Basic facts in functional analysis

Contents

Introduction. The theory about normed vector spaces, and more generally locally convex spaces, owes much to pioneering work by Banach, Krein and Smulian, while the theory about inner product spaces goes back to work by Fredholm and Hilbert. Some crucial facts in topology play an important role, such as Baire's category theorem and Tychonoff's theorem about product of compact spaces. The basic results are proved in a rather straightforward manner but their merit is the generality. The beginner should pay attentions to notion of dual topological vector spaces since duality arguments often are used to prove existence theorems. The determination of the dual space X^* of a given normed vector space X is therefore an important issue.

Remarks about general topology. The notion of compact topological spaces is essential. Here certain phenomena can arise which differ from results in euclidian spaces. For example, start with the set $\mathbf{N}$ of non-negative integers. There exists the commutative Banach algebra B whose vectors are bounded complex valued functions on $\mathbf{N}$. We remark that B often is denoted by ℓ^∞ and equal to the dual of the Banach space ℓ^1. Now there exists the maximal ideal space $\mathfrak{M}(B)$ of B. Its construction is given in § xx and one often refers to $\mathfrak{M}(B)$ as the Gelfand space of B. It is equipped with the Gelfand topology and becomes a compact topological space which in addition is Hausdorff. Here $\mathbf{N}$ appears as a discrete subset of $\mathfrak{M}(B)$. The construction of points in $\mathfrak{M}(B) \setminus \mathbf{N}$ requires the axiom of choice. It turns out that $\mathfrak{M}(B)$ is a compact topological space which fails to be *sequentially compact*. In fact, from the countable subset $\mathbf{N}$ we cannot find a subsequence of integers $1 \leq n_1 < n_2 < \dots$ which converges to a limit point $\xi \in \mathfrak{M}(B)$. To see this we construct the bounded function on f on $\mathbf{N}$ where $f(n_k) = (-1)^k$ while $f = 0$ at integers outside the subsequence. If $\{n_k\}$ converges to a point $\xi \in \mathfrak{M}(B)$ we must have

$$\lim_{k \to \infty} f(n_k) = f(\xi)$$

But this cannot hold since f changes signs along to the subsequence. In addition to this "ugly example" we mention that there exist topological spaces X which are non-compact while every countable subset contains at least one convergent subsequence. Thus, X can be sequentially compact but not compact. We refer to § xx for an example.

The examples above show that one often must be careful about the notion of compactness. An important "positive result" due to Eberlein and Smulian asserts that if X is a Banach space and we equip X with the weak topology, then every weakly compact subset is sequentially compact, and conversely, if we start from a subset A in X with the property that every countable sequence of vectors in A has a subsequence which converges in the weak topology to a limit vector in X, then the closure of A taken in the weak topology is compact. We prove this in § xx and remark that the proof of the Eberlein-Smulian theorem is quite hard compared to more straightforward results in § 4. which are immediate consequences of Baire's category theorem.

Compact metric spaces. To avoid possible confusion we recall the notion of compactness on a metric space which goes back to Heine and Bolzano. Let S be a metric space where d is the distance function. A closed subset K is totally bounded in Heine's sense if there to every $\epsilon > 0$ exists a finite subset $\{x_1, \ldots, x_N\}$ in K such that K is covered by the union of the open balls $B_\epsilon(x_k) = \{x\colon\, d(x, x_k) < \epsilon\}$. A wellknown fact whose verification is left to the reader asserts that a closed subset K is totally bounded if and only if it is compact in the sense of Bolzano, i.e. every sequence $\{x_n\}$ in K contains at least one convergent subsequence.

Advanced results. The reader has mastered the basic theory once proofs of such results as the Hille-Phillips-Yosida theorem sbout infinitesmal generators of strongly continuous semi-groups in § xx and Schauder's fixed point theorem in § xx, have been pursued. Of course there are many results which are not covered in these notes. Non-linear maps between topological vector spaces is an extensive subject, and many important results deal with topological groups which in general need not be commutative. Here is an example of a result due to G. Birkhoff and Kakutani: Let G be a topological group whose topology is metrizable, i.e. there exists some distance function whose metric is equivalent to the given topology. Then there exists a left invariant metric d, i.e.

$$d(gx, gy) = d(x, y)$$

hold for every triple g, x, y in G. A notable point is that the existence of d is valid under the sole hypotheis that the topology is metrizable, i.e. G need not be a locally compact.

Next follow some examples about continuous functions on the real line whose proofs require methods beyond standard functional analysis, i.e. the general theory is not "the end of the story".

The Carleman-Weierstrass approximation theorem. It asserts that if $f(x)$ is a continuous complex-valued function on the real x-line then it can be uniformly approxmated by restrictions of entire functions in the complex z-plane with $z = x + iy$. In other words, for every $\epsilon > 0$ there exists an entire function $G(z)$ such that

$$\max_x |G(x) - f(x)| < \epsilon$$

The proof appears in my notes devoted to the matheramtics by Carleman. Notice that the result above goes beyond the ordinary result by Weierstrass since we obtain a uniform approximation on the whole real line.

Beurling's closure theorem. On the real t-line we have the space of bounded and complex-valued bounded and uniformly continuous functions ψ. It means that to each $\epsilon < 0$ there exists $\delta > 0$ such that

$$\max |\psi(x) - \psi(x')| < \epsilon$$

with the maximum taken over pairs of real numbers such that $|x - x'| < \delta$. Denote this space by $C_*(\mathbf{R})$. If μ is a Riesz measure on the real ξ-line we define the function

$$\mathcal{F}_\mu(x) = \int e^{x\xi}\, d\mu(\xi)$$

It is easily seen that $\mathcal{F}_\mu \in C_*(\mathbf{R})$. Denote by $\mathcal{A}$ be subspace of $C_*(\mathbf{R})$ given by $\mathcal{F}_\mu$-functions as μ varies over all Riesz measures as above. Before we announce Beurling's result below we define the notion of weak-star limits in $\mathfrak{M}$. Let $\{\mu_n\}$ be a bounded sequence of Riesz measures, i.e. there exists a constant such that

$$||\mu_n|| \leq M$$

hold for all n. The sequence $\{\mu_n\}$ converges in the weak-star sense to zero if

$$\lim_{n\to\infty} \int e^{ix\xi} \cdot d\mu_n(x) = 0$$

holds pointwise for every ξ.

Theorem. *A function $\psi \in C_*$ belongs to the closure of $\mathcal{A}$ if and only if*

$$(*) \qquad \lim_{n\to\infty} \int \psi(x) \cdot d\mu_n(x) = 0$$

whenever $\{\mu_n\}$ is a sequence in $\mathfrak{M}$ which converges weakly to zero.

Remark. Beurling's proof relies upon a certain non-linear variational problem which requires a quite deliate analysis and is exposed in my notes entitled *Mathematics by Beurling.*

Neumann's calculus. Analytic function theory appears during the study of spectra associated to linear operators. Here the constructions of resolvents play a crucial role. In § xx we expose the theory which started in Carl Neumann's pioneering work about boundary valuee problems in 1879. So Neumann's calculus is a central topic whose great merit is that it can be applied to densely defined but unbounded linear operators. In specific cases it often requires considerable work to determine the spectrum. An example arises when we take the Laplce operator Δ in $\mathbf{R}^3$ and consider a densely defined linear operator L on the Hilbert space $L^2(\mathbf{R}^3)$ given by

$$L_g(x) = \Delta(g)(x) + c(x) \cdot g(x)$$

where $c(x)$ is a real-vaued and locally square integrable function. The domain of defintion for L contains C^2-functions with compact support. Neccesary and sufficient conditions on the c-function in order that Neumann's spectrum of L is contained in the real line are not known. However, the following suffiency was proved by Carleman in lectures at Sorbonne in 1930:

Theorem. *Assume that there exists a constant M such that*

$$\limsup_{|x|\to\infty} c(x) \leq M$$

Then the spectrum $\sigma(L)$ is a closed subset of the real line.

The proof of this result appears in § xx from my notes entitled *Mathematics by Carleman.*

Summary of the contents.

§ 1 studies normed vector spaces over the complex field **C** or the real field **R**. We explain how each norm is defined by a convex subset of V with special properties. 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 says that the norm is complete and refer to the pair $(X, ||\cdot||)$ as a Banach space.

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^* of unit norm. Then one has the equality

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

The determination of X^* is often an important issue. An example is the dual of the normed space $X = H^\infty(T)$ of bounded Lebesgue measurable functions on the unit circle which are boundary values of bounded analytic functions in the open disc D. A portion of its dual space is given by the quotient space:

$$\text{(ii)} \qquad Y = \frac{L^1(T)}{H_0^1(T)}$$

where $H_0^1(T)$ is the closed subspace of $L^1(T)$ whose functions are boundary values of analytic functions in D which vanish at $z = 0$. However, the dual X^* is considerably larger. In fact, we shall learn that $H^\infty(T)$ is an example of a commutative Banach algebra to which we can assign the maximal ideal space $\mathfrak{M}_X$ and now X^* is the space of Riesz measures on this compact space. But a concrete description of $\mathfrak{M}_X$ is not known. Via point evaluations in D it is clear that D appears as a subset of $\mathfrak{M}_X$. Its position in $\mathfrak{M}_X$ was an open question for some time, until Carleson in an article from 1957 proved that D is a dense subset. This result is known as the Corona Theorem whose proof requires a deep analysis based upon geometric constructions such as Carleson measures.

Reflexive spaces. Starting from a Banach space X we get X^* whose dual is denoted by X^{**} and called the bi-dual of X. There is a natural injective map $\mathfrak{i}_X \colon X \to X^{**}$ and (i) above shows thatit is an isometry, i.e. the norms $||x||$ and $||\mathfrak{i}_X(x)||$ are equal. But in general the bi-dual embedding map is nor surjective. If it is one says that X is reflexive.

Compact operators. Let X and Y be a pair of Banach spaces. A bounded linear operator $T \colon X \to Y$ is compact if the image of the unit ball $B(X)$ in X is relatively compact in Y. In applications one often encounters compact operators and several facts about these operators occur in § 6 where the crucial results deal with Fredholm operators and culminate with Fredholm's index theorem. When the target space $Y = C^0(S)$ a classic result due to Arzela and Vitali asserts that $T(B(X))$ is relatively compact if and only if this family is equi-continuous, i.e. to each $\epsilon > 0$ there exists $\delta > 0$ such that

$$\max_{x \in B} \omega_{Tx}(\delta) \leq \epsilon$$

where we introduced the modulus of continuity, i.e if $f \in C^0(S)$ and $\delta > 0$ then

$$\omega_f(\delta) = \max |f(y_1) - f(y_2)| \quad : \ d(y_1, y_2) < \epsilon$$

In § xx we prove a result due to Schauder which asserts that if $T \colon X \to Y$ is a compact operator between a pair of Banach spaces, then its adjoint $T^* \colon X^* \to Y^*$ is also compact. This is an example of a more involved result whose proof relies upona fundamental theorem due to Krein and Smulian concerned with the weak star topology on dual spaces. Smulian proved this result in 1940 and it ewas not until 1972 that the depth of this result was revealed when Enflo constructed a pair (X, T) where X is a separable Banach space and $T \colon X \to X$ is a compact operator, and yet T cannot be approximated in the operator norm by linear operators with finite dimesnsional range. Enflo's work [Acta 1972] is one of the greatest achievement dealing with geometry on topological vector spaces and has created a veritable industry here one seeks to determine when a given Banach space has the approximation property in the sense that every compact operator can be approximated by finite range operators.

Calculus on Banach spaces. Let X and Y be two Banach spaces. In § 7 we define the differential of a C^1-map $g\colon X \to Y$ where g in general is non-linear. Here the differential of g at a point $x_0 \in X$ is a bounded linear operator from X into Y. This extends the construction of the Jacobian for a C^1-map from $\mathbf{R}^n$ into $\mathbf{R}^m$ expressed by an $m \times n$-matrix. More generally one constructs higher order differentials and refer to C^∞-maps from one Banach space into another. We shall review this in § 7. Let us remark that Baire's category theorem together with the Hahn-Banach theorem show that if S is an arbitrary compact metric space and ϕ is a continuous function on S 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_S is the distance function on the metric space S and in the right hand side we have taken the norm in X. Next, there exists class of differentiable Banach spaces. By definition a Banach space X is differentiable at a point x if there exists a linear functional $\mathcal{D}_x$ on X such that

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

hold for every $y \in X$ where the limit is taken over complex ζ which tend to zero. One says that X is differentiable if $\mathcal{D}_x$ exist for every $x \in X$. In § XX we expose a result due to Beurling and Lorch concerned with certain non-linear duality maps on uniformly convex and differentiable Banach spaces.

Analytic functions. Let X be a Banach space and consider a power series with coefficients in X:

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

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

$$\text{(ii)} \qquad ||b_\nu|| \leq C \cdot R^\nu \quad : \ \nu = 0, 1, \ldots$$

Then the series (i) converges when $|z| < R$ and $f(z)$ is called an X-valued analytic function in the open disc $|z| < R$. 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

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

Using the dual space X^* results about ordinary analytic functions extend 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 $f(z)$ is 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. This gives Cauchy's integral formula

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

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 there exists a constant M such that

$$\max_{0\leq t\leq 1} ||f(t)|| = M$$

Assume in addition that the complex-valued functions $t \mapsto x^*(f(t))$ are Borel functions on $[0,1]$ for every $x^* \in X^*$. Then there exist the Borel-Stieltjes 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^*. This gives 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.

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

This map applies in particular if X is a Hilbert space since they are reflexive.

Operational calculus. Commutative Banach algebras are defined and studied in § 10. 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 one gets the vector-valued resolvent map:

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

If $\lambda_0 \in \mathbf{C} \setminus \sigma(x)$ there exists a local Neumann series which 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 gives vectors 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 $\sigma(x)$. In § 10 we prove that 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

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

This general result is used in many applications. A crucial case occurs when B is the Banach algebra generated by a single bounded linear operator on a Hilbert space.

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 $\langle x, y\rangle$ satisfying the following three conditions:

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

$$(2) \qquad \langle y, x\rangle = \overline{\langle x, y\rangle} \quad : x, y \in \mathcal{H}$$

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

Here (1-3) imply that $\mathcal{H}$ is equipped with a norm defined by $||x|| = \sqrt{\langle x, x\rangle}$ and one easiy verfies that it satisfies

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

Conversely, if (*) holds for a norm then it is defined by an inner product where

$$2 \cdot \mathfrak{Re}\,\langle x, y\rangle = ||x+y||^2 - ||x||^2 - ||y||^2 \quad : 2 \cdot \mathfrak{Im}\,\langle x, y\rangle = ||ix+y||^2 - ||x||^2 - ||y||^2$$

If a norm given via an inner product is complete one says that $\mathcal{H}$ is a Hilbert space. A fundamental fact is that Hilbert spaces are *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.

Orthonormal bases. If $\mathcal{H}$ is separable, i.e. contains a denumerable dense subset, then it is isomorphic to the Hilbert space ℓ^2 whose elements are sequences of complex numbers $\{c_0, c_1, \ldots\}$ for which $\sum |c_n|^2 < \infty$. However, this does not mean that one has a clear picture of a Hilbert space which is given without a specified orthonormal basis. So even if all separable Hiolbert spaces are isomorphic one often needs considerable work to exhibit a sequence" of pairwise orthogonal vectors to exhibit that $\mathcal{H} \simeq \ell^2$. An example occurs when $\mathcal{H}$ is the Hilbert space of square integrable analytic functions in a bounded domain in $\mathbf{C}$. Here one defines Bergman's kernel function and for simply connected domains an orthonormal bases given by polynomials was found by Faber which has consequences in analytic function theory. See § X in Chapter VI for an account.

Some specifix examples.

General results in functional analysis are not always sufficient for more precise conclusions. A typical problem is to describe the range of a bounded linear operator $T\colon X \to Y$ from a Banach space into another. Consider as an example the Banach space $C^0[0,1]$ of complex-valued and continuous functions on the closd unit interval and the bounded linear operator

$$L_g(x) = \int_0^1 \log \frac{1}{|x-t|} \cdot g(t)\, dt$$

whose operator norm is bounded above by

$$\max_{0 \leq x \leq 1} \int_0^1 \log \frac{1}{|x-t|}\, dt$$

Lebesgue theory teaches that

$$\lim_{\delta \to 0} \max_{x_1, x_2} \int_0^1 \big[\log \frac{1}{|x_1 - t|} - \frac{1}{|x_2 - t|}\big|\, dt = 0$$

where the maximum for each $\delta > 0$ is taken over pairs of points in $[0,1]$ for which $|x_1 - x_2| \leq \delta$. This entails that L maps the unit ball in $C^0[0,1$ to an equi-continuous family of functions and as a consequence the Arzela-Ascoli theorem implies that L is a compact operator. It turns out that L is injective and has a dense range which contains all C^2-functions. See § xx for the proof which includes an inversion formula for L which therefore describes the range of L. But this description is rather involved and the precise description can onoly be found via the expålicit inversion forula which was established by Carleman and also trewated in work by Zeilon from the years around 1920.

The Dirichlet problem. Fortunately one is not always obliged to describe the range in detail. Consider as an example a bounded open set Ω in $\mathbf{R}^n$ for some $n \geq 2$. We assume that the boundary $\partial\Omega$ is a finite union of pairwise disjoint and closed C^1-submanifolds. Now there exists a bounded linear operator L on $C^0(\partial\Omega)$ defined by

$$L_g(p) = \int_{\partial\Omega} \frac{g(q)}{|p-q|^{n-2}}\, dA(q)$$

where dA is the area measure. If L has a dense reange in $C^0(\partial\Omega)$ the maximum principle for harminic functions entails that the Dirchlet problem has a solution, i.e. each continuous boundsry function has a unique harmonic extension to Ω. Using a result due to F. Riesz which identifiesz the dual of $C^0(\partial\Omega)$ with Ruesz measures supported by $\partial\Omega$, the range of L is dense if one has proved that if μ is a Riesz measure on $\partial\Omega$ such that

$$(*) \qquad \int_{\partial\Omega} \frac{d\mu(q)}{|p-q|^{n-2}} \quad : p \in \partial\Omega$$

then $\mu = 0$. This ilustrates the usefulness of arguments basied upon duality. The proof that (*) gives $\mu = 0$ is an easy consequence of Green's formula and is exposed in § xx.

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 with elements $\{\partial^2 f/\partial x_j \partial x_k\}$. The function is called strictly convex if $H_f(x)$ is positive for all x, i.e. if the eigenvalues are all > 0. Assume in addition that

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

Under these conditions one has the result below which is due to Lagrange and Legendre.

Theorem. *The function takes its minimum at a unique point in $\mathbf{R}^n$ and the vector valued function*

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

bijective, i.e. $\nabla_f(p) \neq \nabla_f(q)$ hold for all pairs $p \neq q$ in $\mathbf{R}^n$.

Exercise. Prove this classical result.

With f as in the theorem we can move origin and assume that f takes its mimium at $x = 0$. Replacing f by $f(x) - f(0)$ the minimum value is zero. To each positive real number s we put

$$K_s = \{f \leq s\}$$

Then $\{K_s\}$ is an increasing sequence of convex and compact sets whose boundaries $\{\partial K_s\}$ are C^1-submanifolds in $\mathbf{R}^n$. The reader is invited to analyze the ray functions where one for each point ω on the unit sphere S^{n-1} studies the function

$$\rho_\omega(s) = r \quad : r \cdot \omega \in \partial K_s$$

The reader should verify that ρ_ω is a strictly increasing function of s and it is instructive to plot level sets $\{f = a\}$ when $a > 0$ for some different f-functions as above.

