Appendix B: Functional analysis.

Contents.

Introduction.

An important topic is the operational calculus in § 7 and Sections 8 and 9 expose the spectral theorem for self-adjoint operators on Hilbert spaces which is the main issue in this chapter. General results in functional analysis arise via extensions from calculus in several real variables where convexity plays an essential role. In spite of relatively easy proofs, the general facts are very useful and can be applied in many situations. The passage from finite to infinite-dimensional vector spaces can give rise to new phenomena such as Enflo's construction of a separable Banach space B on which there exists a compact operator $T\colon B \to B$ which cannot be approximated in the operator norm by by operators with finite range. This is remarkable since every normed vector space whose unit ball is compact in the norm topology must be finite dimensional. Enflo's example shows therefore that the geometry in infinite dimensional normed spaces can be quite involved. An extensive literature treats various geometric conditions which can be imposed on normed spaces, and more generally on locally convex topological vector spaces We shall not discuss this in detail and examples which exhibit dual spaces will only be mentioned briefly. For a more extensive introduction to functional analysis we recommend the text-book [Taylor] and volume 1 in [Dunford-Schwarts].

The general notion of a normed linear space was put forward by Banach in the monograph [Banach]. Infinite dimensional systems of equations had been studied before, but these considerations were not put in full generality. A notable exception are Hilbert spaces whose abstract definition was given by Hilbert in 1904. Hilbert extended the Gram-Schmidt orthogonalisation to infinite dimensional vector spaces equipped with a hermitain inner product which implies that every separable Hilbert space is isomorphic to ℓ^2 whose elements are sequences of complex numbers $\{c_n\}$

for which $\sum |c_n|^2 < \infty$. A bounded linear operator T on ℓ^2 is represented by a doubly indexed sequence $\{a_{pq}\}$ such that where

$$(*) \qquad Te_p = \sum_{q=1}^{\infty} a_{pq} \cdot e_q$$

and $\{e_q\}$ is the ortohonormal basis where a unit at place q and zeros elsewhere are assigned for every positive integer q. The condition that $Te_p \in \ell^2$ means that

$$\sum_{q=1}^{\infty} |a_{pq}|^2 < \infty \quad \text{for every} \quad p$$

However, this is not sufficient in order that T is a bounded operator. The necessary and sufficient condition for boundedness is that there exists a constant C and

$$\sum_{p=1}^{\infty} \Big|\sum_{q=1}^{\infty} |a_{pq} \cdot x_q \Big|^2 \leq C \cdot \sum |x_n|^2$$

hold for every ℓ^2-sequence $\{x_n\}$. The normed linear space of all bounded linear operators on ℓ^2 is dentoed by $L(\ell^2)$. This is a huge space which is not separable. A non-denumerable set of linear operators with unit norm arises when we for each sequence of signs +1 or -1 associate the linear operator

$$T(e_p) = \epsilon_p \cdot e_p \quad : \quad \epsilon_p = + \text{ or } -1$$

Self-adjoint operators. A bounded linear operator A on ℓ^2 is self-adjoint if the doubly-indexed sequence $\{a_{pq}\}$ is hermitian, i.e.

$$\bar{a}_{pq} = a_{qp}$$

hold for all pairs p, q. Hilbert proved that every self-adjoint operator$A$ has a spectrum $\sigma(A)$ confined to a compact real subset of the real line and constructed a spectral decomposition of A. More precisely, let $\mathcal{B}(\sigma(A))$ denote the algebra of complex-valued and bounded Borel functions on $\sigma(A)$ which was introduced in [Measure appendix]. Then there exists an injective algebra homomorphism

$$g \mapsto g(A)$$

from $\mathcal{B}(\sigma(A))$ into $L(\ell^2)$ and to each pair of vectors x, y in ℓ^2 one has a unique Riesz measure $\mu_{x,y}$ on $\sigma(A)$ such that

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

If $p(t)$ is a polynomial then $p(A)$ is the corresponding polynomial in A taken in $L(\ell^2)$. In § 0.B below we give further comments about Hilbert's result and detailed proofs appear in § 8 after we have presented the operational calculus in § 7.

Let us give an example of a "concrete" result due to F. Riesz which was established before the general notion of normed vector spaces became standard. Consider a doubly indexed sequence $\{c_{n,\nu}\}$ of complex numbers indexed by pair of non-negative integers. Assume that

$$(1) \qquad \lim_{n \to +\infty} \max_{\nu} |c_{n,\nu}| = 0$$

Under this assumption we study solutions to inhomogeneous systems of linear equations

$$(*) \qquad \sum_{n=1}^{\infty} c_{n,\nu} \cdot x_n = y_\nu$$

where $\{y_\nu\}$ is a bounded sequence of complex numbers. The equation is solvable in ℓ^1 if there exists a sequence $\{x_n\}$ which is absolutely convergent, i.e. $\sum |x_n| < \infty$. The following result was proved by Riesz:

Theorem. *Let* $\{c_{n,\nu}\}$ *and* $\{y_\nu\}$ *be such that for every finite sequence* $\lambda_0,\ldots,\lambda_r$ *it holds that*

$$|\sum_{\nu=0}^{\nu=r}\lambda_\nu\cdot y_\nu|\leq\sup_{n\geq 0}|\sum_{\nu=0}^{\nu=r}\lambda_\nu\cdot c_{n,\nu}|$$

Then (*) has an ℓ^1*-solution such that* $\sum|x_n|\leq 1$.

Here follows an example where Riesz' theorem can be applied. Let E be a compact subset of the unit circle T. We have the Banach space $C^0(E)$ of complex-valued continuous functions on E. Measure theory teaches that the dual space consists of Riesz measures μ of finite total variation on E. Let $M(E)$ denote this set of measures. To every $\mu\in M(E)$ and each non-negative integer we set

$$\widehat{\mu}(n)=\int_E e^{in\theta}\cdot d\mu(\theta)$$

We say that E is a Carleson-Kronecker set if there exists $0<p(E)\leq 1$ such that

$$||\mu||\leq\frac{1}{p(E)}\cdot\sup_{n\geq 0}|\widehat{\mu}(n)|\quad\text{hold for all}\quad\mu\in M(E)$$

Theorem. *Let* E *be a Carleson-Kronecker set. Then, for each* $\phi\in C^0(E)$ *there exists an absolutely convergent sequence* $\{x_n\}$ *such that*

$$\phi(e^{i\theta})=\sum_{n\geq 0}x_n\cdot e^{in\theta}\quad\text{holds on}\quad E$$

Remark. We prove this theorem in § XX in *Special Topics*.

L^p**-inequalities.** Results which classically were constrained to rather specified situations extend to far more general cases when notions in functional analysis are adopted. Let us finish this introduction by one example which stems from the article [Hörmander- 1960] where the theory about singular integrals was treated in a general context. Here is the set up: Let B_1 and B_2 be two Banach spaces and $K(x)$ is a function defined for points $x\in\mathbf{R}^n$ with values in the Banach space $L(B_1,B_2)$ of continuous linear operators from B_1 to B_2. Notice that one only assumes that K is a continuous function., i.e. it need not be linear. If $f(x)$ is a continuous B_1-valued function defined in $\mathbf{R}^n$ with compact support there exists the convolution integral

$$\mathcal{K}f(x)=\int_{\mathbf{R}^n}K(x-y)(f(y))\,dy$$

To be precise, with x fixed in $\mathbf{R}^n$ the right hand side is a welldefined B_2-valued integral which arises because $K(x-y)$ as a linear operator sends the B_1-vector $f(y)$ into B_2, i.e. the integrand in the right hand side is a B_2-valued function

$$y\mapsto K(x-y)(f(y))$$

which can be integrated with respoect to y and the resulting integral yields the B_2-vector in the left hand side where $x\mapsto\mathcal{K}f(x)$ becomes a B_2-valued function. Following [ibid] we impose two conditions on K where norms on the Banach spaces $L(B_1,B_2),B_1,B_2$ are used. We say that K satisfies the Hörmander condition if there exist positive constants A and C such that the following hold for every real number $t>0$:

$$(1)\qquad\int_{|x|\geq 1}||K(t(x-y))-K(tx)||_2\,dx\leq C\cdot t^{-n}\quad\text{for all}\quad|y|\leq 1$$

In addition to (1) we impose the L^2-inequality:

$$(2)\qquad\int_{\mathbf{R}^n}||\mathcal{K}f(x)||_2^2\,dx\leq C^2\cdot\int_{\mathbf{R}^n}||f(x)||_1^2\,dx$$

where C is independent of f. Above we use norms on B_1 respectively B_2 during the integration. Hörmander extended Vitali's Covering Lemma to normed spaces and used this to establish the following weak-type inequality:

Theorem. *There exists an absolute constant C_n which depends on n only such that for every pair of Banach spaces B_1, B_2 and every linear operator K which satisfies (1-2), the following hold for every $\alpha > 0$:*

$$\text{vol}_n(\{x \in \mathbf{R}^n : ||\mathcal{K}f(x)||_2 > \alpha\}) \leq \frac{C_n \cdot C}{\alpha} \cdot \int_{\mathbf{R}^n} ||f(x)||_1 \, dx$$

Remark. The merit of Hörmander's result is the generality. It has a wide range of applications when one combines it with interpolation results due to Markinkiewicz and Thorin. Many results which involve L^p-inequalities during the passage to Fourier transforms or other singular kernel functions can be put in a more general frame where one employs vector-valued functions rather than scalar-valued functions. Typical examples occur when K sends real-valued functions to vectors in a Hilbert space, i.e. here B_1 is the 1-dimensional real line and $B_2 = \ell^2$. So Hörmander's result constitutes a veritable propagande for learning general notions in functional analysis. Let us now summarize the contents in this chapter.

A.0 Normed spaces. § 1 studies normed vector spaces over the complex field $\mathbf{C}$ or the real field $\mathbf{R}$. We explain how each norm is defined by a convex subset of V. If X is a normed vector space such that every Cauchy sequence with respect to the norm $|| \cdot ||$ converges to some vector in X one refers to $(X, || \cdot ||)$ as a Banach space. The *Banach-Steinhaus theorem* asserts that if X is a Banach space equipped with the complete norm $|| \cdot ||^*$, then this norm is stronger than any other norm $|| \cdot ||$ on X, i.e. there exists a constant C such that

$$(*) \qquad ||x|| \leq C \cdot ||x||^* \quad : \quad x \in X$$

Thus, up to equivalence, a vector space can only be equipped with one complete norm. The proof of (*) is an immediate consequence of *Baire's category theorem.* But in spite of the trivial proof the result has a wide range of applications. For example, the Banach-Steinhaus theorem gives the *Open Mapping Theorem* and the *Closed Graph Theorem* for linear operators from one Banach space into another.

A.1 Dual spaces. When X is a normed linear space one constructs the linear space X^* whose elements are continuous linear functionals on X. The *Hahn-Banach Theorem* identifies norms of vectors in X via evaluations by X^*-elements. More precisely, denote by S^* the unit sphere in X^*, i.e. linear functionals x^* on X of unit norm. Then one has the equality

$$\text{(i)} \qquad ||x|| = \max_{x^* \in S^*} |x^*(x)| \colon \text{ for all } x \in X\,.$$

Reflexive spaces. Starting from a Banach space X we get X^* whose dual $(X^*)^*$ is denoted by X^{**} and called the bi-dual of X. There is a natural injective map $j \colon X \to X^{**}$ and (i) above shows that j is an isometry, i.e. the norms $||x||$ and $||j(x)||$ are equal. But in general j is nor surjective. If j is surjective so that $X = X^{**}$ one says that X is reflexive. An example of a non-reflexive space is the disc algebra $A(D)$. The Borthers Riesz' Theorem from Special Topics § XX entails that

$$A(D)^* \simeq \frac{L^1(T)}{H_0^1(T)}$$

where we divided out the Hardy space whose functions are zero at the origin. From this one finds that the bi-dual

$$A(D)^{**} = H^\infty(T)$$

which is the Banach space of bounded Lebesgue measurable functions on T which extend to analytic functions in D.

A.3 Calculus on Banach spaces. Results about differentiable maps from $\mathbf{R}^n$ to $\mathbf{R}^m$ extend verbatim to maps from one Banach space X into another Banach space Y. In Section 7 we define the differential of a C^1-map $g\colon\ X \to Y$ which in general need not be linear. By definition g has a differential at a point x if the limit

$$(1) \qquad y(\xi) = \lim_{t\to 0} \frac{g(x+t\xi)-g(x)}{t}$$

exists for every vector $\xi \in X$ and the limit vector y is expressed via a linear operator $D_g(x)$ from X into Y, i.e.

$$D_g(x_0)(\xi) = y(\xi)$$

hold for every $\xi \in X$. Here $D_g(x)$ is a vector in the space $L(X,Y)$ of linear operators fro x to Y. If the map

$$x \mapsto D_g(x)$$

from X into $L(X,Y)$ is continuous one says that g is of class C^1. The reader may notice that this extends the usual notion of C^1-maps from one euclidian space into another where the differential is expressed by a Jacobian matrix. More generally one constructs higher order differentials, and in this way one can refer to C^∞-maps from one Banach space into another. We review this at the end of Section 7.

A.3.1 Exercise. Use Baire's category theorem together with the Hahn-Banach theorem to show that if K is any compact metric space and ϕ is a continuous function with values in a normed space X, then ϕ is *uniformly continuous*, i.e. to every $\epsilon > 0$ there exists $\delta > 0$ such that

$$d_K(p,q) \leq \delta \implies ||\phi(p) - \phi(q)|| \leq \epsilon$$

where d_K is the distance function on the metric space K and the right hand side the norm in X.

A.3.2 Differentiable Banach spaces. A complex Banach space X is differentiable at a point x if there exists a linear map $\mathcal{D}_x$ from X into $\mathbf{R}$ such that

$$(*) \qquad ||x+\zeta\cdot y|| - ||x|| = \mathfrak{Re}\,(\zeta\cdot\mathcal{D}_x(y)) + \text{small ordo}(|\zeta|)$$

hold for every $y \in X$ and the limit is taken over complex numbers ζ which tend to zero. One says that X is differentiable if $\mathcal{D}_x$ exist for every $x \in X$.

Remark. If $\mathcal{D}_x$ exists we can take $y = x$ while ζ are small positive real numbers ϵ. This entails that

$$(\text{A.3.2}) \qquad \lim_{\epsilon\to 0} \frac{\epsilon\cdot||x|| - \epsilon\cdot\mathfrak{Re}\,\mathcal{D}_x(x)}{\epsilon} = 0 \implies \mathcal{D}_x(x) = ||x||$$

The reader may also verify the following:

Exercise. Show that $\mathcal{D}_x$ is unique if it exists and that

$$|\mathcal{D}_x(y)| \leq ||y||$$

hold for every vector y in X.

A.3.3 Uniform convexity. A normed space X is uniformly convex if there to each $\epsilon > 0$ corresponds some $\delta(\epsilon) > 0$ where

$$\lim_{\epsilon\to 0} \delta(\epsilon) = 0$$

and the following implication hold for each pair of vectors x, y of norm one at most:

$$||x+y|| \geq 2(1-\epsilon) \implies ||x-y|| \geq \delta(\epsilon)$$

If X is uniformly convex Banach space one proves easily that every closed and convex subset possesses a unique element of minimal norm.

A.3.4 Duality maps. Let X be a complex Banach space and denote by S the unit sphere of vectors $x \in X$ with $||x|| = 1$. Similarly S^* is the unit sphere in X^*. A pair $x \in S$ and $u \in S^*$ are conjugate if $u(x) = 1$.

A.3.5 Theorem. *Let X be differentiable and uniformly convex Banach space. Then each $x \in S$ has a unique conjugate in S^* given by $\mathcal{D}_x$ and the map from S to S^* defined by $x \mapsto \mathcal{D}_x$ is bijective.*

The proof is given in section 5.xx. where we also study duality maps which arise as follows:

A.3.6 Definition. *Let X be as in Theorem A.3.5. A map $T\colon X \to X^*$ is called a duality map if it is bijective and sends each sphere $S_r = \{||x|| = r\}$ onto a sphere $S^*_{\rho(r)}$ in X^* where the function $r \mapsto \rho(r)$ is strictly increasing. Finally, T maps each ray in X onto the conjugate ray in X^*.*

Remark. The last condition means that if $0 \neq x \in X$ then

$$T(s \cdot x) = \phi(s) \cdot x^* \quad \text{when } s \text{ are real and positive}$$

In § 5.XX we prove the following result which applies to many optimization problems.

A.3.7 Theorem. *Let T be some duality map from X onto X^*. Let Y be a closed subspace of X and $Y^\perp$ its orthogonal complement in X^*. Then, for every pair $x \in X$ and $u \in X^*$ the set-theoretic intersection*

$$Y^\perp + \{u\} \cap T(Y + \{x\})$$

consists of a single point.

A.4 Analytic functions.

Let X be a Banach space and consider a power series of a complex variable z:

(i) $$f(z) = \sum_{\nu=0}^{\infty} z^\nu \cdot b_\nu \cdot \qquad b_0, b_1, \dots \text{ is a sequence in } X\,.$$

Let $R > 0$ and suppose there exists a constant C such that

(ii) $$||b_\nu|| \leq C \cdot R^\nu \quad : \nu = 0, 1, \dots$$

The series (i) converges in the disc $\{|z| < R\}$ and $f(z)$ is called an X-valued analytic function More generally, let Ω be an open set in $\mathbf{C}$. An X-valued function $f(z)$ is analytic if there to every $z_0 \in \Omega$ exists an open disc D centered at z_0 such that the restriction of f to D is represented by a convergent power series

$$f(z) = \sum (z - z_0)^\nu \cdot b_\nu$$

Using the dual space X^* one extends results about ordinary analytic functions to X-valued analytic functions. Namely, for each fixed $x^* \in X^*$ the complex valued function

$$z \mapsto x^*(f(z))$$

is analytic in Ω. From this one recovers the Cauchy formula. For example, let Ω be a domain in the class $\mathcal{D}(C^1)$ and let $f(z)$ be an analytic X-valued function in Ω which extends to a continuous X-valued function on $\bar{\Omega}$. If $z_0 \in \Omega$ there exists the complex line integral

$$\int_{\partial\Omega} \frac{f(z)dz}{z - z_0}$$

It is evaluated by sums just as for a Riemann integral of complex-valued functions. One simply replaces absolute values of complex valued functions by the norm on X in approximating sums which converge to the Riemann integral. From this we obtain Cauchy's formula

$$f(z_0) = \int_{\partial\Omega} \frac{f(z)dz}{z - z_0}\,.$$

A.5 Borel-Stieltjes integrals. Let μ be a Riesz measure on the unit interval $[0, 1]$ and f an X-valued function, which to every $0 \leq t \leq 1$ assigns a vector $f(t)$ in X. Suppose that the X-norm $||f(t)|| \leq M$ hold for some constant M and every t. We say that f is Borel measurable if the complex-valued functions $t \mapsto x^*(f(t))$ are Borel functions on $[0, 1]$ for every $x^* \in X^*$. Then there exists the integral

$$J(x^*) = \int_0^1 x^*(f(t))\dot{d}t$$

for every x^*. The boundedness of f implies that $x^* \mapsto J(x^*)$ is a continuous linear functional on X^* which means that there exists a vector $\xi(f)$ in the bi-dual X^{**} such that

$$(1) \qquad \xi(f)(x^*) = J(x^*) \quad : \quad x^* \in X^*$$

When X is reflexive the f-integral yields a vector in $\mu_f \in X$ which computes (1), i.e.

$$x^*(\mu_f) = \int_0^1 x^*(f(t))\dot{d}t \quad \text{hold for all} \quad x^* \in X$$

Keeping μ fixed this means that $f \mapsto \mu_f$ is a bounded linear operator from the Borel algebra of functions on $[0,1]$ to X. This applies in particular if X is a Hilbert space since they are reflexive.

A.6 Operational calculus. Commutative Banach algebras are studied in Section XX. If B is a semi-simple Banach algebra with a unit element e and $x \in B$, then the spectrum $\sigma(x)$ is a compact subset of $\mathbf{C}$ and in the open complement there exists B-valued resolvent function

$$\text{(i)} \qquad \lambda \mapsto R_x(\lambda) = (\lambda \cdot e - x)^{-1} \quad : \quad \lambda \in \mathbf{C} \setminus \sigma(x)$$

If $\lambda_0 \in \mathbf{C} \setminus \sigma(x)$ a *local Neumann series* represents $R_x(\lambda)$ when λ stays in the open disc of radius $\text{dist}(\lambda_0, \sigma(x))$. It follows that $R_x(\lambda)$ is a B-valued analytic function of the complex variable λ defined in the open complement of $\sigma(x)$. Starting from this, Cauchy's formula is used to construct elements in B for every analytic function $f(\lambda)$ which is defined in some open neighborhood of $\sigma(x)$. More precisely, denote by $\mathcal{O}(\sigma(x))$ the algebra of germs of analytic functions on the compact set $\sigma(x)$. Then there exists an *algebra homomorphism* from $\mathcal{O}(\sigma(x))$ into X which sends $f \in \mathcal{O}(\sigma(x))$ into an element $f(x) \in X$. Moreover, the *Gelfand transform* of $f(x)$ is related to that of x by the formula

$$(*) \qquad \widehat{f}(x)(\xi) = f(\widehat{x}(\xi)) \quad : \quad \xi \in \mathfrak{M}_B$$

A.7 Hilbert spaces. An non-degenerate inner product on a complex vector space $\mathcal{H}$ is a complex valued function on the product set $\mathcal{H} \times \mathcal{H}$ which sends each pair (x, y) into a complex number denoted by $\langle x, y \rangle$ satisfying the following three conditions:

$$(1) \qquad x \mapsto \langle x, y \rangle \text{ is a linear form on } \mathcal{H} \text{ for each fixed } y \in \mathcal{H}$$

$$(2) \qquad \langle y, x \rangle = \text{ the complex conjugate of } \langle x, y \rangle \text{ for all pairs } x, y \in \mathcal{H}$$

$$(3) \qquad \langle x, x \rangle > 0 \text{ for all } x \neq 0$$

Here (1-3) imply that $\mathcal{H}$ is equipped with a norm defined by $||x|| = \sqrt{\langle x, x \rangle}$. If this norm is complete we say that $\mathcal{H}$ is a Hilbert space. A fundamental fact is that a Hilbert space is *self-dual.* This means that if γ is an element in the dual $\mathcal{H}^*$, then there exists a unique vector $y \in \mathcal{H}$ such that

$$\gamma(x) = \langle x, y \rangle \quad \text{for all } x \in \mathcal{H}\,.$$

We prove this in the section devoted to Hilbert spaces.

B. Hilbert's spectral theorem for bounded self-adjoint operators.

The theory about integral equations created by Fredholm led to Hilbert's result from 1904 which we begin to describe. Let $\mathcal{H}$ be a Hilbert space and denote by $L(\mathcal{H})$ the set of all bounded linear operators on $\mathcal{H}$. Every $T \in L(\mathcal{H})$ has its operator norm

$$||T|| = \max_x ||T(x)|| \quad \text{maximum over vectors of norm } \leq 1$$

Next, let A be a bounded self-adjoint operator on a Hilbert space whose compact spectrum is denoted by $\sigma(A)$. The Operational Calculus in Section 7 will show that there exists an algebra isomorphism from the sup-norm algebra $C^0(\sigma(A))$ into a closed subalgebra $\mathcal{A}$ of $L(\mathcal{H})$, i.e. to every continuous function g on the compact spectrum $\sigma(A)$ one gets a bounded linear operator G and $g \mapsto G$ is an algebra isomorphism. Moreover, it is an isometry which means that

$$(*) \qquad |g|_{\sigma(A)} = ||G||$$

where the left hand side is the maximum of $|g|$ over $\sigma(A)$. Next, since A is self-adjoint its spectrum $\sigma(A)$ is a compact subset of the real line where we use t as the variable. If $g(t) = c_0 + c_1 t + \ldots + c_m t^m$ is a polynomial, the operational calculus shows that $G = E + c_1 A + \ldots + c_m A^m$ where E is the identity operator on $\mathcal{H}$. By Weierstrass' theorem the set of polynomials a dense subalgebra of $C^0(\sigma(A))$ and hence $\mathcal{A}$ is the closure in $L(\mathcal{H})$ of the algebra formed by all polynomials of A.

