f has maximum norm and by the above:

$$\int_S f \cdot d\mu_n = e^{in\theta}$$

Hence the total mass $\mu_n(S)$ is at least one. Next, for each $n \geq 2$ we set

$$\pi_n = \frac{1}{n} \cdot (\mu_1 + \ldots + \mu_n)$$

Since $\mu_n(S) \geq 1$ for each n we get $\pi_n(S) \geq 1$. Put

$$\mathcal{A} = \bigcap A_n$$

Above we proved that $\mu_n(A_n) = 0$ hold for every n which gives

(*) $$\pi_n(\mathcal{A}) = 0 \quad : \, n = 1, 2, \ldots$$

Next, when the sets $\{S \setminus A_k\}$ are pairwise disjoint one has the inclusions

$$S \setminus A_k \subset A_\nu \quad \forall \, \nu \neq k$$

Keeping k fixed it follows that $\pi_\nu(S \setminus A_k) = 0$ for every $\nu \geq 0$. So when n is large while k is kept fixed we obtain

$$(**) \qquad \pi_n(S \setminus A_k)) = \frac{1}{n} \cdot \mu_k(S \setminus A_k)) \implies \lim_{n\to\infty} \pi_n(S \setminus A_k)) = 0 \quad : k = 1, 2, \ldots$$

At this stage we use Lemma xx which shows that $R_T(\lambda)$ has at most a simple pole at $\lambda = 1$. With $\epsilon > 0$ the Neumann series expansion gives

$$E + \sum_{k=1}^{\infty} \frac{T^k}{(1+\epsilon)^k} = R_T(1+\epsilon)) = \frac{1}{\epsilon} \cdot Q + W(1+\epsilon))$$

where $W(\lambda)$ is an operator-valued analytic function in an open disc centered at $\lambda = 1$ while Q is a bounded linear operator on $C^0(S)$. Keeping $\epsilon > 0$ fixed we apply both sides to the identity function 1_S on S and the construction of the measures $\{\mu_n\}$ gives

$$1 + \sum_{k=1}^{\infty} \frac{\mu_k(S)}{(1+\epsilon)^k} = \frac{1}{\epsilon} \cdot Q(1_S)(s_0) + W(1+\epsilon))(1_S)(s_0)$$

If $n \geq 2$ is an integer and $\epsilon = \frac{1}{n}$ one gets the inequality

$$\sum_{k=1}^{k=n} \frac{\mu_k(S)}{(1+\frac{1}{n})^k} \leq n \cdot |Q(1_S)(s_0)| + |W(1+1/n))(1_S)(s_0)| \leq n||Q|| + ||W(1+1/n)|| \implies$$

$$\frac{1}{n} \cdot \sum_{k=1}^{k=n} \mu_k(S) \leq (1+\frac{1}{n})^n \cdot (||Q|| + \frac{||W(1+1/n)||}{n}$$

Since Neper's constant $e \geq (1+\frac{1}{n})^n$ for every n we find a constant C which is independent of n such that

$$\frac{1}{n} \cdot \sum_{k=1}^{k=n} \mu_k(S) \leq C$$

Hence the sequence $\{\pi_n(S)\}$ is bounded and we can pass to a subsequence which converges weakly to a limit measure μ_*. For this σ-additive measure the limit formula in (**) above entails that

$$\text{(i)} \qquad \mu_*(S \setminus A_k) = 0 \quad : \ k = 1, 2, \ldots$$

Moreover, by (*) we also have

$$\text{(ii)} \qquad \pi_*(\mathcal{A}) = 0$$

Now $S = \mathcal{A} \cup A_k$ so (i-ii) give:

$$\mu_*(S) = 0$$

But this is impossible for at the same time we have already seen that $\pi_n(S) \geq 1$ for each n and hence also $\mu_*(S) \geq 1$.

Compact pertubations to finish Kakutani-Yosida !!!

In general, consider some complex Banch space X be a Banach space and denote by $\mathcal{F}(X)$ the family of bounded liner operators T on X such that

$$(*) \qquad \lim_{n\to\infty} \frac{1}{n} \cdot x^*(T^n x) = 0$$

hold for all pairs $x \in X$ and $x^* \in X^*$.