Cones in $\mathbf{R}^n$. A subset Γ is a cone if $x \in \Gamma$ implies that the half-ray $\mathbf{R}^+ \cdot x \subset \Gamma$. Suppose that Γ is a closed set in $\mathbf{R}^n$. Let S^{n-1} be the euclidian unit sphere and put

$$\Gamma_* = \Gamma \cap S^{n-1}$$

We say that Γ is *fat* if Γ_* has a non-empty interior in S^{n-1} and Γ is *proper* if

$$\Gamma_* \cap -\Gamma_* = \emptyset$$

The reader may verify that this is equivalent with the condition that Γ does not contain any 1-dimensional subspace. We leave it to the reader to verify that a cone Γ is proper if and only if $\widehat{\Gamma}$ is fat.

Dual cones. The *dual cone* is defined by

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

Exercise. Show the biduality formula, i.e. that Γ is equal to the dual cone of $\widehat{\Gamma}$.

1. Normed spaces.

A norm on a complex vector space X is a map from X into $\mathbf{R}^+$ satisfying:

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

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

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

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 $||\cdot||$. 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

$$\text{(i)}\qquad ||(x,y)|| = |x| + |y|$$

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.

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

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:

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

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

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

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

$$(*)\qquad \rho_K(y) = \min_{s>0} y \in s \cdot K$$

If $y \in K$ then $s = 1$ 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 large s-values to get $y \in s \cdot K$ and therefore $\rho_K(y)$ is large. It is clear that

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

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

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

Next, the convexity of K gives the inclusion

$$s \cdot K + t \cdot K \subset (s+t) \cdot K$$

for every pair of positive numbers. Indeed, wirth x and y in K it follows that

$$\frac{sx}{s+t} + \frac{ty}{s+t} \in K$$

and (iii) follows. Now the reader can check that the construction of ρ_K gives the triangle inequality

$$\text{(iii)} \qquad \rho_K(y_1 + y_2) \leq \rho_K(y_1) + \rho_K(y_2)$$

We conclude that ρ_K yields a norm on X.

1.4 A converse. If $|| \cdot ||$ is a norm on Y we get the 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 the reader should notice 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 (ii-iii) in § 1.2 are said to be 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 $\mathcal{K}$ of convex sets which are symmetric and satisfy (ii-iii) from § 1.2. For each given norm on Y we can assign the largest convex set K^* in the corresponding equivalence class.

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

$$\text{(0.6)} \qquad \frac{1}{C} \cdot ||y||_1 \leq ||y||_2 \leq C \cdot ||y||_1 \quad : \quad y \in Y$$

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

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

1.6 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 have the *euclidian norm* which measures the euclidian length from a vector y to the origin:

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

The reader should verify that (i) satisfies the triangle inequality

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

which amounts to verify the Cauchy-Schwartz inequality. We have also the norm $|| \cdot ||^*$ defined by

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

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

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

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 via convex sets K which not only are symmetric with respect to scalar multiplication with real numbers, but is also invariant under multiplication with complex numbers $e^{i\theta}$ which entails that

$$\rho_K(\lambda \cdot x) = \lambda \cdot \rho_K(x)$$

hold for every complex number λ, i.e. we get a norm on the complex vector space.

2. Banach spaces.

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

(*) $$\lim_{n,m \to \infty} ||x_n - x_m|| = 0$$

We obtain a vector space $\widehat{X}$ whose vectors are defined as equivalence classes of Cauchy sequences. The norm of a Cauchy sequence $\{x_n\}$ is defined by

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

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

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

$$C^{-1} \cdot ||x||^* \leq ||x|| \leq C \cdot ||x||^* \quad : \quad x \in X$$

Remark. Thus, if $|| \cdot ||_1$ and $|| \cdot ||_2$ are two complete norms on the same vector space then they are equivalent. The proof is given in § xx below.

2.2 Separable Banach spaces. A Banach space X which which contain a denumerable and dense subset $\{x_n\}$ is called separable. If this holds we get for each n the finite dimensional subapce X_n generated by $x_1, \ldots, x_n$ By a wellknown procedure from Linear algebra we can construct a basis in each X_n and construct a denumarable sequence of linearly independent vectors $e_1, e_2, \ldots$ such that the increasing sequence of subspaces $\{X_n\}$ are contained in the vector space

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

Then X_* is a dense subspace of X. Of course, there are many ways to construct a denumerable sequence of linearly independent vectors which give a dense subspace of X. One may ask if it is possible to choose a sequence $\{e_n\}$ as above such that every $x \in X$ can be expanded as follows:

2.3 Definition. *A denumerable sequence $\{e_n\}$ of $\mathbf{C}$-linearly independent vectors in a complex vector space X is called a Schauder basis if there to each $x \in X$ exists a unique sequence of complex numbers $c_1(x), c_2(x), \ldots$ such that*

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

2.4 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 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. For the construction we refer to the article [Enflo-Acta Mathematica]. Let us remark that the essential ingredient in Enflo's construction relies upon a study of Fourier series where one has *Rudin-Schapiro* polynomials which consist of trigonometric polynomials

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

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

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

Remark. After [Enflo] it became a veritable industry to verify that various concrete Banach spaces Y do have a Schauder basis and perhaps more important, also 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 disc algebra $A(D)$ of continuous functions on the closed unit disc which are analytic in the interior.

3. Bounded 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. The least constant C for which this holds is denoted by $||T||$ and called the operator norm of T. Denote by $\mathcal{L}(X,Y)$ the set of all continuous linear operators from X into Y. From the above it is equipped with the norm

$$(*) \qquad ||T|| = \max_{||x||=1} ||T(x)|| \quad :$$

Above X and Y are not necessarily Banach spaces. But one verifies easily that if $\widehat{X}$ and $\widehat{Y}$ 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 $\hat{Y}$. Moreover, if Y from the start is a Banach space and $T \in \mathcal{L}(X,Y)$ then it extends in a unique way to a bounded linear operator from $\widehat{X}$ into Y. Finally the reader may verify 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.*

3.2 Null spaces and the range. 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 the kernel is a closed subspace of X and we get the quotient space

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

It is clear 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 range $T(X)$ and notice the equality

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

One says that T has *closed range* if the linear subspace $T(X)$ of Y is closed. When this holds the complete norm on Y induces a complete norm on $T(X)$. In §4. A we prove the *Open Mapping Theorem* which shows that if T has closed range then there exists a constant C such that for every vector $y \in T(X)$ there exists $x \in X$ with $y = Tx$ and

$$\text{(3.2.1)} \qquad ||x|| \leq C \cdot ||y||$$

3.3 The closed graph theorem Let X and Y be Banach spaces and T a linear operator from X into Y. For the moment we do not assume that it is bounded. In the product space $X \times Y$ we get the graph

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

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.*

Proof. We have the surjective map

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

from X onto the graph. By assumption Γ_T is a closed subspace of the Banach space $X \oplus Y$ equipped with the norm

$$||(x, Tx)|| = ||x|| + ||Tx||$$

The Open Mapping Theoren gives a constant C such that

$$||x|| + ||Tx|| \leq C||x|| \implies ||Tx|| \leq (C-1)||x||$$

and hence T is bounded.

3.5 Densely defined and closed operators.

Let $X_* \subset X$ be a dense subspace and $T\colon\ X_* \to Y$ a linear operator where Y is a Banach space. As above we construct the graph

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

If Γ_T is a closed subspace of $X \times Y$ we say that the densely defined operator T is closed.

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. Let $Y = L^1[0,1]$ which gives a linear map $T\colon\ X_* \to Y$ defined by

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

Now T has a graph

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

Let $\overline{\Gamma(T)}$ denote the closure taken in the Banach space $X \times Y$. By definition a pair (f,g) belongs to $\overline{\Gamma(T)}$ if and only if

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

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

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

Hence the limit function f is a primitive integral

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

Conclusion. The linear space $\overline{\Gamma(T)}$ 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 $\widehat{T}$ with a closed graph. More precisely, $\mathcal{D}(\widehat{T})$ consists of functions $f(x)$ which are primitives of L^1-functions. Lebesgue theory this means that that the domain of definition of $\widehat{T}$ consists of *absolutely continuous functions.* Thus, by enlarging the domain of definition the linear operator T is extended to a densely defined and closed linear operator.

3.7 Remark. The example above is typical for many constructions where one starts with some densely defined linear operator T and finds an extension $\widehat{T}$ whose graph is the closure of $\Gamma(T)$. The reader should notice that the choice of the target space Y affects the construction of closed extensions. For example, replace above $L^1[0,1]$ with the Banach space $L^2[0,1]$ of square integrable functions on $[0,1]$. In this case we find a closed graph extension S 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 S is a proper subspace of the linear space of all absolutely continuous functions. In PDE-theory one starts from a differential operator

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

where $x = (x_1, \ldots, x_n)$ 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 we let Y be the Hilbert space $L^2(\Omega)$. 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. With $Y = L^2(\Omega)$ we construct a closed extension as bove. This device is often used in PDE-theory.

§ 4. A. Three basic principles.

A crucial result due to Baire goes as follows: Let X be a metric space whose metric d is complete, i.e. every Cauchy sequence with respect to the distance function d converges.

A.1 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. To say that F^* is meager means that if $x_0 \in X$ and $\epsilon > 0$ then

(i) $$B_\epsilon(x_0) \cap (X \setminus F^*) \neq \emptyset$$

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 $\delta_1 < \epsilon/3$ so that

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

Now $B_{\delta_1/3}(x_1)$ is not contained in F_2 and we find a pair x_2 and $\delta_2 < \delta_1/3$ 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) $$B_{\delta_n}(x_n) \cap F_n = \emptyset \quad : \quad x_n \in B_{\delta_{n-1}}(x_{n-1}) \quad : \quad \delta_n < \delta_{n-1}/3$$

For each pair $M > N \geq 1$ the triangle inequality gives

(iv) $$d(x_M, x_N) \leq \delta_{N+1} + \ldots + \delta_M \leq \delta_N(3^{-1} + \ldots 3^{-M+1}) \leq \frac{2}{3}\delta_N$$

At the same time we notice that

(v) $$\frac{2}{3}\delta_N \leq 3^{-N+1} \cdot \frac{2}{3}\epsilon$$

Hence $\delta_N \to 0$ and (iv) entails that $\{x_n\}$ is a Cauchy sequence. Since the metric is complete there exists a limit point x_* where

$$\lim_{M\to\infty} d(x_M, x_*) = 0$$

To each $N \geq 1$ we apply (iv) for arbitrary large M and obtain

$$d(x_*, x_N) \leq \frac{2}{3}\delta_N \implies x_* \in B_{\delta_N}(x_N) \subset X \setminus F_N$$

At the same time the the last inequality in (iv) applied with arbitrary large N gives

$$d(x_*, x_0) \leq \frac{2}{3}\epsilon \implies x_* \in B_\epsilon(x_0)$$

So x_* gives the requested point which produces a non-empty set in (i).

A.2 The Banach-Steinhaus Theorem. Let X be a vector space equipped with a complete norm $||\cdot||$. It means that X at the same time is a complete metric spacce with the distance function

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

Suppose now that $||\cdot||_*$ is another complete norm on X. To each positive number a we set

$$F_a = \overline{\{x : ||x||_* \leq a\}}$$

where the closure is taken with respect to d. The triangle inequality applied to the pair of norms give the inclusion

$$F_a + F_b \subset F_{a+b}$$

for each pair of positive real numbers. Next, since every vector in X has a finite norm $||x||_*$ it follows that the union of these F-sets taken as a varies over the set of positive integers is equal to X. Baire's theorem applied to the complete metric space (X, d) gives an integer N such that F_N has a non-empty interior. Thus, we find $x_0 \in X$ and $\epsilon > 0$ such that

$$B_\epsilon(x_0) \in F_N$$

Here $||x_0|| < M$ for some integer M and the triangle inequality implies that

$$B_\epsilon(0) \subset F_{N+M}$$

If k is a positive integer such that $k^{-1} \leq \epsilon$ and $K = (N+M)k$ it follows after scaling that the open unit ball

$$B_1 \subset F_K$$

and, again via a scaling

(i) $$B_a \subset F_{aK}$$

hold for every $a > 0$. Let us then take a vector x with $||x|| < 1$. Then we find a vector y_1 with $||y_1||_* \leq K$ while

$$||x - y_1|| < 1/2$$

Appy (i) with $a = 1/2$ which gives a vector y_2 with $||y_2||_* \leq K/2$ and

$$||x - y_1 - y_2|| < 1/4$$

We can continue and obtain a sequence $\{y_n\}$ such that

$$||y_n||_* \leq 2^{-n+1}K \quad : \quad ||x - (y_1 + \ldots + y_n)|| \leq 2^{-n}$$

It follows that $\sum ||y_n||_* < \infty$ and since the norm $||\cdot||_*$ is complete there exists a limit vector $y_* = \sum y_n$ with $||y||_* \leq 2K$. At the same time it is clear from the second inequality in (xx) that $x = y^*$ and hence we have the inclusion

$$B_1 \subset \{y : ||y||_* \leq 2K\}$$

This impies via scaling that

$$||x|| \leq 2K \cdot ||x||_*$$

hold for every $x \in X$. Hence the complete norm $||\cdot||_*$ is stronger than the given complete norm. We can reverse this and conclude

A.2.1 Theorem. *A pair of complete norms $||\cdot||$ and $||\cdot||_*$ on a vector space are always equivalent, i.e there exists a constant $C \geq 1$ such that*

$$C^{-1}||x||_* \leq ||x|| \leq C||x||_*$$

A.3 The Open Mapping theorem. Let $T\colon X \to Y$ be a linear operator where X and Y both are Banach spaces. We assume that T is surjective and bounded, i.e. there exists a constant C such that

(i) $$||Tx|| \leq C \cdot ||x||$$

In X we find the closed subspace X_0 given by the T-kernel and obtain a bijective map

$$S\colon \frac{X}{X_0} \to Y$$

Set $Z = \frac{X}{X_0}$ which is a new Banach space equipped with the quotient norm as explained in § xx. If z denote vectors in Z we get

(ii) $$||Sz||_Y \leq C \cdot ||z||_Z$$

where the subscripts refer to norms on Y and Z respectively. Since S is surjective we get a norm on Y by

$$||y||_* = ||z||_Z : y = Sz$$

and (ii) means that

(iii) $$||y||_Y \leq C \cdot ||y||_*$$

Suppose now that $\{y_n\}$ is a Cauchy sequence with respect to the norm $||\cdot||_*$. Then (iii) entails that it is also a Cauchy sequence with respect to $||\cdot||_Y$ and hence there exists a limit vector $y_* = \lim y_n$ in the Banach space Y. At the same time $y_n = Sz_n$ and the construction of $||\cdot||_*$ entails that $\{z_n\}$ is a Cauchy sequence in Z which therefore converges to a limit vector z_* and it is clear that (ii) gives $y_* = Sz_*$ which means that $\lim ||y_n - y_*||_* = 0$. Hence the norm $||\cdot||_*$ is complete and the Banach-Steinhaus theorem gives a constant C_* such that

$$\text{(iv)} \qquad ||y||_* \leq C_* \cdot ||y||$$

This means that S maps the open ball of radius C_* in Z onto the open unit ball in Y, Returning to X and the construction of the quotient norm the reader can check that this implies that T maps the open ball of radius C_* in X onto the open unit ball in Y which means that T is an open mapping.

A.4 The Hahn-Banach theorem.

Consider a real vector space E. A function $\rho\colon E \to \mathbf{R}^+$ is subadditive and positively homogeneous if the following hold for each pair x_1, x_2 in X and every non-negative real number s:

$$\rho(x_1 + x_2) \leq \rho(x_1) + \rho(x_2) \quad : \rho(sx_1) = s \cdot \rho(x_1)$$

Let E_0 be a subspace and $\lambda_0\colon E_0 \to \mathbf{R}$ a linear functional such that

$$\lambda_0(x_0) \leq \rho(x_0) \quad : x_0 \in E_0$$

Then there exists a linear functional λ on X which extends λ_0 and

$$\text{(A.4.1)} \qquad \lambda(x) \leq \rho(x) \quad : x \in X$$

To prove this we use Zorn's Lemma. Namely consider the partially ordered family of pairs (V, μ) where V is a subspace of E which contains E_0 and $\mu\colon V \to \mathbf{R}$ a linear map which extends λ_0 and satisfies.

$$\text{(i)} \qquad \mu(x) \leq \rho(x) \quad : x \in V$$

Zorn's Lemma gives a maximal pair in this family and there remains to prove that $V = E$. If $V \neq E$ we pick a vector $y \in E \setminus V$ and get a contradiction if we can extend μ to a linear map μ^* on $W = V + \mathbf{R}y$ which satisfies (A.4.1) when $x \in W$. To see that this is possible we set

$$\alpha = \min_{x \in V} \rho(x + y) - \mu(x)$$

$$\beta = \max_{\xi \in V} \mu(\xi) - \rho(\xi - y)$$

Now we show that

$$\text{(ii)} \qquad \alpha \geq \beta$$

To get this we consider a pair of vectors x and ξ in V and then

$$\rho(x + y) - \mu(x) - (\mu(\xi) - \rho(\xi - y)) = \rho(x + y) + \rho(\xi - y)) - \mu(x + \xi)$$

Next, the triangle inequality fo the ρ-function gives

$$\rho(x + \xi) = \rho((x + y) + (\xi - y)) \leq \rho(x + y) + \rho(\xi - y))-$$

At the same time the vector $x + \xi$ belonfgs to V and hence

$$\mu(x + \xi) \leq \rho(x + \xi)$$

This proves (ii) and now we pick a real number a such that $\beta \leq a \leq \alpha$. At this stage the reader can check that if we define μ^* by

$$\mu^*(x + sy) = \mu(x) + sa$$

for every real nuymber s, then μ^* is a linear map on W for which (i) holds.

A.4.2 The case of complex vector spaces. Consider a complex vector space X equipped with a norm $||\cdot||$ as in § xx. Let $X_0 \subset X$ be a complex subspace and μ a $\mathbf{C}$-linear map on X_0 such that

$$|\mu(x)| \leq ||x||$$

hold for every $x \in X_0$. Taking real and imaginary parts of the complex numbers under μ we get a pair of real-valued and $\mathbf{R}$-linear maps f_1, f_2 on X_0 such that

$$\mu(x) = f_1(x) + if_2(x)$$

Since μ is $\mathbf{C}$-linear we have

$$i(f_1(x) + if_2(x)) = i\mu(x) = \mu(ix) = f_1(ix) + if_2(ix)$$

Identifying the real parts we see that

(i) $$f_2(x) = -f_1(ix)$$

Next, with $x \in X_0$ we have

$$f_1(x) \leq |f_1(x) - if_1(ix)| \leq ||x||$$

The real version of the Hahn Banach theorem gives an $\mathbf{R}$-linear map $F \colon X \to \mathbf{R}$ which extends f_1 and satisfies

(ii) $$F_1(x) \leq ||x|| \quad : x \in X$$

Now we get a $\mathbf{C}$-linear map μ^* on X defined by

$$\mu^*(x) = F_1(x) - iF_1(ix)$$

From the above μ^* extends μ and there remains to prove that

(*) $$|\mu^*(x)| = |F_1(x) - iF_1(ix)| \leq ||x||$$

hold for all $x \in X$. To prove this we apply (ii) to vectors $e^{i\theta}x = \cos\theta \cdot x + \sin\theta \cdot ix$ which gives

(iii) $$\cos\theta \cdot F_1(x) + \sin\theta \cdot F(ix) \leq ||e^{i\theta}x|| = ||x||$$

This hold for all $0 \leq \theta \leq 2\pi$ and choose θ so that

$$\cos\,\theta = \frac{F_1(x)}{|F_1(x) - iF_1(ix)|} \quad : \; \sin\,\theta = \frac{F_2(x)}{|F_1(x) - iF_1(ix)|}$$

Then the left hand side in (iii) becomes $|F_1(x) - iF_1(ix)|$ and (*) follows.