B.1 The spectral measure. The algebra isomorphism above gives a map from the product $\mathcal{H} \times \mathcal{H}$ to the space of Riesz measures on $\sigma(A)$ which to every pair (x, y) in $\mathcal{H}$ assigns a Riesz measure $\mu_{x,y}$ such that

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

holds for every $g \in C^0(\sigma(A))$. The isometry (*) implies that the total variation of $\mu_{x,y}$ is bounded by $||x|| \cdot ||y||$ for every pair x, y. Now measure theory is applied to construct a larger subalgebra of $L(\mathcal{H})$. Namely, for every bounded Borel function $g(t)$ on $\sigma(A)$ the integrals in the sense of Borel and Stieltjes exists in the right hand side of (**) for each pair x, y in $\mathcal{H}$. In this way the g-function gives a bounded linear operator G such that

$$(***) \qquad \langle G(x), y \rangle = \int_{\sigma(A)} g(t) \cdot d\mu_{x,y}(t) \quad \text{hold for all pairs} \quad x, y$$

This yields an algebra isomorphism from the algebra $\mathcal{B}^\infty(\sigma(A))$ of bounded Borel functions to a subalgebra of $L(\mathcal{H})$ denoted by $B(\mathcal{A})$. Again the map $g \mapsto G$ is an isometry and in this extended algebra we can construct an ample family of self-adjoint operators. Namely, for every Borel subset γ of $\sigma(A)$ we can take its characteristic function and get the bounded linear operator Γ. By the Operational Calculus the spectrum of Γ is equal to the closure of γ. Moreover, Γ is a self-adjoint operator and commutes with A. In particular we can consider partitions of $\sigma(A)$. Namely, choose $M > 0$ so that $\sigma(A) \subset [-M, M]$ and M is outside $\sigma(A)$. With a large integer N we consider the half-open intervals

$$\gamma_\nu = \big[-M + \frac{\nu}{N} \cdot M, -M + \frac{\nu+1}{N} \cdot M\big] \quad : 0 \leq \nu \leq 2N-1$$

Then $\Gamma_0 + \ldots + \Gamma_{2N-1} = E$ and we also get the decomposition

$$(1) \qquad A = A_0 + \ldots + A_{2N-1} \quad : \; A_\nu = A\Gamma_\nu$$

Above $\{\Gamma_\nu\}$ gives a *resolution of the identity* where (1) means that A is a sum of self-adjoint operators where every individual operator has a spectrum confined to an interval of length $\leq \frac{1}{N}$. This resembles the finite dimensional case and constitutes Hilbert's Theorem for bounded self-adjoint operators.

C. Carleman's theorem for unbounded operators

In a note from May 1920 [Comptes rendus], Carleman indicated a new procedure to handle unbounded self-adjoint operators expressed via integral kernels which do not satisfy the Fredholm conditions. The conclusive theory was presented in the book *Sur les équations singulières à noyau reèl et symmétrique* published by Uppsala University in 1923. Following Carleman we expose the results for unbounded operators. One starts with a linear operator A on $\mathcal{H}$ which only is *densely defined.* That is, the domain of definition $\mathcal{D}(A)$ is a dense subspace of $\mathcal{H}$ while A is unbounded, i.e.

$$\sup_x ||A(x)|| = +\infty$$

with the supremum taken over the unit ball in $\mathcal{H}$. One says that A is *symmetric* if

$$(1) \qquad \langle Ax, y \rangle = \langle x, Ay \rangle \quad \text{hold for all pairs} \quad x, y \in \mathcal{D}(A)$$

The adjoint A^*. Let A be a symmetric operator. Given a vector x_* in $\mathcal{H}$ we define a linear functional on $\mathcal{D}(A)$ by

$$x \mapsto \langle Ax, x_* \rangle$$

Suppose there exists a constant $C(x_*)$ such that

(1) $$|\langle Ax, x_* \rangle| \leq C(x_*) \cdot ||x|| \quad \text{hold for all pairs} \quad x, y \in \mathcal{D}(A)$$

Since $\mathcal{D}(A)$ is dense and $\mathcal{H}$ is self-dual this gives a unique vector y_* such that

(2) $$\langle Ax, x_* \rangle = \langle x, y_* \rangle \quad : x \in \mathcal{D}(A)$$

The set of vectors x_* for which $C(x_*)$ exists is a subspace of $\mathcal{H}$ which we denote by $\mathcal{D}^*$. From (2) we get the linear operator $x_* \mapsto y_*$. It is denoted by A^* and called the adjoint operator of A. So here $\mathcal{D}(A^*) = \mathcal{D}^*$. Next follow some exercises where the reader if necessary can consult XX for a more detailed account.

Exercise A. Show that $\mathcal{D}(A) \subset \mathcal{D}(A^*)$ and that A^* extends A in the sense that $A^*(x) = A(x)$ for every $x \in \mathcal{D}(A)$.

Exercise B. Show that A^* has a closed graph, i.e. put

$$\Gamma(A^*) = \{(x, A^*x) \quad : x \in \mathcal{D}(A^*)\}$$

and verify that $\Gamma(A^*)$ is a closed subspace of $\mathcal{H} \times \mathcal{H}$.

Exercise C. On $\mathcal{D}(A^*)$ we define an inner product by

$$\{x, y\} = \langle x, y \rangle + \langle A^*x, A^*y \rangle$$

Use that $\Gamma(A^*)$ is closed to conclude that this inner product is complete and hence $\mathcal{D}(A^*)$ is a Hilbert space under this inner product.

C.1 The eigenspaces $\mathcal{D}_+$ and $\mathcal{D}_-$. Put

$$\mathcal{D}_+ = \{x \in \mathcal{D}(A^*) \quad : A^*(x) = ix\} \quad \text{and} \quad \mathcal{D}_- = \{x \in \mathcal{D}(A^*) \quad : A^*(x) = -ix\}$$

Since A^* has a closed graph it is obvious that these two subspaces of $\mathcal{H}$ are closed.

C.2 A direct sum decomposition. Recall the inclusion $\mathcal{D}(A) \subset \mathcal{D}(A^*)$. We can therefore construct the closure of $\mathcal{D}(A)$ under the norm defined by the complete inner product in Exercise C. Let $cl(\mathcal{D}(A))$ denote this closure. The following will be proved in Section 9:

C.3 Proposition. *One has the following orthogonal decomposition in the Hilbert space $\mathcal{D}(A^*)$:*

$$\mathcal{D}(A^*) = cl(\mathcal{D}(A)) \oplus \mathcal{D}_+ \oplus \mathcal{D}_-$$

C.4 The self-adjoint case. Following [Carleman] the symmetric operator A gives *Case I* if $\mathcal{D}_+$ and $\mathcal{D}_-$ both are zero spaces. Then Proposition C.3 gives the equality $\mathcal{D}(A^*) = cl(\mathcal{D}(A))$ and since A is symmetric it follows that A^* also is symmetric, i.e.

(*) $$\langle A^*x, y \rangle = \langle x, A^*y \rangle$$

holds for all pairs (x, y) in $\mathcal{D}(A^*)$. Starting from the symmetric and densely defined operator A^* we can construct its adjoint. But this time the process stops, i.e. one finds that $(A^*)^* = A^{**}$.

C.5 The bounded resolvent operator. Let A be a densely defined and self-adjoint operator. Thus, it is symmetric and one has the equality $\mathcal{D}(A) = \mathcal{D}(A^*)$. The extension of Hilbert's theorem for bounded self-adjoint operators relies upon the existence of a bounded resolvent.

C.6 Theorem. *There exists a bounded and normal operator S such that the range $S(\mathcal{H}) = \mathcal{D}(A)$ and*

(*) $$(i \cdot E + A)(S(x)) = x$$

hold for all $x \in \mathcal{H}$. Moreover, the spectrum $\sigma(S)$ is contained in the set

$$\Sigma = \Big\{\frac{1}{a+i} \quad : a \in \mathbf{R} \cup \{0\}\Big\}$$

Remark. We refer to Section 9 for the proof. Next, a bounded linear operator R on $\mathcal{H}$ is normal if it commutes with its adjoint R^*. Hilbert's theorem extends verbatim to the normal operator S above. Namely, if $\sigma(S)$ is the compact spectrum there exists an isometric algebra isomorphism from $C^0(\sigma(S))$ onto the closed subalgebra $\mathcal{S}$ of $L(\mathcal{H})$ generated by S and its adjoint S^*. Moreover, exactly as in the self-adjoint case we use this to construct a map from $\mathcal{H} \times \mathcal{H}$ into Riesz measures on $\sigma(S)$. The isometric algebra isomorphism extends to a map from $\mathcal{B}^\infty(\sigma(S))$ onto a closed subalgebra $B(\mathcal{S})$ of $L(\mathcal{H})$ where each operator in $B(\mathcal{S})$ is normal and commutes with S.

Theorem C.6 implies that the set Σ is a simple closed curve which contains $-i$ and the origin in the complex λ-plane and we can apply the operational calculus to the normal operator S. So for every positive integer N we get the bounded self-adjoint operator Γ_N on $\mathcal{H}$ obtained via the characteristic function of the set

(1) $$\gamma_N = \{\lambda \in \sigma(S) \quad : \mathfrak{Im}(\lambda) \geq \frac{1}{N}\}$$

Since γ_N does not contain $\lambda = 0$ there exists the bounded normal operator :

$$S_N = \int_{\gamma_N} \frac{1 - i\lambda}{\lambda} \cdot d\mathcal{S}$$

The equality (*) in Theorem C.6 gives

(*) $$A\Gamma_N = S_N$$

Moreover, from the equation which defines the set Σ it follows that the spectrum of S_N is *real*. Since every normal operator with a real spectrum is self-adjoint we conclude that S_N is so. Finally, the construction of the γ_N-sets implies that the sequence $\{\Gamma_N\}$ converges to the identity operator. More precisely, the kernels of these bounded self-adjoint operators decrease and the intersection

$$\bigcap_{N \geq 1} \mathrm{Ker}(\Gamma_N) = \{0\}$$

C.7 Conclusion. *The results imply that the sequence $\{S_N\}$ converges to A in the sense that*

$$A(x) = \lim_{N \to \infty} S_N(x) \quad \text{holds for all } x \in \mathcal{D}(A)$$

Moreover, $\mathcal{D}(A)$ is equal to the set of $x \in \mathcal{H}$ for which the limit of $\{S_N(x)\}$ exists.

This is Carleman's theorem for unbounded self-adjoint operators. *Case II* arises when we start from a symmetric operator A where at least one of the eigenspaces $\mathcal{D}_+$ and $\mathcal{D}_-$ of A^* is non-zero was also considered in [Carleman] where it is proved that if they are finite dimensional and have the same dimension, then one can still construct a self-adjoint operator A_0^* attached to the given symmetric operator A and after apply the spectral theorem to A_0^* to investigate A.

D. Application to a dynamical system.

Using the spectral theorem a rigorous proof of the Ergodic Hypothesis in Statistical Mechanics was given by Carleman at seminars held at Institute Mittag-Leffler in May 1931. Here is the situation: There is given an n-tuple of C^1-functions $A_1(x), \ldots, A_n(x)$ where $x = (x_1, \ldots, x_n)$ are points in $\mathbf{R}^n$. Let t be a time variable and consider the differential system

(1) $$\frac{dx_k}{dt} = A_k(x_1(t), \ldots, x_n(t)) \quad : \quad 1 \leq k \leq n$$

Assume that there exists a compact hypersurface S in $\mathbf{R}^n$ such that if $p \in S$ and $\mathbf{x}_p(t)$ is the vector-valued solution to (1) with initial condition $\mathbf{x}_p(0) = p$, then $\mathbf{x}_p(t)$ stay in S for every t. The uniqueness for solutions to the differential systems gives for every t a bijective map $p \mapsto \mathbf{x}_p(t)$ from S onto itself. It is denoted by $\mathcal{T}_t$ and we notice that

$$\mathcal{T}_s \circ \mathcal{T}_t = \mathcal{T}_{s+t}$$

In addition we assume that there exists an invariant measure σ on S for the $\mathcal{T}$-maps given by a non-negative measure σ such that

$$\sigma(\mathcal{T}_t(A)) = \sigma(A)$$

hold for every σ-measurable set. For applications to classical differential systems it suffices to consider the case when σ is a density expressed by a positive continuous function times the area measure on S. We have the Hilbert space $L^2(\sigma)$ of complex-valued measurable functions U on S for which

$$\int_S |U(p)|^2 \cdot d\sigma(p) < \infty$$

On $L^2(\sigma)$ there exists the densely defined symmetric operator:

$$(*) \qquad U \mapsto i \cdot \sum_{\nu=1}^{\nu=n} A_\nu \cdot \frac{\partial U}{\partial x_\nu}$$

It is easy to verify that *Case 1* holds for this operator and hence the spectral theorem applies. To each a pair of L^2-functions U and V one considers the following mean-value integrals over time intervals $[0, T]$:

$$(*) \qquad J_T(U, V) = \frac{1}{T} \cdot \int_0^T \big[\int_S \langle\, U(\mathcal{T}_t(p)) \cdot V(p) \cdot d\sigma(p) \rangle \,\big] \cdot dt$$

Next, let $\{\omega_\nu\}$ be an orthonormal basis in $\mathcal{H}$ and each L^2-function U has an expansion

$$U = \sum \langle \omega_\nu, U \rangle \cdot \omega_\nu \quad : \quad \langle \omega_\nu, U \rangle = \int_S \omega_\nu(p) \cdot U(p) \cdot d\sigma(p)$$

Theorem. *Let $\{\omega_\nu\}$ be an orthonormal basis in $\mathcal{H}$. For each pair U, V in $L^2(\sigma)$ one has the equality*

$$\lim_{T \to \infty} J_T(U, V) = \sum_{\nu=1}^{\infty} \langle \omega_\nu, U \rangle \cdot \langle \omega_\nu, V \rangle$$

Remark. Let $\mathcal{H}_*$ be the space of $L^2(\sigma)$-functions which are $\mathcal{T}$-invariant, i.e. L^2-functions ω satisfying

$$(2) \qquad \mathcal{T}_t(\omega) = \omega \quad \text{for all} \quad t$$

Here $\mathcal{H}_*$ is a closed subspace of $L^2(\sigma)$. In the case when $\mathcal{H}_*$ is reduced to the one-dimensional space of constant functions, Theorem D.1 implies that almost every trajectory which comes from the differential system comes close to every point in S which confirms the original assertions about returining points by Liouviulle and Poincaré. Let us remark that in Eergodic Theory one refers to Theorem D.1 as a mean-value result. A more precise result about almost everywhere convergence was settled by Birkhoff in the article [Birkhoff xx] which superseeds Theorem D.1 above. See also § xx in [Functional Analysis] for further material related to Ergodic Theorey.

E. Schrödinger's equation.

In 1923 quantum mechanics had not yet appeared so the studies in [Car] were concerned with singular integral equations, foremost inspired from previous work by Fredholm and Volterra. The creation of quantum mechanics led to new applications of the spectral theorem. The interested reader can consult the lecture held by Niels Bohr at the Scandianavian congress in mathematics held in Copenhagen 1925 where he speaks about the interplay between the new physics and pure mathematics. Bohr's lecture presumably inspired Carleman when he some years later resumed his work in [Car 1923]. Recall that the fundamental point in Schrödinger's theory is the hypothesis on energy levels which correspond to the possible orbits in Bohr's theory of atoms which are described by Bohr in his plenary talk when he received the Nobel Prize in physics 1923. For a single particle where one actually seeks solution to a wave equation, the first mathematical problem is to study the PDE-equation

$$(*) \qquad \Delta\phi + 2m \cdot \big(E - U\big)\big(\frac{2\pi}{h}\big)^2 \cdot \phi = 0$$

Here Δ is the Laplace operator in the 3-dimensional (x, y, z)-space, m the mass of a particle and h Planck's constant while $U(x, y, z)$ is a potential function. Finally E is a parameter and one seeks values on E such that (*) has a solution ϕ which belongs to $L^2(\mathbf{R}^3)$. If the spectrum, i.e. those e-values which produce non-zero ϕ-solutions have been determined one can after solve the corresponding wave equatiion where a time paramater appears, i.e. just in the case of a classical heat equation as explained in § XX. Concerning the mathematical interest in (*) we cite an excerpt from Carlemans lectures in Paris at Institut Henri Poincaré held in 1931:

Dans ces dernières années l'ínteret de la question qui nous occupe a considérablement augmenté. Cést en effet in instrument mathématique indispensable pour development de la mechanique moderne crée par M.M. de Brogile, Heisenberg et Schrödinger. Etude de l'équation integrale:

$$\phi(x) = \lambda \cdot \int_a^b K(x, y)\phi(y)dy + f(x) \quad : \lambda \in \mathbf{C} \setminus \mathbf{R}$$

The basic mathematical issue is to consider a second order differential operator

$$(*) \qquad L = \Delta + c(x, y, z) \quad : \Delta = \partial_x^2 + \partial_y^2 + \partial_z^2$$

where $c(x, y, z)$ is a real-valued function. Here L is densely defined on $L^2(\mathbf{R}^3)$ and since c is real-valued also symmetric, i.e. it suffices to notice that

$$\iiint L(\phi) \cdot \psi \, dxdydz = \iiint L(\psi) \cdot \phi \, dxdydz =$$

hold for a pair of C^2-functions which both have compact support in $\mathbf{R}^3$. A first problem is to find conditions on the c-function in order that the favourable Case 1 occurs. The following sufficiency result was established in [ibid]:

E.1 Theorem *Let $c(x, y, z)$ be a continuous and real-valued function such that there is a constant M for which*

$$\limsup_{x^2+y^2+z^2 \to \infty} c(x, y, z) \leq M$$

Then the favourable Case 1 holds for the operator $\Delta + c(x, y, z)$.

Remark. When the c-function satisfies (*) in Theorem E.1 one can proceed and obtain exhibit eigenfunctions via a limit process where solutions to Neumann's classical boundary value problem are determined on an increasing sequence of balls in $\mathbf{R}^3$. The interested reader can consult Carleman's article [car][for details where we remark that the analysis uses the wellknown potentia intergal of Laplace operator to rewrite the PDE-equation into an integral equation and then the theory from [Car: 1923] applies.

E.2 A special case. Here one considers a potential function:

$$W(p) = \sum \frac{\alpha_k}{|p - q_k|} + \beta$$

where $\{q_k\}$ is a finite subset of $\mathbf{R}^3$ and the α-numbers and β are real and positive. With $c = W$ we get the favourable case and hence this central case for Schrödinger equations is covered by Theorem D.1 above. We refer to [Carleman] for a detailed proof of Theorem D.1 which also describes how to attain solutions via a limit process where Neumann's boundary value problem is considered on an increasing sequence of balls in $\mathbf{R}^3$.

E.3 Remark. The literature about the Schrödinger equation and other equations which emerge from quantum mechanics is very extensive. Numerical solutions to the special equation considered above can be obtained of computers. But already the determination of some initial spectral values when W is a Newtonian potential and the number of mass-points is some finite number ≥ 3 is quite involved. For sources of quantum mechanics the reader should first of all consult the plenary talks by Heisenberg, Dirac and Shrödinger when they received the Nobel prize in physics. Apart from physical considerations the reader will find expositions where explanations are given in a mathematical framework. Actually Heisenberg was sole winner 1931 while Dirac and Shrödinger shared the prize in 1932. But they visited Stockholm together in December 1932.

For mathematician who wants to become acquainted with aspects of quantum physics the eminent text-books by Lev Landau are recommended. Volume 3 is entitled *Quantum mechanics : Non-relativistic theory*. In the relativist case which is treated in later volums of [L-L] one must employ Heisenberg's matrix representation and Dirac equations are used to study radiation phenomena in quantum physics. In the introduction to [ibid: Volume 3] Landau inserts the following remark: *It is of interest to note that the complete mathematical formalisim of quantum mechanics was constructed by W. Heisenberg and E. Schrödinger in 1925-26, before the discovery of the uncertainty principle which revealed the physical contents of this formalism.*

Staying in the non-relativistic situation one studies wave equations of Schrödinger's type whose mathematical foundations were laid in Schrödinger's article *Quantizierung als Eigenwertproblem* from 1926 In § XX from special topic we describe another of Schrödinger's equations which led to a veritable challenge in the "world of mathematics".

1. Normed spaces.

A normed space over the complex field is a complex vector space X equipped with a norm $||\cdot||$ expressed by a map from X into $\mathbf{R}^+$ satisfying:

(*) $$||x+y|| \leq ||x|| + ||y|| \quad \text{and} \quad ||\lambda \cdot x|| = |\lambda \cdot ||x|| \quad : \quad x,y \in X \quad : \ \lambda \in \mathbf{C}$$

Moreover $||x|| > 0$ holds for every $x \neq 0$. A norm gives a topology on X defined by the distance function

(**) $$d(x,y) = ||x-y||$$

1.1 Real versus complex norms. The real numbers appear as a subfield of $\mathbf{C}$. Hence every complex vector space has an underlying structure as a vector space over $\mathbf{R}$. A norm on a real vector space Y is a function $y \mapsto ||y||$ where (*) holds for real numbers λ. Next, let X be a complex vector space with a norm $||\cdot||$ satisfying (*) above. Since we can take $\lambda \in \mathbf{R}$ in (*) the complex norm induces a real norm on the underlying real vector space of X. Complex norms are more special than real norms. For example, consider the 1-dimensional complex vector space given by $\mathbf{C}$. When the point 1 has norm one there is no choice for the norm of any complex vector $z = a + ib$, i.e. its norm becomes the usual absolute value. On the other hand we can define many norms on the underlying real (x,y)-space. For example, we may take the norm defined by

(i) $$||(x,y)|| = |x| + |y|$$

It fails to satisfy (*) under complex multiplication. For example, with $\lambda = e^{\pi i/4}$ we send $(1,0)$ to $p = (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$ whose norm from (i) becomes $\sqrt{2}$ while it should remain with norm one if (*) holds.

1.2 Convex sets. We shall work on real vector spaces for a while. Let Y be a real vector space. A subset K is convex if the line segment formed by a pair of points in K stay in K, i.e.

(i) $$y_1, y_2 \in K \implies s \cdot y_2 + (1-s) \cdot y_1 \in K \quad : \quad 0 \leq s \leq 1$$

Let $\mathbf{o}$ denote the origin in Y. Let K be a convex set which contains $\mathbf{o}$ and is symmetric with respect to $\mathbf{o}$:

$$y \in K \implies -y \in K$$

The symmetric convex set K is called *absorbing* if there to every $y \in Y$ exists some $t > 0$ such that $ty \in K$. Suppose that K is symmetric and absorbing. To every $s > 0$ we set

$$sK = \{sx : \ x \in K\}$$

Since $\mathbf{o} \in K$ and K is convex these sets increase with s and since K is absorbing we have:

(ii) $$\bigcup_{s>0} sK = Y$$

Next, we impose the condition that K does not contain any 1-dimensional subspace, i.e. whenever $y \neq 0$ is a non-zero vector there exists some large t^* such that $t^* \cdot y$ does not belong to K. The condition is equivalent with

(iii) $$\bigcap_{s>0} s \cdot K = \mathbf{o}$$

1.3 The norm ρ_K. Let K be convex and symmetric and assume that (ii-iii) hold. To each $y \neq 0$ we set

(*) $$\rho_K(y) = \min_{t>0} \frac{1}{t} \quad : \ t \cdot y \in K$$

Notice that if $y \in K$ then $t = 1$ is competing when we seek the minimum and hence $\rho_K(y) \leq 1$. On the other hand, if y is "far away" from K we need small t-values to get $t \cdot y \in K$ and therefore $\rho_K(y)$ is large. It is also clear that

(i) $$\rho_K(ay) = a \cdot \rho_K(y) \quad : \ a \text{ real and positive}$$

Finally , since K is symmetric we have $\rho_K(y) = \rho_K(-y)$ and hence (i) gives

(ii) $$\rho_K(ay) = |a| \cdot \rho_K(y) \quad : \ a \text{ any real number}$$

1.4 Proposition. *By (*) we get a norm which is called the K-norm defined by the convex set K.*

Proof. The verification of the triangle inequality:

$$\rho_K(y_1 + y_2) \leq \rho_K(y_1)\rho_K(y_2)$$

is left as an exercise. The hint is to use the convexity of K.