1. Exercise. Apply the Banach-Steinhaus theorem to show that if $T \in \mathcal{F}(X)$ then there exists a constant M such that the operator norms satisfy

$$||T^n|| \leq M \cdot n \quad : \ n = 1, 2, \ldots$$

Since the n:th root of $M \cdot n$ tends to one as $n \to +\infty$, the spectral radius formula entails that the spectrum $\sigma(T)$ is contained in the closed unit disc of the complex λ-plane. So in the exterior disc $\{|\lambda| > 1\}$ there exists the the resolvent

$$R_T(\lambda) = (\lambda \cdot E - T)^{-1}$$

2. The class $\mathcal{F}_*$. It consists of those T in $\mathcal{F}(X)$ for which there exists some $\alpha < 1$ such that $R_T(\lambda)$ extends to a meromorphic function in the exterior disc $\{|\lambda| > \alpha\}$. Since $\sigma(T) \subset \{|\lambda| \leq 1\}$ it follows that when $T \in \mathcal{F}_*$ then the set of points in $\sigma(T)$ which belongs to the unit circle in the complex λ-plane is empty or finite and after we can always choose $\alpha < 1$ such that

$$\sigma(T) \cap \{\alpha < |\lambda| < 1\} = \emptyset$$

3. Proposition. *If $T \in \mathcal{F}_*$ and $e^{i\theta} \in \sigma(T)$ for some θ, then Neumann's resolvent $R_T(\lambda)$ has a simple pole at $e^{i\theta}$.*

Proof. Replacing T by $e^{-i\theta} \cdot T$ it suffices to prove the result when $e^{i\theta} = 1$. If $R_T(\lambda)$ has a pole of order ≥ 2 at $\lambda = 1$ we know from § XX that there exists $x \in X$ such that

(i) $$Tx \neq x \quad \text{and} \quad (E - T)^2 x = 0$$

The last equation means that $T^2 + x = 2Tx$ and an induction over n gives

(ii) $$\frac{1}{n} \cdot T^n x = \frac{1}{n} \cdot x + (E - T)x$$

Since $T \in \mathcal{F}$ we have

(iii) $$\lim_{n \to \infty} \frac{1}{n} \cdot x^*(T^n x) = 0 \quad : \forall, x^* \in X^*$$

Then (ii) entails that $x^*(E - T)(x) = 0$. Since x^* is arbitrary we get $Tx = x$ which contradicts (i) and hence the pole is simple.

4. Theorem. *Let $T \in \mathcal{F}(X)$ be such that there exists a compact operator K where $||T + K|| < 1$. Then $T \in \mathcal{F}_*$ and for every $e^{i\theta} \in \sigma(T)$ the eigenspace $E_T(\theta) = \{x \in X : Tx = e^{i\theta} x\}$ is finite dimensional.*

Proof. Set $S = T + K$ and for a complex number λ we write $\lambda \cdot E - T = \lambda \cdot E - T - K + K$. Outside $\sigma(S)$ we get

(i) $$R_S(\lambda)(\lambda \cdot E - T) = E + R_S(\lambda) \cdot K$$

The Neumann series for large absolute values $|\lambda|$ applied to $R_S(\lambda)$ gives some $\rho > 0$ and

(ii) $$(E + R_S(\lambda) \cdot K)^{-1} = E + R_S(\lambda) \cdot K\dot{(}E + R_S(\lambda) \cdot K)^{-1} \quad : |\lambda| > \rho$$

Next, when $|\lambda|$ is large we notice that (i) gives

(iii) $$R_T(\lambda) = (E + R_S(\lambda) \cdot K)^{-1} \cdot R_S(\lambda)$$

Together with (ii) we obtain

(iv) $$R_T(\lambda) = R_S(\lambda) + R_S(\lambda) \cdot (E + R_S(\lambda) \cdot K)^{-1} \cdot R_S(\lambda)$$

Set $\alpha = ||S||$ which by assumption is < 1. Now $R_S(\lambda)$ is analytic in the exterior disc $\{\lambda| > \alpha\}$ so in this exterior disc $R_\lambda(T)$ differs from the analytic function $R_\lambda(S)$ by

(v) $$\lambda \mapsto R_S(\lambda) \cdot (E + R_S(\lambda) \cdot K)^{-1} \cdot R_S(\lambda)$$

Here K is a compact operator so the result in § XX entails that this function extends to be meromorphic in $\{|\lambda| > \alpha\}$. There remains to prove that eigenspaces at spectral points on the unit circle are finite dimensional. To prove this we use (iv). Let $e^{i\theta} \in \sigma(T)$. By Proposition 3 it is a simple pole so we have a Laurent series expansion

$$R_T(e^{i\theta} + z) = \frac{A_{-1}}{z} + A_0 + A_1 z + \ldots$$

By the general results from §§ there remains to show that A_{-1} has finite dimensional range. To see this we apply (iv) which gives the equation

$$R_S(e^{i\theta}+z)+R_S(e^{i\theta}+z)\cdot(E+R_S(e^{i\theta}+z)\cdot K)^{-1}\cdot R_S(e^{i\theta}+z)$$