A.4.3 Separating hyperplanes. Consider a convex set K in a real normed vector space X which contains the origin. With $\delta > 0$ we put

$$K_\delta = K + B(\delta) = \{x + \xi : x \in K \, ||x|| < \delta\}$$

As explained in § xx we get a function ρ defined by

$$\rho(x) = \inf_{s>0} \, x \in s \cdot K_\delta$$

where ρ satisfies (A.4.0). Let x_0 be a vector such that

$$\min_{x \in K} ||x - x_0|| = \delta$$

This entails that

$$\rho(x) = 1$$

On the 1-dimensional real subspace of X generated by x_0 we define the linear functional λ_0 where $\lambda_0(x) = 1$. Now the real version of the Hahn Banach theorem gives an extension λ to X such that

$$\lambda(x) \leq \rho(x)$$

hold for all $x \in X$. In particular we take vectors in the open ball $B(\delta)$ for these vectors the inclusion $B(\delta) \subset K_\delta$ gives $\rho(x) \leq 1$. After scaling we conclude that

$$|\lambda(x)| \leq \delta^{-1} \cdot ||x||$$

hold for all vectors in X which means that the linear functional λ is continuous. Next, since $\rho(x) \leq 1$ for all vectors $x \in K_\delta$ we have

$$\lambda(k) + \lambda(\xi) = \lambda(k + \xi) \leq 1$$

for pairs $k \in K$ and $\xi \in B(\delta)$. Here ξ can be chosen so that

$$\lambda(\xi) = -\delta \cdot ||\lambda||$$

We conclude that

$$\max_{k \in K} \lambda(k) \leq 1 - \delta \cdot ||\lambda||$$

At the same time $\lambda(x) = 1$ which means that the half-space

$$\{\lambda \leq 1 - \delta \cdot ||\lambda||\}$$

contains K while x has distance $\geq \delta \cdot ||\lambda||$ to this half-space.

A.4.4 Example. One is often confronted with severe difficulties in specific cases because the Hahn Banach theorem is not constructive. Consider for example a positive integer N and let $\mathcal{P}_N$ denote the $(N+1)$-dimensional real vector space of polynomials $P(t)$ of degree $\leq N$. Define a linear functional λ_0 on $\mathcal{P}_N$ by

(i)
$$\lambda_0(P) = \sum_{\nu=1}^{\nu=M} c_\nu \cdot P(t_\nu)$$

where $\{t_\nu\}$ and $\{c_\nu\}$ are real numbers. We can identify $\mathcal{P}_N$ with a subspace of $X = C^0[0,1]$ where $C^0[0,1]$ is the vector space of real-valued and continuous functions on the unit interval $\{0 \leq t \leq 1\}$. On X we get a ρ-function via the maximum norm, i.e.

$$\rho(f) = \max_{0 \leq x \leq 1} |f(x)|$$

Next, there exists a unique smallest constant $C > 0$ such that

$$|\lambda_0(P)| \leq C \cdot \rho(P) \quad : P \in \mathcal{P}_N$$

Notice that C exists even in the case when some of the t-points in (i) are outside $[0,1]$. By the result above we can extend λ_0 to a linear functional λ on X and the Riesz representation theorem gives a real-valued Riesz measure μ on $[0,1]$ such that

$$\lambda_0(P) = \int_0^1 P(t) \cdot d\mu(t) \quad : P \in \mathcal{P}_N$$

and at the same time the total variation of μ is equal to the constant C above. However, to determine C and to find μ requires a further analysis which leads to a rather delicate problem in optimization thoery. Already a case when $M = 1$ and we take $t_1 = 2$ and $c_1 = 1$ is highly non-trivial. Here we find a constant C_N for each $N \geq 2$ such that

$$|P(2)| \leq C_N \cdot \rho(P) \qquad : P \in \mathcal{P}_N$$

I do not know how to determine C_N and how to find a Riesz measure μ_N whose total variation is C_N while

$$P(2) = \int_0^1 P(t) \cdot d\mu_N(t) \quad : P \in \mathcal{P}_N$$

A.6 On finite dimensional normed spaces.

Let E be a finite dimensional subspace of some dimension n with a basis $e_1, \ldots, e_n$ and let $|| \cdot ||$ be a norm on E. On the unit sphere S^{n-1} in $\mathbf{R}^n$ we get the continuous function

$$\phi(x_1, \ldots, x_n) = ||x_1 e_1 + \ldots + x_n e_n||$$

Since it is everywhere positive there exist constants $0 < a < A$ such that the range $\phi(S^{n-1})$ is contained in $[a, A]$. From this we conclude that

$$a \cdot \sqrt{\sum |x_k|^2} \leq ||x_1 e_1 + \ldots + x_n e_n|| \leq A \cdot \sqrt{\sum |x_k|^2}$$

Hence the induced norm-topology on E is equivalent to the euclidian topology expressed via a chosen basis in E. It means that on a finite dimensional vector space every pair of norms are equivalent in the sense that they are compared as above by some constant as above.

A special construction. Let X be a vector space of infinite dimension equipped with a norm and consider an infinite sequence $\{x_1, x_2, \ldots\}$ of linearly independent vectors in X. To each $n \geq 1$ we get the n-dimensional subspaces

$$E_n = \{x_1, \ldots, x_n\}$$

generated by the first n vectors. For a fixed n and every vector $y \in X \setminus E_n$ we set

$$d(y; E_n) = \min_{x \in E_n} ||y - x||$$

A.6.1 Exercise. Show that there exists at least one vector $x_* \in E_n$ such that

$$d(y; E_n) = ||y - x_*||$$

Next, construct via an induction over n a sequence of vectors $\{y_n\}$ where each $||y_n|| = 1$ with the following properties. For every n one has

$$E_n = \{y_1, \ldots, y_n\} \quad : \ d(y_{n+1}, E_n) = 1$$

In particular $||y_n - y_m|| = 1$ when $n \neq m$ which entails that the sequence $\{y_n\}$ cannot contain a convergent subsequence and hence the unit ball in X is non-compact. Thus, only finite dimensional normed spaces have a unit ball which is compact with respect to the metric defined by the norm.

A.6.2 A direct sum decomposition. Let X be a normed space and V some finite dimensional subspace of dimension n. Choose a basis $v_1, \ldots, v_n$ in V and via the Hahn-Banach theorem we find an n-tuple $x_1^*, \ldots, x_n^*$ such that $x_j^*(v_k)$ is Kronecker's delta-function. Set

$$W = \{x \in X : x_\nu^*(x) = 0 : \ 1 \leq \nu \leq n\}$$

Then W is a closed subspace of X and we have a direct sum decompotion

$$X = W \oplus V$$

Thus, every finite dimensional subspace of a normed space X has a closed complement. On the other hand a closed infinite dimensional subspace X_0 has in general not closed complenent.

A.7 Frechet spaces.

A pseudo-norm on a complex vector space X is a map ρ from X into the non-negative real numbers with the properties

$$\text{(A.7.1)} \qquad \rho(\alpha \cdot x) = |\alpha| \cdot \rho(x) \quad : \rho(x_1 + x_2) \leq \rho(x_1) + \rho(x_2)$$

where x, x_1, x_2 are vectors in X and $\alpha \in \mathbf{C}$. Notice that (A.7.1) entails that the kernel of ρ is a subspace of X denoted by $\text{Ker}(\rho)$. If it is reduced to the zero vector one says that ρ is a norm.

Suppose now that $\{\rho_n\}$ is a denumerable sequence of pseudo-norms and put

$$\text{(iii)} \qquad d(x, y) = \sum_{n=1}^{\infty} 2^{-n} \cdot \frac{\rho_n(x - y)}{1 + \rho_n(x - y)}$$

Assume that the intersection

$$\cap \ \ker(\rho_n) = \{0\}$$

Then it ks clear that d is a metric. If the resultning metric space is comple, i.e. if every Cauchy-sequence with respect to d converges to a vector in X one refers to X as a Frechet space

The class $\mathcal{K}_X$. It consists of convex subsets K which contrain the origin and are absorbing in the sense that for every vector $x \in X$ there exists some $a > 0$ such that $ax \in K$. If N is a positive integer we set

$$NK = \{Nx : x \in K\}$$

A.7.2 Theorem. *Let (X, d) be a Frechet space. Then every d-closed set $K \in \mathcal{K}_X$ contains an open neighborhood of the origin.*

Proof. To each integer $N \geq 1$ we have the closed set $F_N = N \cdot K$. Since K is absorbing the union of these F-sets is equal to X. Baire's theorem applied to the complete metric space (X, d) yields some N and $\epsilon > 0$ such that F_N contains an open ball of radius ϵ centered at some $x_0 \in F_N$. If $d(x) < \epsilon$ we write

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

Now F_N contains $x_0 + x$ and $x_0 - x$. By symmetry it also contains $-(x_0 - x)$ and the convexity entails that $x \in F_N$. Hence one has the implication

$$d(x, 0) < \epsilon \implies x \in N \cdot K$$

Now the d-metric is defined by a sequence of pseudo-norms $\{\rho_n\}$. Choose an integer M where $2^{-M} < \epsilon/2$. The construction of d shows that if $\rho_n(x) < \frac{\epsilon}{2M}$ hold for $1 \leq n \leq M$, then $d_\rho(x) < \epsilon$. Hence (i) gives

$$\max_{1 \leq n \leq M} \rho_n(x) < \frac{\epsilon}{2M} \implies x \in N \cdot K$$

After a scaling one has

$$\max_{1 \leq n \leq M} \rho_n(x) < \frac{\epsilon}{2MN} \implies x \in K$$

The reader can check that there exists $\epsilon_* > 0$ such that

$$d(x, 0) < \epsilon_* \implies \max_{1 \leq n \leq M} \rho_n(x) < \frac{\epsilon}{2MN}$$

and conclude that K contains an open neighborhood of the origin.

A.7.3 The open mapping theorem. *Consider a pair of Frechet spaces X and Y and let*

$$u \colon X \to Y$$

be a continuous linear and surjective map. Then u is an open mapping, i.e. for every $\epsilon > 0$ the u-image of the ϵ-ball with respect to the Frechet metric on X contains an open neighborhood of the origin in Y.

The proof is left as an exercise to the reader where the hint is to employ Baire's theorem and similar arguments as in the proof of the open mapping theoren for Banach spaces.

A.7.4 Closed Graph Theorem. *Let X and Y be Frechet spaces and $u \colon X \to Y$ is a linear map. Set*

$$\Gamma(u) = \{(x, u(x))\}$$

and suppose it is a closed subspace of the Frechet space $X \times Y$. Under this hypothesis it follows that u is continuous.

Exercise. Verify the closed graph theorem. A hint is that one has the bijective linear map from X onto $\Gamma(u)$ defined by $x \mapsto (x, u(x))$. Since $\Gamma(u)$ is closed in $X \times Y$ it is a Frechet space so the mapping above is open and from this one easily checks that u is continuous.

§ 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:

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

The smallest constant C above is the norm of γ. In this way the continuous linear forms give vectors in a normed vector space denoted by X^*. 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. The reader may verify that of $\widehat{X}$ is the completition of a normed space X, then its dual is equall to that 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. one has the restriction map

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

Since a restricted linear form cannot increase the norm one has the 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 gives a subspace of X^* denoted by $Y^\perp$. It can be identfied with the dual of a new normed space. Namely, consider the quotient space

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

Vectors in Z are images of vectors $x \in X$ where a pair of 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

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

From the constructions above the reader can verify that one has a canonical isomorphism

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

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

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

One refers to γ^* as a norm-preserving extension of γ and we recall that this result was proved in § 4.A.x.

5.2 An exact sequence. Let $Y \subset X$ be a closed subspace and put $Z = \frac{X}{Y}$. The Hahn-Banach Theorem identifies Z^* with $Y^\perp$ and gives an exact sequence

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

where the restriction map $X^* \to Y^*$ sends the unit ball in X^* onto the unit ball of Y^*.

5.3 Example Let $X = L^1(T)$ be the normed space of integrable functions on the unit circle. Recall from measure theory 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 that

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

In other words

(*) $$H^\infty(T) \simeq \frac{L^\infty(T)}{H_0^1(T)}$$

Consider now some $g \in L^\infty(T)$ and put

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

The Hahn-Banach theorem gives some $h \in H^\infty(T)$ such that the L^∞-norm norm $||g - h||_\infty = C$. Thus, we can write

(1) $$g = h + f$$

where $h \in H^\infty$ and the L^∞-function f has norm C. The function h is not determined because there exist several decompositions in (1). However, in § XX we shall find a specific decompositions of g in certain cases.

5.4 The weak star topology

Let X be a normed space. On the dual X^* there is a topology where a fundamental system of open neighborhood of the origin in the vector space X^* consista of sets

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

Let Y be the finite dimensional subspace of X generated by $x_1, \ldots, x_n$. It is clear that the kernel of $\mathfrak{res}_Y$ is contained in the U-set above. If k is the dimension of Y, then linear algebra entials 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. Thus, weak-star open neighborhoods of the origin in X^* contain subspaces with a finite codimension which for an infinite dimensional space means that the weak-star topology is coarse. However, thanks to the Hath-Banach theorem it yields a Hausdorff topology. Next, let S^* be the unit ball in X^*, i.e. the set of linear functionals on X with norm ≤ 1. Using Tychonoff's theorem in general toplogy the reader should confirm :

5.4.1 Theorem. *The unit ball S^* is compact in the weak star topology.*

The case when X is separable. This means that there exists a denumerable dense set $x_1, x_2, \ldots$ in X. On S^* we define a metric by

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

Exercise. Verify that the weak-star topology on S^* is equal to the topology defined by the metric above. So when X is separable, then S^* is a compact metric space in the weak topology. Notice also that the topology defined by the metric in (5.4.2) does not depend upon the chosen dense subsequence in X.

5.4.3 A warning. If X is not separable the weak star topology can be "nasty". More precisely, we can find a non-separable Banach space X such that S^* contains a denumerable sequence $\{x_n^*\}$ which does not contain any convergent subsequence. In other words, S^* is not sequentially compact. To get such an example one employs a wellknown construction in general topology which produces a compact Hausdorff space Z containing a denumerable sequence of points $\{z_n^*\}$ which has no convergent subsequence. Let X be the Banach space of bounded complex-valued functions on Z. The dual X^* is the space of Riesz measures on Z. In S^* we find the sequence of unit point masses $\{\delta(z_n)\}$.

5.4.4 Exercise. Use the fact from topology which assers that a compact Hausdorff space is normal to show that if $\{z_n\}$ is a sequence of points in Z such that

$$\lim_{n\to\infty} |, f(z_n) = \int_Z f\, d\mu$$

hold for a Riesz measure μ and all $f \in C^0(Z)$, then $\{z_n\}$ must converge to a point z_* in Z and $\mu = \delta(z_*)$. So with an "ugly compact space" Z as above we get an example of a Banach space X for which S^* is not sequentially compact.

5.4.5 Weak hulls in X^*. Assume now that X is 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-star topology. This is proved in the next exercise.

5.4.6 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 should construct 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-star topology.

5.4.7 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 $x \in X$.

5.4.8 An 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

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

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

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

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.

5.5 The weak topology on X

.

Let X be a Banach space. In the weak topology on X a fundaemtnal system of open neighborhoods of the origin consist of sets

$$U(x_1^*, \ldots, x_N^*; \epsilon) = \{x \in X : |x_\nu^*(x)| < \epsilon\}$$

where $\{x_\nu^*\}$ are finite subsets of X^*.

5.5.1 Weakly convergent sequences. A sequence $\{x_m\}$ in X converges weakly to a limit vector x if

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

5.5.2 Exercise. Apply Baire's theorem and show that a weakly convergent sequence $\{x_k\}$ is bounded, i.e. there exists a constant C such that

$$||x_k|| \leq C \quad : \; k = 1, 2, \ldots$$

5.5.3 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

$$\lim_{n\to\infty} \int_0^1 x_n(t)\cdot d\mu(t) = 0 \quad \text{hold for every Riesz measure} \quad \mu$$

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. One can construct many such pointwise convergent sequences which fail to converge in the maximum norm.

5.5.4 Remark. Let X be an infinite dimensional Banach space. Then the norm-topology is always strictly stronger than the weak topology. To prove this we argue as follows. By the construction of the weak topology on X its equality with the norm topology gives a finite subset $x_1^*, \ldots, x_N^*$ of X^* and a constant C such that one has the implication

$$\max_\nu |x_\nu^*(x)| < C \implies ||x|| < 1 \quad : x \in X$$

But then the Hahn-Banach theorem implies that the complex vector space X^* is generated by the n-tuple $x_1^*, \ldots, x_N^*$ which entails that X has dimension N at most and contradicts the assumption that X has infinite dimension.

5.6 The bidual X^{**} and reflexive spaces.

Let X be a normed space. The dual of X^* is denoted by X^{**} and called the bidual of X. Each $x \in X$ yields a bounded linear functional on X^* by

$$\widehat{x}(x^*) = x^*(x)$$

The Hahn Banach theorem implies that when x is given, then there exists $x^* \in X*$ such that $x^*(x) = ||x||$. This implies that the norm of $\widehat{x}$ taken in X^{**} is equal to x. We can express this by saying that the bidual embedding $x \mapsto \widehat{x}$ is norm preserving and the image of X in X^{**} is denoted by $\mathfrak{i}(X)$. If $\mathfrak{i}$ is surjective one says that X is reflexive. Recall that dual spaces always are complete. So every reflexive normed space must be a Banach space. But the conveerse is not true, i.e there exists non-reflexive Banach spaces. An example is to take $X = \mathbf{c}_0$ in which case $X^* = \ell^1$ and $X^{**} = \ell^\infty$. An example of a reflexive Banach space is ℓ^p where $1 < p < \infty$. The vectors are sequences of complex numbers $x_1, x_2, \ldots$ for which

$$||x||_p = \big(\sum_{\nu=1}^\infty |x_\nu|^p\big)^{\frac{1}{p}} < \infty$$

Hölder's inequality entails that the dual space of ℓ^p is ℓ^q where $q = \frac{p-1}{p}$, and from which it is clear that ℓ^p is reflexive.

5.6.1 Condition for X to be reflexive. Let X be a normed space whose unit ball is denoted by S. The bidual embedding identifies S with a subset of S^{**}. We shall analyze how S can deviate from S^{**}. For this purpose we consider a finite set $x_1^*, \ldots, x_n^*$ in X^* and if $\gamma \in S^{**}$ we get complex numbers $\{c_i = \gamma(x_i^*)\}$. Then, if $\alpha_1, \ldots, \alpha_n$ is some n-tuple of complex numbers we have

$$|\sum \alpha_i c_i| = |\gamma(\sum \alpha_i x_i^*)| \leq ||\gamma|| \cdot ||\sum \alpha_i x_i^*|| \leq ||\sum \alpha_i x_i^*||$$

where the last inequality follows since γ has norm one at most. The general result in § xx gives for each $\epsilon > 0$ some $x \in X$ with $||x|| < 1 + \epsilon$ and

$$\widehat{x}(x_i^*) = x_i^*(x) = c_i = \gamma(x_i^*)$$

Let us now assume that S equipped with the weak topology is compact. The same holds for $2S$ and taking from (x) we find a sequence $\{x_n\}$ wih norms $||x_n|| < 1 + 1/n$. The assumed compactness gives a limit vector $x_* \in X$ of norm ≥ 1 where $x_n \overset{w}{\to} x$. Moreover, this weak convergence entials that

$$\widehat{x_*}(x_i^*) = \gamma(x_i^*)$$

hold for each $1 \leq i \leq n$. Here we started with a finite subset of X^*. At this stage the reader should verify that when S is compact in the weak topology then we can find $x \in S$ such that $\widehat{x} = \gamma$. This proves that the bidual embdedding is surjective, i.e. X is reflexive. Summing up we have proved:

5.6.2 Theorem. *If S is compact with respec to the weak toplogy, then X is reflexive.*

5.7 The Eberlein-Smulian theorem.

Let X be a Banach space. A subset A is called weakly sequentially compact if every countable sequence $\{x_n\}$ in A contains at least one subsequence which converges weakly to a limit vector x. Here it is not required that x belongs to A. Now we prove a highly non-trivial result due to Eberlein and Smulian.