1.5 A converse. Let $|| \cdot ||$ be a norm on Y. Then we get a convex set

$$K^* = \{y \in Y \quad : ||y| \leq 1\}$$

It is clear that $\rho_{K^*}(y) = ||y||$ holds, i.e. the given norm is recaptured by the norm defined by K^*. We can also regard the set

$$K_* = y \in Y \quad : ||y| < 1\}$$

Here $K_* \subset K^*$ but he reader should notice that one has the equality

$$\rho_{K_*}(y) = \rho_{K^*}(y)$$

Thus, the two convex sets define the same norm even if the set-theoretic inclusion $K_* \subset K^*$ may be strict. In general, a pair of convex sets K_1, K_2 satisfying (i-ii) above are equivalent if they define the same norm. Starting from this norm we get K_* and K^* and then the reader may verify that

$$K_* \subset K_\nu \subset K^* \quad : \nu = 1, 2$$

Summing up we have described all norms on Y and they are in a 1-1 correspondence with equivalence classes in the family $\mathcal{K}$ which consists of all convex sets which are symmetric, absorbing and satisfy (iii) above, i.e. when $K \in \mathcal{K}$ then K does not contain any 1-dimensional subspace. For each specific norm on Y we can assign the largest convex set K^* from the corresponding equivalence class.

1.6 Equivalent norms. Two norms $|| \cdot ||_1$ and $|| \cdot ||_2$ are equivalent if there exists a constant $C \geq 1$ such that

$$\frac{1}{C} \cdot ||y||_1 \leq ||y||_2 \leq C \cdot ||y||_1 \quad : \quad y \in Y \tag{0.6}$$

Notice that if the norms are defined by convex sets K_1 and K_2 respectively, then (0.6) means that there exists some $0 < t < 1$ such that

$$tK_1 \subset K_2 \subset t^{-1}K_1$$

The case $Y = \mathbf{R}^n$. If Y is finite dimensional all norms are equivalent. To see this we consider the euclidian basis $e_1, \ldots, e_n$. To begin with we get the *euclidian norm* which by definition measures the euclidian length from a vector y to the origin. It means that

$$||y||_e = \sqrt{\sum_{\nu=1}^{\nu=n} |a_\nu|^2} \quad : \quad y = a_1 e_1 + \ldots + a_n e_n \tag{i}$$

The reader should verify that the norm satisfies the triangle inequality

$$||y_1 + y_2||_e \leq ||y_1||_e + ||y_2||_e$$

which amounts to verify the Cauchy-Schwartz inequality. In the euclidian norm the unit sphere S^{n-1} corresponds to vectors whose euclidian norm os one. We also define the norm $|| \cdot ||^*$ by

$$||y||^* = \sum_{\nu=1}^{\nu=n} |a_\nu| \quad : \quad y = a_1 e_1 + \ldots + a_n e_n \tag{ii}$$

This norm is equivalent to the euclidian norm. More precisely the reader may verify the inequality

$$\frac{1}{\sqrt{n}} \cdot ||y||_e \leq ||y||^* \leq \sqrt{n} \cdot ||y||_e \tag{iii}$$

Next, let $||\cdot||$ be some arbitrary norm. Put

(iv) $$C=\max_{1\leq\nu\leq n}\,||e_\nu||$$

Then (ii) and the triangle inequality for the norm $||\cdot||$ gives

(v) $$||y||\leq C\cdot||y||^*$$

By the equivalence (iii) the norm topology defined by $||\cdot||^*$ is the same as the usual euclidian topology in $Y=\mathbf{R}^n$. Next, notice that (v) implies that the sets

$$U_N=\{y\in Y\quad:\quad||y||<\frac{1}{N}\}\quad:\quad N=1,2,\ldots$$

are *open* sets when Y is equipped with its usual euclidian topology. Now $\{U_N\}$ is an increasing sequence of open sets and their union is obviously equal t oY. in particular this union covers the compact unit sphere S^{n-1}. This gives an integer N such that

$$S^{n-1}\subset U_N$$

This inclusion gives

$$||y||_e\leq N\cdot||y||$$

Together with (iii) and (v) we conclude that $||\cdot||$is equivalent with $||\cdot||_e$. Hence we have proved

1.7 Theorem. *On a finite dimensional vector space all norms are equivalent.*

1.8 The complex case. If X is a complex vector space we obtain complex norms when we restrict the attention to convex sets K which not only are symmetric with respect to scalar multiplication with real numbers but is also invariant under i. To be precise, one requires that

$$\lambda\cdot K\subset K\quad:\quad\forall\,\lambda\in\mathbf{C}\quad:|\lambda|\leq 1$$

Here a similar result as in Theorem 1.7 holds for complex norms on $\mathbf{C}^n$, i.e. they are all equivalent.

1.9 Non-linear convexity.

Let $f(x)$ be a real-valued function in $\mathbf{R}^n$ of class C^2. To every point x we assign the hessian $H_f(x)$ which is the symmetric matrix whose elements are $\{\partial^2 f/\partial x_j\partial x_k\}$. The function is strictly convex if $H_f(x)$ is positive for all x, i.e. if the eigenvalues are all >0. Assume in addition that

(1) $$\lim_{|x|\to+\infty}\,f(x)=+\infty$$

Under these conditions one has the classic results below which are due to Lagrange and Legendre:

1.10 Theorem. *The vector valued function*

$$x\mapsto\nabla_f(x)$$

is a C^1-differomorphism of $\mathbf{R}^n$ onto itself.

Next, with f still as above one defines the function below for each $y\in\mathbf{R}^n$:

(*) $$\mathcal{L}_f(y)=\max_x\,\langle x,y\rangle-f(x)$$

1.11 Theorem. *For each y the maximum in (*) is taken at a unique pointt $x^*(y)$ and one has the equality*

(**) $$y=\nabla_f(x^*(y)$$

Moreover, $\mathcal{L}_f$ is again strictly convex and one has the biduality formula:

(***) $$f=\mathcal{L}\circ\mathcal{L}_f$$

1.12 Exercise. Prove the two theorems above. Legendre's biduality means that $\mathcal{L}$ is a bijective map on the class of strictly convex functions which satisfy (1) and the composed operator $\mathcal{L} \circ \mathcal{L}$ is the identty.

1.13 On cones in $\mathbf{R}^n$. Here follow an exercise which helps the reader to grasp some geometry in $\mathbf{R}^n$. A subset Γ is a cone if $x \in \Gamma$ implies that the half-ray $\mathbf{R}^+ \cdot x \subset \Gamma$. We suppose that the cone is closed. In particular the origin is included and notice that Γ is determined by the compact subset $\Gamma_* = \Gamma \cap S^{n-1}$ where S^{n-1} is euclidian the unit sphere. We say that Γ is *fat* if Γ_* has a non-empty interior in the unit sphere and Γ is *proper* if $\Gamma_* \cap -\Gamma_* = \emptyset$, i.e. equivalently Γ does not contain any 1-dimensional subspace. next, the *dual cone* is defined by

$$\widehat{\Gamma} = \{x \quad : \langle x, \Gamma \rangle \leq 0\}$$

1.14 Exercise. Show that a cone Γ is proper if and only if $\widehat{\Gamma}$ is fat and show also that Γ is equal to the dual of $\widehat{\Gamma}$.

1.15 Remark. Legendre used Theorem 1.11 to study extremal solutions in the calculus of variation. The constructions of quantized contact transformations were later introduced in by Hamilton and many specific examples were treated by Jacobi to solve Hamiltonian systems of non-linear systems of first order differential equations. For this more advanced material related to convexity, the reader may consult Chapter XX from volume ! on classical mechanics in the eminent text-book series Lev Landau, which is devoted to theoretical physics but also offer very instructive and rigorous material in pure mathematics of high standard and great interest.

2. Banach spaces.

Let Y be a normed space over $\mathbf{C}$ or over $\mathbf{R}$. A sequence of vectors $\{y_n\}$ is called a Cauchy sequence if

$$(*) \qquad \lim_{n,m\to\infty} ||y_n - y_m|| = 0$$

We obtain a vector space $\widehat{Y}$ whose vectors are defined as equivalence classes of Cauchy sequences. The norm of a Cauchy sequence $\hat{y} = \{y_n\}$ is defined by

$$||\hat{y}|| = \lim_{n\to\infty} ||y_n||$$

One says that the norm on Y is complete if every Cauchy sequence converges, or equivalently $Y = \widehat{Y}$. A complete normed space is called a *Banach space* as an attribution to Stefan Banach whose pioneering article [Ban] introduced the general concept of normed vector spaces.

2.1 The Banach-Steinhaus theorem. *Let X be a Banach space equipped with the complete norm $||\cdot||^*$. Then for every other norm $||\cdot||$ there exists a constant C such that*

$$||x|| \leq C \cdot ||x||^* \quad : \quad x \in X$$

Remark. In particular we see that if $||\cdot||_1$ and $||\cdot||_2$ are two complete norms on the same vector space then they are equivalent in the sense that there exists a constant C such that

$$C^{-1} \cdot ||\cdot||_2 \leq ||\cdot||_1 \leq C \cdot ||\cdot||_2$$

The proof of Theorem 2.1 relies upon a result due to Baire which we recall below.

The Baire category theorem. Let X be a metric space whose metric d is complete, i.e. every Cauchy sequence with respect to the distance function d converges.

2.2 Theorem. *Let $\{F_n\}$ is an increasing sequence of closed subsets of X where each F_n has empty interior. Then the union $F^* = \cup F_n$ is meager, i.e. F^* does not contain any open set.*

Proof. Let $x_0 \in X$ and $\epsilon > 0$ be given. It suffices to show that $B_\epsilon(x_0)$ contains a point x_* outside F^* for every $\epsilon > 0$. To show this we first use that F_1 has empty interior which gives some $x_1 \in B_{\epsilon/2}(x_0) \setminus F_1$ and we choose $\delta_1 < \epsilon/2$ so that

(i) $$B_{\delta_1}(x_1) \cap F_1 = \emptyset$$

Now $B_{\delta_1/2}(x_1)$ is not contained in F_2 and we find a pair x_2 and $\delta_2 < \delta_1/2$ such that

(ii) $$B_{\delta_2}(x_2) \cap F_2 = \emptyset$$

We can continue in this way and to every n find a pair x_n, δ_n such that

(iii) $$B_{\delta_n}(x_n) \cap F_n = \emptyset \quad : \quad x_n \in B_{\delta_{n-1}}(x_{n-1}) \quad : \quad \delta_n < \delta_{n-1}/2$$

Since X by assumption is complete and $\{x_n\}$ by the construction is a Cauchy sequence there exists a limit $x_n \to x^*$. The rapid decrease of the δ-numbers gives $x^* \in B_\epsilon(x_0)$ and the inductive construction shows that x^* does not belong to the union F^*.

2.3 Proof of the Banach-Steinhaus theorem. Let X be a Banach space equipped with the complete norm $||\cdot||^*$ and let $||\cdot||$ be some other norm. To each positive integer N we put

$$F_N = \text{The closure of the set } \{x : \ ||x|| \leq N\} \text{ with respect to } ||\cdot||^* - \text{topology}$$

Obviously $\cup F_N = X$ and Baire's category theorem gives the existence of some $N \geq 1$, a point $x_0 \in X$ and some $\delta > 0$ such that the open ball

(i) $$B_\delta(x_0) = \{x : \ ||x - x_0||^* < \delta\} \subset F_N$$

Next, notice that F_N is convex and symmetric. So if $||x|| < \delta$ we get

$$x = \frac{x_0 + x}{2} + \frac{-x_0 + x}{2} \in F_N$$

Hence we get the implication:

(ii) $$||x|| \leq \delta \implies ||x||^* \leq N$$

But this means precisely that

$$||x|| \leq \frac{N}{\delta} \cdot ||x||^*$$

This finishes the proof of the Banach-Steinhaus theorem.

2.4 Separable Banach spaces. This is the class of Banach spaces which contain a denumerable and dense subset. Let Y be a separable Banach space and $\{y_n\}$ a dense subset indexed by positive integers $n = 1, 2, \ldots$. To every n we get the finite dimensional vector space Y_n generated by $y_1, \ldots, y_n$ and by the procedure in Linear algebra we can construct a basis in Y_n and when $Y_n \subset Y_{n+1}$ get a new basis vector. In this way one arives at a denumarable sequence of linearly independent vectors $e_1, e_2, \ldots$ such that the increasing sequence of subspaces $\{Y_n\}$ are all contained in the vector space

(i) $$Y_* = \oplus \mathbf{R} \cdot \mathbf{e_n}$$

By the construction Y_* is a dense subspace of Y. Of course, there are many ways to construct a denumerable sequence of linearly independent vectors which by (i) give a dense subspace of Y.

2.5 Schauder basis. One may ask if it is possible to choose a sequence $\{e_n\}$ as above such that every $y \in Y$ can be expanded in this basis as follows:

2.6 Definition. *A denumerable sequence $\{e_n\}$ of $\mathbf{R}$-linearly independent vectors is called a Schauder basis if there to each $y \in Y$ exists a unique sequence of real numbers $c_1(y), c_2(y), \ldots$ such that*

$$\lim_{N \to \infty} ||y - \sum_{n=1}^{n=N} c_\nu(y) \cdot e_\nu|| = 0$$

2.7 Per Enflo's example. The existence of a Schauder basis in every separable Banach space appears to be natural and Schauder constructed such a basis in several cases, such as the Banach space $\mathbf{C}^0[0,1]$ of continuous functions on the closed unit interval equipped with the maximum norm. For several decades the question of existence of a Schauder basis in *every* separable Banach space was open until Per Enflo at seminars in Stockholm University during the autumn in 1972 presented an example where a Schauder basis does not exist. Actually Enflo also gave a counter-example concerning compact operators. More precisely, to every $2 < p < \infty$ he constructed a closed subspace Y of the Banach space ℓ^p on which there exists a *compact* linear operator T which cannot be approximated in the operator norm by linear operators on Y with finite dimensional range. One verifies easily that the failure of such an approximation implies that Y cannot have a Schauder basis. So Enflo constructed a very "ugly" separable Banach space. For the detailed construction we refer to his article [En-Acta Mathematatica]. Let us remark that the essential ingredient in Enflo's construction relies upon a study of Fourier series where the efficient tool is to employ *Rudin-Schapiro* polynomials which consist of trigonometric polynomials

$$(*) \qquad P_N(x) = \epsilon_0 + \epsilon_1 e^{ix} + \ldots + \epsilon_N \cdot e^{iNx}$$

where each ϵ_ν is +1 or -1. For any such sequence Plancherel's equality gives

$$\frac{1}{2\pi}\int_{=}^{2\pi} |P_n(x)|^2 \cdot dx = 2^{N+1}$$

This implies that the maximum norm of $|P(x)|$ is at least $2^{\frac{N+1}{2}}$. In [Ru-Sch] it is shown that there exists a fixed constant C such that to every $N \geq 1$ there exists at least one choice of signs of the $\epsilon_\bullet$-sequence so that

$$\max_{0 \leq x \leq 2\pi} |P_N(x)| \leq C \cdot 2^{\frac{N+1}{2}}$$

A Remark. After Enflo's work [En] it became a veritable industry to *verify* that various "concrete" Banach spaces Y do have a Schauder basis and perhaps more important, enjoy the approximation property , i.e. that the class of linear operators on Y with finite dimensional range is dense in the linear space of all compact operators on Y. Fortunately most Banach spaces do have a Schauder basis. But the construction of a specific Schauder basis is often non-trivial. It requires for example considerable work to exhibit a Schauder basis in the Banach space $A(D)$ of continuous functions on the closed unit disc which are analytic in the interior.

3. Linear operators.

Let X and Y be two normed spaces and $T\colon\ X \to Y$ a linear operator. We say that T is continuous if there exists a constant C such that

$$||T(x)|| \leq C \cdot ||x||$$

where the norms on X respectively Y appear. Denote by $\mathcal{L}(X,Y)$ the set of all continuous linear operators from X into Y. This yields a vector space equipped with the norm:

$$(*) \qquad ||T|| = \max_{||x||=1} ||T(x)|| \quad : T \in \mathcal{L}(X,Y)$$

Above X and Y are not necessarily Banach spaces. But one verifies easily that if $\hat{X}$ and $\hat{Y}$ are their completitions, then every $T \in \mathcal{L}(X,Y)$ extends in a unique way to a continuous linear operator $\hat{T}$ from $\hat{X}$ into $\hat{Y}$. One refers to $\hat{T}$ as the completion of T. Let us also notice the following:

3.1 Proposition. *If Y is a Banach space then the norm on $\mathcal{L}(X,Y)$ is complete, i.e. this normed vector space is a Banach space.*

The easy verification is left to the reader.

3.2 The open mapping theorem. Let X and Y be two Banach spaces and $T \in \mathcal{L}(X,Y)$. In X we get the subspace

$$\mathcal{N}(T) = \{x\colon\ T(x) = 0\}$$

Since T is continuous it is obvious that the kernel is a closed subspace of X. So by the general construction in XX we get the quotient space

$$\bar{X} = \frac{X}{\mathcal{N}(T)}$$

One verifies that T yields a linear operator $\bar{T}$ from $\bar{X}$ into Y which by the construction of the quotient norm on $\bar{X}$ has the same norm as T. Next, consider the image $T(X)$. It is obvious that

(i) $$T(X) = \bar{T}(\bar{X})$$

One says that T has *closed range* if the linear subspace $T(X)$ of Y is closed. Assume this holds. Then the complete norm on Y induces a complete norm on the closed subspace $T(X)$. In addition to this complete norm on $T(X)$ we have the norm defined by

$$||y||^* = ||\bar{x}|| \quad : \quad y = \bar{T}(\bar{x})$$

The Banach-Steinhaus theorem gives a constant C such that

$$||y||^* \leq C \cdot ||y||$$

This means that if $y \in T(X)$, then there exists $x \in X$ such that

(*) $$y = T(x) \quad : \quad ||x|| \leq C \cdot ||y||$$

Remark. One refers to (*) as the Open Mapping Theorem. The terminology is perhaps a bit confusing since (*) means that given a vector y in the closed range of T one can always find $x \in X$ such that $y = T(x)$ and at the same time choose x so that its norm in X does not exceed the constant C times $||y||$.

3.3 The closed graph theorem Let X and Y be Banach spaces. Consider a linear operator T from X into Y. In the product space $X \times Y$ we get the graph

$$\Gamma_T = \{(x, T(x)) \quad : x \in X\}$$

Now we can impose the condition that Γ_T is a closed subset of the Banach space $X \times Y$. Notice that

$$\mathcal{N}(T) = \{x\colon\ (x, 0) \in \Gamma_T\}$$

The hypothesis that Γ_T is a closed subset of $X \times Y$ obviously implies that $\mathcal{N}(T)$ is a closed subspace of X. Now we get the Banach space $X_* = \frac{X}{\mathcal{N}(T)}$ and obtain a *bijective* linear map:

(i) $$\mathfrak{i}\colon\ \bar{x} \mapsto (\bar{x}, T(\bar{x}))$$

from X_* into Γ_T. The induced complete norm on the closed graph Γ_T is defined by

(ii) $$||(\bar{x}, T(\bar{x})|| = ||\bar{x}|| + ||T(\bar{x})||$$

Theorem xx applies to $\mathfrak{i}$ and proves that the inverse map is continuos. This gives a constant C such that

(iii) $$||\bar{x}|| + ||T(\bar{x}]|| \leq C \cdot ||\bar{x}|| \implies ||T(\bar{x}]|| \leq C \cdot ||\bar{x}||$$

This implies that T has finite norm. Hence we have proved the following:

3.4 Theorem. *Let T be a linear operator from one Banach space X into another Banach space Y with a closed graph Γ_T. Then T is continuous.*

3.5 Densely defined operators.

Let $X_* \subset X$ be a dense subspace and $T\colon X_* \to Y$ a linear operator where Y is a Banach space. We get the linear subspace of $X \times Y$ defined by

$$\Gamma_T = \{(x, y)\colon \quad x \in X_* : \ y = T(x)\}$$

We can impose the condition that Γ_T is a closed subspace of $X \times Y$. When it holds we say that the densely defined operator T has a closed graph. Let us give

3.6 Example. Let $X = C^0_*[0,1]$ be Banach space whose elements are continuous functions $f(x)$ on the closed interval $[0,1]$ with $f(0) = 0$. The space $X_* = C^1_*[0,1]$ of continuously differentiable functions appears as a dense subspace of X. Next, let $Y = L^1[0,1]$. We get a linear map T from X_* into Y defined by

(i) $$T(f) = f' \quad : \quad f \in C^1_*[0,1]$$

In other words, we take the derivative $f'(x)$ which belongs to Y since it is a continuous function. Now T has a graph

(ii) $$\Gamma_T = \{(f, f') \quad : f \in C^1_*[0,1]\}$$

Here Γ_T is no a closed subspace of $X \times Y$. But we can construct its closure which yields a closed subspace denoted by Γ^*_T. By definition a pair (f, g) belongs to Γ^*_T if and only if

$$\exists \{f_n\} \in C^1_*[0,1] \quad : ||f - f_n|| \to 0 \quad : \int_0^1 |f'_n(t) - g(t)| \cdot dt = 0$$

The last limit means that the derivatives f'_n converge to an L^1-function g. Since $f_n(0) = 0$ are assumed we have

$$f_n(x) = \int_0^x f'_n(t \cdot dt \to \int_0^x g(t) \cdot dt$$

It follows that the continuous limit function f is equal to the primitive integral

(iii) $$f(x) = \int_0^x g(t) \cdot dt$$

3.7 Conclusion. The linear space Γ^*_T consists of pairs (f, g) with $g \in L^1[0, T)$ and f is the g-primitive defined by (iii). In this way we obtain a linear operator T^* with a closed graph. More precisely, T^* is defined on the linear subsapce of X given by functions $f(x)$ which are primitives of L^1-functions. This means by Lebesgue theory that the domain of definition of T^* consists of *absolutely continuous functions*. Thus, by enlarging the domain of definition the linear operator T is extended to a linear operator T^* whose graph is closed in $X \times Y$. One refers to T^* as a closed graph extension of T.

The example above is typical for many constructions where one starts with some densely defined linear operator T and finds an extension T^* whose graph is the closure of Γ_T. Notice that the choice of the target space Y affects the situation. As a further illustration, replace $L^1[0,1]$ with the Banach space $L^2[0,1]$ of square integrable functions on $[0,1]$. In this case we find a closed graph extension T^{**} whose domain of definition consists of continuous functions $f(x)$ which are primitives of L^2-functions. Since the inclusion $L^1[0,1] \subset L^2[0,1]$ is strict the domain of definition for T^{**} is a proper subspace of the linear space of all absolutely continuous functions. At the same time one gets a complete linear space given by

$$\mathcal{D}_{T^{**}} = \{f \in C^0_*[0,1] \quad : f(x) = \int_0^x g(t) \cdot dt \quad : g \in L^2[0,1]\}$$

This linear space is indeed complete when it is equippped with the norm

$$||f|| = ||g||_2 = \sqrt{\int_0^1 |g(t)|^2 \cdot dt}$$

This is an example of a Sobolev space. Constructions as above are often used in PDE-theory where one in general starts from a differential operator

(*) $$P(x, \partial) = \sum p_\alpha(x) \cdot \partial^\alpha$$

Here $x = (x_1, \ldots, x_n)$ are coordinates in $\mathbf{R}^n$ and ∂^α denote the higher order differential operators expressed by products of the first order operators $\{\partial_\nu = \partial/\partial x_\nu\}$. The coefficients $p_\alpha(x)$ are in general only continuous functions defined in some open subset Ω of $\mathbf{R}^n$, though the case when p_α

are C^∞-functions is the most frequent. Depending upon the situation one takes various target spaces Y, for example the Hilbert space $L^2(\Omega)$ of functions which are square integrable over Ω. To begin with one restricts $P(x, \partial)$ to the linear space $C_0^\infty(\Omega)$ of test-functions in Ω and constructs the corresponding graph. Then one seeks for extensions of this linear operator to larger subspace of functions on Ω and in favourable cases there exists a densely defined linear operator with a closed graph. We cannot enter this in more detail since this is a subject within PDE-theory. Let us only mention that the us of "abstract functional analysis" in this context is quite useful in PDE-theory. A result of this nature is *Gårding's inequality* established by Lars Gårding in [Gå] and later extended to the so called sharp Gårding inequality by L. Hörmander in [Hö]. This illustrates the usefulness of functional analysis, though one must not forget that delicate parts in the proofs rely upon "hard analysis".

4. Hilbert spaces.

Introduction. First we recall some geometric facts in the finite dimensional case which later on clarify properties of Hilbert spaces in the infinite dimensional case. A result in euclidian geometry asserts that if A is some invertible $n \times n$-matrix whose elements are real numbers and we regard A as a linear map from $\mathbf{R}^n$ into itself, then the image of the euclidian unit sphere S^{n-1} is an ellipsoid $\mathcal{E}_A$, and conversely if $\mathcal{E}$ is an ellipsoid then there exists an invertible matrix A such that $\mathcal{E} = \mathcal{E}_{\mathcal{A}}$.