To simplify notations we set $B(z)=R_S(e^{i\theta}+z)$ which by assumption is analytic in a neighborhood of $z=0$. Moreover, the operator $B(0)$ is invertible. So now one has

$$\frac{A_{-1}}{z}+A_0+A_1z+\ldots=B(z)+B(z)(E+B(z)\cdot K)^{-1}B(z)$$

Since $B(0)$ is invertible we have a Laurent series expansion

$$(E+B(z)\cdot K)^{-1}=\frac{A^*_{-1}}{z}+A^*_0+A^*_1z+\ldots$$

and identying the coefficient of z^{-1} gives

$$A_{-1}=B(0)A^*_{-1}B(0)$$

Next, from (xx) one has

$$E=(E+B(z)\cdot K)(\frac{A^*_{-1}}{z}+A^*_0+A^*_1z+\ldots)\implies(E+B(0)\cdot K)A^*_{-1}=0$$

Here $B(0)\cdot K$ is a compact operator and hence Fredholm theory implies that A^*_{-1} has a finite dimensional range. Since $B(0)$ is invertible the same is true for A_{-1} which finishes the proof of Theorem 4.

5. Proposition. *If $T\in\mathcal{F}$ is such that $T^N\in\mathcal{F}_*$ for some integer $N\geq 2$. Then $T\in\mathcal{F}_*$.*

Proof. We have the algebraic equation

$$\lambda^N\cdot E-T^N=(\lambda\cdot E-T)(\lambda^{N-1}\cdot E+\lambda^{N-2}\cdot T+\ldots+T^{N_1})$$

It follows that

$$R_T(\lambda)=(\lambda^{N-1}\cdot E+\lambda^{N-2}\cdot T+\ldots+T^{N_1})\cdot R_{T^N}(\lambda^N)$$

Since $T^NB\in\mathcal{F}_*$ there exists $\alpha<1$ such that

$$\lambda\mapsto R_{T^N}(\lambda^N)$$

extends to be meromorphic in $\{|\lambda|>\alpha\}$. At the same time $(\lambda^{N-1}\cdot E+\lambda^{N-2}\cdot T+\ldots+T^{N_1})$ is a polynomial and hence $R_T(\lambda)$ also extends to be meromorphic in this exterior disc so that $T\in\mathcal{F}_*$.

12.3 Factorizations of non-symmetric kernels.

Recall that the Neumann-Poincaré kernel $K(p,q)$ of a plane C^1-curve $\mathcal{C}$ is given by

$$K(p,q) = \frac{\langle p-q, \mathbf{n}_i(p)\rangle}{|p-q|^2}$$

This kernel function gives the integral operator $\mathcal{K}$ defined on $C^0(\mathcal{C})$ by

$$\mathcal{K}_g(p) = \int_C K(p,q) \cdot g(q)\, ds(q)$$

where ds is the arc-length measure on C. Let M be a positive number which exceeds the diameter of $\mathcal{C}$ so that $|p-q| < M :\ p,q \in \mathcal{C}$. Set

$$N(p,q) = \int_{\mathcal{C}} K(p,\xi) \cdot \log \frac{M}{|q-\xi|} \cdot ds(\xi)$$

Exercise. Verify that N is symmetric, i.e. $N(p,q) = N(q,p)$ hold for all pairs p,q in $\mathcal{C}$. Moreover,

$$S(p,q) = \log \frac{M}{|p-q|}$$

is a symmetric and positive kernel function and since $\mathcal{C}$ is of class C^1 the reader should verify that it gives a Hilbert-Schmidt kernel, i.e.

$$\iint_{\mathcal{C}\times\mathcal{C}} S(p,q)^2\, ds(p)ds(q) < \infty$$

Hence the Neuman-Poincaré operator $\mathcal{K}$ appears in an equation

$$(*) \qquad \mathcal{N} = \mathcal{K} \circ \mathcal{S}$$

where $\mathcal{S}$ is defined via a positive symmetric Hilbert-Schmidt kernel and $\mathcal{N}$ is symmetric. Following [Carleman: § 11] we give a procedure to determine the spectrum of $\mathcal{K}$.