5.7.1 Theorem. *Let A be a subset of a Banach space X. Then A is weakly sequentially compact if and only if its closure taken in the weak topology is weakly compact.*

Before we enter the proof we notice that the example in § xx. shows that a similar result does not hold when we regard the weak star topology on X^*. Let us also point out that the theorem above is valid in general, i.e. we have not assumed that X is separable and the weak topology on X need not be metrizable. The proof of Theorem 5.7.1 requires several steps. Assume first that A is weakly sequentally compact. We leave it to the reader that this implies that A is bounded. So without loss of generality we can assume that A is contained in the unit ball S in X. Next, denote by $w(A)$ its weak closure. To prove that $w(A)$ is compact in the weak topology on X we shall use the bidual embedding. By definition the weak star topology on X^{**} restricts to the weak topology on the subsapace $\mathfrak{i}(X)$. In particular the bi-dual embedding is continuous when X is equipped with the weak topology and X^{**} with the weak star topology. It follows that

$$\text{(i)} \qquad \mathfrak{i}(w(A)) \subset \overline{\mathfrak{i}(A)}$$

where the right hand side is the weak star closure of $\mathfrak{i}(A)$. Since $A \subset S$ and $\mathfrak{i}$ is norm preserving we have

$$\text{(ii)} \qquad \overline{\mathfrak{i}(A)} \subset S^{**}$$

By Theorem xx the unit ball S^{**} is compact in the weak star topology. So if we have proved the inclusion

$$\text{(iii)} \qquad \overline{\mathfrak{i}(A)} \subset \mathfrak{i}(X)$$

it follows that $\overline{\mathfrak{i}(A)}$ is weakly compact in X and together wtih the inclusion (i) it follows that $w(A)$ is weakly compact.

Proof of (iii). Let λ be a vector in $\overline{\mathfrak{i}(A)}$ and set

$$N(\lambda) = \{x^* \in X^* : \lambda(x^*) = 0\}$$

This is a hyperplane in X^*. If it is closed with respect to the weak star topology on X^* then the observation in § xx gives a vector $x_0 \in X$ such that $\lambda = \widehat{x_0}$ and (iii) follows. So there remains to show that $N(\lambda)$ is weak star closed in X^*. To prove this we shall use a general result which goes as follows where Z is an arbitrary bounded set in X.

Lemma. *Let λ a vector in $\overline{\mathfrak{i}(Z)}$ with norm ≤ 1 and y_0^* a vector in the weak star closure of $N\lambda)$. Then, for each $\epsilon > 0$ there exist sequences $y_1^*, y_2^* \dots$ in $N(\lambda)$ and $x_1, x_2 \dots$ in Z such that the following hold for every $n \geq 1$:*

$$(1) \qquad |y_n^*(x_k) - \lambda(y_0^*)| < \epsilon \quad : 1 \leq k \leq$$

$$(2) \qquad |y_n^*(x_k)| < \epsilon \quad : 1 \leq n < k$$

Exercise. Prove this via an inductive construction.

Apply the Lemma with $Z = A$. The hypothesis on A entails that $\{x_n\}$ has a convergent subsequence with a limit vector x_*. Passing to a subsequence we may assume that (1-2) hold above and $x_n \overset{w}{\to} x_*$. The weak convergence entails that x_* is a strong limit of convex combinations of $\{x_\nu\}$. So with $\epsilon > 0$ kept fixed we find a large positive integer N and real non-negative numbers $a_1, \dots, a_N$ whose sum is one such that

$$\text{(iv)} \qquad ||\xi - x_*|| < \epsilon \quad : \xi = a_1 x_1 + \dots + a_N x_N$$

Now (1) gives

$$\text{(v)} \qquad |\lambda(y_0^*) - y_N^*(\xi)| \leq \sum_{k=1}^{k=N} a_k||\lambda(y_0^*) - y_N^*(x_k)| < \epsilon$$

Next, since y_N^* has norm ≤ 1 we have

$$|y_N^*(\xi) - y_N^*(x_*)| \leq ||\xi - x_*|| < \epsilon$$

From (iv-v) the triangle inequality gives

$$\text{(vi)} \qquad |\lambda(y_0^*)| < 2\epsilon + |y_N^*(x_*)|$$

Funally, since $x_k \overset{w}{x}_*$ it follows that we can take $k > N$ so large that

$$|y_N^*(x_*) - |y_N^*(x_k)| < \epsilon$$

Above $k > N$ and (2) entails that $|y_N^*(x_k)| < \epsilon$. So another application of the triangle inequality gives

$$|\lambda(y_0^*)| < 2\epsilon + 2\epsilon = 4\epsilon$$

In Lemma xx we can take ϵ arbitrarily small and conclude that $\lambda(y_0^* = 0$. Hence $y_0^* \in N(\lambda)$ and since y_0^* was an arbitrary vector in the weak star closure of $N(\lambda)$ we have proved that this hyperplane is weak star closed.

Proof of the converse. There remains to show that if $w(A)$ is weakly compact then it is sequentially compact. DO IT ...

5.7.2 Applications of Theorem 5.7.1 Let A be a subset of the Banach space X. We construct its convex hull $\text{co}(A)$ and pass to its closure in the norm topology, i.e. we get the norm-closed set $\overline{\text{co}(A)}$. With these notations one has the result below which also is due to Eberlein and Smulian.

5.7.3 Theorem. *If A is weakly compact it follows that $\overline{\text{co}(A)}$ also is weakly compact.*

Proof. By xx the norm closed convex set $\overline{\text{co}(A)}$ is weakly closed and even equal to the weak closure of $\text{co}(A)$. By Theorem 5.7.1 there remains to show that $\text{co}(A)$. is weakly sequentially compact. So let u sconsider a suence of points $\{p_n\}$ in $\text{co}(A)$. Each p_n is a finite convex combination of points in A denoted by B_n and we set

$$B^* = \cup B_n$$

The countable set B^* generates a separable closed subspace X_0 of X. CONTINUE PROOF ....

5.8 The Krein-Smulian theorem.

Articles by these authors from the years around 1940 contain a wealth of results. A major theorem from their work goes as follows. Let X be a Banach space and X^* its dual. In (B.2) we constructed the weak star topology. Next, the *bounded weak-star topology* is defined as follows. Let S^* be the open ball of vectors in X^* with norm < 1. If n is a positive integer we get the ball nS^* of vectors with norm $< n$. A subset V of X^* is open in the bounded weak-star topology if and only if the interesections $V \cap nS^*$ are weak-star open for every positive integer n. In this way we get a new topology on X^* whose corresponding topological vector space is denoted by X^*_{bw}, while X^*_w denotes the topological vector space when X^* is equipped with the weak topology. Notice that the family of open sets in X^*_{bw} contains the open sets in X^*_w, i.e. the bounded weak-star topology is stronger. Examples show that the topologies in general are not equal.

Next, let λ be a linear functional on X^* which is continuous with respect to the weak-star topology. This gives by definition a finite set $x_1, \ldots, x_M$ in X such that if $|x^*(x_\nu)| < 1$ for each ν, then $\lambda(x^*)| < 1$. This implies that the subspace of X^* given by the common kernels of $\widehat{x}_1, \ldots \widehat{x}_M$ contains the λ-kernel and linear algebra gives an M-tuple of complex numbers such that

$$\lambda = \sum c_\nu \cdot \widehat{x}_\nu$$

We can express this by saying that the dual space of X_w^* is equal to $\widehat{X}$, i.e. every linear functional on X^* which is continuous with respect to the weak-star topology is of the form $\widehat{x}$ for a unique $x \in X$. Less obvious is the following:

5.8.1. Theorem. *The dual of X_{bw}^* is equal to $\widehat{X}$.*

Proof. For each finite subset A of X we put

$$\widehat{A} = \{x^* : \max_{x \in A} |x^*(x)| \leq 1\}$$

Let U be an open set in X_{bw}^* which contains the origin and S^* is the closed unit ball in X^*. The construction of the bounded weak-star topology gives a finite set A_1 in X such that

(i) $$S^* \cap \widehat{A}^0 \subset U$$

Next, let $n \geq 1$ and suppose we have constructed a finite set A_n where

(ii) $$nS^* \cap \widehat{A_n} \subset U$$

To each finite set B of vectors in X with norm $\leq n^{-1}$ we notice that

(iii) $$\widehat{A_n \cup B} \subset \widehat{A_n}$$

Put

$$F(B) = (n+1)S^* \cap \widehat{A_n \cup B} \cap (X^* \setminus U)$$

It is clear that $F(B)$ is weak-star closed for every finite set B as above. If these sets are non-empty for all B, it follows from the weak-star compactness of $(n+1)S^*$ that the whole intersection is non-empty. So we find a vector

$$x^* \in \bigcap_B F(B)$$

Notice that $F(B) \subset \widehat{B}$ for every finite set B as above which means that $|x^*(x)||leq1$ for every vector x in in X of norm $\leq n^{-1}$. Hence the norm

$$||x^*|| \leq n$$

But then (iii) gives the inclusion

(iv) $$x^* \in nS^* \widehat{A_n}) \cap (X \setminus U)$$

This contradicts (ii) and hence we have proved that there exists a finite set B of vectors with norm $\leq n^{-1}$ such that $F(B) = \emptyset$.

From the above it is clear that an induction over n gives a sequence of sets $\{A_n\}$ such that (ii) hold for each n and

(v) $$A_{n+1} = A_n \cup B_n$$

where B_n is a finite set of vectors of norm $\leq n^{-1}$.

Final part in the roof of the Krein-Smulian theorem. Let θ be a linear functional on X^* which is continuous with respect to the bounded weak-star topology. This gives an open neighborhood U in X_{bw}^* such that

(i) $$|\theta(x^*) \leq 1 \quad : \; x^* \in U$$

To the set U we find a sequence $\{A_n\}$ as above. Let us enumerate the vectors in this sequence of finite sets by $x_1, x_2, \ldots$, i.e. start with the finite string of vectors in A_1, and so on. By the inductive construction of the A-sets we have $||x_n|| \to 0$ as $n \to \infty$. If x^* is a vector in X^* we associate the complex sequence

$$\ell(x^*) = \{x^*(x_n)\}$$

which tends to zero since $||x_n|| \to 0$ as $n \to \infty$. Then

$$x^* \mapsto \ell(x^*)$$

is a linear map from X^* into the Banach space $\mathbf{c}_0$. If

$$\max_n |x^*(x_n)| \leq 1$$

we have by definition $x^* \in A_n^0$ for each n. Choose a positive integer N so that $||x^*|| \leq n$. Thus entails that

$$x^* \in NS^* \cap A_N^0$$

From (ii) during the inductive construction of the A-.sets, the last set is contained in U. Hence $x^* \in U$ which by (i) gives $\theta(x^*)| \leq 1$. We conclude that θ yields a linear functional on on the image space of the ρ-map with norm one at most. The Hahn-Banach theoren gives $\lambda \in \mathbf{c}_0^*$ of norm one at most such that

$$\theta(x^*) = \lambda(\ell(x^*))$$

Next, by a wellknown result due to Banach the dual of $\mathbf{c}_0$ is ℓ^1. Hence there exists a sequence $\{\alpha_n\}$ in ℓ^1 such that

$$\theta(x^*) = \sum \alpha_n \cdot x^*(x_n)$$

In X we find the vector $x = \sum \alpha_n \cdot x_n$ and conclude that $\theta = \widehat{x}$ which proves the Krein-Smulian theorem.

5.9 A result by Pietsch .

The result below illustrates the usefulness of regarding various weak topologies. Let T be a bounded linear operator on a Banach space X and $\{p_n(z)\}$ is a sequence of polynomials with complex coefficients where $p_n(1) = 1$ for each n. We get the bounded operators

$$A_n = p_n(T) \quad : \; n = 1, 2, \ldots$$

Suppose that

(i) $$\lim_{n \to \infty} A_n(x) - A_n(T(x)) \xrightarrow{w} 0$$

hold for every $x \in X$ where the superscript w mens that we regard weak convergence. In addition to (i) we assume that for every $x \in X$, the sequence $\{A_n(x)\}$ is relatively compact with respect to the weak topology. Under these two assumptions one has:

5.9.1 Theorem. *For every $x \in X$ the sequence $\{A_n(x)\}$ converges weakly to a limit vector $B(x)$ where B is a bounded linear operator on X. Moreover B is an idempotent, i.e. $B = B^2$ and one has a direct sum decomposition*

$$X = \overline{(E-T)(X)} \oplus \ker(E-T)$$

where E is the identity operator on X and $\overline{(E-T)(X)}$ is the closure taken in the norm topology of the range of $E-T$. Finally, $\ker(E-T)$ *is equal to the range $B(X)$ while* $\ker(B) = \overline{(E-T)(X)}$.

The proof in § xx and gives an instructive lesson of "duality methods" while infinite dimensional normed spaces are considered.

6. Fredholm theory.

From now on X and Y are Banach spaces with dual spaces 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

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

Exercise. Show that the Hahn-Banach theorem gives the equality of operator norms:

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

6.2 The operator $\bar{u}$. The bounded linear operator u has a kernel denoted by N_u in X. One also refers to N_u as the null space of u. Since u is bounded it is clear that N_u is a closed subspace of X and there exists the Banach space $\frac{X}{N_u}$. Now one has the induced linear operator

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

By construction $\bar{u}$ is an *injective* linear operator with the same range as u:

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

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

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

Consider a pair $y^* \in Y^*$ and $x \in N_u$. Then

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

The proves the inclusiuon

(ii) $$u^*(Y^*) \subset N_u^\perp$$

Next, the Hahn-Banach theorem gives the canonical isomorphism

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

Now we consider the linear operator $\bar{u}$ from (6.2.1) whose adjoint $\bar{u}^*$ maps Y^* into the dual of $\frac{X}{N_u}$. The canonical isomorphism (iii) gives a linear map

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

From this and (ii) we get the equality

(6.3.1) $$\text{Im}(\bar{u}^*) = \text{Im}(u^*)$$

where both sides appear as subspaces of $N_u^\perp$.

6.4 The closed range property

A bounded linear operator $u\colon X \to Y$ is said to have closed range if $u(X)$ is a closed subspace of Y. When this holds

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

is a bijective map between Banach spaces and the The Open Mapping Theorem implies that this is an isomorphism of Banach spaces. We use this to prove:

6.4.1 Proposition. *If u has closed range then u^* has closed range and one has the equality*

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

Proof. Using (6.3.1) we can replace u by $\bar{u}$ and assume that $u\colon X \to Y$ is injective. Then $N_u^\perp = Y^*$ and we must show that u^* is surjective, i.e.

(i) $$u^*(Y^*) = X^*$$

To prove (i) this we use the assumption that u has closed range which by the Open Mapping theorem gives a constant $c > 0$ such that

(ii) $$||u(x)|| \geq c \cdot ||x|| \quad : x \in X$$

If $x^* \in X^*$ the injectivity of u gives a linear functional ξ on $u(X)$ defined by

(iii) $$\xi(u(x)) = x^*(x)$$

Now (ii) entails that ξ belong to $u(X)^*$ with norm $\leq c \cdot ||x^*||$. The Hahn-Banach theorem applied to the subspace $u(X)$ of Y gives a norm preserving extension $y^* \in Y^*$ where (iii) entails that

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

This means that $u^*(y^*) = x^*$ and the requested surjectivity follows.

6.4.2 The converse result. Let $u\colon X \to Y$ be a bounded linear operator and assume that u^* has closed range. Then we shall prove that u has closed range. To begin with we reduce the proof to the case when u is injective. For if $X_0 = \frac{X}{\ker(u)}$ we have the induced linear operator

$$u_0\colon X_0 \to Y$$

where $u_0(X_0) = u(X)$ and at the same time

$$u_0^*\colon Y \to X_0^*$$

where we recall that

$$X_0^* = \ker(u)^\perp = \{x^* \in X^* \colon x^*(\ker(u)) = 0\}$$

Here $u_0^*(Y^*)$ can be identified with the closed subspace $u^*(Y)$ in X_0^* which entails that we reduce the proof to the case when u is injective. From now on u is injective and consider the image space $u(X)$ whose closure yields a Banach space $\overline{u(X)}$. Here

$$u\colon X \to \overline{u(X)}$$

is a linear operator whose range is dense. Let us denote this operator with T. The adjoint

$$T^*\colon\ \overline{u(X)} \to X^*$$

and we have seen that

$$\overline{u(X)} = \frac{Y^*}{\ker(u^*)}$$

In particular the T^*-image is equal to $u^*(Y^*)$ and hence T^* has a closed range. The requuested closedness of $u(X)$ follows if we show that

$$T\colon X \to \overline{u(X)}$$

is surjective. Hence, we have reduced the proof of to the following:

6.4.3 Proposition. *Let $T\colon X \to Y$ be injective where $T(X)$ is dense in Y and T^* has closed range. Then $T(X) = Y$.*

Proof. Let y be a non-zero vector in Y and put

$$\{y\}^\perp = \{y^* \in Y^* \colon\ y^*(y) = 0\}$$

Consider also the image space

$$V = T^*(\{y\}^\perp)$$

Let us first show that

(i) $$V \neq X^*$$

To prove (i) we choose y^* in Y^* such that $y^*(y) = 1$ and get the vector $T^*(y^*)$. If $V = X^*$ this gives some $\eta \in \{y\}^\perp$ such that $T^*(y) = T^*(\eta)$. This means that

$$y^*(Tx) = \eta(Tx) \quad : x \in X$$

The density of $T(X)$ implies that $y^* = \eta$ which is a contradiction since $\eta(y) = 0$.

Next we show that V is closed in the weak-star topology on X^*. By the Krein-Smulian theorem the weak-star closedness follows if V is closed in X^*_{bw}. So let S be the unit ball in X and $\{\xi_n\}$ is a sequence in $V \cap S^*$ where $\xi_n \overset{w}{\to} x^*$ for some limit vector x^*. The Open Mapping Theorem applies to the operator $T^* \colon \{y\}^\perp \to X^*$ and gives a constant C and a sequence $\{y_n^* \in \{y\}^\perp\}$ such that $||y_n^*|| \leq C$ and $T^*(y_n) = \xi_n$. By weak-star compactness for bounded sets in $\{y\}^\perp$ we can pass to a subsequence and assume that y_n^* converge in $\{y\}^\perp$ to a limit vector y^* in the weak star topology. In particular we can apply this to every vector Tx with $x \in X$ and get

$$y^*(Tx) = \lim y_n^*(Tx)$$

This entails that

$$T^*(y^*)(x) = y^*(Tx) = \lim y_n^*(Tx) = \lim T^*(y_n)(x) = \lim \xi_n(x) = x^*(x)$$

Hence $x^* = T^*(y^*)$ which proves that V is weak-star closed in X^*.

We have proved that V is closed in the weak-star topology on X^* and not equal to the whole of X^*. This gives the existence a non-zero vector $x \in X$ such that $\widehat{x}(V) = 0$. So if $y^* \in Y^*$ is such that $y^*(y) = 0$ we have by definition $T^*(y^*) \in V$ and obtain

(ii) $$y^*(Tx) = \widehat{x}(T^*(y^*)) = y^*(Tx) = 0$$

Hence we have the implication:

(iii) $$y^*(y) = 0 \implies y^*(Tx) = 0$$

Finally, since T is injective we have $Tx \neq 0$ and then (iii) gives a complex number α such that $y = \alpha \cdot T(x)$, i.e. the vector y belongs to $T(X)$ as requested.