0.1 The case $n = 2$. Already this case is instructive and the reader is invited to contemplate upon the two-dimensional case and study specific examples. For example, 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 $\mathcal{E}_A$.

4.2 A Historic Remark. The fact that $\mathcal{E}_A$ is an ellipsoid was wellknown in the Ancient Greek mathematics when $n = 2$ and $n = 3$. Moreover, the geometric constructions by Appolonius can be used to determine $\mathcal{E}_A$ when the linear map A is given. After general matrices and their determinants were introduced, the spectral theorem for symmetric matrices was established by A. Cauchy around 1810 under the assumption that the eigenvalues are different. Later Weierstrass found the proof in the general case. Independently Gram-Schmidt and Weierstrass also gave a method to produce an orthonormal basis of eigenvectors for a given symmetric $n \times n$-matrix B. An eigenvector with largest eigenvalue is found when 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 study the restricted extremal problem where x say in this orthogonal complement. Here we find a new 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.

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.

4.3 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 was evident by via 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.

4.4 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.

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

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

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.*

4.6 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.

4.7 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 showed the importance to study systems of linear equations in an infinite number of variables with certain bounds. For this purpose Fredholm constructed infinite families of pairwise orthogonal functions attached to a concrete inner product space. His procedure was to regard 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 we get the ellipsoid $\mathcal{E}_N = \mathcal{E}_{A_N}$ on V_N where it 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 and an Exercise. Certain 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. An excellent source is the introduction to integral equations by the former professor at Harvard University Maxime Bochner [Cambridge University Press: 1914):

4.8 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

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

4.9 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. Following the Gram-Schmidt construction it follows that 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 it follows that 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 an orthonormal family therefore 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 turns out to be very useful since inner products on specific linear spaces appear in many different situations. For example, in complex analysis an example occurs when we regard the space of analytic functions which are square integrable on a domain or whose boundary values are square integrable. Here is the inner product is given in advance but it can be a highly non-trivial affair to exhibit an orthonormal basis.

4.10 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 put:

(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 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 he norm of the T-images of the orthonormal basis vectors. So when T is bounded with some 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

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

So we encounter a rather 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.

4.11 General results on Hilbert spaces.

Let $\mathcal{H}$ for a while be a real Hilbert space. A fundamental result is that if K is a closed convex subset of $\mathcal{H}$ and if $\xi \in \mathcal{H} \setminus K$, then there exists a unique $k_* \in K$ such that

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

The proof is easy. For let ρ be the minimal distance. We find a sequence $\{k_n\}$ in K such that $||\xi - k_n|| \to \rho$. Now we shall prove that $\{k_n\}$ is a Cauchy sequence. To show this we let $\epsilon > 0$ and find first N_* such that

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

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

(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 we conclude that $\{k_n\}$ 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_* is a direct consequence of (XX).

4.12 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.

4.13 Complex Hilbert spaces. On a complex vector space similar results as above hold provided that we regard convex sets which are **C**-invariant. We leave details to the reader and refer to the literature for a more detailed account about general properties on Hilbert spaces. See for example the text-book [Hal] by P. Halmos - a former student to J. von Neumann - which in addition to theoretical results contains many interesting exercises.

4:B. Eigenvalues of matrices.

Using the Hermitian inner product on $\mathbf{C}^n$ we establish results about eigenvalues of an $n \times n$-matrices A with complex elements. The spectrum $\sigma(A)$ is the n-tuple of roots $\lambda_1, \dots, \lambda_n$ of the characteristic polynomial $P_A(\lambda) = \det(\lambda \cdot E_n - A)$, where eventual multiple eigenvalues are repeated. We also have the Hermitian matrix A^*A. Recall from (*) that $\sigma(A^*A)$ consists of non-negative real numbers denoted by $\{\mu_k\}$ which are arranged so that $\mu_1 \geq \mu_2 \geq \ldots \geq \mu_n$. In particular one has

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

4:B.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 $\xi_1, \dots, \xi_n$ 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 $e_1, \dots, e_n$ 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. $|\lambda_1| \geq \ldots \geq |\lambda_n|$. With these notations the following holds for an arbitrary $n \times n$-matrix A:

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

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

Before we begin the proof for a general p we consider the special case $p = 1$ and prove:

B.3 Proposition. *One has the inequality*

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

Proof. 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 of Weyl's theorem. The proof employs a construction of independent interest. Let $e_1, \dots, e_n$ be some orthonormal basis in $\mathbf{C}^n$. For every $p \geq 2$ we get the inner product space V^p whose vectors are

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

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. Next, 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 $1 \leq j_1 < \ldots < j_p$.

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

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

Hint. First, the eigenvalues 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 $A^{(p)}$. Next, using a Polarisation we may assume from the start that A is an upper triangular matrix and at this stage the reader can verify (*).

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

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

At this stage the reader can apply the Exercise and finish the proof of Weyl's theorem.

B.4 An inequality by Pick. Let C be a skew-symmetric $n \times n$-matrix, i.e. here $C^* = -C$. Notice that it implies that the eigenvalues of C are pure imaginary. Denote by g the maximum of the absolute values of the matrix elements of C. With these notations we have

B.5 Theorem. *One has the inequality*

$$\max_{|x|=1} \left|\langle Ax, x\rangle\right| \leq g \cdot \cot(\frac{\pi}{2n}) \cdot \sqrt{n(n-1)/2}$$

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

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

It follows that

$$\langle Ax, x\rangle = \sum\sum a_{ik} x_i \bar{x}_k = \sum_{i<k} a_{ik} x_i \bar{x}_k + \sum_{i>k} a_{ik} x_i \bar{x}_k$$

where the last equality simply follows when i and k are interchanged in the last sum on the first line. Since A is skew-symmetric the last term becomes

$$\sum_{i<k} a_{ik}[x_i \cdot \bar{x}_k - \bar{x}_i \cdot x_k]$$

B.5 Results by A. Brauer.

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

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

B.6 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

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

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

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

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

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

Exactly as above we get

$$|\lambda - a_{kk}| = |\bar{\lambda} - \bar{a}_{kk}| \leq \sum_{j \neq k} |a_{jk}| \tag{2}$$

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

B.7 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.

B.8 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 $x_k \geq 0$. The following result was originally established by Perron in [xx]:

B.9 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.*

Remark. We refer to $\mathbf{x}^*$ as the Perron vector of A. In [Frob] a proof is given by an induction over n which leads to the following addendum of Theorem B.9:

B.10 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 $|b_{pq}| \leq a_{pq}$ hold for all pairs p, q, it follows that every root of $P_B(\lambda)$ has absolute value $\leq s^*$ and equality holds if and only if $B = A$.*

Proof of Theorem B.9. First we establish the existence part. In $\mathbf{C}^n$ we have the norm defined by

$$||y|| = |y_1| + \ldots + |y_n| \tag{i}$$

Next, for each $\mathbf{x} \in \Delta_+^n$ we set

$$\phi(\mathbf{x}) = s \quad \text{where} \quad s = \max_{\xi > 0} \text{ such that } \quad \sum a_{pq} \cdot x_q \leq \xi \cdot x_p \quad : 1 \leq p \leq n$$

It is clear that ϕ is a continuous function on Δ_+^n and hence it takes is maximum at some point $\mathbf{x}^*$. Next, let $\lambda \in \sigma(A)$ have a maximal absolute value and let $\mathbf{y}$ be an eigenvector of norm one. The triangle inequality gives

$$||A(\mathbf{y})|| = \sum_{p=1}^{p=n} \Big|\sum_{q=1}^{q=n} a_{pq} \cdot y_q\Big| \leq \sum_{p=1}^{p=n} \sum_{q=1}^{q=n} a_{pq} \cdot |y_q| \leq s^* \cdot ||y|| = s^*$$

Hence we have the inequality

$$\text{(ii)} \qquad s^* \geq |\lambda| \quad \text{for all} \quad \lambda \in \sigma(A)$$

On the other hand we notice that if $N \geq 2$, then the definition of s^* gives

$$||A^N(\mathbf{x}^*)|| \leq (s^*)^N$$

It follows that

$$\text{(iii)} \qquad s^* \leq \lim_{N\to\infty} [||A^N||]^{\frac{1}{N}}$$

Hence the spectral radius formula implies that equality holds in (ii) and that s^* must be an eigenvalue for A which gives an eigenvector $\mathbf{x}^* \in \Delta_+^n$.

The uniqueness. There remains to prove that $\mathbf{x}^*$ is unique, or equivalently that the ϕ-function above attains its maximum at a unique point on Δ_+^n. This is left as an exercise to the reader. If necessary, consult the literature where the most elegant proofs occur in the collected work by Frobenius.

B.11 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 the 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 $s_* = 1/4$ with the eigenvector is $(1, -2)$.

5. Dual vector spaces

Let X be a normed space over the complex field. A continuous linear form on X is a $\mathbf{C}$-linear map γ from X into $\mathbf{C}$ such that there exists a constant C with:

$$\text{(iii)} \qquad \max_{||x||=1} |\gamma(x)| \leq C$$

The set of these continuous linear forms is denoted by X^*. It is obvious that X^* is a vector space and that (iii) defines a norm on X^*. Moreover, since Cauchy-sequences of complex numbers converge it follows that X^* is a Banach space. Notice that this holds even if X from the start is not complete. One refers to X as the dual of X. Next, let Y be a subspace of X. Every $\gamma \in X^*$ can be restricted to Y and gives an element of Y^*, i.e. we have the restriction map

$$\text{(i)} \qquad \mathfrak{res}_Y : X^* \to Y^*$$

Since a restricted linear form cannot increase the norm on X one has the trivial inequality

$$||\mathfrak{res}_Y(\gamma)| \leq ||\gamma|| \quad : \quad \gamma \in X^*$$

The kernel of $\mathfrak{res}_Y$. The kernel is by definition the set of X^*-elements which are zero on Y. This is a closed subspace of X^* which can be identfied with the dual of a new normed space. Namely, consider the linear quotient space

$$Z = \frac{X}{Y}$$

Thus, elements in Z are images of vectors $x \in X$. Here two vectors x_1 and x_2 give the same vector in Z if and only if $x_2 - x_1 \in Y$. Let $\pi_Y(x)$ denote the image of $x \in X$. Now Z is is equipped with a norm defined by

$$||z|| = \min ||x|| \quad : \quad z = \pi_Y(x)$$

This gives a norm on Z and by he construction above one has a canonical isomorphism

$$Z^* \simeq \mathrm{Ker}(\mathfrak{res}_y)$$

Thus, the dual space Z^* can be identified with a closed subspace of X^*.

5.1 The Hahn-Banch Theorem. It asserts that every continuous linear form on a subspace Y of X has a *norm preserving extension* to a linear form on X. Thus, if $\gamma_* \in Y^*$ has some norm C, then there exists $\gamma \in X$ such that

$$\mathfrak{res}_Y(\gamma) = \gamma_*$$

One refers to γ^* as a norm-åreserving extension of γ.

5.2 Exercise. Consult a text-book for the proof or give alternatively the details using the following hint. Given the pair (Y, γ_*) we consider all pairs (Z, ρ) where $Y \subset Z \subset X$ and $\rho \in Z^*$ is such that its norm is C and $\rho|Y = \gamma_*$.Thus, ρ is a norm preserving extension of γ_* to Z. By *Zorn's Lemma* there exists a maximal pair (Z, ρ) in this family. There remains only to show that $Z = X$ for then ρ gives the required norm-preserving extension. To prove that $Z = X$ one argues by contradiction. Namely, suppose $Z \neq X$ and choose a vector $x_0 \in X \setminus Z$ of norm one. Next, if α is a complex number we get a linear form on $Z^* = Z + \mathbf{C} \cdot x_0$ defined by

$$\rho_\alpha(ax_0 + z) = a \cdot \alpha + \rho(z)$$

where $a \in \mathbf{C}$ and $z \in Z$ are arbitrary. The contradiction follows if we can find α so that the norm of ρ_α again is $\leq C$. It is clear that $\rho_\alpha|| \leq C$ holds if and only if

$$(*) \qquad |\alpha + \rho(z)| \leq C \cdot ||x_0 + z|| \quad \text{hold for all } z \in Z$$

At this stage the reader should be able to finish the proof.

5.3 A separation theorem. Above we studied complex normed spaces. We can also consider a normed space X over the real numbers in which case the dual X^* consists of bounded $\mathbf{R}$-linear forms. Let us now consider some closed and convex subset K of X. Then, if $p \in X \setminus K$ is outside K there exists $x^* \in X^*$ which separates p from K in the sense that there is some $\delta > 0$ so that

$$x^*(p) \geq \delta + x^*(x) \quad \text{for all} \quad x \in K$$

An exact sequence. Let $Y \subset X$ as above be a subspace and $Z = \frac{X}{Y}$ the quotient space. The Hahn-Banach Theorem shows that there exists an exact sequence

$$0 \to Z^* \to X^* \to Y^* \to 0$$

Moreover, the restriction map $X^* \to Y^*$ sends the unit ball in X^* onto the unit ball of Y^*, Notice that this precisely is the assertion of the Hahn-Banach Theorem.

An example Let X, Y, Z be as above and consider some $\gamma \in X^*$. Now γ can be restricted to Y where we get the norm

$$A = ||\mathfrak{res}_Y(\gamma)||$$

By the Hahn-Banach theorem there exists some $\hat{\gamma} \in X^*$ of norm A whose image under $\mathfrak{res}_Y$ is the same as for γ. Identifying Z^* with a subspace of X^* this means that

$$\gamma - \hat{\gamma} \in Z^*$$

Let us give a specific example which will be applied to Hardy spaces. Here $X = L^1(T)$ is the normed space of integrable functions on the unit circle. Recall that the dual space $X^* = L^\infty(T)$. Next, we have the subspace $H^\infty(T)$ of X^* of those Lebesgue measurable and bounded functions on T which are boundary values to analytic functions in the unit disc D. We have also the subspace $Y = H_0^1(T)$ of L^1-functions which are boundary values of analytic functions which are zero at the origin. As explained in XXX expansions in Fourier series show that if $g \in L^\infty(T)$ then

$$\int_0^{2\pi} g \cdot h \cdot d\theta = 0 \quad \text{for all} \quad f \in H_0^1$$

holds if and only if $g \in H^\infty(T)$. This means precisely that

$$H^\infty(T) = \text{Ker}(\mathfrak{res}_Y) \quad : Y = H_0^1(T) \subset L^1(T) = X$$

Consider now some $g \in L^\infty(T)$ and define the constant C by:

$$C = \max_h |\int_0^{2\pi} g \cdot h \cdot d\theta \quad : \quad h \in H_0^1(T) \text{ and } ||h||_1 = 1$$

The general result above applies and gives the existence of some $h \in H^\infty(T)$ such that the L^∞-norm norm $||g - h||_\infty = C$. Thus, we have

$$g = h + f \tag{1}$$

where $h \in H^\infty$ and the L^∞-function f has norm C. Notice that $||g||_\infty \geq C$ holds here. The norm of h is not determined because one may have several decompositions in (1). However, in XX we shall find a specific decompositions of g in certain cases.

5.4 Weak Convergence.

Let X be a normed space. On the dual X^* one can define a topology where convergence only has to be pointwise. It means that a fundamental system of open neighborhood of the origin in the vector space X^* is given by

$$U(x_1, \ldots, x_N; \epsilon) = \{\gamma \in X^* \quad : |\gamma(x_\nu)| < \epsilon \quad : x_1, \ldots, x_N \text{ finite set}\} \tag{*}$$

Notice that each such U-set is a convex subset of X^*. Let Y be the finite dimensional subspace of X generated by $x_1, \ldots, x_n$. Then it is clear that the kernel of $\mathfrak{res}_Y$) is contained in the U-set above.

If k is the dimension of the vector space generated by $x_1, \ldots, x_n$ then Linear Algebra implies that the kernel of $\mathfrak{res}_Y$) has codimension k in X. So the U-set in (*) contains a subspace of X^* with finite codimension, i.e. roughly speaking the open U-set in X^* is quite large.

5.5 The case when X is separable. Suppose that a sequence $x_1, x_2, \ldots$ is a dense subset of X. Let $B(X^*)$ denote the unit ball in X^*, i.e. elements $\gamma \in X^*$ of norm ≤ 1. On this unit ball we define a metric by

$$d(\gamma_1, \gamma_2) = \sum_{n=1}^{\infty} 2^{-n} \cdot \big|\gamma_1(x_n) - \gamma_2(x_n)\big|$$

Exercise. Verify that the metric above gives the induced weak topology on $B(X^*)$ defined via the U-sets in (*).

5.6 Theorem. *The metric space $B(X^*)$ is compact.*

Proof. Let $\{\gamma_k\}$ be a sequence in $B(X^*)$. To every j we get the bounded sequence of complex numbers $\{\gamma_k(x_j)\}$. By the wellknown diagonal construction there exists a strictly increasing sequence $k_1 < k_2 < \ldots$ such that if $\rho_\nu = \gamma_{k_\nu}$ then

(*)
$$\lim_{\nu \to \infty} \rho_\nu(x_j)$$

exists for every j. Since every ρ_j has norm ≤ 1 and $\{x_j\}$ is dense in X it follows that

$$\lim_{\nu \to \infty} \rho_\nu(x) \quad \text{exist for all } x \in X$$

This gives some $\rho \in X^*$ such that $\rho(x)$ is the limit value above for every x. It is clear that the norm of ρ is ≤ 1 and by the construction of the distance function on $B(X^*)$ we have:

$$\lim_{\nu \to \infty} d(\rho_\nu, \rho) = 0$$

This proves that the given γ-sequence contains a convergent subsequence. So by the definition of compact metric spaces Theorem 6.1 follows.

5.7 Weak hulls in X^*. Let X be separable and choose a denumerable and dense subset $\{x_n\}$. Examples show that in general the dual space X^* is no longer separable in its norm topology. However, there always exists a denumerable sequences $\{\gamma_k\}$ in X^* which is dense in the weak topology.

Exercise. For every $N \geq 1$ we let V_N be the finite dimensional space generated by $x_1, \ldots, x_N$. It has dimension N at most. Applying the Hahn-Banch theorem the reader finds a sequence $\gamma_1, \gamma_2, \ldots$ in X^* such that for every N the restricted linear forms

$$\gamma_\nu | V_N \quad 1 \leq \nu \leq N$$

generate the dual vector space V_N^*. Next, let Q be the field of rational numbers. Show that if Γ is the subset of X^* formed by all finite Q-linear combinations of the sequence $\{\gamma_\nu\}$ then this denumerable set is dense in X^* with respect to the weak topology.

Another exercise. Let X be a separable Banach space and let E be a subspace of X^*. We say that E point separating if there to every $0 \neq x \in X$ exists some $e \in E$ such that $e(x) \neq 0$. Show first that every such point-separating subspace of X^* is dense with respect to the weak topology. This is the easy part of the exercise. The second part is less obvious. Namely, put

$$B(E) = B(X^*) \cap E$$

Prove now that $B(E)$ is a dense $B(X^*)$. Thus, if $\gamma \in B(X^*)$ then there exists a sequence $\{e_k\}$ in $B(E)$ such that

$$lim_{k \to \infty} e_k(x) = \gamma(x)$$

hold for all $x \in X$.

5.8 Example from integration theory. An example of a separable Banach space is $X = L^1(\mathbf{R})$ whose elements are Lebesgue measurable functions $f(x)$ for which the L^1-norm

$$\int_{-\infty}^{\infty} |f(x)| \cdot dx < \infty$$

If $g(x)$ is a bounded continuous functions on $\mathbf{R}$, i.e. there is a constant M such that $|g(x)| \leq M$ for all x, then we get a linear functional on X defined by

$$g^*(f) = \int_{-\infty}^{\infty} g(x) \cdot f(x) \cdot dx < \infty$$

Let E be the linear space of all bounded and continuous functions. By the previous exercise it is a dense subspace of X^* with respect to the weak topology. Moreover, by the second part of the exercise it follows that if $\gamma \in X^*$ has norm one, then there exists a sequence of continuous functions $\{g_n\}$ of norm one at most such that $g_\nu \to \gamma$ holds weakly. Let us now find γ. For this purpose we define the functions

$$\text{(i)} \qquad G_n(x) = \int_0^x g_n(t) \cdot dt \quad : \quad x \geq 0$$

These primitive functions are continuous and enjoy a further property. Namely, since the maximum norm of every g-function is ≤ 1 we see that

$$\text{(ii)} \qquad |G_n(x) - G_n(x')| \leq |x - x'| \quad : \quad x, x' \geq 0$$

This is means that whenever $a > 0$ is fixed, then the sequence $\{G_n\}$ restricts to an *equi-continuous* family of functions on the compact interval $[0, a]$. Moreover, for each $0 < x \leq a$ since we can take $f \in L^1(\mathbf{R})$ to be the characteristic function on the interval $[0, x]$, the weak convergence of the g-sequence implies that there exists the limit

$$\text{(iii)} \qquad \lim_{n \to \infty} G_n(x) = G_*(x)$$

Next, the equi-continuity in (ii) enable us to apply the classic result due to C. Arzéla in his paper *Intorno alla continua della somma di infinite funzioni contionue* from 1883 and conclude that the point-wise limit in (iii) is uniform. Hence the limit function $G_*(x)$ is continuous on $[0, a]$ and it is clear that G_* also satisfies (ii), i.e. it is Lipschitz continuous of norm ≤ 1. Since the passae to the limit can be carried out for every $a > 0$ we conclude that G_* is defined on $[0, +\infty >)$. In the same way we find G_* on $(-\infty, 0]$. Next, by the result in [XX-measure] there exists the Radon-Nikodym derivative $G_*(x)$ which is a bounded measurable function $g_*(x)$ whose maximum norm is ≤ 1. So then

$$G_*(x) = \int_0^x g_*(t) \cdot dt = \lim_{n \to \infty} G_n(x) = \lim_{n \to \infty} \int_0^x g_n(t) \cdot dt$$

holds for all x. Since finite $\mathbf{C}$-linear sums of characteristic functions is dense in $L^1(\mathbf{R})$ we conclude that the limit functional γ is determined by the L^∞-function g_*. So this shows the equality

$$L^1(\mathbf{R})^* = L^\infty(\mathbf{R})$$

Remark. The result above is of course wellknown. But it is interesting to see how the last duality formula is derived from studies of the Lebesgue integral.

5.9 The weak topology on X. Let X be a Banach space. Here we do not assume that X is separable. A sequence $\{x_k\}$ in X converges weakly to a limit vector x if

$$\lim_{k \to \infty} x^*(x_k) = x^*(x) \quad \text{hold for all } x^* \in X^*$$

Exercise. Show that when $\{x_k\}$ is a weakly convergent sequence then it must be bounded, i.e. there exists a constant C such that

$$||x_k|| \leq C$$

hold for all k.

5.10 Weak versus strong convergence. A weakly convergent sequence need not be strongly convergent. An example is when $X = C^0[0, 1]$ is the Banach space of continuous functions on the closed unit interval. By the Riesz representation theorem the dual space X^* consists of Riesz measures. A sequence $\{x_n(t)\}$ of continuous functions converge weakly to zero if

$$\lim_{n\to\infty} \int_0^1 x_n(t)\cdot d\mu(t) = 0$$

hold for every Riesz measure μ. By the result from [Measure] this holds if and only if the maximum norms of the x-functions are uniformly bounded and the sequence converges pointwise to zero. We can construct many such pointwise convergent sequences which fail to converge in the maximum norm. So in this example the weak convergence is *strictly weaker* than the topology defined by the maximum norm on X.

5.11 Remark. The example below was not so special. Namely, if X is an arbitrary infinite dimensional Banach space then the norm-topology is always strictly stronger than the weak topology. The proof is very easy for by the definition of the weak topology an equality with the norm topology implies that there exists a finite subset $x_1^*,\ldots,x_N^*$ of X^* and a constant C such that one has the implication

$$\max_\nu |x_\nu^*(x)| < C \implies ||x|| < 1 \quad \text{for all} \quad x \in X$$

But then it is clear that X^* as a complex vector space is generated by the n-tuple $x_1^*,\ldots,x_N^*$ and via the Hahn-Banach theorem it follows that X has dimension N at most.