12.3.1 Spectral properties of non-symmetric kernels.

Let $K(x,y)$ be a continuous real-valued function on the closed unit square $\square = \{0 \leq x, y \leq 1\}$. We do not assume that K is symmetric but there exists a positive definite Hilbert-Schmidt kernel $S(x,y)$ such that

$$N(x,y) = \int_0^1 S(x,t)K(t,y)\,dy$$

yields a symmetric kernel function, i,e, $N(x,y) = N(y,x)$. The Hilbert-Schmidt theory gives an orthonormal basis $\{\phi_n\}$ in $L^2[0,1]$ formed by eigenfunctions to $\mathcal{S}$ where

$$\mathcal{S}\phi_n = \kappa_n\phi_n \tag{1}$$

where the positive κ-numbers tend to zero, and each $u \in L^2[0,1]$ has a Fourier-Hilbert expansion

$$u = \sum \alpha_n \cdot \phi_n \tag{2}$$

We seek eigenfunctions of the integral operator $\mathcal{K}$. Let u be a function in $L^2[0,1]$ such that:

$$u = \lambda \cdot \mathcal{K}u \tag{3}$$

where λ in general is a complex number. It follows that

$$\lambda \cdot \int N(x,y)u(y)\,dy = \lambda \iint S(x,t)K(t,y)u(y)\,dtdy = \int S(x,t)u(t)\,dt \tag{4}$$

Take some p and multiply both sides with $\phi_p(x)$ and then an integration gives

$$\lambda \cdot \int \phi_p(x)N(x,y)u(y)\,dxdy = \iint \phi_p(x)S(x,t)u(t)\,dxdt = \kappa_p \int \phi_p(t)u(t)\,dt \tag{5}$$

Set

$$c_{qp} = \iint \phi_q(x)\phi_p(x)N(x,y)\,dxdy$$

Since $N(x,y) = N(y,x)$ the doubly indexed c-sequence is symmetric. Next, the expnasion of u in (2) shows that the left hand side in (5) becomes

$$\lambda \cdot \sum \alpha_q \cdot c_{pq}$$

At the same time we nortice that since the ϕ-functions is an orthonormal basis, the right hand side in (5) is $\kappa_p \cdot \alpha_p$. Hence the α-sequence satisfies the system of equations

$$\kappa_p \cdot \alpha_p = \lambda \cdot \sum \alpha_q \cdot c_{pq}$$

Next, the expansion of u from (2) gives the equations:

$$\sum_{q=1}^{\infty} \alpha_q \cdot \iint \phi_q(x)\phi_p(x)N(x,y)\,dxdy = \kappa_p\alpha_p \quad : \; p = 1, 2, \ldots \tag{6}$$

$$\beta_p = \sqrt{\kappa_p} \cdot \alpha_p \implies \beta_p = \lambda \cdot \sum_{q=1}^{\infty} \frac{c_{pq}}{\sqrt{\kappa_p} \cdot \sqrt{\kappa_q}} \cdot \beta_q \tag{8}$$

Next, put

$$k_{p,q} = \iint K(x,y)\phi_p(x)\phi_q(y)\,dxdy \tag{9}$$

From the above the following hold for each pair p, q:

$$c_{pq} = \iiint \phi_q(x)\phi_p(y)S(x,t)K(t,y)\,dxdydt = \kappa_q k_{p,q} = \kappa_p k_{q,p} \implies$$

$$\frac{c_{p,q}^2}{\kappa_p \cdot \kappa_q} \leq |k_{p,q} \cdot k_{q,p}| \leq \frac{1}{2}(k_{p,q}^2 + k_{q,p}^2) \tag{10}$$

Here $\{k_{p,q}\}$ are the Fourier-Hilbert coefficients of $K(x,y)$ which entails that

$$\sum\sum k_{p,q}^2 \leq \iint K(x,y)^2\, dxdy$$

Hence the symmetric and doubly indexed sequence

$$\frac{c_{p,q}}{\sqrt{\kappa_p \cdot \kappa_q}} \tag{11}$$

is of Hilbert-Schmidt type.

11.6.2 Conclusion. The eigenfunctions u in $L^2[0,1]$ associated to the $\mathcal{K}$-kernel have Fourier-Hilbert expansions via the $\{\phi_n\}$-basis which are determined by α-sequences satisfying the system (7)

11.6.3 Remark. When a plane curve $\mathcal{C}$ has corner points the Neumann-Poincaré kernel is unbounded. Here the reduction to the symmetric case is more involved and leads to quite intricate results which appear in Part II from [Carleman]. The interplay between singularities on boundaries in the Neumann-Poincaré equation and the corresponding unbounded kernel functions illustrates the general theory densely defined self-adjoint operators. Much analysis remains to be done and open problems about the Neumann-Poincaré equation remains to be settled in dimension three. So far it appears that only the 2-dimensional case is properly understood via results in [Car:1916]. See also § xx for a studiy of Neumann's boundary value problem both in the plane and $\mathbf{R}^3$.