6.5 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. By the general result about compact metric spaces in § xx an equivalent condition for T to be compact is 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 $y \in Y$.

6.5.1 Exercise. 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 $L(X,Y)$. Verify that T also is a compact operator.

6.5.2 Theorem. *A bounded linear operator T is compact if and only if its adjoint T^* is compact.*

Proof. Assume first that T is compact and let B be the unit ball in X. From the material in § xx this entails that for each positive integer N there exists a finite set F_N in B which is N^{-1}-dense in $T(B)$, i.e. for each $y \in T(B)$ there exists $x \in F_N$ and

(i) $$||T(x) - y|| < N^{-1}$$

Let us then consider a sequence $\{y_n^*\}$ in the unit ball of Y^*. By the standard diagonal procedure we find a subsequence $\{\xi_j = y_{n_j}^*\}$ such that

(ii) $$\lim_{j \to \infty} \xi_j(Tx) \quad : x \in \bigcup_{N \geq 1} F_N$$

Next, if $x \in B$ and $\epsilon > 0$ we choose N so large that $N^{-1} < \epsilon/3$. Since (i) hold for the finite set of points in F_N there exists an integer w such that

(iii) $$|\xi_j(Tx) - \xi_i(Tx)| < \epsilon/3 \quad : j, i \geq w$$

hold for each $x \in F_N$. Since the ξ-vectors have unit norm it follows from (i) and the triangle inequality that (iii) hold for each $x \in B$. Above $\epsilon > 0$ is arbitrary small which entials that $\{\xi_j(Tx)\}$ is a Cauchy seqence of complex numbers for every $x \in B$ and then the same hold for each $x \in X$. Since Cauchy sequences of complex numbers converge there exist limits:

(iv) $$\lim_{j \to \infty} \xi_j(Tx) \quad : x \in X$$

It is clear that these limits values are linear with respect to x. So by the construction of T^* there exist the pointwise limits

(v) $$\lim_{j \to \infty} T^*\xi_j(x) \quad : x \in X$$

which means that there exists $x^* \in X^*$ such that

(vi) $$x^*(x) = \lim_{j \to \infty} T^*\xi_j(x) \quad : \in x \in X$$

The requested compactness of T^* follows if we have proved that the pointwise convergence in (vi) is uniformwhen x stays in B, i.e. that

(vii) $$\lim_{j \to \infty} ||T^*\xi_j - x^*|| = 0$$

To prove that (vi) gives (viii) we apply the Arzela-Ascoli theorem which shows that pointwise convergence on the relatively compact set $T(B)$ of the equicontinuous family of functions $\{\xi_j\}$ gives the uniform convergence in (vii).

Above we proved that if T is compact, so is T^*. To prove the the opposed implication we employ the bi-dual space X^{**}. From the above the compactness of T^* implies that T^{**} is compact. At this stage the reader can check that the restriction of T^{**} to the the closed subspace $j(X)$ of X^{**} under the bi-dual embedding is compact which entails that T is compact.

6.5.3 Operators with finite dimensional range. Suppose that the image space $u(X)$ has a finite dimension N and choose an N-tuple $x_1, \ldots, x_N$ in X such that $\{u(x_k)\}$ is a basis for $u(X)$. In Y^* we can find an N-tuple $y_1^*, \ldots, y_N^*$ such that

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

So if u^* is the adjoint operator then

$$u^*(y_j^*)(x_k) = \text{Kronecker's delta function}$$

If $y^* \in Y^*$ we can therefore find an N-tuple of complex numbers such that

$$u^*(y^* - \sum c_j \cdot y_j^*)(x_k) = 0 \quad : k = 1, \ldots, N$$

This entails that the vector $y^* - \sum c_j \cdot y_j^*)$ belongs to the kernel of u^* and hence the range of u^* is the N-dimensional subapace of X^* generated by $\{u^*(y_j^*)\}$. in particular the adjoint u^* has finite dimensional range.

6.6 Compact pertubations.

We shall prove the following:

6.6.1 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. To begin with the Open Mapping Theoren gives gives a poaitive number c such that

(i) $$||u(x)|| \geq c \cdot ||x||$$

Now we show that the null space N_{u+T} is finite dimensional. By the result in § xx it suffices to show that the set below is relatively compact:

$$V = N_{u+T} \cap B(X)$$

where $B(X)$ is the unit ball in X. Let $\{x_n\}$ be a sequence in V. Since T is compact there is a subsequence $\{\xi_j = x_{n_j}\}$ and some vector y such that $\lim T\xi_j = y$. Since $u(\xi_j) = -T(\xi_j)$ it follows that $\{u(\xi_j)\}$ is a Cauchy sequence and (i) entails that $\{\xi_j\}$ is a Cauchy sequence and hence has a limit vector. This proves that V is relatively compact.

The closedness of Im$(u+T)$. Since N_{u+T} is finite dimensional the result in § xx gives a direct sum decomposition

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

Here $(u+T)(X) = (u+T)(X_*)$ so it suffices that the last image is closed and we can restrict both u and T to X_* where T_* again is compact. Hence we may assume that the operator $u+T$ is *injective*. If a vector y belongs to the closure of $\text{Im}(u+T)$ there exists a sequence ξ_n in X such that

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

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 $\xi_n = \frac{x_n}{||x_n||}$ it follows that

$$\text{(iii)} \qquad \lim\, u(\xi_n) + T(\xi_n) = 0$$

Now $\{\xi_n\}$ is bounded and since T is compact we can pass to another subsequence and assume that $T(\xi_n) \to y$ holds for some $y \in Y$. Then (ii) entails that $u(\xi_n)$ also has a limit and (i) implies that $\{\xi_n\}$ converges in X to a limit vector ξ_*. Here $\xi_* \neq 0$ since $||\xi_n|| = 1$ for all n. Moreover, (iii) entails that $u(\xi_*) + T(\xi_*) = 0$. This gives a contradiction since N_{u+T} was assumed to be the zero space.

Returning to (i) we now have a bounded sequence $\{x_n\}$. Since T is compact we can pass to a subsequence and assume that $T(x_n) \to \eta$ holds for some $\eta \in Y$. But then (i) entails that the sequence $\{u(x_n)\}$ converges to $y - \eta$. Again (i) implies that $\{x_n\}$ converges to some limit vector ξ. Passing to the limit in (ii) we get

$$u(\xi) + T(\xi) = y$$

Hence y belongs to $\text{Im}(u+T)$ and Theorem 6.6.1 is proved.

6.7 Spectra and resolvents of compact operators.

Let $T\colon X \to X$ be a compact operator. For each complex number $\lambda \neq 0$ we set

$$\mathcal{N}(\lambda) = \{x : Tx = \lambda \cdot x\}$$

6.7.1 Theorem. *The set of non-zero λ for which $N(\lambda)$ contains a non-zero vector is discrete.*

Proof. Suppose that there exists a non-zero cluster point $\lambda_0 \neq 0$, i.e. a sequence $\{\lambda_n\}$ where $\lambda_n \to \lambda_0$ and for each n a non-zero vector $x_n \in \mathcal{N}(\lambda_n)$. Since the numbers $\{\lambda_n\}$ are distinct and $T(x_n) = \lambda_n \cdot x_n$ hold one easily verifies that the vectors $\{x_n\}$ are linearly independent. By the result in § xx we find a sequence of unit vectors $\{y_n\}$ satisfying the separation in Theorem § xx and

$$\{x_1, \ldots, x_n\} = \{y_1, \ldots, y_n\}$$

hold for each n. Now

$$\text{(i)} \qquad T(\lambda_n^{-1} y_n) - \lambda_m^{-1} T(y_m) = y_n - y_m \in \{y_1, \ldots, y_{n-1}\} \quad : n > m$$

At the same time the separation gives

$$\text{(ii)} \qquad ||y_n - y_m|| \geq 1 \quad : n > m$$

Since λ_n converge to the non-zero number λ_0 this entails that $\{\lambda_n^{-1} y_n\}$ is a bounded sequence and from (i-ii) we see that $\{T(\lambda_n^{-1} y_n)\}$ cannot contain a convergent subsequence. This contradiction proves Theorem 6.7.1

6.7.2 Theorem. *The spectrum of a compact operator T is discrete outside the origin.*

Proof. Consider a non-zero $\lambda_0 \in \sigma(T)$. By Schauder's result in (xx) the adjoint T^* is also compact. Hence Theorem 6.7.1 applies to T and T^* which gives a small punctured disc $\{0 < |\lambda - \lambda_0| < \delta\}$ such that $\lambda \cdot E - T$ and $\lambda \cdot E^* - T^*$ both are injective when λ belongs to the punctured disc. Next, $\lambda \cdot E - T$ has a closed range by Theorem 5.X. If it is a proper subspace of X find a non-zero $x^* \in X^*$ which vanishes on this range. The construction of T^* entails that $T^*(x^*) = \lambda \cdot x^*$. But this was not the case and hence $\lambda \cdot E - T$ is surjective which shows that λ is outside $\sigma(T)$ and finishes the proof of Theorem 6.7.2.

6.7.3 Spectral projections. Let λ_0 be non-zero in $\sigma(T)$. By Theorem 6.7.2 it is an isolated poiunt in $\sigma(T)$ which gives the spectral projection $E_T(\lambda_0)$ where we recal, from § xx that this operator commutes with T. So if $V = E_T(\lambda_0)(X)$ then T restricts to a bounded linear operator on V denoted by T_V where we recall from § xx that the spectrum of T_V is reduced to the singleton set $\{\lambda_0\}$. Since $\lambda_0 \neq 0$ it means that T_V is an invertible and compact operator on V. So by the result in § xx V is finite dimensional. This finiteness and linear algebra applied to T_V gives an integer $m \geq 1$ such that

$$(Tx - \lambda_0 x)^m = 0 \quad : x \in V \tag{6.7.3.1}$$

6.8 Fredholm operators.

A bounded linear operator $u\colon X \to Y$ is called a Fredholm operator if it has closed range and the kernel and the cokernel of u are both finite dimensional. When u is Fredholm its index is defined by:

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

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

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

The proof requires several steps where the crucial point is to regard the case $X = Y$ and a compact pertubation of the identity operator. Thus we begin with:

6.8.2 Theorem. *Let $T\colon X \to X$ be compact. Then $E - T$ is Fredholm and has index zero.*

Proof. Apply (6.7.3) with $\lambda_0 = 1$ which gives the decompostion

$$X = E_T(1)(X) \oplus (E - E_T(1))(X) \tag{i}$$

Theorem 6.6.1 implies that $E - T$ has closed range. Next, from the spectral decomposition in (6.7.3) it follows that $E - T$ restricts to a bijective operator on $(E - E_T(1))(X)$ which by (i) entails that the codimension of $(E - T)(X)$ is at most the dimension of the finite dimensional vector space $V = E_T(1)(X)$. Moreover, the kernel of $E - T$ is finite dimensional by Theorem 6.6.1.

Hence we have proved that $E - T$ is a Fredholm operator and there reamins to show that its index is zero. To obtain this we notice again that the decomposition (i) impies that this index is equal to that of the restricted operator $E - T$ to the finite dimensional vector space V. Finally, recall from linear algebra that the index of a linear operator on a finite dimensional vector space always is zero which finishes the proof.

6.8.3 The general case. Consider first the case when the kernel of the Fredholm operstor u is zero. Now there exists a finite dimensional subspace W of Y such that

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

and here $u\colon X \to u(X)$ is an isomorphisn between the Banach spaces X and $u(X)$ which gives the existence of a bounded inverse operator

$$\phi\colon u(X) \to X$$

So here $\phi \circ u$ is the identity on X. Next, given a compact operator T we consider the projection operator $\pi\colon Y \to u(X)$ whose kernel is W and notice that $\pi \circ T$ is a compact operator. Now we can regard the operator

$$u + \pi \circ T\colon X \to u(X)$$

From the above and Theorem 6.8.2 the reader can verify that this Fredholm operator has index zero. Next, we notice that

$$\mathfrak{ind}(u) = -\dim W \tag{1}$$

We have also the operator

$$T_* = (E_Y - \pi) \circ T\colon X \to W$$

The direct sum decomposition (xx) entails that

$$\ker(u + T) = \ker(u + \pi \circ T\colon X) \cap \ker T_* \tag{2}$$

$$\frac{Y}{(u+T)(X)} = \frac{Y}{(u + \pi \circ T)(X)} \oplus \frac{W}{T_*(X)} \tag{3}$$

From (1-3) we leave it as an exercise to show that the index of $u + T$ is equal to that if u given by (1).

Above we treated the case when u is injective. Since N_u is finite dimensional we have a decompostion

$$X = N_u \oplus X_*$$

and now the restricted operator $u\colon X_* \to Y$ is injective. For a given compact operator T we consider the restrcited compact operator $T_*|colon X_* \to Y$ and form the previous special case we have

$$\mathfrak{ind}(u_* + T_*) = \mathfrak{ind}(u_*)$$

Next, let $\pi\colon X \to N_u$ be the projection with kernel X_* which gives

$$T = T_* + T \circ \pi$$

At this stage we leave it as an exercise to verify that (x) gives the requested index formula

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

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 and we suppose that the Banach spaces are real so the dual space Y^* consists of continuous $\mathbf{R}$-linear maps from Y into $\mathbf{R}$. Every $y^* \in Y^*$ yields the real-valued function $y^* \circ g$ on X. With x_0 kept fixed we can impose the condition that there exist limits

$$(1) \qquad \lim_{\epsilon \to 0} \frac{y^* \circ g(x_0 + \epsilon \cdot x) - y^*(g(x_0))}{\epsilon}$$

for each vector $x \in X$. These limits resemble directional derivatives in calculus and we can impose the extra condition that the limits above depend linearly upon x. Thus, assume that each $y^* \in Y$ yields a linear form $\chi(y^*)$ on X such that (1) is equal to $\chi(y^*)(x)$ for every $x \in X$. When this holds it is clear that

$$(2) \qquad y^* \mapsto \chi(y^*)$$

is a linear mapping from Y^* into X^*. When both Y and X are finite dimensional real vector spaces this linear operator corresponds to the usual Jacobian in calculus. In the infinite-dimensional case it is not always true that (2) is continuous with respect to the norms on the dual spaces. As an extra condition for differentaiblity at x_0 we impose the condition that there exists a constant C such that

$$(3) \qquad ||\chi(y^*)|| \leq C \cdot ||y^*||$$

When (3) holds we have a bounded linear operator $\chi\colon Y^* \to X^*$ associated to g and the given point $x_0 \in X$. It may occur that χ is the adjoint of a bounded linear operator from X into Y which means that there exists a bounded linear operator $L\colon X \to Y$ such that

$$(*) \qquad \lim_{\epsilon \to 0} \frac{||g(x_0 + \epsilon \cdot x) - g(x_0) - \epsilon \cdot L(x)||}{\epsilon} = 0$$

Concerning the passage to the limit the weakest condition is that it holds pointwise, i.e. (*) holds for every vector x. A stronger condition is to impose that the limits above hold uniformly which means that

$$(**) \qquad \lim_{\epsilon \to 0} \max_{x \in B(X)} \frac{||g(x_0 + \epsilon \cdot x) - g(x_0) - \epsilon \cdot L(x)||}{\epsilon} = 0$$

where the maximum is taken over X-vectors with norm ≤ 1. In applications the the condition (**) is often taken as a definition for g to be differentiable at x_0 and the uniquely determined linear map L above is denoted by $D_g(x_0)$ and called the differential of g at x_0. If g is a map from some open subset Ω of X with values in Y we can impose the condition that g is differentiable at each $x_0 \in \Omega$ and add the condition that $x \mapsto D_g(x)$ is continuous in Ω where the values are taken in the Banach space of continuous linear maps from X into Y. When this holds we get another map $x \to D_g(x)$ from Ω into $\mathcal{L}(X, Y)$ and can impose the condition that it also is differentiable in the strong sense above. This leads to the notion of k-times continuosly differentiable maps from a Banach space into another for every positive integer k.

Remark. We shall not dwell upon a general study of differentiable maps between Banach spaces which is best illustrated by various examples. For a concise treatment we refer to Chapter 1 in Hörmander's text-book [PDE:1] which contains a proof of the implicit function theorem for differentiable maps between Banach spaces in its most general set-up.

7.1 Line integrals

Let Y be a Banach space. Consider a continuous map g from 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 there exists the Y-valued line integral

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

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 and since Y is complete this gives a unique limit vector in Y.

7.2 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

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

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

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

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 Banach-space version of Cauchy's theorem.

7.3 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 $\bar\Omega$. Then*

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

Moreover, we have the vanishing integral

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

Exercise. Let $g(z)$ be as above and suppose that $\phi(z)$ is a complex-valued analytic function in Ω which extends to a continuous function on its closure. Multiplying the Y-vectors $g(z)$ with the complex scalars $\phi(z)$ we get the Y-valued funtion $z\mapsto\phi(z)g(z)$. Show that this Y-valued function is analytic and verify also the Cauchy formula

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

7.4 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

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

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 $\bar{\Omega}$. We impose the condition

(ii) $$\partial\Omega \cap \sigma(A) = \emptyset$$

Then we can construct the line integral

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

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

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

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:

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

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:

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

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

Recall 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 the operator $T_A \colon f \mapsto f(A)$.

7.6 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. 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 $\bar{\Omega} \setminus \Omega_*$ does not intersect $\sigma(A)$. Consider pair f, g in $\mathcal{A}(\Omega)$. The careful choice of Ω_* and Cauchy's integral formula give the equality

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

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

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

Next we use the Nemann's equation

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

The double integral in (ii) 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:

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

Here ζ stays *outside* the domain Ω_* and hence (iv) is zero by Cauchy's vanishing theorem. So $C_2 = 0$ and (iii) gives the equality in Theorem 7.6.

7.7 The Banach algebra $\mathcal{A}$. This is the closed subagebra of Y generated by A and all the resolvent operators $R_A(z)$ as z moves outside $\sigma(A)$. Let $\mathfrak{M}_A$ denote its Gelfand space.

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

Proof. As explained in § XX the points in $\mathfrak{M}_A$ correspond to multiplicative and linear forms on $\mathcal{A}$. Let γ be such a multiplicative linear functional. If $\gamma(A) = z_*$ for some complex number z_* then the γ-values are determined on all resolvents. In fact, this holds since

$$1 = \gamma((\lambda \cdot E - A) R_A(\lambda)) = (\lambda - z_*) \cdot \gamma(R_A(\lambda))$$

hold when $\lambda \in \mathbf{C} \setminus \sigma(A)$ Notce that this also entails that z_* must belong to $\sigma(A)$. There remains to show that for each point $z_* \in \sigma(A)$ there exists some γ such that $\gamma(A) = z_*$. To prove this we notice that the hypothesis on z_* means that the operator $z_* \cdot E - A$ is not inverible on X. If z_* is outside the range of the Gelfand transform of A in the Banach algebra $\mathcal{A}$, then $z_* \cdot E - A)$ would be invertible in $\mathcal{A}$ whuch gives a contradiction and finishes the proof of Proposition 7.8.

7.9 The sup-norm case. Suppose that $\mathcal{A}$ is a sup-norm algebra and put $K = \sigma(A)$. Let Ω be an open set Ω which contains K. The operational calculus gives an algebra homomorphism

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

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

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

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

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

Moreover it is an isometry, i.e.

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

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

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

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

For example, Theorem §§ in chapter XX 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 $P(z)$. So in this case polynomials in A generate a dense subalgebra of $\mathcal{A}$.