5.12 Further results. Much more could have been said about topologies on X and its dual. For example, we have not defined *reflexive* Banach spaces which are characterised by the condition that the natural map from X into its bi-dual X^{**} is surjective. Other results deal with various separation theorems. A major result asserts that if K is a closed and convex set in X then it is also closed with respect to the weak topology on X. For proofs we refer to the extensive literature devoted to functional analysis. An outstanding reference for the foundations in functional analysis is the text-book Linear Operators Volume 1 by Dunford and Schwarz which covers all essential results in functional analysis. with elegant and very detailed proofs, including very many instructive exercises and an extensive list of references covering all discoveries before 1960. For a more recent account we refer to the text-book [Lax] by Peter Lax who received the Abel Prize in 2005 for his contributions in non-linear PDE-theory.

Duality mappings in Banach spaces.

The notion if uniform convexity and differentiability was defined in A.X from the introduction. Theorem XX below was proved in the article [Beurling-Livingston] and the subsequent material follows [ibid].

5.13 Strictly convex Banach spaces. A Banach space B is strictly convex if

$$||x+y|| < ||x|| + ||y||$$

for all pairs x, y except when $y = a\cdot x$ for some real and positive a. The stronger condition of of uniformly convex Banach spaces from XX in the introduction was given by Clark in 1936. Next, ecall from § XX that the Banach space X is differentiable at a point x if there for every $y \in B$ exists a real number $\mathcal{D}_x(y)$ such that

$$\lim_{\zeta\to 0} ||x+\zeta\cdot y|| - ||x|| = \mathfrak{Re}(\,\zeta\cdot\mathcal{D}_x(y)) + \text{small ordo}(|\zeta|)$$

5.14 Exercise. Verify that when B is differentiable at some x, then

$$y \mapsto \mathcal{D}_x(y)$$

is a linear functional on X whose norm is one and that

$$\mathcal{D}_x(x) = ||x||$$

5.15 Conjugate vectors. Let S be the unit sphere in X and S^* the unit sphere in X^*. A pair $x \in S$ and $u \in S^*$ are said to be conjugate if

$$u(x) = 1$$

5.16 Theorem. *Assume that X is uniformly convex and differentiable. Then, for every $x \in S$, the vector $\mathcal{D}_x$ is the unique conjugate of x. Moreover, the map $x \to \mathcal{D}_x$ from S to S^* is bijective.*

Proof. Let $x \in S$. Then (ii) in Exercise 5.14 shows that $\mathcal{D}_x$ is a conjugate vector in S^*. To prove the uniqueness we suppose that (x, u) are conjugate for some $u \in S^*$. For every vector $y \in X$ and complex number ζ we have

$$0 \leq ||x + \zeta \cdot y|| - \mathfrak{Re}\, u(x + \cdot y) = ||x + \zeta \cdot y|| - 1 - \mathfrak{Re}(\,\zeta \cdot u(y))$$

The construction of $\mathcal{D}_x$ and a passage to the limit in (i) when $\zeta \to 0$ entails that

$$\mathcal{D}_x(y) = u(y)$$

This proves that $\mathcal{D}_x$ is the unique conjugate vector to x. There remains to prove that the map

$$x \to \mathcal{D}_x$$

is bijective. To prove this we use the strict convexity of X which implies that for each $u \in S^*$ there exists a unique $x \in S$ such that $u(x) = 1$. Applied to a pair $\mathcal{D}_x$ and $\mathcal{D}_y$ with x and y in S, we conclude that the map (xx) is injective. Finally, let $u \in S^*$ be given. Again, by the result in XX we find $x \in S$ such that $u(x) = 1$ and then we have seen that $u = \mathcal{D}_x$ which proves that (xx) is surjective.

5.17 Duality maps. Let X as above be uniformly convex and differentiable and $\phi(r)$ a strictly increasing and continuous function defined on $[0, +\infty)$ where

$$\phi(0) = 0 \quad \text{and} \quad \lim_{r \to +\infty} \phi(r) = +\infty$$

A map Φ from X into X^* is called an associated duality map if $\Phi(0) = 0$ and

$$\Phi(x) = \phi(||x||) \cdot (\frac{x}{||x||})^*$$

for every non-zero vector x.

Notice that this entails that Φ is a bijective map from x onto X^*. Next, let C be a closed subspace in X and set:

$$C^\perp = \{\xi \in X^* : \xi(C) = 0\}$$

5.18 Theorem. *For each closed subspace C of X and every pair of vectors $x_0 \in X$ and $y_0 \in X^*$ the intersection*

$$\Phi(C + x) \cap \{C^\perp + y_0\}$$

is non-empty and consists of a single point ξ.

Proof. Let us introduce the primitive function

$$G(r) = \int_0^r \phi(s) \cdot ds$$

Notice that (xx) above gives

$$\lim_{r \to \infty} \frac{G(r)}{r} = +\infty$$

Define the functional F on X by

$$F(x) = G(||x||) - \mathfrak{Re}\, y_0(x))$$

If $||x|| = r$ we notice that

$$F(x) \geq G(||x||) - r \cdot ||y_0||$$

Hence (xx) shows that $F(x) \to +\infty$ as $||x|| \to +\infty$ and there exists a finite minimum

$$M = \inf_{x \in C + x_0} F(x)$$

Let $\{x_n\}$ be a sequence such that $F(x_n) \to M$. Let $|epsilon > 0$ and choose n_* such that

$$n \geq n_* \implies F(x_n) < M + \epsilon$$

Since the set C is convex we have

$$F(\frac{x_n + x_m}{2}) \geq M$$

when n and m both are $\geq n_*$. Next, since the function G is convex we conclude that if n, m both are $\geq n_*$ then

$$0 \leq \frac{1}{2}[G(||x_n||) + G(||x_m||)] - G(||\frac{x_n + x_m}{2}||) = \frac{1}{2}[F(x_n) + F(x_m)] - F(\frac{x_n + x_m}{2}) < \epsilon$$

Next, passing to a subsequence if necessary we may assume that there exists a limit

$$\lim_{n \to \infty} ||x_n|| = \alpha$$

It follows from the above that

$$\lim_{n,m \to \infty} G(||\frac{x_n + x_m}{2}|| = G(\alpha)$$

Since G is strictly increasing this entails that

$$\lim_{n,m \to \infty} ||\frac{x_n + x_m}{2}|| = \alpha$$

Now (xx) and the uniform convexity implies that $\{x_n\}$ is a Cauchy sequence and hence it converges to a limit vector ξ. Since the set $C + x_0$ is closed we have

$$\xi \in C + x_0$$

There remains to show the inclusion

$$\Phi(\xi) \in C^\perp + y_0$$

To prove (xx) we notice that since F achieves a minimum at ξ one has

$$F(\xi + s \cdot x) \geq F(\xi)$$

when $x \in C^\perp$ and s are complex numbers. (C= real subspace assumed). Now we have

$$\frac{F(\xi + s \cdot x) - F(\xi)}{s} = \frac{G(||\xi + s \cdot x)|| - G(||\xi||)}{s} - \mathfrak{Re}\, y_0(x)$$

Next, the G-function is differentiable with derivstive ϕ and since X is differentiable at ξ we see that

$$\lim_{s \to 0} \frac{G(||\xi + s \cdot x)|| - G(||\xi||)}{s} = \phi(||\xi||) \cdot \mathcal{D}_\xi(x)$$

By (xx) the left hand side in (xx) is ≥ 0 and hence we get the inequality

$$\phi(||\xi||) \cdot \mathcal{D}_\xi(x) \geq \mathfrak{Re}\, y_0(x) \quad : \quad x \in C$$

Since C is a real vector space this entails that

$$\phi(||\xi||) \cdot \mathcal{D}_\xi(x) = \mathfrak{Re}\, y_0(x) \quad : \quad x \in C$$

In other words, we have the inclusion

$$\phi(||\xi||) \cdot \mathcal{D}_\xi - y_0 \in C^\perp$$

FINISH

6. Fredholm theory.

Introduction. We prove some general results about bounded operators which go back to Ivar Fredholm's article [Fredholm] about integral equations from 1901. Fredholm's work was restricted to special Banach spaces but the proofs in the general case are verbatim except for Theorem 6.12 which is due to Laurent Schwartz. Here is the set-up for this section. Let X and Y be two Banach spaces. Their dual spaces are denoted by X^* and Y^*.

6.1 Adjoint operators. Let $u\colon X \to Y$ be a bounded linear operator . The adjoint u^* is the linear operator from Y^* to X^* defined by

$$u^*(y^*)\colon\, x \mapsto y^*(u(x)) \quad : \quad y^* \in Y \quad : x \in X \tag{1}$$

Exercise. Show the equality of operator norms:

$$||u|| = ||u^*||$$

The hint is to apply the Hahn-Banach theorem.

6.2 Operators with finite dimensional range. The range is the image space $u(X)$. Suppose the range is finite dimensional and let N be the dimension of the vector space $u(X)$. Then we can choose an N-tuple $x_1, \ldots, x_N$ in X such that the vectors $\{u(x_k)\}$ is a basis for $u(X)$. Notice that this implies that $x_1, \ldots, x_N$ are linearly independent in X. Hence we get the N-dimensional subspace of X :

$$V = \mathbf{C}x_1 + \ldots + \mathbf{C}x_N$$

Consider the u-kernel

$$N_u = \{x \quad : \, u(x) = 0\}$$

The reader should verify that

$$X = N_u \oplus V$$

Next, consider the adjoint operator u^*. In Y^* we can find an N-tuple $y_1^*, \ldots, y_N^*$ such that

$$j \neq k \implies y_j^*(u(x_k)) = 0 \quad \text{and} \quad y_j^*(u(x_j)) = 1$$

If N_{u^*} is the kernel of u^* the reader may verify that

$$Y^* = N_{u^*} \oplus \mathbf{C}y_1^* + \ldots + \mathbf{C}y_N^*$$

Conclude from this that the range of u^* also is an N-dimensional vector space.

6.3 The operator $\bar{u}$. Let $u\colon X \to Y$ which gives the closed null space N_u in X. Now $\frac{X}{N_u}$ is a new Banach space and we get the induced linear operator

$$\bar{u}\colon\, \frac{X}{N_u} \to Y$$

By construction $\bar{u}$ is an *injective* linear operator and it is clear that it has the same range as u, i.e. one has the equality

$$u(X) = \bar{u}(\frac{X}{N_u}) \tag{2}$$

6.4 The image of u^*. In the dual space X^* we get the subspace

$$N_u^\perp = \{x^* \in X^* \quad : \, x^*(N_u) = 0\} \tag{3}$$

Next, let $y^* \in Y^*$ and consider the image $u^*(y^*)$. If $x \in N_u$ we have by (1)

$$u^*(y^*)(x) = y^*(u(x)) = 0$$

Hence we get the following inclusion:

$$u^*(Y^*) \subset N_u^\perp \tag{4}$$

Next, recall the canonical isomorphism

$$(5) \qquad [\frac{X}{N_u}]^* \simeq N_u^\perp$$

Now we consider the linear operator $\bar{u}$ whose adjoint $\bar{u}^*$ maps Y^* into the dual of $\frac{X}{N_u}$. Using the canonical isomorphism (5) this means that

$$\bar{u}^*\colon\ Y^* \mapsto N_u^\perp$$

At this stage the reader should verify the equality

$$(6) \qquad \mathrm{Im}(\bar{u}^*) = \mathrm{Im}(u^*)$$

The closed range property

A bounded linear operator $u\colon X \to Y$ has closed range if $u(X)$ is a closed subspace of Y. Suppose this holds. By the constructions above the linear operator $\bar{u}$ is injective and its image space is $u(X)$. So when $u(X)$ is closed it follows that $\bar{u}\colon X \to u(X)$ is a bijective map between Banach spaces. The Open Mapping Theorem applies and implies that $\bar{u}$ is an invertible linear operator between $\frac{X}{N_u}$ and $u(X)$. Passing to its adjoint we get a bijective and bi-continuous map

$$(i) \qquad \bar{u}^*\colon\ u(X)^* \to N_u^\perp$$

where $N_u^\perp$ is identified with the dual of $\frac{X}{N_u}$. Using the equality (6) above we conclude that the image space

$$u^*(Y^*) = N_u^\perp$$

Here $N_u^\perp$ is a closed subspace of X^* and hence we have proved:

6.5 Proposition. *Assume that u has closed range. Then u^* has closed range and one has the equality*

$$\mathrm{Im}(u^*) = N_u^\perp$$

A converse result. Let $u\colon X \to Y$ be a bounded linear operator. But this time we do not assume that it has a closed range from the start. Instead we assume that the adjoint operator u^* has a closed range. By the equality (xxx) this implies that the injective linear operator $\bar{u}$ is such that its adjoint $\bar{u}^*$ has closed range. Using this the reader should verify the following converse to Proposition 6.5.

6.6 Proposition. *If u^* has closed range it follows that u also has closed range.*

6.7 Compact operators.

A linear operator $T\colon X \to Y$ is compact if the the image under T of the unit ball in X is relatively compact in Y. In other words, compactness means that if $\{x_k\}$ is an arbitrary sequence in the unit ball $B(X)$ then there exists a subsequence of $\{T(x_k)\}$ which converges to some $y \in Y$. Next, let $\{T_n\}$ be a sequence of compact operators which converge to another operator T, i.e.

$$\lim_{n\to\infty} ||T_n - T|| = 0$$

where we employ the operator norm on the Banach space $L(X, Y)$. Then the reader may verify that T also is a compact operator.

6.8 Operators with finite-dimensional range. If $T\colon X \to Y$ is such that $T(X)$ is a finite dimensional subspace of Y then it is easily seen that T is compact. Denote by $\mathcal{F}(X, Y)$ the linear space of operators from X to Y with finite dimensional range. So now $\mathcal{F}(X, Y)$ is a subspace of the linear space $\mathcal{C}(X, Y)$ of all compact operators.

6.9 Enflo's example. The question arises if $\mathcal{F}(X, Y)$ is a dense subspace of $\mathcal{C}(X, Y)$. This was an open problem for many decades until Per Enflo in a seminar at Stockholm University in 1972 constructed an example of a separable Banach space X and a compact operator $T \in \mathcal{C}(X, X)$ which cannot be approximated by operators from $\mathcal{F}(X,X)$. This example has led to a veritable industry where one seeks to determine "good pairs" of Banach spaces X and Y for which $\mathcal{F}(X,Y)$ is dense in $\mathcal{C}(X,Y)$. We shall not dwell upon this but remark only that for most of the "familiar" Banach spaces one has the density which therefore means that a compact operator can be approximated in the operator norm by operators having finite dimensional range.

Before we announce the result below we notice that if T belongs to $\mathcal{F}(X,Y)$ then it has closed range. Indeed, this follows since every finite dimensional subspace of Y is closed. Moreover, the reader should verify that the image space $T^*(Y^*)$ also is finite dimensional and hence the adjoint T^* is a compact operator. However, taking Enflo's example into the account this special case does not cover the result below.

6.10 Theorem. *Let T be compact. Then the adjoint T^* is also compact.*

Exercise. Prove this result.

6.11 Stable range.

Now we study compact pertubations of linear operator. The main result goes as follows:

6.12 Theorem. *Let $u\colon X \to Y$ be an injective operator with closed range and $T\colon X \to Y$ a compact operator. Then the kernel of $u+T$ is finite dimensional and $u+T$ has closed range.*

Proof. First we show that N_{u+T} is finite dimensional. By XX it suffices to show that the set

$$B = \{x \in N_{u+T} \quad : \; ||x| \leq 1\}$$

is compact. So let $\{x_n\}$ be a sequence in B. Since T is compact there is a subsequence $\{\xi_j = x_{n_j}\}$ such that lim $T\xi_j = y$. It follows that

$$u(\xi_j) = -T(\xi_j) \to y$$

Now u is injective and has closed range so by the Open Mapping Theorem it is bi-continuous. So from the Cauchy sequence $\{u(\xi_j)\}$ produced via the limit in (i),it follows that the sequence $\{\xi_j\}$ converges to a vector ξ^* in X. But then it is clear from (i) that $u(\xi_*) = -T(\xi_*$ and hence $\xi_* \in B$. This proves that B is compact.

The closedness of Im$(u+T)$. Since N_{u+T} is finite dimensional we have a direct sum decomposition

$$X = N_{u+T} \oplus X_*$$

Now $(u+T)(X) = (u+T)(X_*)$ so it suffices that the last image is closed and we can restrict both u and T to X_* where we notice that the restricted operator T_* again is compact. Hence we may assume that the operator $u+T$ is *injective*. Next, let y be in the closure of Im$(u+T)$. It means that there is a sequence ξ_n in X such that

$$\text{(i)} \qquad \lim\,(u+T)(x_n) = y$$

Suppose first that the norms of $\{x_n\}$ are unbounded. Passing to a subsequence if necessary we may assume that $||x_n|| \to \infty$. With $x_n^* = \frac{x_n}{||x_n||}$ it follows that

$$\text{(ii)} \qquad \lim\, u(x_n^*) + T(x_n^*) = 0$$

Now $\{x_n^*\}$ is bounded and since T is compact we can pass to another subsequence and assume that $T(x_n^*) \to y$ holds for some $y \in Y$. But then (x) entails that $u(x_n^*)$ also has a limit and since u is ionjective it follows as above that $\{x_n^*\}$ is convergent. It $x_n^* \to x_*$. Here $x^* \neq 0$ since $||x_n^*|| = 1$ for all n. We see that (xx) entails that $u(x_*) + T(x_*) = 0$ and this is contradiction since N_{u+T} is assumed to be the null space.

So in (i) we now have that the sequence $\{x_n\}$ is bounded. Since T is compact we can pass to a subsequence and assume that $T(x_n) \to \xi$ holds for some $\xi \in Y$. But then (i) entails that the sequence $\{u(x_n)\}$ converges to $y - \xi$. Now u is injective so the Open Mapping Theorem implies that ¨$\{x_n\}$ is a Cauchy sequence in X and hence converges to some x^*. Passing to the limit in (i) we then get

$$u(x_*) + T(x_*) = y$$

Hence y belongs to $\mathrm{Im}(u+T)$ and Theorem 6.12 is proved.

6.13 Fredholm operators.

Let $u\colon X \to Y$ be a bounded linear operator. It is called a Fredholm operator if the kernel and the cokernel of u both are finite dimensional. In particular $\frac{Y}{u(X)}$ is a finite dimensional space and therefore $u(X)$ is closed, i.e. every Fredholm operator has a closed range. When u is a Fredholm operator we set

$$\mathfrak{ind}(u) = \dim N_u - \dim\left[\frac{Y}{u(X)}\right]$$

6.14 Theorem. *Let u be a Fredholm operator and $T\colon X \to Y$ a compact operator. Then $u+T$ is Fredholm and one has the equality*

$$\mathfrak{ind}(u) = \mathfrak{ind}(u+T)$$

Remark. When T has finite dimensional range this is an easy exercise and the equality for the indices follows from the Fredholm index formula in Linear algebra.

Proof in the general case.

Since $u(X)$ has finite codimension there exists a closed complement in Y, i.e.

$$Y = u(X) \oplus W$$

where W is a finite dimensional. Let $\pi\colon Y \to u(X)$ be the projection. Now $\pi \circ T$ is a compact operator from X into $u(X)$. If $\epsilon > 0$ we get the induced linear operator Next, we have the bijective operator

$$\bar{u}\colon \frac{X}{N_u} \to u(X)$$

Moreover, since N_u is finite dimensional we have another direct sum

$$X = N_u \oplus X_*$$

where X_* now is a finite dimensional subspace of X. Then u restricts to a bijective linear operator

$$u_*\colon X_* \to u(X)$$

Next, we can restrict T to the subspace X_* which yields an operator T_* from X_* to Y. Then we regard the composed operator $\pi \circ T_*$. With these notations we obtain for every $\epsilon > 0$ a linear operator $S_\epsilon = u_* + \epsilon \cdot \pi \circ T_*$. Here

$$(*) \qquad S_\epsilon\colon\ X_* \to u(X)$$

Now u_* is an isomorphism. By the general result in XX it first follows that the null space of S_ϵ is zero if ϵ is small. Notice that this only uses that the operator T is bounded. Next, since T is compact it follows from XX that S_ϵ has a closed range. Next, since the adjoint S_ϵ^* is injective when ϵ is small, it follows from XX that S_ϵ is an isomorphism, i.e. this conclusion holds for sufficiently small ϵ. Finally, since W and N_u are finite dimensional it follows via Linear algebra that $u + \epsilon T$ is Fredholm and has the same index as u. Now the reader can finish the proof using a homotopy argument over ϵ.

7. Calculus on Banach spaces.

Let X and Y be two Banach spaces and $g\colon X \to Y$ some map. Here g is not assumed to be linear. But just as in calculus one can impose the condition that when $x_0 \in X$ is a given then the difference $g(x) - g(x_0)$ is approximated in a linear way as the norms of $x - x_0$ becomes small. This leads to:

7.1 Definition. *We say that g is differentiable at x_0 if there exists a linear operator $L \in \mathcal{L}(X, Y)$ such that*

$$\lim_{||x-x_0||\to 0} \frac{||g(x) - g(x_0) - L(x - x_0)||}{||x - x_0||} = 0 \tag{*}$$

Remark. One verifies easily that L is unique if it exists. It is denoted by $D_g(x_0)$ and called the differential of g at x_0. If g has a derivative everywhere we get a new function

$$x \mapsto D_g(x) \tag{i}$$

with values in the Banach space $\mathcal{L}(X, Y)$. If D_g also has derivatives one says that g is twice differentiable and we get its second order differential defined by

$$D_g^2 = D_{D_g}$$

One continues in this way and for each $k \geq 1$ we get the class $C^k(X, Y)$ of k-times differentiable functions from X onto Y. Notice that the higher order differential maps have target manifolds which are iterated constructions of $\mathcal{L}(X, Y)$.

7.1.B Exercise. Let X be a Banach space and $g\colon X \to X$ a C^1-map such that D_g is the identity at the origin So the assumption is that

$$\lim_{||x||\to 0} \frac{||g(x) - x||}{||x||} = 0$$

Show that g is a local diffeomorphism, i.e. there exists some $\epsilon > 0$ such that g yields a bijective map from the open ball $B(\epsilon) = \{||x|| < \epsilon\}$ onto an open neighborhood U of the origin and $g^{-1}\colon U \to B(\epsilon)$ is a C^1map.

Remark. We shall not enter a more detailed discussion of the differential calculus of Banach-space valued functions but refer to the concise presentation of basic facts from Chapter 1 in Hörmander's text-book [Hö] where the reader also can find a proof of the exercise above.

7.2 Line integrals

Let $X = \mathbf{C}$ equipped with its usual norm given by absolute values of complex numbers. Let Y be a Banach space. Consider continuous maps g defined on some open set Ω in $\mathbf{C}$ with values in Y. Let $t \mapsto \gamma(t)$ be a parametrized C^1-curve whose image is a compact subset of Ω. Then we can define the line integral

$$\int_\gamma g \cdot dz = \int_0^T g(\gamma(t)) \cdot \dot{\gamma}(t) \cdot dt \tag{*}$$

The evaluation is performed exactly as for ordinary Riemann integrals, Namely, one uses the fact that the Y-valued function

$$t \mapsto g(\gamma(t))$$

is uniformly continuous with respect to the norm on Y, i.e. the Bolzano-Weierstrass theorem gives:

$$\lim_{\epsilon\to 0} \max_{|t-t'|\leq\epsilon} ||g(t) - g(t')|| = 0$$

Then (*) is approximated by Riemann sums exactly as in ordinary Calculus.

7.3 Analytic functions.

Let $g(z)$ be a continuous map from the open set Ω into the Banach space Y. We say that $g(z)$ is analytic a point $z_0 \in \Omega$ if there exists some $\delta > 0$ and a convergent power series expansion

$$(*) \qquad g(z) = g(z_0) + \sum (z - z_0)^\nu \cdot y_\nu : \qquad \sum ||y_\nu|| \cdot \delta^\nu < \infty$$

The last condition implies that the power series $\sum (z - z_0)^\nu \cdot y_\nu$ converges in the Banach space Y when $z \in D_\delta(z_0)$. Notice that if $\gamma \in Y^*$ then (*) gives an ordinary complex-valued analytic function

$$(**) \qquad \gamma(g)(z) = \gamma(g(z_0) + \sum c_\nu \cdot (z - z_0)^\nu : \qquad c_\nu = \gamma(y_\nu)$$

Since elements y in Y are determined when we know $\gamma(y)$ for every $\gamma \in Y^*$ we see that (**) entails that the sequence $\{y_\nu\}$ in (*) is unique, i.e. Y-valued analytic functions have unique power series expansions. Moreover, using (**) the reader may verify the following Bamach-space version of Cauchy's theorem.