7.11 Uniformly convex Banach spaces.

A Banach space X is uniformly convex if there corresponds to each $0 < \epsilon < 1$ some $\delta(\epsilon)$ tending to zero with ϵ such that

$$\frac{||x+y||}{2} \geq 1 - \epsilon \implies ||x-y|| \leq \delta(\epsilon)$$

for all pair of vectors of norm one at most. This condition was introduced by Clarkson in the article [Clarkson] from 1936.

Exercise. Show that in a uniformly convex Banach space each closed convex set contains a unique vector of minimal norm.

Next, let p be a non-zero vector in X. If $x \neq 0$ is another vector we get the function of a real variable a defined by:

$$a \mapsto ||p + ax|| - ||p||$$

We say that X has directional x-derivative at p if there exists the limit

$$\lim_{a \to 0} \frac{||p+ax|| - ||p||}{a} = D_p(x)$$

Notice that in this limit a can tend to zero both from the negative and the positive side. The following result is due to Clarkson:

Theorem *Let p be a non-zero vector in X such that the directional derivatives above exist for every $x|nX$. Then $x \mapsto D - p(x)$ is linear and the vector $D_p \in X^*$ has norm one with $D_p(p) = 1$.*

Exercise. Prove Clarkson's result.

7.11.2 Conjugate vectors. For brevity we say that X is differentiable if the directional derivatives above exist for all pairs p, x. Let S be the unit sphere in X and S^* the unit sphere in X^*. A pair $x \in S$ and $x^* \in S^*$ are said to be conjugate if

$$x^*(x) = 1$$

7.11.3 Theorem. *Let X be uniformly convex and differentiable. Then every $x \in S$ has a unique conjugate given by $x^* = D_x(x)$ and the map $x \to x^*$ from S to S^* is bijective.*

Exercise. Prove this result or consult Clarkson's article or some text-book.

7.11.4 Duality maps. Let X be uniformly convex and differentiable and $\phi(r)$ is a strictly increasing and continuous function on $r \geq 0$ where $\phi(0) = 0$ and $\phi(r) \to +\infty$ when $r \to +\infty$. Each verctor in X is of the form $r \cdot x$ with $x|inS$ while $r \geq 0$. The duality map above yields a function $\mathcal{D}_\phi$ from X into X^* defined by

$$\mathcal{D}_\phi(rx) = \phi(r) \cdot x^* \quad \text{when} \quad x \in S \quad \text{and} \quad r \geq 0$$

If C is a closed subspace in X we put:

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

7.14 Theorem. *For each closed and proper subspace $C \neq X$ the following hold: For every pair of vectors $x_* \in X$ and $y^* \in X^*$ the intersection*

$$\mathcal{D}_\phi(C + x_*) \cap \{C^\perp + y*\}$$

is non-empty and consists of a single point in X^*.

Proof. Introduce the function

$$\Phi(r)=\int_0^r \phi(s)\,ds$$

Since ϕ is strictly increasing, Φ is a strictly convex function and since $\phi(r)\to+\infty$ we have $\frac{\Phi(r)}{r}\to+\infty$. Consider the functional defined on $C+x_*$ by

$$F(x)=\Phi(||x||)-y^*(x)$$

If $||x||=r$ we have

$$F(x)\geq \Phi(r)-r||y^*||$$

The right hand side is a strictly convex function of r which tends to $+\infty$ and is therefore bounded below. Hence there exists a number

$$\delta=\inf_{x\in C+x_*} F(x)$$

Let $\{x_n\}$ be a minimizing sequence for F. The strict convexity of Φ entails that the norms $\{||x_n||\}$ converge to some finite limit α. Since the set $C+x_*$ is convex we have

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

Next, the convexity of ϕ entails that

$$0\leq\frac{1}{2}\big[\Phi(||x_n||)+\Phi(||x_m||)-\Phi(||\frac{x_n+x_m}{2}||)=$$
$$\frac{1}{2}\big(F(x_n)+F(x_m)\big)-F(\frac{x_n+x_m}{2})$$

Since $\{x_n\}$ is F-minimizing the last terms tend to zero when n and m increase which gives

$$\lim_{n,m\to\infty}\Phi(||\frac{x_n+x_m}{2}||)=\Phi(\alpha) \tag{1}$$

where we recall that

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

Now (1) and the strict convexity of Φ gives

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

The unifom convexity entails that $\{x_n\}$ is a Cauchy sequence which gives a limit point p where $F(p)=\delta$ and consequently

$$F(p+tx)-F(p)\geq 0 \quad : \quad x\in C \tag{2}$$

Since p belongs to $C+x_*$ the existence part in Theorem 7.14 follows if we have proved the inclusion

$$\mathcal{D}_\phi(p)\in C^\perp+y^* \tag{3}$$

To get (3) we use that the Banach space is differentiable and since the Φ-function as a primitive of a continuous function is of class C^1 one has

$$\Phi(||p+tx||)-\Phi(||p||)=\Phi'(||p||)\cdot\mathfrak{Re}\,tD_p(x)+o(|t|)=\phi(\alpha)\cdot\mathfrak{Re}\,tD_p(x)+o(|t|)$$

where t is a small real or complex number. Together with (2) this gives

$$\phi(\alpha)\cdot\mathfrak{Re}\,tD_p(x)+o(|t|)\geq\mathfrak{Re}\,y^*(tx) \quad : \quad x\in C$$

By linearity it is clear that this implies that

$$\phi(\alpha)\cdot D_p(x)=y^*(x)$$

This means precisely that the linear functional $y^*-\phi(\alpha)\cdot D_p$ belongs to $C^\perp$, or equivalently that

$$\phi(||p||)\cdot D_p\in C^\perp+y^*$$

Finally we recall from (7.13) that

$$\mathcal{D}_\phi(p) = \phi(||p||) \cdot D_p(p)$$

and the requested inclusion (3) is proved.

The uniqueness part. Above we proved the existence of at least one point in the intersection from Theorem 7.12. The verification that this set is reduced to a single point is left as an exercise to the reader.

§ 8. Locally convex spaces

Introduction. We shall consider real vector spaces, i.e. a space X whose elements are called vectors which can be added so that X is an abelian group. In addition real numbers yield scalar multiples of vectors. We shall expose basic facts about real vector spaces equipped with a locally convex topology. The definition is given in § 2. A crucial result is the Hahn-Banach theorem for lcoally convex spaces which we describe in § 1. It has several important consequences such as Hörmander's result in Theorem 3.6.

Fixed point theorems. A compact topological space S has the fixed point property if every continuous map $T\colon S \to S$ has at least one fixed point. An example is the closed unit ball in $\mathbf{R}^n$ whose fixed-point property is proved in § 5.xx. More generally, consider a locally convex vector space X. In § 5 we explain the construction of its dual space X^* whose vectors are continuous linear functionals on X. Now one equips X with the weak topology whose open sets are generated by pairs $x^* \in X^*$ and positive numbers δ) of the form:

$$B_\delta(x^*) = \{x \in X : \, |x^*(x)| \leq \delta\}$$

Denote by $\mathcal{K}(X)$ the family of convex subsets of X which are compact with respect to the X^*-topology. In § xx we prove the following two results.

The Schauder-Tychonoff fixed point theorem. *Each K in $\mathcal{K}(X)$ has the fixed point property.*

The merit of this result is of course that one allows non-linear maps. The next result is due to Kakutani and goes as follows: By a group of linear transformations $\mathbf{G}$ on a real vector space X we mean a family of bijective linear maps $g\colon X \to X$ such that composed maps $g_2 \circ g_1$ again belong to the group as well as the inverse of every g.

Kakutani's theorem. *Let $K \in \mathcal{K}(X)$ be $\mathbf{G}$-invariant, i.e. $g(K) \subset K$ hold for every $g \in \mathbf{G}$. Assume in addition that the family of the restricted $\mathbf{G}$-maps to K is equicontinuous. Then there exists at least some vector $k \in K$ such that $g(k) = k$ for every $g \in \mathbf{G}$.*

Remark. The equicontinuous assumption means that to each pair every (x^*, ϵ) with $x^* \in X^*$ and $\epsilon > 0$, there exists a finite family $x_1^*, \ldots, x_M^*$ and some $\delta > 0$ such that the following hold: If p and q is a pair of points in K such that $p - q$ belongs to $\cap B_\delta(x_\nu^*)$, then

$$g(p) - g(q) \in B_\epsilon(x^*)$$

hold for all $g \in \mathbf{G}$.

Haar measures. Let G be a compact topological group which means that the group is equipped with a Hausdorff topology where the group operations are continuous, i.e the map from $G \times G$ into G which sends a pair of group elements g, h to the product gh is continuous, and the inverse map $g \mapsto g^{-1}$ is bi-continuous. Now there exists the Banach apace $C^0(G)$ of continuous real-valued functions on G. Recall from basic measure theory that the dual space consists of Riesz measures. Denote by $P(G)$ the family of non-negative measures with total mass one, i.e. probability measures on G. If $\phi \in C^0(G)$ and $g \in G$ we get the new continuous function $S_g(\phi)$ defined by

$$S_g(\phi)(h) = \phi(gh) \quad : h \in G$$

Next, if $\mu \in P(G)$ we get the new probability measure $T_g(\mu)$ given by the linear functional

$$\phi \mapsto \int_G S_g(\phi)\, d\mu$$

In this way G is identified with a group of transformations on $P(G)$. Next, $P(G)$ is equipped with the weak-star topology where open neighborhoods of a given $\mu \in P(G)$ consists of finite intersections of sets $\{\gamma \in P(G) : |\gamma(\phi) - \mu(\phi)| < \delta\}$ for pairs $\delta > 0$ and $\phi \in C^0(G)$. The uniform continuity of every $\phi \in C^0(G)$ entails that that the group action on $P(G)$ is equi-continuous on $P(G)$ with respect to the weak-star topology. As explained in § xx, Kakutani's theorem also applies in this situation which yields a fixed point. Hence there is a probability measure μ such that

$$(*) \qquad \int_G \phi(gh)\, d\mu(h) = \int_G \phi(h)\, d\mu(h)$$

hold for every pair $g \in G$ and $\phi \in C^0(G)$. In § xx we show that μ is uniquely determined by (*), i.e. only one probability measure enjoys the invariance above. Moreover, starting with the operators

$$S_g^*(\phi)(h) = \phi(hg) \quad : h \in G$$

one finds a probability measure μ^* such that

$$(**) \qquad \int_G \phi(hg)\, d\mu(h) = \int_G \phi(h)\, d\mu(h)$$

hold for every pair $g \in G$ and $\phi \in C^0(G)$. In § xx we prove that $\mu = \mu^*$ which means that the unique Haar measure is both left and right invariant.

§ 1. Convex sets and their ρ-functions.

Let E be a real vector space. A convex set U which contains the origin is said to be absorbing if there for each vector $x \in E$ exists some real $s > 0$ such that $s \cdot x \in U$. It may occur that the whole line $\mathbf{R}x$ is contained in U, and then we say that x is fully absorbed by U. The convexity of U entails that the set of fully absorbed vectors is a linear subspace of E which we denote by $\mathcal{L}_U$.

1.1 The function ρ_U. Let x be a non-absorbed vector x. Then there exists a positive real number

$$\mu(x) = \max\{s : sx \in U\}$$

If x is absorbed we put $\mu(x) = +\infty$ and for every non-zero vector x we set

$$\rho_U(x) = \frac{1}{\mu(x)}$$

it is clear that if $x \in U$ then $\mu(x) \geq 1$ and hence $\rho_U(x) \leq 1$. Notice that we also have

$$\rho_U(x) = \min\{s : \ x \in s^{-1}U\}$$

1.2 Exercise. Show that the convexity of U entails that ρ_U satisfies the triangle inequlity

$$(1.2.1) \qquad \rho_U(x_1 + x_2) \leq \rho_U(x_1) + \rho_U(x_2)$$

for all pairs of vectors in E. Moreover, check also that $\rho_U(x) = 0$ if and only if x belongs to $\mathcal{L}_U$ and that ρ_U is positively homogeneous, i.e. the equality below holds when a is real and positive:

$$(1.2.2) \qquad \rho_U(ax) = a\rho_U(x) \quad : a > 0$$

1.3 The Hahn-Banach theorem. Keeping U fixed we set $\rho(x) = \rho_U(x)$. An $\mathbf{R}$-linear map λ from E to the 1-dimensional real line is majorised by ρ if

$$(1.3.1) \qquad \lambda(x) \leq \rho(x)$$

hold for every vector x. More generally, let E_0 be a subspace of E and $\lambda_0 \colon E_0 \to \mathbf{R}$ a linear map such that (1.3.1) hold for vectors in E_0. Then there exists a linear map $\lambda \colon E \to \mathbf{R}$ which extends λ_0 and is again majorised by ρ. This result was proved in § 4.A.4.

8.2 Locally convex topologies.

Denote by $\mathcal{C}_E$ the family of convex sets U as in § 1. Let $\mathfrak{U} = \{U_\alpha\}$ be a family in $\mathcal{C}_E$ such that

$$\bigcap \mathcal{L}_{U_\alpha} = \{0\}$$

i.e. the intersction is reduced to the origin. Now there exists a topology on E where a basic for open neighborhoods of the origin consists of sets:

$$(1) \qquad \cap \{\rho_{U_{\alpha_i}}(x) < \epsilon\}$$

where $\epsilon > 0$ and $\{\alpha_1, \ldots, \alpha_k\}$ is a finite set of indices defining the U-family. If x_0 is a vector in XS, then a basis for its open neihghborhoods are given by sets of the for $x_0 + U$ where U is a set from (1). In general, a subset Ω in E is open if there to eah $x_0 \in \Omega$ exists some U from (1) such that $x_0 + U \subset \Omega$. it is clear that this gives a topology and (1) entaills that it is separated, i.e. a Huasdorff topology on E. Notice also that eqch set in 82) is convex. One therefore refers to a locally con vex topology on E.

2.1 Remark. The locally convex topology above depends upon the family $\mathfrak{U}$-toplogy. Its topology is not changed if we enlarge the family to consist of all finite intersection of its convex subsets. When this has been done we notice that if $U_1, \ldots, U_n$ is a finite family in $\mathfrak{U}$ then the norm defined by $U = U_1 | \cap \ldots \cap U_n$ is stronger than the individual ρ_{U_i}-norms. Hence a fundamental system of neighborhoods consists of single ρ-balls:

$$\Omega = \{\rho_U < \epsilon\} \quad : U \in \mathfrak{U}$$

2.2 The dual space E^*. Let E be equipped with a locally convex $\mathfrak{U}$-topology. As above $\mathfrak{U}$ has been enlarged so that the balls above give a basis for neighborhoods of the origin. A linear functional ϕ on E is $\mathfrak{U}$-continuous if there exists some $U \in \mathfrak{U}$ and a constant C such that

$$|\phi(x)|| \leq C \cdot \rho_U(x)$$

2.3 Closed half-spaces. To each pair $\phi \in E^*$ and a real number a one assigns the closed half-space

$$H = \{x \in X : \phi(x) \leq a\}$$

Notice that $a < 0$ can occur in which case H does not contain the origin.

2.4 The separation theorem. *Each closed convex set K in E is the intersection of closed half-spaces.*

2.5 Exercise. Show (2.4) using the Hahn-Banach theorem.

Next, let K_1 and K_2 be a pair of closed and disjoint convex sets. Then they can be spearated by a hyperplane. More precisley, there exists some $\phi \in E^*$ and a positive number δ such that

$$\max_{x \in K_1} \phi(x) + \delta \leq \min_{x \in K_2} \phi(x) \tag{2.6}$$

Again we leave the proof as an exercise to the reader.

8.3. Support functions of convex sets.

Let E be a locally convex space as above. Vectors in E are denoted by x, while y denote vectors in E^*. To each closed and convex subset K of E we define a function $\mathcal{H}_K$ on the dual E^* by:

$$\mathcal{H}_K(y) = \sup_{x \in K} y(x)$$

Notice that $\mathcal{H}_K$ take values in $(-\infty, +\infty]$, i.e. it may be $+\infty$ for some vectors $y \in E^*$. For example, let $K = \{\mathbf{R}^+ x_0$ be a half-line. Then $\mathcal{H}_K(y) = +\infty$ when $y(x_0) > 0$ and otherwise zero. So here the range consisists of 0 and $+\infty$. It is clear that

$$\mathcal{H}_K(sy) = s\mathcal{H}_K(y)$$

hold when s is a positive real number, i.e $\mathcal{H}_K$ is positively homogeneous.

3.1 Exercise. Show that the convexity of K entails that

$$\mathcal{H}_K(y_1 + y_2) \leq \mathcal{H}_K(y_1) + \mathcal{H}_K(y_2)$$

for each pair of vectors in E^*.

3.2 Upper semi-continuity. For each fixed vector $x \in E$ the function

$$y \mapsto y(x)$$

is weak-star continuous on E^*. Since the supremum function attached to an arbitrary family of weak-star continuoes functions is upper semi-continuous, it follows that $\mathcal{H}_K$ is upper semi-contiuous.

3.4 Exercise. Let K and K_1 be a pair of closed convex sets such that $\mathcal{H}_K = \mathcal{H}_{K_1}$. Show that this entails that $K = K_1$. The hint is to use the separation theorem.

3.5 The class $\mathcal{S}(E)$. It consids of all all upper semi-continuous functions G on E^* with values in $(-\infty, +\infty]$ which satisfy (x) and (xx). The next result was proved by Hörmander in the article *Sur la fonction d'appui des ensembles convexes dans un espaces localementt convexe* [Arkiv för mat. Vol 3: 1954].

3.6 Theorem. *Each $G \in \mathcal{S}(E)$ is of the form $\mathcal{H}_K$ for a unique closed convex subset K in E.*

Remark. As pointed out by Hörmander in [ibid] this result is closely related to earlier studies by Fenchel in the article *On conjugate convex functions* Canadian Journ. of math. Vol 1 p. 73-77) where Legendre transforms are studied in infinite dimensional topological vector spaces. The novely in Theorem 3.3 is the generality and we remark that various separation theorems in text-books dealing with notions of convexity are easy consequences of Theorem 5.C.2.

Proof of Theorem 3.6 Put $F = E \oplus \mathbf{R}$ which is a new vector space where the 1-dimensional real line is added. It dual space $F^* = E^* \oplus \mathbf{R}$. We are given $G \in \mathcal{S}(E)$ and put

(i) $$G_* = \{(y, \eta) \in E^* \oplus \mathbf{R} \quad : G(y) \leq \eta\}$$

Condition in (*) entails that G_* is a convex cone in F^* and the semi-continuous hypothesis on G implies that G_* is closed with respect to the weak-star toplogy on F^*. Next, in F we define the set

(ii) $$G_{**} = \{(x, t) \in E \oplus \mathbf{R}^+ : \ y(x) \leq \eta t : \ (y, \eta) \in G_*\}$$

This gives a set $\widehat{C}$ in F^* which consists of vectors (y, η) such that

$$\max_{(x,t) \in G_{**}} y(x) - \eta t \leq 0$$

It is clear that $G_* \subset \widehat{C}$. Now we prove the equality

(*) $$G_* = \widehat{C}$$

To get (*) we use Theorem 2.4. Namely, since the two sets in (*) are weak-star closed a strict inequality gives a separating vector $(x_*, t_*) \in E$, i.e. there exists $(y_*, \eta_*) \in \widehat{C}$ and a real number α such that

(iv) $$y_*(x_*) - \eta_* t_* > \alpha \quad \text{and} \quad (y, \eta) \in D_K \implies y(x_*) - \eta t_* \leq \alpha$$

Since G_* contains $(0, 0$ we have $\alpha \leq 0$. and since it also is a cone the last implication gives $(x_*, t_*) \in G_{**}$. Now the construction of $\widehat{C}$ in (iii) contradicts the strict inequality in the left hand side of (iv). Hence there cannot exist a separating vector and (*) follows.

Next, in E we consider the convex set

$$K = \{x : (x, 1) \in G_{**}\}$$

Using (*) the reader can check that

$$\mathcal{H}_K(y) = G(y)$$

for all $y \in E^*$ which proves that G has the requested form. The uniqueness of K follows from Exercise 3.4.

3.7 The case of normed spaces. If X is a normed vector space Theorem 3.6 leads to a certain isomorphism of two families. Denote by $\mathcal{K}$ the family of all convex subsets of E which are closed with respect to the norm topology. A topology on $\mathcal{K}$ is defined when we for each $K_0 \in \mathcal{K}$ and $\epsilon > 0$ declare an open neighborhod

$$U_\epsilon(K_0) = \{K \in \mathcal{K} : \operatorname{dist}(K, K_0) < \epsilon\}$$

where the norm defines the distance between K and K_0 in the usual way. Denote by $\mathfrak{H}$ the family of all functions G on E^* which satisfy (*) in 5.B.1 and are continuous with respect to the norm topology on E^*. A subset M of $\mathfrak{H}$ is equi-continuous if there to each $\epsilon > 0$ exists $\delta > 0$ such that

$$||y_2 - y_1|| < \delta \implies ||G(y_2) - G(y_1)|| < \epsilon$$

for every $G \in M$ and all pairs y_1, y_2 in E^*. The topology on $\mathfrak{H}$ is defined by uniform convergence on equi-continuous subsets.

3.8 Theorem. *If E is a normed vector space the set-theoretic bijective map $K \to \mathcal{H}_K$ is a homeomorphism when $\mathcal{K}$ and $\mathfrak{H}$ are equipped with the described topologies.*

3.9 Exercise. Deduce this result from Theorem 3.6

5. Fixed point theorems.

A topological space S has the fixed-point property if every continuous map $f\colon S\to S$ has at least one fixed point. To begin with we apply Stokes theorem to prove the classical result:

5.1 Theorem. *The closed unit ball in* $\mathbf{R}^n$ *has the fixed point property.*

Proof. By Weierstrass approximation theorem every continuous map from B into itself can be approxiated unifomly by a C^∞-map. Together with the compactness of B the reader should conclude that it suffices to prove every C^∞-map, $\phi\colon B\to B$ has at least one fixed point. We are going to derive argue by contradiction, i.e suppose that $\phi(x)\neq x$ for all $x\in B$. Each $x\in B$ gives the quadratic equation in the variable a

(i) $$1=|x+a(x-\phi(x)|^2=|x|^2+2a(1-\langle x,\phi(x)\rangle)+a^2|x-\phi(x)|^2$$

Exercise 1. Use that $\phi(x)\neq x$, to check that (i) has two simpe roots for each $x\in S$, and if $a(x)$ is the larger, then the function $x\mapsto a(x)$ belongs to $C^\infty(B)$. Moreover

(E.1) $$a(x)=0 \quad : x\in S$$

Next, for each real number t we set

$$f(x,t)=x+ta(x)(x-\phi(x))$$

This is a vector-valued function of the $n+1$ variables $t,x_1,\ldots,x_n$ where x varies in B. With $f=(f_1,\ldots,f_n)$ we set

$$g_i(x)=a(x)(x_i-\phi_i(x)$$

Taking partial derivatives with respect to x we get

(ii) $$\frac{\partial f_i}{\partial x_k}=e_{ik}+t\frac{\partial g_i}{\partial x_k}$$

where $e_{ii}=1$ and $e_{ik}=0$ if $i\neq k$. Let $D(x;t)$ be the determinant of the $n\times n$-matrix whose elements are the partial derivatives in (ii) and put

(iii) $$J(t)=\int_B D(x;t)\,dx$$

When $t=0$ we notice that the $n\times n$-matrix above is the identy matrix and hence $D(x;0)$ has constant value one so that $J(0)$ is the volume of B. Next, (i) entails that $x\mapsto f(x;1)$ satisfies the functional equation

$$|f(x;1)|^2=1$$

which implies that $x\mapsto D(x;1)$ is identically zero and hence $J(1)=0$. The requested contradiction follows if $t\mapsto J(t)$ is a constant function of t. To attain this we shall need:

2. Exercise. Use Leibniz's rule and that determinants of matrices with two equal columns are zero to conclude that

(E.2) $$\frac{d}{dt}(D(x;t)=\sum\sum(-1)^{j+k}\cdot\frac{\partial g_i}{\partial x_k}$$

where the double sum extends over all pairs $\leq j,k\leq n$.

Next, for all pairs $i.k$, Stokes theorem gives

(iv) $$\int_B\frac{\partial g_i}{\partial x_k}\,dx=\int_S g_i\cdot\mathbf{n}_k\,d\omega$$

where ω is the area measure on the unit sphere. From (E.1) we have $g_i=0$ on S for each i. Hence (E.2) and (iv) imply that

$$\frac{dJ}{dt}=\int_B\frac{d}{dt}(D(x;t)\,dx=0$$

So $t\mapsto J(t)$ is constant which is impossible because $J(0)=1$ and $J(1)=0$ which finishes the proof.

5.2 The Hilbert cube $\mathcal{H}_\square$. It is the closed subset of the Hilbert space ℓ^2 which consists of vectors $x = (x_1, x_2, \ldots)$ such that $|x_k| \leq 1/k$ for each k.

5.3 Proposition. *Every closed and convex subset of $\mathcal{C}$ has the fixed point property.*

5.4 Exercise. Deduce this result from Theorem 5.1.

Next, let X be a locally convex vector space and X^* its dual. Denote by $\mathcal{K}(X)$ the family of convex subsets which are compact with respect to the weak topology on X. Let $K \in \mathcal{K}(X)$ and $T\colon K \to K$ a continuous map with respect to the weak topology. For each fixed $f \in X^*$, it follows from our assumptions that the complex-valued function on K defined by

$$p \mapsto f(T(p))$$

is uniformly contnuus with respect to the weak topology. So for each positive integer n there exists a finite set $G_n = (x_1^*, \ldots, x_N^*)$ and some $\delta > 0$ such that the following implication holds for each pair of points p, q in K:

$$\text{(i)} \qquad p - q \in \cap B_\delta(x_\nu^*) \implies |f(T(p)) - f(T(q))| \leq n^{-1}$$

We can attain this for each positive integer n and get a denumerable set

$$G = \cup G_n$$

From (i) it is clear that if p, q is a pair in K and $g(p) = g(q)$ hold for every $g \in G$, then $x^*(T(p)) = x^*(T(q)$. We refer to G as a determining set for the map T. In a similar way we find a denumerable determining set $G^{1)}$ for g_1, By a standard diagonal argument the reader may verify the following:

5.5 Proposition. *There exists a denumerable subset G in X^* which contains f and is self-determining in the sense that it determines each of its vectors as above.*

5.6 An embedding into the Hilbert cube. During the construction of the finite G_m-sets which give (i), we can choose small δ-numbers and take $\{x_\nu^*\}$ such that the maximum values

$$\max_{p \in K} |x_\nu^*(p)|$$

are small. From this observation the reader should confirm that in Proposition 5.5 we can construct the sequence $G = (g_1, g_2, \ldots)$ in such a way that

$$\max_{p \in K} |g_n(p)| \leq n^{-1}$$

hold for every n. Hence each $p \in K$ gives the vector $\xi(p) = (g_1(p), g_2(p), \ldots)$ in the Hilbert cube and now

$$K_* = \{\xi(p) :\ p \in K\}$$

yields a convex subset of $\mathcal{C}$. Since G is self-determininig we have $T(p) = T(q)$ whenever $\xi(p) = \xi(q)$. Hence there exists a map from K_* into itself defined by

$$\text{(4.1)} \qquad T_*(\xi(p)) = \xi(T(p)$$

5.7 Exercise. Use the compact property of K to show that K_* is closed in the Hibert cube and that T_* is a continuous map with respect to the induced strong norm topology on K_* derived from the complete norm on ℓ^2.

5.8 Consequence. Suppose from the start that we are given a pair of points p_1, p_2 in K and some $f \in X^*$ where $f(p_1) \neq f(p_2)$. From the above f appears in the G-sequence and put

$$K_0 = \{p \in K\colon\ \xi(p) = \xi(p_1)\}$$

Then K_0 is a convex subset of K, and since f appears in the G-sequence it follows that p_2 does not belong to K_0. Moreover, since G is self-determining with respect to T it is clear that

$$T(K_0) \subset K_0$$

Hence we have proved:

5.9 Proposition. *For each pair K and T as above where K is not reduced to a single point, there exists a proper and X^*-closed convex subset K_0 of K such that $T(K_0) \subset K_0$.*

5.10 Proof of the Schauder-Tychonoff theorem.

Let $T\colon K \to K$ be a continuous map where K belongs to $\mathcal{K}(X)$. Consider the family $\mathcal{F}$ of all closed and convex subsets which are T-invariant. It is clear that intersections of such sets enjoy the same property. So we find the minimal set

$$K_* = \bigcap K_0$$

given by the intersection of all sets K_0 in $\mathcal{F}$. If K_* is not reduced to a single point then Proposition 6.1 gives a proper closed subset which again belongs to $\mathcal{F}$. This is contradicts the minimal property. Hence $K_* = \{p\}$ is a singleton set and p gives the requested fixed point for T.

5.11 Proof of Kakutani's theorem.

With the notations from the introduction we are given a group $\mathbf{G}$ where each element g preserves the convex set K in $\mathcal{K}(X)$. Zorn's lemma gives a minimal closed and convex subset K_* of K which again is invariant under the group. Kakutani's theorem follows if K_* is a singleton set. To prove that this indeed is the case we argue by contradiction. For K_* is not a singleton set then

$$K_* - K_* = \{p - q\colon p, q \in K_*\}$$

contains points outside the origin, and we find a convex open neighborhood V of the origin such that

$$\text{(i)} \qquad (K_* - K_*) \setminus \overline{V} \neq \emptyset$$

Since $\mathbf{G}$ is equicontinuous on K and hence also on K_* there exists an open convex neighborhood U of the origin such that whenever k_1, k_2 is a pair in K_* such that $k_1 - k_2 \in U$, then the orbit

$$\text{(ii)} \qquad \mathbf{G}(k_1 - k_2) \subset V$$

Set

$$U^* = \text{convex hull of } \ \mathbf{G}(U)$$

Since the $\mathbf{G}$-maps are linear, the set U^* is invariant and continuity entails the equality

$$\text{(iii)} \qquad \mathbf{G}\overline{U^*}) = \overline{U^*}$$

We find the unique positive number δ such that the following hold for every $\epsilon > 0$:

$$\text{(iv)} \qquad K_* - K_* \subset (1+\epsilon) \cdot U^* \quad : \ (K_* - K_*) \setminus (1-\epsilon) \cdot \delta \cdot \overline{U^*} \neq \emptyset$$

Next, $\{k + \frac{\delta}{2} \cdot U\colon \ k \in K_*\}$ is an open covering of the compact set K_*. Hence Heine-Borel's Lemma gives a finite set $k_1, \ldots, k_n$ in K_* such that

$$\text{(v)} \qquad K_* \subset \ \cup (k_\nu + \frac{\delta}{2} \cdot U)$$

Put

$$\text{(vi)} \qquad K_{**} = K_* \cap \bigcup_{k \in K_*} (k + (1 - 1/4n)\delta \cdot \overline{U})$$

Since $\overline{U}$ is $\mathbf{G}$-invariant and the intersection above is taken over all k in the invariant set K_*, we see that K_{**} is a closed convex and $\mathbf{G}$-invariant set. The requested contradiction follows if we prove that K_{**} is non-empty and is strictly contained in K_*. To get the strict inclusion follows we take some $0 < \epsilon < 1/4n$. Then (iv) gives a pair k_1, k_2 in K_* such that $k_1 - k_2$ does not belong to $(1-\epsilon)\delta \cdot \overline{U^*}$. At the same time the inclusion $k_1 \in K_{**}$ entails that

$$\text{(v)} \qquad k_1 \in (k_2 + (1 - 1/4n)\delta \cdot \overline{U}) \implies k_1 - k_2 \in (1 - 1/4n)\delta \cdot \overline{U}$$

which cannot hold since $1 - 1/4n < 1 - \epsilon$. The proof of Kakutani's theorem is therefore finished if we have shown that

$$\text{(vi)} \qquad p \in K_{**}$$

To see this we take an arbitrary $k \in K_*$. From (v) we find some $1 \leq i \leq n$ such that

$$\text{(vii)} \qquad k_i - k \in \frac{\delta}{2} \cdot U$$

Without loss of generality we can assume that $i = 1$ and get a vector $u \in U$ such that

$$\text{(viii)} \qquad k_1 = k + \frac{\delta}{2} \cdot u$$

It follows that

$$\text{(ix)} \qquad p = \frac{k_1 + \ldots + k_n}{n} = k + \frac{\delta}{2n} \cdot u + \sum_{i=2}^{i=n} \frac{1}{n}(k_i - k)$$

Next, for each $\epsilon > 0$ the left hand inclusion in (iv) and the convexity of U give

$$\text{(x)} \qquad \sum_{i=2}^{i=n} \frac{1}{n}(k_i - k) \subset \frac{n-1}{n}(1+\epsilon) \cdot \delta \cdot U$$

It follows that

$$\text{(xi)} \qquad \frac{\delta}{2n} \cdot u + \sum_{i=2}^{i=n} \frac{1}{n}(k_i - k) \in (\frac{n-1}{n}(1+\epsilon)\delta + \frac{\delta}{2n}) \cdot U$$

Above we can choose ϵ so small that

$$\frac{n-1}{n}(1+\epsilon) + \frac{1}{2n}) < 1 - 1/4n$$

and then we see that

$$p \in k + (1 - 1/4n)\delta \cdot \overline{U}$$

Since $k \in K_*$ was arbitrary the requested inclusion $p \in K_{**}$ follows.

§ 9. Neumann's resolvent operators

Introduction. we expose constructions which go back to work by Carl Neumann. A major result appears in Theorem 0.6.3 where analytic function theory is used to perform an operational calculus using resolvent operators. From now on X denotes a Banach space.

0.1 The class $\mathcal{I}(X)$. It consists of bounded linear operators R on X with the property that R is injective and the range $R(X)$ is a dense subspace of X. We do not exclude the possibility that R is surjective. Each such operator R gives a densely defined operator T as follows: If $x \in R(X)$ the injectivity of R gives a unique vector $\xi \in X$ such that $R(\xi) = x$ and we set

(i) $$T(x) = \xi$$

It means that the composed operator $T \circ R = E$, where E is the identity operator on X. Here the domain of defintion for T is equal to the range $R(X)$ and this dense subspace of X is denoted by $\mathcal{D}(T)$. By construction we have

$$R \circ T(x) = x \quad : x \in \mathcal{D}(T)$$

Next, the bounded operator R has a finite operator norm $||R||$ and (i) entails that

(ii) $$||x||| \leq ||R|| \cdot ||T(x)||$$

Thus, with $c = ||R||^{-1}$ one has

(iii) $$||T(x)|| \geq c \cdot ||x||| \quad : x \in \mathcal{D}(T)$$

The graph $\Gamma(T)$. It is the subset of $X \times X$ given by $\{(x, Tx) : x \in \mathcal{D}(T)\}$. The construction of T gives

$$\Gamma(T) = \{(Rx, x) \colon x \in X\}$$

Since R is a bounded inear operator it is clear that the last set is closed in $X \times X$, i.e. $\Gamma(T)$ is closed which means that T is a densely defined and closed linear operator on X. The inequality (iii) shows that T is injective and since

$$T(Rx)) = x \quad : x \in X$$

the range of T is equal to X.

A converse result. Assume that T is a densely defined and closed operator such that (iii) holds and in addition the range of T is dense in X. It turns out that this gives the equality

(1) $$T(\mathcal{D}(T)) = X$$

For if $y \in X$ the density of the range gives a sequence $\{x_n\}$ in $\mathcal{D}(T)$ such that

(2) $$\lim_{n \to \infty} ||T(x_n) - y|| = 0$$

Now

$$||x_n - x_m|| \leq c^{-1} \cdot ||T(x_n) - T(x_m)||$$

and (2) entails that $\{T(x_n)\}$ is a Cauchy sequence. Since X is a Banach space it follows that $\{x_n\}$ converges to a limit vector x. Now $\Gamma(T)$ is closed which implies that (x, y) belongs to the graph, i.e. $x\mathcal{D}(T)$ and $T(x) = y$ which proves (1).

Exercise. Let T be densely defined and closed where (iii) holds and $T(\mathcal{D}(T)) = X$. Show that there exists a unique bounded operator $R \in \mathcal{I}(X)$ such that T is the attached operator as in 0.1 above.

0.2 Spectra of densely defined operators.

Let T be a densely defined and closed linear operator. Each complex number λ gives the densely defined operator $\lambda \cdot E - T$. We say that λ is a resolvent value of T if $\lambda \cdot E - T$ is surjective and there exists a positive constant c such that

$$||\lambda \cdot x - T(x)|| \geq c \cdot ||x||$$

The set of resolvent values is denoted by $\rho(T)$. Its closed complement is called the spectrum of T and we put

$$\sigma(T) = \mathbf{C} \setminus \rho(T)$$

Each $\lambda \in \rho(T)$ gives a unique bounded operator $R_T(\lambda) \in \mathcal{I}(X)$ such that

$$(\lambda \cdot E - T) \circ R_T(\lambda)(x) = x$$

Since $\mathcal{D}(T) = \mathcal{D}(\lambda \cdot E - T)$ it follows that the range of $R_T(\lambda)$ is equal to $\mathcal{D}(T)$.

0.2.1 Definition. *Tbe family $\{R_T(\lambda) : \lambda \in \rho(T)\}$ are called Neumann's resolvents of T.*

An example. Let X be the Hilbert space ℓ^2 whose vectors are complex sequences $\{c_1, c_2, \ldots\}$ for which $\sum |c_n|^2 < \infty$. We have the dense subspace ℓ^2_* vectors such that $c_n \neq 0$ only occurs for finitely many integers n. If $\{\xi_n\}$ is an arbitrary sequence of complex numbers there exists the densely defined operator T on ℓ^2 which sends every sequence vector $\{c_n\} \in \ell^2_*$ to the vector $\{\xi_n \cdot c_n\}$. If λ is a complex number the reader may check that (i) holds in (0.0.1) if and only if there exists a constant C such that

$$\text{(v)} \qquad |\lambda - \xi_n| \geq C \quad : n = 1, 2, \ldots$$

Thus, $\lambda \cdot E - T$ has a bounded left inverse if and ony if λ belongs to the open complement of the closure of the set $\{\xi_n\}$ taken in the complex plane. Moreover, if (v) holds then $R_T(\lambda)$ is the bounded linear operator on ℓ^2 which sends $\{c_n\}$ to $\{\frac{1}{\lambda - \xi_n} \cdot c_n\}$. Since every closed subset of $\mathbf{C}$ is equal to the closure of a denumerable set of points our construction shows that the spectrum of a densely defined operator $\sigma(T)$ can be an arbitrary closed set in $\mathbf{C}$.

0.3 Neumann's equation.

Let T be closed and densely defined. Assume that $\rho(T) \neq \emptyset$. The equation below was discovered by Neumann:

For each pair $\lambda \neq \mu$ in $\rho(T)$ the operators $R_T(\lambda)$ and $R_T(\mu)$ commute and

$$(*) \qquad R_T(\mu) R_T(\lambda) = \frac{R_T(\mu) - R_T(\lambda)}{\lambda - \mu}$$

Proof. Notice that

$$(\mu \cdot E - T) \cdot \frac{R_T(\mu) - R_T(\lambda)}{\lambda - \mu} =$$

$$\text{(i)} \qquad \frac{E}{\lambda - \mu} - (\mu - \lambda) \cdot \frac{R_T(\lambda)}{\lambda - \mu} - (\lambda \cdot E - T) \cdot \frac{R_T(\lambda)}{\lambda - \mu} = R_T(\lambda)$$

Multiplying to the left by $R_T(\mu)$ gives (*) which at the same time shows that Neumann's resolvent operators commute.