7.4 Theorem. *Let $\Omega \in \mathcal{D}^1(\mathbf{C})$ and $g(z)$ is an Y-valued function which is analytic in Ω and extends to a continuous function on $\bar{\Omega}$. Let $f(z)$ be an ordinary analytic function in Ω which extends continuously to $\bar{\Omega}$. Then*

$$f(z_0) \cdot g(z_0) = \int_{\partial\Omega} \frac{f(\zeta) g(\zeta) d\zeta}{\zeta - z_0} \quad : \quad z_0 \in \Omega$$

Similarly, with the assumptions as above on f and g we have the vanishing result

$$\int_{\partial\Omega} f(\zeta) \cdot g(\zeta) d\zeta = 0$$

Remark. The results above show that analytic function theory can be applied in a quite general context. In these notes we have illustrated this in a section devoted to an existence proof of a non-linear PDE-quation where the strategy of the proof is to reduce everything to solutions of linear PDE-equations and use convergent series expansions with values in a suitable Banach space.

7.5 Resolvent operators

Let A be a continuous linear operator on a Banach space X. In XX we defined the spectrum $\sigma(A)$ and proved that the resolvent function

$$\text{(i)} \qquad R_A(z) = (z \cdot E - A)^{-1} \quad : \quad z \in \mathbf{C} \setminus \sigma(A)$$

is an analytic function, i.e. the local Neumann series from XX show that $R_A(z)$ is an analytic function with values in the Banach space $Y = \mathcal{L}(X, X)$. Let us now consider a connected bounded domain $\Omega \in \mathcal{D}^1(\mathbf{C})$ whose boundary $\partial\Omega$ is a union of smooth and closed Jordan curves $\Gamma_1, \ldots, \Gamma_p$. Let $f(z)$ be an analytic function in Ω which extends to a continuous function on $\bar{\Omega}$. Assume that

$$\text{(ii)} \qquad \partial\Omega \cap \sigma(A) = \emptyset$$

Then we can construct the line integral

$$(*) \qquad \int_{\partial\Omega} f(\zeta) \cdot R_A(\zeta) \cdot d\zeta$$

This yields an element of Y denoted by $f(A)$. Thus, if $\mathcal{A}(\Omega)$ is the space of analytic functions with continuous extension to $\bar{\Omega}$ then (*) gives a map

$$(**) \qquad T_A : \mathcal{A}(\Omega) \to Y$$

Let us put

$$\delta = \min \{|z - \zeta| \quad : \zeta \in \partial\Omega : \ z \in \sigma(A)\}$$

By the result in XX there is a constant C which depends on A only such that the operator norms:

$$(***) \qquad ||R_A(\zeta)|| \leq \frac{C}{\delta} \quad : \quad \zeta \in \partial\Omega$$

From (***) and the construction in (*) we conclude that the linear operators $T_A(f)$ have norms which are estimated by

$$||T_A(f)|| \leq \frac{C}{\delta} \cdot \ell(\partial\Omega) \cdot |f|_{\partial\Omega}$$

where $\ell(\partial\Omega)$ is the total arc-length of the boundary. Hence we have proved:

7.6 Theorem. *With Ω as above there exists a continuos linear map $f \mapsto T_A(f)$ from the Banach space $\mathcal{A}(\Omega)$ into Y and one has the norm inequality*

$$||T_A|| \leq \frac{C}{\delta} \cdot \ell(\partial\Omega)$$

The range of T_A. There remains to describe the range of the linear operator T_A and to discover further properties. Recall first that the resolvent operators $R_A(z)$ commute with A in the algebra of linear operators on X. Since $f(A)$ is obtained by a Riemann sum of resolvent operators, it follows that $f(A)$ commutes with A for every $f \in \mathcal{A}(\Omega)$. At the same time $\mathcal{A}(\Omega)$ is a *commutative Banach algebra*. It turns out that one has a multiplicative formula for T_A. More precisely one has:

7.7 Theorem. *T_A yields an algebra homomorphism from $\mathcal{A}(\Omega)$ into a commutative subalgebra of Y, i.e.*

$$T_A fg) = T_A(f) \cdot T_A(g) \quad : f, g \in \mathcal{A}(\Omega)$$

Proof of Theorem 7.7

The proof requires several steps. To begin with, in Y we get the closed subalgebra $\mathbf{A}$ generated by A and all the resolvent operators $R_A(z)$ as z moves outside $\sigma(A)$. Then $\mathbf{A}$ is a commutative Banach algebra whose Gelfand space is denoted by $\mathfrak{M}$. The first step towards the proof of Theorem 7.7. is:

7.8 Proposition *The Gelfand space $\mathfrak{M}$ can be identified with the compact set $\sigma(A)$.*

Proof. Let λ be a multiplicative linear functional on $\mathbf{A}$. By the definition of $\sigma(A)$ we must have

$$\lambda(A) = z_* \quad : z_* \in \sigma(A)$$

Now z_* determines λ. For if $R_A(z)$ is a resolvent operator we have

$$R_A(z) \cdot (z \cdot E - A) = E$$

where E is the identity in $\mathbf{A}$. Since λ is multiplicative this entails that

$$\text{(i)} \qquad 1 = \lambda(R_A(z) \cdot (z - z_*) \implies \lambda(R_A(z) = \frac{1}{z - z_*}$$

Hence z_* determines λ. Conversely, if we take $z_* \in \sigma(A)$ then we *define* λ such that $\lambda(A) = z_*$ and (i) holds for every $z \in \mathbf{C} \setminus \sigma(AS)$ and the reader may verify that this yields a multiplicative functional.

Remark. Recall that $\mathfrak{M}$ is the maximal ideal space of $\mathbf{A}$. If $z_* \in \sigma(A)$ and regard the *non-closed* algebra $\mathbf{A}_*$ generated by A and its resolvent operators, then it is obvious that we get the maximal ideal

$$\mathfrak{m}_*(zE - AS) \cdot \mathbf{A}_*$$

Taking its closure in $\mathbf{A}$ we get a maximal ideal in this commutative Banach algebra which corresponds to the point in $\mathfrak{M}$) determined by z_*.

Final part in the proof of Theorem 7.7. In addition to the given domain Ω we construct a slightly smaller domain Ω_* which also is bordered by p many disjoint and closed Jordan curves $\Gamma_1^*, \ldots, \Gamma_p^*$ where each single Γ_ν^* is close to Γ_ν and $\partial\Omega^*$ stays so close to $\partial\Omega$ that does not intersect $\sigma(A)$. Let us then consider pair f, g in $\mathcal{A}(\Omega)$. Now $\partial\Omega \cup \partial\Omega_*$ border a small domain where all functions are analytic. By analyticity and line integrals over $\partial\Omega$ or $\partial\Omega_*$ are equal. In particular we get:

$$T_A(g) = \int_{\partial\Omega_*} g(\zeta_*) \cdot R_A(\zeta_*) \cdot d\zeta_*$$

where we use ζ_* as a variable to distinguish from the subsequent integration along $\partial\Omega$. To compute $T_A(f)$ we keep integration on $\partial\Omega$ and obtain

(i) $$T_A(f) \cdot T_A(g) = \int_{\partial\Omega_*} \int_{\partial\Omega} g(\zeta_*) \cdot f(\zeta) \cdot R_A(\zeta_*) \cdot R(\zeta) \cdot d\zeta_* d\zeta$$

Next we use the resolvent equation

(ii) $$R_A(\zeta_*) \cdot R(\zeta) = \frac{R(\zeta_*) - R(\zeta)}{\zeta - \zeta_*}$$

The double integral in (i) therefore becomes a sum of two integrals

$$C_1 = \int_{\partial\Omega_*} \int_{\partial\Omega} g(\zeta_*) \cdot f(\zeta) \frac{R(\zeta_*)}{\zeta - \zeta_*} \cdot d\zeta_* d\zeta$$

$$C_2 = -\int_{\partial\Omega_*} \int_{\partial\Omega} g(\zeta_*) \cdot f(\zeta) \frac{R(\zeta)}{\zeta - \zeta_*} \cdot d\zeta_* d\zeta$$

To find C_1 we first perform integration with respect to ζ. Since every ζ_* from the inner boundary $\partial\Omega_*$ belongs to the domain Ω Cauchy's formula applied to the analytic function f gives:

$$f(\zeta_*) = \frac{1}{2\pi i} \int_{\partial\Omega} \frac{f(\zeta) d\zeta}{\zeta - \zeta_*}$$

Inserting this in the double integral defining C_1 we get

(iii) $$C_1 = \frac{1}{2\pi i} \int_{\partial\Omega_*} \int_{\partial\Omega_*} f(\zeta_*) g(\zeta_*) \cdot R(\zeta_*) \cdot d\zeta_* = T_A(fg)$$

To evaluate C_2 we first perform integration along $\partial\Omega_*$, i.e. we regard:

(iv) $$\int_{\partial\Omega_*} \frac{g(\zeta_*)}{\zeta - \zeta_*} \cdot d\zeta_*$$

Now ζ stays *outside* the domain Ω_* and hence (iv) is zero by Cauchy's vanishing theorem. So $C_2 = 0$ and then (i) gives the equality in Theorem 7.7.

7.9 The sup-norm case. Assume now that $\mathbf{A}$ is a sup-norm algebra and put $K = \sigma(A)$. Consider an open set Ω which contains the compact set K. By the previous results there exists an algebra homomorphism

$$T_A \colon \mathcal{O}(\Omega) \to \mathbf{A}$$

Let $f \in A(\Omega$. The spectrum of the $\mathbf{A}$-element $T_A(f)$ is equal to $f(\sigma(A))$. Since $\mathbf{A}$ is assumed to be a sup-norm algebra it follows that

(*) $$\max_{z \in K} |f(z)| = ||T_A(f)||$$

Here Ω is an arbitrary open neighborhood of K. Since $\mathbf{A}$ is a Banach algebra we can therefore perform a limit as the open sets Ω shrink to K and obtain another algebra homomorphism as follows: First we have the sup-norm algebra $\mathcal{A}(K)$ which consists of continuous functions on K which an be uniformly approximated b on K by analytic functions defined in small open neighborhoods. Then (*) implies that we have an algebra homomorphism

$$T_A \colon f \mapsto T_A(t) \quad : \quad f \in \mathcal{A}(K)$$

Moreover it is an isometry, i.e.

$$\max_{z \in K} |f(z)| = ||T_A(f)$$

In this way the Banach algebra $\mathbf{A}$ is identified with the sup-norm algebra $\mathcal{A}(K)$.

7.10 A special case. If K is "thin" one has the equality

$$(*) \qquad \mathcal{A}(K) = C^0(K)$$

For example, Theorem XXX shows that if the 2-dimensional Lebesgue measure of K is zero then all continuous functions on K can be uniformly approximated by rational functions with poles outside K and then (*) holds. If we also assume that $\mathbf{C} \setminus K$ is connected then Runge's Theorem from XX shows that $C^0(K)$ is equal to the closure of analytic polynomials $P(z)$. Passing to $\mathfrak{A}$ this implies that polynomials in A generate a dense subalgebra of $\mathbf{A}$.

8. Bounded self-adjoint operators.

Introduction. Let $\mathcal{H}$ be a complex Hilbert space. A bounded linear operator S on $\mathcal{H}$ is called is called self-adjoint if

$$(*) \qquad \langle x, Sy \rangle = \text{the complex conjugate of } \langle Sx, y \rangle \quad : \quad x, y \in \mathcal{H}$$

If S is self-adjoint we have the equality of operator norms:

$$(1) \qquad ||S||^2 = ||S^2||$$

To see this we notice that if $x \in \mathcal{H}$ has norm one then

$$(i) \qquad \langle Sx, Sx \rangle = \langle x, S^*Sx \rangle = \langle x, S^2x \rangle$$

By the Cauchy-Schwarz inequality the last term is $\leq ||x| \cdot ||S^2||$. Since (i) holds for every x of norm one we conclude that

$$||S||^2 \leq ||S^2||$$

Now (1) follows from the multiplicative inequality for operator norms. Next, by induction over n we get the equalities

$$||S||^{2n} = ||S^n||^2 \quad : \ n \geq 1$$

Taking the n:th root and passing to the limit the spectral radius formula gives

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

It follows that if $\mathbf{S}$is the closed subalgebra of Y generated by S and the identity, then it becomes a sup-norm algebra, i.e. isometric to a closed subalgebra of $C^0(\sigma(S))$. We can say more because one has:

8.1 Theorem. *The spectrum of a bounded self-adjoint operator is a compact real interval.*

Proof. If $\mathfrak{Im}(\lambda \neq 0$ there cannot exists a non-zero vector x in $\mathcal{H}$ such that

$$Sx = \lambda \cdot x$$

For this would give

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

which cannot hold when $\lambda \neq \bar{\lambda}$. So when $\mathfrak{Im}(\lambda) \neq 0$ we have an injective linear operator

$$T\colon \ x \to \lambda x - Sx$$

There remains to show that T also is surjective which means that $\lambda \cdot E - S$ is invertible and hence λ is outside $\sigma(S)$ as required. First we show that T has closed range. To obtain this we consider some x and set

$$y = \lambda x - Sx$$

It follows that

$$||y||^2 = |\lambda|^2 \cdot ||x||^2 + ||Sx|^2 + \lambda \cdot \langle x, Sx \rangle + \bar{\lambda} \cdot \langle Sx, x \rangle$$

Since S is self-adjoint we get

$$\lambda \cdot \langle x, Sx \rangle + \bar{\lambda} \cdot \langle Sx, x \rangle = 2 \cdot \mathfrak{Re}(\lambda) \cdot \langle Sx, x \rangle$$

Now $|\langle Sx, x \rangle| \leq ||Sx|| \cdot ||x||$ so the triangle inequality gives

$$||y||^2 \geq |\lambda|^2 \cdot ||x||^2 + ||Sx|^2 - 2|\mathfrak{Re}(\lambda)|| \cdot ||Sx|| \cdot ||x||$$

From this the reader easily shows that

$$||y||^2 \geq \mathfrak{Im}(\lambda)^2 \cdot ||x||^2$$

So we have proved that

$$(*) \qquad |\lambda \cdot x - Sx|| \geq |\mathfrak{Im}(\lambda)| \cdot ||x||$$

This implies that T has a closed range. To prove surjectivity it suffices to show that the orthogonal complement of $T(\mathcal{H})$ is zero. To see this we suppose that y is a vector such that

$$\langle \lambda \cdot x - Sx, y \rangle = 0$$

for all x. Since S is self-adjoint it follows that

$$\langle \lambda \cdot x, y \rangle = \langle x, Sy \rangle$$

This holds for every x and therefore $Sy = \lambda \cdot y$. But we have already seen that this gives $y = 0$ and Theorem XX is proved.

A consequence. Theorem 8.1 together with the general result from 7.XX gives the following:

8.2 Theorem. *Let S be a self-adjoint operator. Then the closed subalgebra of $L(\mathcal{H}, \mathcal{H})$ generated by S is a sup-norm algebra which is isomorphic to $C^0(\sigma(S))$.*

8.3 Normal operators.

A bounded linear operator A is normal if it commutes with its adjoint A^*.

Exercise. Let A be a normal operator. Show that the operator A^*A is self-adjoint. The hint is to use the general equality:

$$B^*S^* = (SB)^*$$

for an arbitrary pair of linear operators.

Next, let A be normal and set $S = A^*A$ which is self-adjoint by the exercise. It follows that $S^2 = A^2(A^*)^2$ and the multiplicative inequality for operator norms gives:

$$(1) \qquad ||S^2|| \leq ||A^2|| \cdot ||(A^*)^2|| = ||A^2||^2$$

where the last equality follows since the norm of an operator is equal to the norm of its adjoint. Next, since S is self-adjoint we have already proved that

$$(2) \qquad ||S^2|| = ||S||^2 = ||AA^*||^2 = ||A||^4$$

where the last equality follows from the general identity

$$||T||^2 = ||T^*T||$$

when T is an arbitrary operator on $\mathcal{H}$. From (1-2) we conclude that

$$||A||^2 = ||A^2||$$

We can take higher powers and exactly as in XX the spectral radius formula gives the equality:

$$(*) \qquad ||A|| = \max_{z \in \sigma(A)} |z|$$

Since every polynomial in A again is a normal operator for which (*) holds we have proved the following:

8.4 Theorem *Let A be a normal operator. Then the closed subalgebra* **A** *generated by A in $L(\mathcal{H}, \mathcal{H})$ is a sup-norm algebra.*

Remark. The spectrum $\sigma(A)$ is some compact subset of **C**. In general we cannot affirm that $\mathcal{A}(\sigma(A)) = C^0(\sigma(A))$. To overcome this we shall also use the adjoint operator A^* and consider the closed subalgebra of $L(\mathcal{H}, \mathcal{H})$ which is generated by A and A^*. Notice that every polynomial in A and A^* again is a normal operator and it is celar that if a sequenceof normal operators converge in the operator norm then the limit is again a normal perator. So if $\mathcal{B}$ is the closed subalgebra of $L(\mathcal{H}, \mathcal{H})$ then every operator in B is normal. We conclude as above that $\mathcal{B}$ is a sup-norm algebra. There remains to prove the following concluive result:

8.5 Theorem. *The sup-norm algebra $\mathcal{B}$ is via the Gelfand transform isomorphic with $C^0(\mathfrak{M}_{\mathcal{B}})$.*

Proof. If $S \in \mathcal{B}$ is self-adjoint then we know from the previous section that its Gelfand transform is real-valued. Next, let $Q \in \mathcal{B}$ be arbitrary. Now $S = Q + Q^*$ is self-adjoint. So if $p \in \mathfrak{M}_{\mathcal{B}}$ it first follows that $\hat{Q}(p) + \hat{Q}^*(p)$ is real, i.e. with $\hat{Q}(p) = a + ib$ we must have $\hat{Q}^* = a_1 - ib$ for some real number a_1. But now QQ^* is also sefl.-adjoint and hence $(a + ib)(a - 1 - ib)$ is real. This gives $a = a_1$ and hence we have proved that the Gelfand transform of Q^* is the complex conjugate function of $\hat{Q}$. Hence the Gelfand transforms of $\mathcal{B}$-elements is a self-adjoint algebra and the theorem follows from the general fact that a self-adjoint and point separating sup-norm algebra on a compact space X is equal to $C^0(X)$.

Remark. Since $\hat{A}^*$ is the complex conjugate function of $\hat{A}$ it follows that $\hat{A}$ alone separates points on $\mathfrak{M}_{\mathcal{B}}$. We conclude that the Gelfandspace of $\mathcal{B}$ can be identified with $\sigma(A)$.

8.6 Spectral measures.

Given $\mathcal{B}$ and $\sigma(A)$ as above we can construct certain Riesz measures on $\sigma(A)$. Namely, let x, y be a pair of vectors in $\mathcal{H}$. Now we get a linear functional on the Banach space $\mathcal{B}$ defined by

$$T \mapsto \langle Tx, y \rangle$$

The Riesz representation formula gives a *unique* Riesz measure $\mu_{x,y}$ on $\sigma(A)$ such that

$$(*) \qquad \langle Tx, y \rangle = \int \hat{T}(z) \cdot d\mu_{x,y}(z)$$

hold for every $T \in \mathcal{B}$. Since $\hat{A}(z) = z$ is the identity function we have in particular

$$\langle Ax, y \rangle = \int z \cdot d\mu_{x,y}(z)$$

Similarly we get

$$\langle A^*x, y \rangle = \int \bar{z} \cdot d\mu_{x,y}(z)$$

8.7 The operators $E(\delta)$. Notice that (*) implies that the map from $\mathcal{H} \times \mathcal{H}$ into the space of Riesz measures on $\sigma(A)$ is bi-linear. We have for example:

$$\mu_{x_1+x_2,y} = \mu_{x_1,y} + \mu_{x_2,y}$$

Moreover, since $\mathcal{B}$ is a sup-norm algebra it follows from (*) that one has the inequality

$$||\mu_{x,y}|| \leq \max, |\langle Tx, y \rangle|$$

Here $||\nu_{x,y}||$ is the total variation of the complex-valued Riesz measure and the maximum is taken over all $T \in \mathcal{B}$ with operator norm ≤ 1. It follows that

$$(*) \qquad ||\mu_{x,y}|| \leq ||x|| \cdot ||y||$$

Next, let δ be a Borel subset of $\sigma(A)$. Keeping y fixed in $\mathcal{H}$ we obtain a linear functional on $\mathcal{H}$ defined by

$$x \mapsto \int_\delta d\mu_{x,y}(z) = \mu_{x,y}(\delta)$$

By (*) it has norm $\leq ||y||$ and itis represented by a vector $E(\delta)x$ in $\mathcal{H}$. More precisely we have

$$\langle E(\delta)x, y \rangle = \int_\delta d\mu_{x,y}(z) = \mu_{x,y}(\delta)$$

Finally, using the additivity in (xx) once more we see that

$$x \mapsto E(\delta)(x)$$

is linear and hence we obtain the linear operator $E(\delta)$.

Exercise. Show that $E(\delta)$ commutes with all operators in $\mathcal{B}$ and that it is a self-adjoint projection, i.e.

$$E(\delta)^2 = E\delta) \quad \text{and} \quad E(\delta)^* = E(\delta)$$

Show also that the spectrum of this linear operator is contained in the closure of the Borel set δ. Finally, show that

$$E(\delta_1 \cap \delta_2) = E(\delta_1)E(\delta_2)$$

holds for every pair of Borel subsets and if we take $\delta = \sigma(A)$ we get the identity operator.

8.8 Resolution of the identity. The self-adjoint projection operators above enable us to decompose the identity on $\mathcal{H}$. Namely, if $\delta_1, \ldots, \delta_N$ is any finite family of disjoint Borel sets whose union is $\sigma(A)$ then

$$1 = E(\delta_1) + \ldots + E(\delta_N)$$

At the same tine we get a decomposition of the operator A, i.e.

$$A = A_1 + \ldots + A_N \quad \text{where} \quad A_k = E(\delta_k) \cdot A$$

For each k the spectrum $\sigma(A_k)$is contained in the closure of δ_k. So the normal operator is represented by a sum of normal operators where the individual operator has a small spectrum when the δ-partition is fine.

9. Unbounded self-adjoint operators.

First we prove some general results about densely defined linear operators. A linear operator T on $\mathcal{H}$ is densely defined if there exists a dense subspace $\mathcal{D}(T)$ on which T is defined, i.e. to every $x \in \mathcal{D}(T)$ we get an image vector Tx. For the moment no further assumption is imposed on T. In particular it may be unbounded, i.e. if $\Sigma(T) = \mathcal{D}(T) \cap \Sigma$ where Σ is the unit ball in $\mathcal{H}$ then it can occur that

$$\max_{x \in \Sigma(T)} ||Tx|| = +\infty$$

9.1 Constructions of graphs. The product $\mathcal{H} \times \mathcal{H}$ is a Hilbert space whose inner product is defined by

$$\langle (x, y), (x_1, y_1) \rangle = \langle x, x_1 \rangle + \langle y, y_1 \rangle$$

If T is densely defined we set

$$\Gamma(T) = \{(x, Tx) \quad : \, x \in \mathcal{D}(T)\}$$

This graph is a subspace of $\mathcal{H} \times \mathcal{H}$. Its closure consists of points (x_*, y_*) for which there exists a sequence $\{x_n\}$ in $\mathcal{D}(T)$ such that

$$\lim x_n = x_* \quad \text{and} \quad \lim Tx_n = y_*$$

It is an easy exercise to verify that we obtain a linear operator T_c whose graph is the closure of $\mathcal{D}(T)$. Thus, $\mathcal{D}(T_c)$ consists of all x_* for which a limit as above exists. So T_c is a extension of T in the sense that $\Gamma(T) \subset \Gamma(T_c)$. In this way the study of densely defined linear operator is essentially reduced to operators with a closed graph.

9.2 Inverse operators. Let T be a densely defined operator. We do not assume that it has a closed graph. We say that T is injective if $Tx \neq Ty$ when $x \neq y$ and both x and y belong to $\mathcal{D}(T)$. Assume this and suppose also that the range $T(\mathcal{D}(T))$ is a dense subspace of $\mathcal{H}$. Then we define the inverse operator T^{-1} where $\mathcal{D}(T^{-1}) = T(\mathcal{D}(T))$ and

$$Tx = y \implies T^{-1}y = x \quad : \, x \in \mathcal{D}(T)$$

9.3 A useful graph map. On $\mathcal{H} \times \mathcal{H}$ there exists the isometry defined by

$$\mathcal{A}_1(x, y) = (y, x)$$

The construction of T^{-1} gives the equality

(*) $$\mathcal{A}_1(\Gamma(T)) = \Gamma(T^{-1})$$

Exercise. Prove (*) and conclude that if T has a closed graph so has T^{-1}.

9.4 Adjoint operators. Let T be a densely defined operator. Given a vector $y \in \mathcal{H}$ we define a linear functional on $D(T)$ by

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

Suppose there exists a constant $C(y)$ such that

(i) $$|\langle Tx, y \rangle| \leq C(y) \cdot ||x|| \quad \text{for all} \quad x \in \mathcal{D}(T)$$

This densely defined linear functional has a unique extension to $\mathcal{H}$ and since a Hilbert space is self-dual there exists a unique vector y^* such that

(ii) $$\langle Tx, y \rangle| = \langle x, y^* \rangle \quad \text{for all} \quad x \in \mathcal{D}(T)$$

The set of all y for which a constant $C(y)$ exists is a subspace of $\mathcal{H}$ which we for the moment denote by $\mathcal{H}_*$. It is clear that the map

$$y \mapsto y^*$$

gives a linear operator from $\mathcal{H}_*$ into $\mathcal{H}$. It is denoted by T^* and is called the adjoint of T. So here $\mathcal{D}(T^*) = \mathcal{H}_*$ holds.

9.5 Another graph equality. On $\mathcal{H} \times \mathcal{H}$ we have the isometry defined by

$$\mathcal{A}_2(x, y) = (y, -x)$$

Exercise. Let T be densely defined. Show that

$$(*) \qquad \Gamma(T^*) = \big[\mathcal{A}_2(\Gamma(T))\big]^\perp$$

In other words, the graph of T^* is the orthogonal complement of $\mathcal{A}_2(\Gamma(T))$. We remark that this equality holds in general, i.e. even if $\mathcal{D}(T^*)$ is not dense. Since the orthogonal complement of an arbirary subspace of a Hilbert space is closed, it follows from (*) that an adjoint operator T^* always has a closed graph.

Next, assume that T is such that T^* also is densely defined. Hence we can construct its inverse $(T^*)^{-1}$. We have also the operator T^{-1} and again we assume that it is densely defined which is equivalent to the condition that the range of T is a dense subspace of $\mathcal{H}$. Now we also get the adjoint operator $(T^{-1})^*$ and with these notations one has

9.6 Theorem. *One has the equality*

$$(T^{-1})^* = (T^*)^{-1}$$

Proof. We must prove that the two operators have the same graph, To get the equality we use the two $\mathcal{A}$-operators. First

$$(1) \qquad \Gamma((T^*)^{-1}) = \mathcal{A}_1(\Gamma(T^*)) = \mathcal{A}_1\big(\big[\mathcal{A}_2(\Gamma(T))\big]^\perp\big)$$

Since $\mathcal{A}_1$ is an isometry the last term is equal to

$$(2) \qquad \big[\mathcal{A}_1\big(\mathcal{A}_2(\Gamma(T))\big]^\perp$$

Next, we notice that the composed operator $\mathcal{A}_1 \circ \mathcal{A}_2 = -\mathcal{A}_2 \circ \mathcal{A}_1$ and while we regard images of subspoaces in $\mathcal{H} \times \mathcal{H}$ the sign does not matter. So (2) becomes

$$(3) \qquad \big[\mathcal{A}_2\big(\mathcal{A}_1(\Gamma(T))\big]^\perp = \big[\mathcal{A}_2\big(\Gamma(T^{-1})\big)\big]^\perp$$

Finally, by another application of (*) from the Exercise above we see that (3) is equal to $\Gamma\big((T^{-1})^*\big)$ and Theorem 9.6 follows.

9.7 Symmetric operators

A symmetric operator is a densely defined operator T such that

$$\langle Tx, y\rangle = \langle x, Ty\rangle$$

hold for each pair x, y in $\mathcal{D}(T)$. It is easily seen that the symmetry is preserved by the operator T_c whose graph is the closure of T. So without loss of generality we consider symmetric operators with a closed range. Next, the symmetry obviously implies that the adjoint T^* is an extension of T, i.e. one has the inclusion

$$(*) \qquad \Gamma(T) \subset \Gamma(T^*)$$

Now we shall find a condition in order that equality holds in (*).

The spaces $\mathcal{D}_+$ and $\mathcal{D}_-$. Let T be a symmetric operator and use its adjoint to define the following two eigenspaces:

$$\mathfrak{D}_+ = \{x \in \mathcal{D}(T^*) \quad : T^*x = ix\} \quad \text{and} \quad \mathfrak{D}_- = \{x \in \mathcal{D}(T^*) \quad : T^*x = -ix\}$$

With these notations one has

9.8 Theorem. *If* $\mathfrak{D}_+ = \mathfrak{D}_- = 0$ *it follows that* $T_c = T^*$

Remark. So when the two $\mathfrak{D}$-spaces are zero we obtain the natural self-adjoint extension of T given by its closure T_c. However, this is not the only case when T has a self-adjoint extension. Namely, the following more general existence result holds:

9.9 Theorem. *Assume that the two linear spaces $\mathfrak{D}_+$ and $\mathfrak{D}_-$ are finite dimensional complex vector spaces of the same dimension. Then the symmetric operator T has a self-adjoint extension.*

We refer to section XX for the proof of this result. It is illustrated by an example below. But the reader who is content with the case in Theorem 9.8 can proceed directly to its proof.

9.10 Example. Let us give an example of a symmetric operator T which has a self-adjoint extension but not given by T_c. Let $\mathcal{H}$ be the Hilbert space $L^2[0,1]$, i.e. the elements are square-integrable functions on the unit interval $[0,1]$ where the coordinate is denoted by t. A dense subspace $\mathcal{H}_*$ consists of functions $f(t) \in C^1[0,1]$ such that $f(0) = f(1) = 0$. On this dense subspace we define the operator T by

$$T(f) = if'(t)$$

A partial integration gives

$$\langle T(f), g\rangle = i\int_0^1 f'(t)\cdot \bar{g}(t)\cdot dt = \int_0^1 \bar{g}'(t)\cdot f(t)dt = \langle f, T(g)\rangle$$

Hence T is symmetric. Next, an L^2-function h belongs to $\mathcal{D}(T^*)$ if and only if there exists a constant $C(h)$ such that

$$\Big|\int_0^1 if'(t)\cdot \bar{h}(t)dt\Big| \leq C(h)\cdot ||f||_2$$

hold for all in $f \in \mathcal{H}_*$. By elementary distribution theory this means that $\mathcal{D}(T^*)$ consists of all L^2-functions h for which the distribution derivative $\frac{dh}{dt}$ again belongs to L^2. Let us then consider the operator T^*. Notice that $\mathcal{D}(T^*)$ contains *all* C^1-functions f, i.e. with no constraint upon the end-values $f(0)$ and $f(1)$. For such pair f, g a partial integration gives

$$\langle T^*(f), g)\rangle - \langle f, T^*(g)\rangle = i\cdot(f(1)\bar{g}(1) - f(0)\bar{g}(0))$$

Hence the left hand side can be $\neq 0$, i.e. choose for example $f(t) = g(t) = t$. Next, we notice that

$$\mathfrak{D}_+ = \{h \in L^2 \quad : \frac{dh}{dt} = h\}$$

This is a 1-dimensional vector space generated by the exponential function e^t. Similarly

$$\mathfrak{D}_+. = \{h \in L^2 \quad : \frac{dh}{dt} = -h\}$$

is 1-dimensional and generated by e^{-t}.

The self-adjoint extension of T. Let $\bar{T}$ be the closure of T. By XX it is again a symmetric operator. Next, consider the exponential function e^t. It belongs to $\mathcal{D}(T^*)$ and satisfies

$$T^*(e^t) = i\cdot e^t$$

Thus, e^t belongs to $\mathfrak{D}_+$. The reader may verify that e^t does not belong to $\mathcal{D}(\bar{T})$. So we get a new subspace of $\mathcal{H}$ generated $\mathcal{D}(\bar{T})$ and e^t. On this dense subspace we define the linear operator

$$S(f + ce^t) = \bar{T}(f) + ice^t$$

when $f \in \mathcal{D}(\bar{T})$ and c is a complex constant.

Exercise Prove that S is symmetric and that $S = S^*$, i.e. S gives a self-adjoint extension of T.

Proof of Theorem 9.8

Recall that $\Gamma(T^*)$ is a closed subspace of $\mathcal{H}\times\mathcal{H}$. It follows that $\mathcal{D}(T^*)$ is equipped with a complete inner product defined by

$$(1) \qquad \{x, y\} = \langle x, y\rangle + \langle T^*x, T^*y\rangle$$

defined for pairs x, y in $\mathcal{D}(T^*)$. Since T^* is an extension of T, the graph $\Gamma(T)$ appears as a closed subspace of $\Gamma(T^*)$ which via (1) is identified with a closed subspace of $\mathcal{D}(T^*)$. To prove the equality in Theorem 9.8 suffices to show that the orthogonal complement of $\mathcal{D}(T)$ is zero. Suppose that some $\xi \in \mathcal{D}(T^*)$ is $\perp$ to $\mathcal{D}(T)$. This means that

$$\text{(i)} \qquad \langle \xi, x\rangle + \langle T^*\xi, Tx\rangle = 0 \quad \text{for all} \quad x \in \mathcal{D}(T)$$

From this it is clear that $\xi \in \mathcal{D}(T)$ which gives

$$0 = \langle \xi, x \rangle + \langle T\xi, Tx \rangle = \langle \xi, x \rangle + \langle T^2\xi, x \rangle$$

where the last equality holds by the symmetry of T. Since $\mathcal{D}(T)$ is dense it follows that

(ii) $$0 = T^2(\xi) + \xi = (T + iE)(T - iE)(\xi) = 0$$

Now the hypothesis that $\mathcal{D}_+ = \mathcal{D}_- = 0$ give $\xi = 0$ and Theorem 9.8 is proved.

9.11 Resolvents of self-adjoint operators.

Let A be a densely defined self-adjoint operator. If $x \in \mathcal{D}(A)$ we get the vector $y = ix - Ax$. Then we obtain

$$||y||^2 = ||x||^2 + ||Ax||^2 - i\langle x, Ax \rangle - \langle Ax, ix \rangle$$

Here we notice that

$$-\langle Ax, ix \rangle = i\langle Ax, x \rangle = i\langle x, Ax \rangle$$

where the last equality holds since A is symmetric. We conclude that

(*) $$||ix - Ax||^2 = ||x||^2 + ||Ax||^2 \quad \text{when} \quad x \in \mathcal{D}(A)$$

9.12 Proposition. *By $x \mapsto ix - Ax$ we get a bijective linear map from $\mathcal{D}(A)$ onto $\mathcal{H}$.*

Proof. Let ρ denote this map. By XX above it is injective. To prove surjectivity we set $Y = \rho(\mathcal{D}(A))$. First we show that the orthogonal complement $Y^\perp = 0$ which means that Y is a dense subspace of $\mathcal{H}$. To see this we consider some vector $\xi \in \mathcal{H}$ such that

(i) $$\langle \xi, i \cdot x - Ax \rangle = 0 \quad \text{for all} \quad x \in \mathcal{D}(A)$$

This implies that the linear functional on $\mathcal{D}(A)$ defined by

$$x \mapsto \langle \xi, Ax \rangle = \langle \xi, i \cdot x \rangle$$

is bounded, i.e. we see that $C(\xi) \leq ||\xi||$. So by definition ξ belongs to $\mathcal{D}(A^*)$ and since $A = A^*$ we have $\xi \in \mathcal{D}(A)$. Then the symmetry of A and (i) give:

$$\langle A\xi, x \rangle = \langle \xi, i \cdot x \rangle = -i \cdot \langle \xi, x \rangle$$

This hold for all x in the dense space $\mathcal{D}(A)$ which gives $A(\xi) = -i \cdot \xi$. But this contradicts the result in XX and hence $Y^\perp = 0$. There remains to show that Y is closed. But this follows easily from (*) above. For if $\{x_n\}$ is a sequence in $\mathcal{D}(A)$ and $y_n = ix_n - A(x_n)$ converge to some y_* then (*) entails

$$||x_n - x_m||^2 \leq ||y_n - y_m||^2$$

for all pair n, m. Since $\{y_n\}$ by hypothesis is a convergent sequence it is a Cauchy sequence and hence $\{x_n\}$is also a Cauchy sequence. Therefore $x_n \to x_*$ hold for some $x_* \in \mathcal{H}$ and at this stage the reader may verify that x_* belongs to $\mathcal{D}(A)$ and that $y_* = \rho(x_*)$.

9.13 The operator R. Since the ρ-image is $\mathcal{H}$ we get a linear operator R defined on the whole Hilbert space such that

$$(iE - A) \circ R(x) = x \quad \text{for all} \quad x \in \mathcal{D}(A)$$

Moreover, by the inequality (*) it follows that R is a bounded linear operator whose operator norm is ≤ 1 and we notice that the range

(1) $$R(\mathcal{H}) = \mathcal{D}(A)$$

Next, from the proof of Proposition 9.12 it is clear that the densely defined operator $iE + A$ has a bounded inverse which we deonte by S. So here

$$(iE + A) \circ S(x) = x \quad \text{for all} \quad x \in \mathcal{D}(A)$$

9.14 An adjoint formula. Above R is the inverse of the densely defined operator $iE - A$. Since A is self-adjoint we have

$$(iE - A)^* = -iE - A = -(iE + A)$$

Now $-S$ is the inverse operator of $-(iE + A)$ and hence Theorem XX gives

(*) $$R^* = -S$$

Using this equality we can prove the following:

9.15 Proposition. *The operator R is normal.*

Exercise. Prove this result where the hint is to use the equality (*) above.

9.16 The spectrum of S. Above we have found the normal and bounded operator R. We also get the normal operator S and from now on we prefer to work with S instead of R and establish the following inclusion for the spectrum $\sigma(S)$:

9.17 Theorem. *The spectrum $\sigma(S)$ contains 0 and is otherwise contained in the set*

$$\Sigma = \{\frac{1}{a+i} \quad : a \in \mathbf{R}\}$$

Proof. Since S is the inverse of $iE + A$ it follows from XX that

$$\Gamma(S) = \{(ix + Ax, x) \quad : x \in \mathcal{D}(A)\}$$

So if λ is a non-zero complex number we get

(1) $$\Gamma(\lambda \cdot E - S) = \{(ix + Ax, -x + \lambda(ix + Ax)) \quad : x \in \mathcal{D}(A)\}$$

Suppose now that λ is outside the set Σ We must show that that $\lambda \cdot E - S$ is invertible. First we prove that the range of $\lambda \cdot E - S$ is dense. For otherwise the formula for its graph in (1) above gives the existence of a non-zero vector y such that

(2) $$\langle -x + \lambda(ix + Ax), y\rangle = 0 \quad : x \in \mathcal{D}(A)$$

Since $A = A^*$ and $\lambda \neq 0$ hold, it is clear that (2) implies that y belongs to $\mathcal{D}(A)$. Now $\langle Ax, y\rangle = \langle x, Ay\rangle$ hold for all $x \in \mathcal{D}(A)$ and hence (2) gives the equality

(3) $$\frac{1 - i\lambda}{\lambda}\langle x, y\rangle = \langle x, Ay\rangle$$

If we set $\mu = \frac{1-i\lambda}{\lambda}$ the density of $\mathcal{D}(A)$ implies that

(4) $$Ay = \bar{\mu} \cdot y$$

By the result in (xx) this is only possible if $\mu = a$ is real and this entails that $\lambda = \frac{1}{a+i}$ which contradicts the hypothesis that λ is outside Σ. Hence the range of $\lambda \cdot E - S$ is dense. To finish the proof we consider the vector:

(5) $$\xi(x) = -x + \lambda(ix + Ax) = \lambda \cdot (\frac{i\lambda - 1}{\lambda} - Ax)$$

Next, put

$$\frac{i\lambda - 1}{\lambda} = a + ib$$

Notice that $b \neq 0$ since λ is outside Σ. Now we get

(6) $$||\xi(x)||^2 = |\lambda|^2 \cdot ||ibx + ax - Ax||^2$$

Next, $aE - A$ is self-adjoint which by XX gives the equality

$$||ibx + ax - Ax||^2 = b^2||x||^2 + ||ax - Ax||^2$$

Next, with the notations above we notice that $\lambda \cdot E - S = ix + Ax \mapsto \xi(x)$ So the required invertibility of $\lambda \cdot E - S$ follows if we can find a constant M such that

(*) $$||x||^2 + ||Ax||^2 = ||x + Ax||^2 \leq M \cdot ||\xi(x)||^2$$

The existence of a constant M follows easily because we have already seen that

$$||\xi(x)||^2 = |\lambda|^2 \cdot \left[b^2||x||^2 + ||ax - Ax||^2\right]$$

This finishes the proof of Theorem 9.17.

10. Commutative Banch algebras

Contents

Introduction

Consider a complex Banach space B equipped with a commutative product such that the norm satisfies the multiplicative inequality

(*) $$||xy|| \leq ||x|| \cdot ||y|| \quad : \; x, y \in B$$

We also assume that B has a multiplicative unit element e where $ex = xe$ hold for all $x \in B$ and $||e|| = 1$. When this holds we refer to B as a commuative Banach algebra with a multiplicative unit. A **C**-linear form λ on B is called multiplicative if:

(**) $$\lambda(xy) = \lambda(x) \cdot \lambda(y) \quad \text{for all pairs} \quad x, y \in B$$

When λ satisfies (**) and is not identically zero it is clear that $\lambda(e) = 1$ must hold.

0.1 Theorem. *Every multiplicative functional λ on B is automatically continuous, i.e. an element in the normed dual space B^* and its norm is equal to one.*

The proof in A.1 below uses analytic function theory via a study of certain Neumann series. The crucial point is that when $x \in B$ has a norm strictly less than one, then $e - x$ is invertible in B whose inverse is given by the B-valued power series

(1) $$(e - x)^{-1} = e + x + x^2 + \ldots$$

The spectral radius formula. Given $x \in B$ we can take its powers and for each n set

$$\rho_n(x) = ||x^n||^{\frac{1}{n}}$$

In XX we show that these ρ-numbers have a limit as $n \to \infty$, i.e. there exists

$$\rho(x) = \lim_{n \to \infty} \rho_n(x)$$

Using Hadamard's formula for the radius of convergence of power series we prove the following in XX:

0.2 Theorem. For each $x \in B$ one has the equality

$$\rho(x) = \max_{\lambda \in \mathcal{M}(B)} |\lambda(x)|$$

where $\mathcal{M}(B)$ denotes the set of all multiplicative functionals on B.

0.3 The Gelfand transform. Keeping an element $x \in B$ fixed we get the complex-valued function on $\mathcal{M}(B)$ defined by:

$$\lambda \mapsto \lambda(x)$$

The resulting function is denoted by $\widehat{x}$ and called the Gelfand transform. Since $\mathcal{M}(B)$ is a subset of the dual space B^* it is equipped with the weak-star topology which is called the Gelfand topology. By definition this is the weakest topology on $\mathcal{M}(B)$ for which every Gelfand transform

$\widehat{x}$ becomes a continuous function. In particular there exists an algebra homomorphism from B into the commutative algebra $C^0(\mathcal{M}(B))$:

(*) $$x \mapsto \widehat{x}$$

0.4 Semi-simple algebras. The spectral radius formula shows that $\widehat{x}$ is the zero function if and only if $\rho(x) = 0$. One says that the Banach algebra B is *semi-simple* if (*) is injective. An equivalent condition is that

$$0 \neq x \implies \rho(x) > 0$$

0.5 Uniform algebras. If B is semi-simple the Gelfand transform identifies B with a subalgebra of $C^0(\mathcal{M}(B))$. In general this subalgebra is not closed. The reason is that there can exist B-elements of norm one while the ρ-numbers can be arbitrarily small. If the equality below holds for every $x \in B$:

(*) $$||x|| = \rho(x) = |\widehat{x}|_{\mathcal{M}(B)}$$

one says that B is a uniform algebra.

Remark. Multiplicative functionals on specific Banach algebras were used by Norbert Wiener and Arne Beurling where the focus was on Banach algebras which arise via the *Fourier transform.* Later Gelfand, Shilov and Raikow established the abstract theory which has the merit that it applies to quite general situations such as Banach algebras generated by linear operators on a normed space. Moreover, Shilov applied results from the theory of analytic functions in several complex to construct *joint spectra* of several elements in a commutative Banach algebra. See [Ge-Raikov-Shilov] for a study of commutative Banach algebras which include results about joint spectra. One should also mention the work by J. Taylor who used integral formulas in several complex variables to analyze the topology of Gelfand spaces which arise from the Banach algebra of Riesz measures with total bounded variation on the real line, and more generally on arbitrary locally compact abelian groups.

A. Neumann series and resolvents

Let B be a commutative Banach algebra whose identity element is denoted by e. The set of elements x whose norms have absolute value < 1 is denoted by $\mathfrak{B}$ and called the open unit ball in B.

A.1 Neumann series. Let us prove that $e-x$ is invertible for every $x \in \mathfrak{B}$. We have $||x|| = \delta < 1$ and the multiplicative inequality for the norm gives:

(1) $$||x^n|| \leq ||x||^n = \delta^n \quad : \quad n = 1, 2, \ldots$$

If $N \geq 1$ we set:

(2) $$S_N(x) = e + x + \ldots + x^N$$

For each pair $M > N$ the triangle inequality for norms gives:

(3) $$||S_M(x) - S_N(x)|| \leq ||x^{N+1}|| + \ldots + ||x^M|| \leq \delta^{N+1} + \ldots + \delta^M$$

It follows that

$$||S_M(x) - S_N(x)|| \leq \frac{\delta^{N+1}}{1-\delta} \quad : M > N \geq 1$$

Hence $\{S_N(x)\}$ is a Cauchy sequence and is therefore convergent in the Banach space. For each $N \geq 1$ we notice that

$$(e-x)S_N(x) = e - x^{N+1}$$

Since $x^{N+1} \to 0$ we conclude that if $S_*(x)$ is the limit of $\{S_N(x)\}$ then

(*) $$(e-x)S_*(x) = e$$

This proves that $e-x$ is an invertible element in B whose inverse is the convergent B-valued series

(**) $$S_*(x) = e + \sum_{k=1}^\infty x^k$$

We refer to (**) as the Neumann series of x. More generally, let $0 \neq x \in B$ and consider some λ such that $|\lambda| > ||x||$. Now $\lambda^{-1} \cdot x \in \mathfrak{B}$ and from (**) we conclude that $\lambda \cdot e - x = \lambda(e - \lambda^{-1} \cdot x)$ is invertible where

(***) $$(\lambda \cdot e - x)^{-1} = \lambda^{-1} \cdot \big[\, e + \sum_{k=1}^\infty \lambda^{-k} \cdot x^k \,\big]$$

Exercise. Deduce from (***) that one has the inequality

$$||(\lambda \cdot e - x)^{-1}|| \leq \frac{1}{|\lambda| - ||x||}$$

A.2. Local Neumann series expansions. To each $x \in B$ we define the set

(*) $$\gamma(x) = \{\lambda :\ e - x \quad \text{is invertible}\}$$

Let $\lambda_0 \in \gamma(x)$ and put

(1) $$\delta = ||(\lambda_0 \cdot e - x)^{-1}||$$

To each complex number λ we set

(2) $$y(\lambda) = (\lambda_0 - \lambda) \cdot (\lambda_0 \cdot e - x)^{-1}$$

If $|\lambda - \lambda_0| < \delta$ we see that $y(\lambda) \in \mathfrak{B}$ and hence $e - y(\lambda)$ is invertible with an inverse given by the Neumann series:

(3) $$(e - y(\lambda))^{-1} = e + \sum_{\nu=1}^\infty (\lambda_0 - \lambda)^\nu \cdot (\lambda_0 \cdot e - x)^{-\nu}$$

Next, for each complex number λ we notice that

$$(\lambda \cdot e - x) \cdot (\lambda_0 \cdot e - x)^{-1} =$$
$$\big[(\lambda_0 \cdot e - x) + (\lambda - \lambda_0) \cdot e\big](\lambda_0 \cdot e - x)^{-1} = e - y(\lambda) \implies$$

(4) $$(\lambda \cdot e - x) = (\lambda_0 \cdot e - x)^{-1} \cdot (e - y(\lambda))$$

So if $|\lambda - \lambda_0| < \delta$ it follows that $(\lambda \cdot e - x)$ is a product of two invertible elements and hence invertible. Moreover, the series expansion from (3) gives:

(**) $$(\lambda \cdot e - x)^{-1} = (\lambda_0 \cdot e - x) \cdot [\, e + \sum_{\nu=1}^\infty (\lambda_0 - \lambda)^\nu \cdot (\lambda_0 \cdot e - x)^{-\nu} \,]$$

We refer to (**) as a local Neumann series. The triangle inequality gives the norm inequality:

$$||(\lambda \cdot e - x)^{-1}|| \leq ||(\lambda_0 \cdot e - x)|| \cdot \big[1 + \sum_{\nu=1}^\infty |\lambda - \lambda_0|^\nu \cdot \delta^\nu \,\big] =$$

(***) $$||(\lambda_0 \cdot e - x)|| \cdot \frac{1}{1 - |\lambda - \lambda_0| \cdot \delta}$$

A.3. The analytic function $R_x(\lambda)$. From the above we see that $\gamma(x)$ is an open subset of $\mathbf{C}$. Let us put:

$$R_x(\lambda) = (\lambda \cdot e - x)^{-1} \quad :\ \lambda \in \gamma(x)$$

The local Neumann series (**) shows that $\lambda \mapsto R(\lambda)$ is a B-valued analytic function in the open set $\gamma(x)$. We use this analyticity to prove:

A.4 Theorem. *The set* $\mathbf{C} \setminus \gamma(x) \neq \emptyset$.

Proof. If $\gamma(x)$ is the whole complex plane the function $R_x(\lambda)$ is entire. When $\lambda| > ||x||$ we have seen that the norm of $R_x(\lambda)$ is $\leq \frac{1}{|\lambda|-||x||}$ which tends to zero as $\lambda \to \infty$. So if ξ is an element in the dual space B^* then the entire function

$$\lambda \mapsto \xi(R_x(\lambda))$$

is bounded and tends to zero and hence identically zero by the Liouville theorem for entire functions. This would hold for every $\xi \in B^*$ which clearly is impossible and hence $\gamma(x)$ cannot be the whole complex plane.

A.5 Definition *The complement* $\mathbf{C} \setminus \gamma(x)$ *is denoted by* $\sigma(x)$ *and called the spectrum of* x.

A.5 Exercise. Let λ_* be a point in $\sigma_B(x)$. Show the following inequality for each $\lambda \in \gamma(x)$:

$$||(\lambda \cdot e - x)^{-1}|| \geq \frac{1}{|\lambda - \lambda_*|}$$

The hint is to use local Neumann series from A.2.

B. The Gelfand transform

Put

$$\mathfrak{r}(x) = \max_{\lambda \in \sigma(x)} |\lambda| \tag{*}$$

We refer to $\mathfrak{r}(x)$ as the spectral radius of x since it is the radius of the smallest closed disc which contains $\sigma(x)$. The next result shows that the spectral radius is found via a limit of certain norms.

B.1 Theorem. *There exists the limit* $\lim_{n\to\infty} ||x^|| ^{\frac{1}{n}}$ *and it is equal to* $\mathfrak{r}(x)$.

Proof. Put

$$\xi(n) = ||x^n||^{\frac{1}{n}} \quad n \geq 1\,.$$

The multiplicative inequality for the norm gives

$$\log \xi(n+k) \leq \frac{n}{n+k} \cdot \log \xi(n) + \frac{k}{n+k} \cdot \log \xi(k) \quad \text{for all pairs } n, k \geq 1\,.$$

Using this convexity it is an easy exercise to verify that there exists the limit

$$\lim_{n\to\infty} \xi(n) = \xi_* \tag{i}$$

There remains to prove the equality

$$\xi_* = \mathfrak{r}(x)\,. \tag{ii}$$

To prove (ii) we use the Neumann series expansion for $R_x(\lambda)$. With $z = \frac{1}{\lambda}$ this gives the B-valued analytic function

$$g(z) = z \cdot e + \cdot \sum_{\nu=1}^{\infty} z^\nu \cdot x^\nu$$

which is analytic in the disc $|z| < \frac{1}{\mathfrak{rad}(x)}$. The general result about analytic functions in a Banach space from XX therefore implies that when $\epsilon > 0$ there exists a constant C_0 such that

$$||x^n|| \leq C_0 \cdot (\mathfrak{r}(x) + \epsilon)^n \quad n = 1, 2, \ldots \implies$$

$$\xi(n) \leq C_0^{\frac{1}{n}} \cdot (\mathfrak{r}(x) + \epsilon))$$

Since $C_0^{\frac{1}{\nu}} \to 1$ we conclude that

$$\limsup_{n\to\infty} \xi(n) \leq \mathfrak{r}(x) + \epsilon)$$

Since $\epsilon > 0$ is arbitrary and the limit (i) exists we get

$$\xi_* \leq \mathfrak{r}(x) \tag{iii}$$

To prove the opposite inequality we use the definition of the spectral radius which to begin with shows that the B-valued analytic function $g(z)$ above cannot converge in a disc whose radius

$$r^* > \frac{1}{\mathfrak{r}(x)}$$

Hence Hadamard's limit formula for B-valued power series in XX gives

$$\limsup_{n\to\infty} \xi(n) \geq \mathfrak{r}(x) - \epsilon \quad \text{for every } \epsilon > 0\,.$$

Since the limit in (i) exists we conclude that $\xi_* \geq \mathfrak{r}(x)$ and together with (iii) above we have proved Theorem B.1.

B.2 The Gelfand space $\mathcal{M}_B$.

Let B be a commutative Banach algebra with a unit element e. As a commutative ring we can refer to its *maximal ideals.* Thus, a maximal ideal $\mathfrak{m}$ is $\neq B$ and not contained in any strictly larger ideal. The maximality means that every non-zero element in the quotient ring $\frac{B}{\mathfrak{m}}$ is invertible, i.e. this quotient ring is a *commutative field.* Since the maximal ideal $\mathfrak{m}$ cannot contain an invertible element it follows from A.1 that

(i) $$x \in \mathfrak{m} \implies ||e - x|| \geq 1$$

Hence the closure of $\mathfrak{m}$ in the Banach space is $\neq B$. So by maximality $\mathfrak{m}$ is a *closed subspace* of B and hence there exists the Banach space $\frac{B}{\mathfrak{m}}$. Moreover, the multiplication on B induces a product on this quotient space and in this way $\frac{B}{\mathfrak{m}}$ becomes a new Banach algebra. Since $\mathfrak{m}$ is maximal this Banach algebra cannot contain any non-trivial maximal ideal which means that when ξ is any non-zero element in $\frac{B}{\mathfrak{m}}$ then the principal ideal generated by ξ must be equal to $\frac{B}{\mathfrak{m}}$. In other words, every non-zero element in $\frac{B}{\mathfrak{m}}$ is *invertible.* Using this we get the following result.

B.3 Theorem. *The Banach algebra $\frac{B}{\mathfrak{m}} = \mathbf{C}$, i.e. it is reduced to the complex field.*

Proof. Let e denote the identity in $\frac{B}{\mathfrak{m}}$. Let ξ be an element in $\frac{B}{\mathfrak{m}}$ and suppose that

(i) $$\lambda \cdot e - \xi \neq 0 \quad \text{for all} \quad \lambda \in \mathbf{C}$$

Now all non-zero elements in $\frac{B}{\mathfrak{m}}$ are invertible so (i) would entail that the spectrum of ξ is empty which contradicts Theorem 3.1. We conclude that for each element $\xi \in \frac{B}{\mathfrak{m}}$ there exists a complex number λ such that $\lambda \cdot e = \xi$. It is clear that λ is unique and that this means precisely that $\frac{B}{\mathfrak{m}}$ is a 1-dimensional complex vector space generated by e.

B.4 The continuity of multiplicative functionals. Let $\lambda \colon B \to \mathbf{C}$ be a multiplicative functional. Since $\mathbf{C}$ is a field it follows that the λ-kerenl is a maximal ideal in B and hence closed. Recall from XX that every linear functional on a Banach space whose kernel is a closed subspace is automatically in the continuous dual B^*. This proves that every multiplicative functional is continuous and as a consequence its norm in B^* is equal to one.

B.5 The Gelfand transform. Denote by $\mathcal{M}_B$ the set of all maximal ideals in B. For each $\mathfrak{m} \in \mathcal{M}_B$ we have proved that $\frac{B}{\mathfrak{m}}$ is reduced to the complex field. This enable us to construct complex-valued functions on $\mathcal{M}_B$. Namely, to each element $x \in B$ we get a complex-valued function on $\mathcal{M}_B$ defined by:

$$\hat{x}(\mathfrak{m}) = \text{ the unique complex number for which } x - \hat{x}(\mathfrak{m}) \cdot e \in \mathfrak{m}$$

One refers to $\hat{x}$ as the Gelfand transform of x. Now we can equip $\mathcal{M}_B$ with the *weakest topology* such that the functions $\hat{x}$ become continuous.

B.6 Exercise. Show that with the topology as above it follows that $\mathcal{M}_B$ becomes a *compact Hausdorff space.*

B.7 Semi-simple algebras. The definition of $\sigma(x)$ shows that this compact set is equal to the range of $\hat{x}$, i.e. one has the equality

$$(*) \qquad \sigma(x) = \hat{x}(\mathcal{M}_B)$$

Hence Theorem 4.1 gives the equality:

$$(**) \qquad \lim_{n\to\infty} ||x^n||^{\frac{1}{n}} = \max_{\mathfrak{m}} \hat{x}(\mathfrak{m}) = |\hat{x}|_{\mathcal{M}_B}$$

where the right hand side is the maximum norm of the Gelfand transform. It may occur that the spectral radius is zero which by (**) means that the Gelfand transform $\hat{x}$ is identically zero. This eventual possibility leads to:

B.8 Definition. *A Banach algebra B is called semi-simple if $\mathfrak{r}(x) > 0$ for every non-zero element.*

B.9 Remark. So when B is semi-simple then the Gelfand map $x \mapsto \widehat{x}$ from B into $C^0(\mathcal{M}_B)$ is injective. In this way B is identified with a subalgebra of all continuous and complex-valued functions on the compact Hausdorff space $\mathcal{M}_B$. Moreover one has the inequality

$$(*) \qquad |\widehat{x}|_{\mathcal{M}_B} \leq ||x||$$

It is in general strict. When equality holds one says that B is a *uniform algebra*. In this case the Gelfand transform identifies B with a closed subalgebra of $C^0(\mathcal{M}_B)$. For an extensive study of uniform algebras we refer to the books [Gamelin] and [Wermer].

C. Examples of Banach algebras.

Below we illustrate the general theory by some examples which appear in applications. Let us start with:

1. Operator algebras. Let B be a Banach space and T is a bounded linear operator on B. Together with the identity operator we construct the subalgebra of $\mathcal{L}(B)$ expressed by polynomials in T and take the closure of this polynomial T-algebra in the Banach space $\mathcal{L}(B)$. In this way we obtain a Banach algebra $\mathcal{L}(T)$ So if $S \in \mathcal{L}(T)$ then $||S||$ is the operator norm taken in $\mathcal{L}(B)$. Here the Gelfand space of $\mathcal{L}(T)$ is identified with a compact subset of $\mathbf{C}$ which is the spectrum of T denoted by $\sigma(T)$. By definition $\sigma(T)$ consists of those complex numbers λ such that the operator $\lambda \cdot E - T$ fails to be invertible in $\mathcal{L}(T)$.

1.0 Permanent spectrum. Above $\sigma(T)$ refers to the spectrum in the Banach algebra $\mathcal{L}(T)$. But it can occur that $\lambda \cdot e - T$ is an invertible linear operator on B even when $\lambda \in \sigma(T)$. To see an example we let $B = C^0(T)$ be the Banach space of continuous functions on the unit circle. Let T be the linear operator on B defined by the multiplication with z, i.e. when $f(\theta)$ is some 2π-periodic function we set

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

If λ belongs to the open unit disc we notice that for any polynomial $Q(\lambda)$ one has

$$|Q(\lambda)| \leq \max_{\theta} |Q(e^{i\theta})| = ||Q(T)||$$

It follows that the spectrum of T in $\mathcal{L}(T)$ is identified with the closed unit disc $\{|\lambda| \leq 1\}$. For example, $\lambda = 0$ belongs to this spectrum. On the other hand T is invertible as a linear operator on B where T^{-1} is the operator defined by

$$T^{-1}(f)(\theta) = e^{-i\theta} \cdot f(e^{i\theta})$$

So in this example the spectrum of T taken in the space of all continuous linear operators on B is reduced to the unit circle $\{|\lambda| = 1\}$.

In general, let B be a commutative Banach algebra which appears as a closed subalgebra of a larger Banach algebra B^*. If $x \in B$ we have its spectrum $\sigma_B(x)$ relative B and the spectrum $\sigma_{B^*}(x)$ relative the larger algebra. The following inclusion is obvious:

$$\sigma_{B^*}(x) \subset \sigma_B(x) \tag{1}$$

The example above shows that this inclusion in general is strict. However, one has the opposite inclusion

$$\partial(\sigma_B(x)) \subset \sigma_{B^*}(x) \tag{2}$$

In other words, if λ belongs to the boundary of $\sigma_B(x)$ then $\lambda \cdot e - x$ cannot be inverted in any larger Banach algebra. It means that λ is a permanent spectral value for x. The proof of (2) is given in XX using Neumann series.

2. Finitely generated Banach algebras. A Banach algebra B is finitely generated if there exists a finite subset $x_1, \ldots, x_k$ such that every B-element can be approximated in the norm by polynomials of this k-tuple. Since every multiplicative functional λ is continuous it is determined by its values on $x_1, \ldots, x_k$. It means that we have an injective map from $\mathcal{M}(B)$ into the k-dimensional complex vector space $\mathbf{C}^k$ defined by

$$\lambda \mapsto (\lambda(x_1), \ldots \lambda(x_k)) \tag{1}$$

Since the Gelfand topology is compact the image under (1) yields a compact subset of $\mathbf{C}^k$ denoted by $\sigma(x_\bullet)$. This construction was introduced by Shilov and one refers to $\sigma(x_\bullet)$ as the joint spectrum of the k-tuple $\{x_\nu\}$. It turns out that such joint spectra are special. More precisely, they are polynomially convex subsets of $\mathbf{C}^k$. Namely, let $z_1, \ldots, z_k$ be the complex coordinates in $\mathbf{C}^k$. If z_* is a point outside $\sigma(x_\bullet)$ there exists for every $\epsilon > 0$ some polynomial $Q[z_1, \ldots, z_k]$ such that $Q(z_*) = 1$ while the maximum norm of Q over $\sigma(x_\bullet)$ is ϵ. To see this one argues by a contradiction, i.e. if this fails there exists a constant M such that

$$|Q(z^*)| \leq M \cdot |Q|_{\sigma(x_\bullet)}$$

for all polynomials Q. Then the reader may verify that we obtain a multiplicative functional λ^* on B for which

$$\lambda^*(x_\nu) = z^*_\nu \quad : 1 \leq \nu \leq k$$

By definition this would entail that $z^* \in \sigma(x_\bullet)$.

Remark. Above we encounter a topic in several complex variables. In contrast to the case $n = 1$ it is not easy to describe conditions on a compact subset K of $\mathbf{C}^k$ in order that it is polynomially convex, which by definition means that whenever z^* is a point in $\mathbf{C}^k$ such that

$$|Q(z^*)| \leq |Q|_K$$

then $z^* \in K$.

3. Examples from harmonic analysis.

The measure algebra $M(\mathbf{R}^n)$. The elements are Riesz measures in $\mathbf{R}^n$ of finite total mass and the product defined by convolution. The identity is the Dirac measure at the origin. Set $B = M(\mathbf{R}^n)$. The Fourier transform identifies the n-dimensional ξ-space with a subset of $\mathcal{M}(B)$. In fact, this follows since the Fourier transform of a convolution $\mu * \nu$ is the product $\widehat{\mu}(\xi) \cdot \widehat{\nu}(\xi)$. In this way we have an embedding of $\mathbf{R}^n_\xi$ into $\mathcal{M}(B)$. However, the resulting subset is not dense in $\mathcal{M}(B)$. It means that there exist Riesz measures μ such that $|\widehat{\mu}(\xi)| \geq \delta > 0$ hold for all ξ, and yet μ is not invertible in B. An example of such a measure was discovered by Wiener and Pitt and one therefore refers to the *Wiener-Pitt phenomenon* in B. Further examples occur in [Gelfand et. all]. The idea is to construct Riesz measurs μ with independent powers, i.e. measures μ such that the norm of a μ-polynomial

$$c_0 \cdot \delta_0 + c_1 \cdot \mu + \ldots + c_k \cdot \mu^k$$

is roughly equal to $\sum |c_k|$ while $||\mu|| = 1$. In this way one can construct measurs μ for which the spectrum in B is the unit disc while the range of the Fourier transform is a real interval. Studies of $\mathcal{M}(B)$ occur in work by J. Taylor who established topological properties of $\mathcal{M}(B)$. The proofs rely upon several complex variables and we shall not try to expose material from Taylor's deep work. Let us only mention one result from Taylor's work in the case $n = 1$. Denote by $i(B)$ the multiplicative group of invertible measures in B where $B = \mathcal{M}(B)$ on the real line. If $\nu \in B$ we construct the exponential sum

$$e^{\nu} = \delta_0 + \sum_{k=1}^{\infty} \frac{\nu^k}{k!}$$

In this way e^B appears as a subgroup of $i(B)$. Taylor proved that the quotient group

$$\frac{i(B)}{e^B} \simeq \mathbf{Z}$$

where the right hand side is the additive group of integers. More precisely one finds an explicit invertible measure μ_* which does not belong to e^B and for any $\mu \in i(B)$ there exists a unique integer m and some $\nu \in B$ such that

$$(*) \qquad \mu = e^{\nu} * \mu_*^k$$

The measure μ_* is given by

$$\mu_* = \delta_0 + f$$

CONTINUE...

3.1 Wiener algebras. We can ask for subalgebras of $M(\mathbf{R}^n)$ where the Wiener-Pitt phenomenon does not occur, i.e. subalgebras B where the Fourier transform gives a dense embedding of $\mathbf{R}^n_\xi$ into $\mathcal{M}(B)$. A first example goes as follows: Let $n \geq 1$ and consider the Banach space $L^1(\mathbf{R}^n)$ where convolutions of L^1-functions is defined. Adding the unit point mass δ_0 at the origin we get the commutative Banach algebra

$$B = \mathbf{C} \cdot \delta_0 + L^1(\mathbf{R}^n)$$

Here the Fourier transform describes $\mathcal{M}(B)$. More precisely, if λ is a multiplicative functional on B whose restriction to $L^1(\mathbf{R}^n)$ is not identically zero, then one proves that there exists a unique point $\xi \in \mathbf{R}^n$ such that

$$\lambda(f) = \widehat{f}(\xi) \quad : \quad f \in L^1(\mathbf{R}^n)$$

In this way the n-dimensional ξ-space is identified with a subset of $\mathcal{M}(B)$. An extra point λ^* appears in $\mathcal{M}(B)$ where $\lambda^*(\delta_0) = 1$ while its restriction to $L^1(\mathbf{R}^n)$ vanishes. Hence the compact Gelfand space $\mathcal{M}(B)$ corresponds to the one-point compactification of the ξ-space. Here the continuity of Fourier transforms of L^1-functions correspond to the fact that their Gelfand transforms are continuous. An important consequence of this is that when $f(x) \in L^1(\mathbf{R}^n)$ is such that $\widehat{f}(\xi) \neq 1$ for every ξ, then the B-element $\delta_0 - f$ is invertible, i.e. there exists another L^1-function g such that

$$\delta_0 = (\delta_0 - f) * (\delta_0 + g) \implies f = g - f * g$$

The equality

$$(*) \qquad \mathcal{M}(B) = \mathbf{R}^n_\xi \cup \{\lambda^*\}$$

was originally put forward by Wiener prior to the general theory about Banach algebras. Another Banach algebra is $M_d(\mathbf{R})^{\mathbf{n}}$ whose elements are discrete measures with a finite total variation. Thus, the elements are measures

$$\mu = \sum c_\nu \cdot \delta(p_\nu)$$

where $\{p_\nu\}$ is a sequence of points in $\mathbf{R}^n$ and $\{c_\nu\}$ a sequence of complex numbers such that $\sum |c_\nu| < \infty$. Here the Gelfand space is more involved. To begin with the Fourier transform identifies $\mathbf{R}^n_\xi$ with a subset of $\mathcal{M}(B)$. But the compact space $\mathcal{M}(B)$ is considerably and given by a compact abelian group which is called the Bohr group after Harald Bohr whose studies of almost periodic functions led to the description of $\mathcal{M}(B)$. However one has the following result:

3.2 Bohr's Theorem. *The subset* $\mathbf{R}^n_\xi$ *is dense in* $\mathcal{M}(B)$.

Remark. See XX for an account about almost periodic functions which proves Bohr's theorem in the case $n = 1$.

3.3 Beurling's density theorem. Consider the Banach algebra B generated by $M_d(\mathbf{R}^n)$ and $L^1(\mathbf{R}^n)$. So its elements are measures of the form

$$\mu = \mu_d + f$$

where μ_d is discrete and f is absolutely continuous. Here the Fourier transform identifies $\mathbf{R}^n_\xi$ with an open subset of $\mathcal{M}(B)$. More precisely, a multiplicative functional λ on B belongs to the open set $\mathbf{R}^n_\xi$ if and only if $\lambda(f) \neq 0$ for at least some $f \in L^1(\mathbf{R}$. The remaining part $\mathcal{M}(B) \setminus \mathbf{R}^n_\xi$ is equal to the Bohr group above.. It means that when λ is an arbitrary multiplicative functional on B then there exists $\lambda_* \in \mathcal{M}(B)$ such that λ_* vanishes on $L^1(\mathbf{R}^n$ while $\lambda_*(\mu) = \lambda(\mu)$ for every discrete measure. The density of $\mathbf{R}^n_\xi$ follows via Bohr's theorem and the fact that Fourier transforms of L^1-functions tend to zero as $|\xi| \to +\infty$.

3.4 Varopoulos' density theorem. For each linear subspace Π of arbitrary dimension $1 \leq d \leq n$ we get the space $L^1(\Pi)$ of absolutely continuous measures supported by Π and of finite total mass. Thus, we identify $L^1(\Pi)$ with a subspace of $M(\mathbf{R}^n)$. We get the closed subalgebra of $M(\mathbf{R}^n)$ generated by all these L^1-spaces and the discrete measures. It is denoted by $\mathcal{V}(\mathbf{R}^n)$ and called the Varopoulos measure algebra in $\mathbf{R}^n$. In [Var] it is proved that the Fourier transform identifies $\mathbf{R}^n_\xi$ with a dense subset of $\mathcal{M}(\mathcal{V}(\mathbf{R}^n))$.

3.5 The extended $\mathcal{V}$-algebra. In $\mathbf{R}^n$ we can consider semi-analytic strata which consist of locally closed real-analytic submanifolds S whose closure $\bar{S}$ is compact and the relative boundary $\partial S = \bar{S} \setminus S$ is equal to the zero set of a real analytic function. On each such stratum we construct measures which are absolutely continuous with respect to the area measure of S. Here the dimension of S is between 1 and $n - 1$ and now each measure in $L^1(S)$ is identified with a Riesz measure in $\mathbf{R}^n$ which happens to be supported by S, One can easily prove that every $\mu \in L^1(S)$ has a power which belongs to the Varopolulos algebra and from this deduce that if $\mathcal{V}^*$ is the closed subalgebra of $M(\mathbf{R}^n)$ generated by the family $\{L^1(S)\}$ and $V(\mathbf{R}^n))$ then one gets a new Wiener algebra.

3.6 Olofsson's example. Above real analytic strata were used to obtain $\mathcal{V}^*$. That real-analyticity is essentially necessary was proved by Olofsson in [Olof]. For example, he found a C^∞-function $\phi(x)$ on $[0, 1]$ such that if μ is the measure in $\mathbf{R}^2$ defined by

$$\mu(f) = \int_0^1 f(x, \phi(x)) \cdot dx$$

then μ has independent powers and it cannot belong to any Wiener subalgebra of $M(\mathbf{R}^n)$. Actually [Olofson] constructs examples as above on curves defined by C^∞-functions outside the Carleman class of quasi-anaytic functions.