0.4 Neumann series.

If $\lambda_0 \in \rho(T)$ we construct the operator valued series

$$\text{(1)} \qquad S(\zeta) = R_T(\lambda_0) + \sum_{n=1}^{\infty} (-1)^n \cdot \zeta^n \cdot R_T(\lambda_0)^{n+1}$$

It is clear that (1) converges in the Banach space of bounded linear operators when

$$|\zeta| < \frac{1}{||R_T(\lambda_0)||}$$

Moreover we see that

$$(2) \qquad (\lambda_0 + \zeta - T) \cdot S(\zeta) = (\lambda_0 - T) \cdot S(\zeta) + \zeta S(\zeta) = E$$

where the last equality follows via the series expansion (1). Hence

$$S(\zeta) = R_T(\lambda_0 + \zeta)$$

give resolvent operators. This proves that the set $\rho(T)$ is open. Moreover, the operator-valued function $\lambda \mapsto R_T(\lambda)$ is an analytic function of the complex varaible λ in defined in $\rho(T)$. If $\lambda \in \rho(T)$ we can pass to the limit as $\mu \to \lambda$ in Neumann's equation and conclude that the complex derivative becomes

$$(**) \qquad \frac{d}{d\lambda}(R_T(\lambda) = -R_T^2(\lambda)$$

Thus, Neumann's resolvent operator satisfies a specifiic differential equation for every densely defined and closed operator T with a non-empty resolvent set.

0.5 The position of $\sigma(T)$. Assume that $\rho(T) \neq \emptyset$. For a pair of resolvent values of T we can write Neumann's equation in the form

$$(1) \qquad R_T(\lambda)(E + (\lambda - \mu)R_T(\mu)) = R_T(\mu)$$

Keeping μ fixed we conclude that $R_T(\lambda)$ exists if and only if $E + (\lambda - \mu)R_T(\mu))$ is invertible. This gives the set-theoretic equality

$$(0.5.1) \qquad \sigma(T) = \{\lambda : \frac{1}{\mu - \lambda} \in \sigma(R_T(\mu))\}$$

Hence one recovers $\sigma(T)$ via the spectrum of any given resolvent operator. Notice that (0.5.1) holds even when the open component of $\sigma(T)$ has several connected components.

0.5.2 Example. Suppose that $\mu = i$ and that $\sigma(R_T(i))$ is contained in a circle $\{|\lambda + i/2| = 1/2\}$. If $\lambda \in \sigma(T)$ the inclusion (0.0.5.1) gives some $0 \leq \theta \leq 2\pi$ such that

$$\frac{1}{i - \lambda} = -i/2 + 1/2 \cdot e^{i\theta} \implies 1 - i \cdot e^{i\theta} = \lambda(e^{i\theta} - i)$$

The last equation entails that

$$\lambda = \frac{2 \cdot \cos\theta}{|e^{i\theta} - i|^2}$$

and hence λ is real.

0.5.3 The case when resolvent operators are compact. Let T be such that $R_T(\lambda_0)$ is a compact operator for some resolvent value. We assume of course that the Banach space X is not finite dimensional. In §§ we shall learn that the spectrum of a compact operator always contains zero and outside the origin the spectrum is a discrete set with a sole cluster point at the origin. From (0.5.1) it follows that $\sigma(T)$ is a discrete set in $\mathbf{C}$, i.e. its intersection with every disc $\{|\lambda| \leq R\}$ is finite.

0.5.4 Remark. In § xx we shall learn that if S is a compact operator then $S \circ U$ and $U \circ S$ are compact for every bounded operator U. Applying Neumann's equation (*) in (0.3) it follows that if one resolvent operator $R_T(\lambda_0)$ is compact, then all resolvents of T are compact.

0.5.5 Adjoint operators and closed extensions.

Let T be densely defined. But for the monent we do not assune that it is closed. In the dual space X^* we have the family of vectors y for which there exists a constant $C(y)$ such that

$$(i) \qquad |y(Tx)| \leq C(y) \cdot ||x|| \quad : x \in \mathcal{D}(T)$$

It is clear that the set of such y-vectors is a subspace of X^*. Moreover, when (i) holds the density of $\mathcal{D}(T)$ gives a unique vector $T^*(y)$ in X^* such that

$$(ii) \qquad y(Tx) = T^*(y)(x) \quad : x \in \mathcal{D}(T)$$

One refers to T^* as the adjoint operator of T whose domain of definition is denoted by $\mathcal{D}(T^*)$.

Exercise. Show that the graph of T^* is closed in $X^* \times X^*$. However, $\mathcal{D}^*(T)$ is in general not a dense subspace of X^*. See § xx for an example.

Closed extensions. There may exist several closed operators S with the property that

$$\Gamma(T) \subset \Gamma(S)$$

When this holds we refer to S as a closed extension of T. Notice that the inclusion above is strict if and only if $\mathcal{D}(S)$ is strictly larger than $\mathcal{D}(T)$.

Exercise. Use the density of $\mathcal{D}(T)$ to show that

$$T^* = S^*$$

hold for every closed extension S of T.

The case when $\mathcal{D}(T^*)$ is dense. Let T be densely defined and assume that its adjoint has a dense domain of definition. In this situation the following holds:

0.5.6 Theorem. *If $\mathcal{D}(T^*)$ is dense then there exists a closed operator $\widehat{T}$ whose graph is the closure of $\Gamma(T)$.*

Proof. Consider the graph $\Gamma(T)$ and let $\{x_n\}$ and $\{\xi_n\}$ be two sequences in $\mathcal{D}(T)$ which both converge to a point $p \in X$ while $T(x_n) \to y_1$ and $T(\xi_n) \to y_2$ hold for some pair y_1, y_2. We must rove that $y_1 = y_2$. To achieve this we take some $x^* \in \mathcal{D}(T^*)$ which gives

$$x^*(y_1) = \lim\, x^*(Tx_n) = \lim\, T^*(x^*)(x_n) = T^*(x^*)(p)$$

In the same way we get $x^*(y_2) = T^*(x^*)(p)$. Now the density of $\mathcal{D}(T^*)$ gives $y_1 = y_2$ which proves that the closure of $\Gamma(T)$ is a graphic subset of $X \times X$ and gives the closed operator $\widehat{T}$ with

$$\Gamma(\widehat{T}) = \overline{\Gamma(T)}$$

The case when X is reflexive. Assume this and let T be a densely defined and closed operator. Suppose in addition that T^* also is densely defined. Now we can construct the adjoint of T^* which is denoted by T^{**}. Since X is reflexive it follows that T^{**} is a closed and densely defined operator on X. If $x \in \mathcal{D}(T)$ and $y \in \mathcal{D}(T^*)$ we have the vector $\widehat{x} \in X^{**}$ and

$$\widehat{x}(T^*(y)) = T^*(y)(x) = y(T(x))$$

From this it is clear that $\widehat{x} \in \mathcal{D}(T^{**})$ and one has the equality

$$T^{**}(\widehat{x}) = T(x)$$

Hence the graph of T is contained in that of T^{**}, i.e. T^{**} is a closed extension of T.

0.5.7 The spectrum of T^*. Let X and T be as above. Then one has the inclusion

$$(*) \qquad \rho(T) \subset \rho(T^*)$$

Proof. By translations it suffices to show that if the origin belongs to $\rho(T)$ then it also belongs to $\rho(T^*)$. So now the resolvent $R_T(0)$ exists which means that T is surjective and there is a constant $c > 0$ such that

$$\text{(i)} \qquad ||x|| \leq c^{-1} \cdot ||Tx|| \quad : x \in \mathcal{D}(T)$$

Consider some $y \in \mathcal{D}(T^*)$ of unit norm. Since T is surjective we find $x \in \mathcal{D}(T)$ with $||Tx|| = 1$ and

$$\text{(ii)} \qquad |y(Tx)| \geq 1/2$$

Now

$$\text{(iii)} \qquad y(Tx) = T^*(y)(x)$$

and from (i) we have

(iv) $$||x|| \leq c^{-1} \cdot ||Tx|| = c^{-1}$$

Then (ii) and (iv) entail that

$$||T^*(y)|| \geq c/2$$

This proves that

(v) $$||T^*(y)|| \geq c/2 \cdot ||y|| \quad : y \in \mathcal{D}(T^*)$$

Hence the origin belongs to $\rho(T^*)$ if we prove that T^* has a dense range. If the density fails there exists a non-zero linear functional $\xi \in X^{**}$ such that

$$\xi(T^*(y)) = 0 \quad : y \in \mathcal{D}(T^*)$$

Since X is reflexive we have $\xi = \mathfrak{i}_X(x)$ for some vector x and obtain

$$y(Tx) = 0 \quad : y \in \mathcal{D}(T^*)$$

The density of $\mathcal{D}(T^*)$ gives $Tx = 0$ which contradicts the hypothesis that T is injective and (*) follows.

0.5.8 The case when X is a Hilbert space. In this case we shall prove that when both T and T^* are closed and densely defined, then one has the equality

$$\sigma(T) \subset \sigma(T^*)$$

We refer to § xx for the proof.

0.6 Operational calculus.

Let T be a densely defined and closed operator on a Banach space X. To each pair (γ, f) where γ is a rectifiable Jordan arc contained in $\mathbf{C} \setminus \sigma(T)$ and $f \in C^0(\gamma)$, there exists the bounded linear operator

(0.6.1) $$T_{(\gamma,f)} = \int_\gamma f(z) R_T(z)\, dz$$

The integral is calculated via a Riemann sum where the integrand has values in the Banach space of bounded linear operators on X. More precisely, let $s \mapsto z(s)$ be a parametrisation with respec to arc-length. If L is the arc-length of γ we get Riemann sums

$$\sum_{k=0}^{k=N-1} f(z(s_k)) \cdot (z(s_{k+1}) - z(s_k)) \cdot (s_{k+1} - s_k) \cdot R_T(z(s_k))$$

where $0 = s_0 < s_1 < \ldots s_N = L$ is a partition of $[0, L]$. These Riemann sums converge to a limit when $\{\max (s_{k+1} - s_k\} \to 0$ with respect to the operator norm and give the T-operator in (0.6.1). The triangle inequality entails that

$$T_{(\gamma,f)} \leq L \cdot |f|_\gamma \cdot \max_{z \in \gamma} ||R_T(z)||$$

where $|f|_\gamma$ is the maximum norm of f on γ..

Neumann's equation in (0.3) entails that $R_T(z_1)$ and $R_T(z_2)$ commute for all pairs z_1, z_2 on γ. It follows that if g is another function in $C^0(\gamma)$ then the operators $T_{f,\gamma}$ and $T_{g,\gamma}$ commute. Moreover, for each $f \in C^0(\gamma)$ the reader may verify that the closedeness of T implies that the range of $T_{f,\gamma}$ is contained in $\mathcal{D}(T)$ and one has

$$T_{f,\gamma} \circ T(x) = T \circ T_{f,\gamma}(x) \quad : x \in \mathcal{D}(T)$$

Next, let Ω be an open set of class $\mathcal{D}(C^1)$, i.e. $\partial\Omega$ is a finite union of closed differentiable Jordan curves. When $\partial\Omega \cap \sigma(T) = \emptyset$ we construct line integrals as in (0.6.1) for continuous functions on the boundary. Consider the algebra $\mathcal{A}(\Omega)$ of analytic functions in Ω which extend to be continuous on the closure. Each $f \in \mathcal{A}(\Omega)$ gives the operator

$$T_f = \int_{\partial\Omega} f(z) R_T(z)\, dz \tag{0.6.2}$$

0.6.3 Theorem. *The map $f \mapsto T_f$ is an algebra homomorphism from $\mathcal{A}(\Omega)$ into a commutative algebra of bounded linear operators on X whose image is a commutative algebra of bounded linear operators denoted by $T(\Omega)$.*

Proof. Let f, g be a pair in $\mathcal{A}(\Omega)$. To show that $T_{gf} = T_f \circ T_g$ we consider a slightly smaller open set Ω_* which again is of class $\mathcal{D}(C^1)$ and each of it bounding Jordan curve is close to one boundary curve in $\partial\Omega$ and $\Omega \setminus \Omega_*$ does not intersect $\sigma(T)$. By Cauchy's theorem we can shift the integration to $\partial\Omega_*$ and get

$$T_g = \int_{\partial\Omega_*} g(z) R_T(z_*)\, dz_* \tag{i}$$

where we use z_* to indicate that integration takes place along $\partial\Omega_*$. Now

$$T_f \circ T_g = \iint_{\partial\Omega_* \times \partial\Omega} f(z) g(z_*) R_T(z) \circ R_T(z_*)\, dz_* dz \tag{ii}$$

Neumann's equation (*) from (0.0.3) entails that the right hand side in (ii) becomes

$$\iint_{\partial\Omega_* \times \partial\Omega} \frac{f(z) g(z_*) R_T(z_*)}{z - z_*}\, dz_* dz + \iint_{\partial\Omega_* \times \partial\Omega} \frac{f(z) g(z_*) R_T(z)}{z - z_*}\, dz_* dz = A + B \tag{iii}$$

Here A is evaluated by first integrating with respect to z and Cauchy's theorem gives

$$f(z_*) = \frac{1}{2\pi i} \cdot \iint_{\partial\Omega} \frac{f(z)}{z - z_*} \quad : z_* \in \partial\Omega_*\, dz$$

It follows that

$$A = \frac{1}{2\pi i} \cdot \iint_{\partial\Omega_* \times \partial\Omega} f(z_*) g(z_*) R_T(z_*)\, dz_* = T_{fg}$$

Next, B is evaluated when we first integrate with respect to z_*. Here

$$\iint_{\partial\Omega} \frac{g(z_*)}{z - z_*} \quad : z \in \partial\Omega$$

which entails that $B = 0$ and the theorem follows.

Spectral gap sets. Let K be a compact subset of $\sigma(T)$ such that $\sigma(T) \setminus K$ is a closed set in $\mathbf{C}$. This implies that if V is an open neighborhood of K, then there exists a relatively compact subdomain $U \in \mathcal{D}(C^1)$ which contains K as a compact subset. To every such domain Ω we can apply Theorem 0.0.6.3. If $U_* \subset U$ for a pair of such domains we can restrict functions in $\mathcal{A}(U)$ to U_* which yields an algebra homomorphism

$$\mathcal{T}(U) \to \mathcal{T}(U_*)$$

Next, denote by $\mathcal{O}(K)$ the algebra of germs of analytic functions on K. So each $f \in \mathcal{O}(K)$ comes from some analytic function in a domain U as above. The resulting operator $T_U(f)$ depends on the germ f only. In fact, this follows because if $f \in \mathcal{A}(U)$ and $U_* \subset U$ is a similar $\mathcal{D}(C^1)$-domain which again contains K, then Cauchy's vanishing theorem from § xxx is applied to $f(z) R_T(z)$ in $U \setminus \bar{U}_*$ and entails that

$$\int_{\partial U_*} f(z) R_T(z)\, dz = \int_{\partial U} f(z) R_T(z)\, dz$$

Hence there exists an algebra homorphism from $\mathcal{O}(K)$ into bounded linear operators on X whose image is denoted by $\mathcal{T}(K)$. The identity in $\mathcal{T}_K$ is denoted by E_K and called the spectral projection operator attached to the compact set K in $\sigma(T)$. By this construction one has

$$E_K = \frac{1}{2\pi i} \cdot \int_{\partial U} z \cdot R_T(z)\, dz$$

for every open domain U around K as above.

0.0.6.4 The operator T_K. When K is a compact spectral gap set of T we set

$$T_K = TE_K$$

This bounded linear operator is given by

$$\frac{1}{2\pi i} \cdot \int_{\partial U} z \cdot R_T(z)\, dz$$

where U is a domain as above containing K.

0.0.6.4.1 *Identify T_K with a densely defined operator on the space $E_K(X)$. Then one has the equality*

$$\sigma(T_K) = K$$

Proof. If λ_0 is outside K we can choose U so that λ_0 is outside $\bar{U}$ and construct the operator

$$S = \frac{1}{2\pi i} \cdot \int_{\partial U} \frac{1}{\lambda_0 - z} \cdot R_T(z)\, dz$$

The operational calculus gives

$$S(\lambda_0 E_K - T) = E_K$$

here E_K is the identity operator on $E_K(X)$ which shows that $\sigma(T_K) \subset K$.

0.0.6.5 Discrete spectra. Consider a spectral set reduced to a singleton set $\{\lambda_0\}$, i.e. λ_0 is an isolated point in $\sigma(T)$. The associated spectral projection is denoted by $E_T(\lambda_0)$ and expressed

$$E_T(\lambda_0) = \frac{1}{2\pi i} \cdot \int_{|\lambda - \lambda_0| = \epsilon} R(\lambda)\, d\lambda$$

for all sufficiently small ϵ. Now $R_T(\lambda)$ is an analytic function defined in some punctured disc $\{0 < \lambda - \lambda_0| < \delta\}$ with a Laurent expansion

$$R_T(\lambda) = \sum_{-\infty}^{\infty} (\lambda - \lambda_0)^k \cdot B_k$$

where $\{B_k\}$ are bounded linear operators obtained by residue formulas:

$$B_k = \frac{1}{2\pi i} \cdot \int_{|\lambda| = \epsilon} \frac{R_T(\lambda)}{(\lambda - \lambda_0)^{k+1}}\, d\lambda \quad : \epsilon < \delta$$

Exercise. Show that $R_T(\lambda)$ is meromorphic, i.e. $B_k = 0$ hold when $k << 0$, if and only if there exists a constant C and some integer $M \geq 0$ such that the operator norms satisfy

$$||R_T(\lambda)|| \leq C \cdot |\lambda - \lambda_0)^{-M}$$

Suppose now that R_T has a pole of some order $M \geq 1$ which gives an expansion

$$R_T(\lambda) = \sum_{1}^{M} \frac{B_{-k}}{(\lambda - \lambda_0)^k} + \sum_{0}^{\infty} (\lambda - \lambda_0)^k \cdot B_k$$

Here $B_{-1} = E_T(\lambda_0)$ and if $M \geq 2$ the negative indexed operators satisfy

$$B_{-k} = B_{-k} E_T(\lambda_0) \quad 2 \leq k \leq M$$

In the case of a simple pole, i.e. when $M = 1$ the operational calculus gives

$$(\lambda_0 E - T)E_T(\lambda_0) = \lim_{\epsilon \to 0} \frac{1}{2\pi i} \cdot \int_{|\lambda - \lambda_0| = \epsilon} (\lambda_0 - \lambda)R(\lambda)\, d\lambda = 0$$

which implies hat the range of the projection operator $E_T(\lambda_0)$ is equal to the kernel of $\lambda_0 \cdot E - T$.

0.0.6.6 The case $M \geq 2$. Now one has a a non-decreasing family of subspaces

$$N_k(\lambda_0) = \{x \colon (\lambda_0 E - T)^k(x) = 0\} \quad : 1 \leq k \leq M$$

Let us analyzie the special case when the range of $E_T(\lambda_0)$ has finite dimension. Here the operator $T(\lambda_0) = TE_T(\lambda_0)$ acts on this finite dimensional vector space and the B-matrices with negative indices can be expressed as in linear algebra via a Jordan decomposition of $T(\lambda_0)$. More precisely Jordan blocks of size > 1 may occur which occurs of the smallest positive integer m such that

$$(\lambda_0 E - T)^m(x) = \quad : x \in E_T(\lambda_0)(X)$$

is strictly larger than one. Moreover, $E - E_T(\lambda_0)$ is a projection operator and one has a direct sum decompostion

$$X = E_T(\lambda_0)(X) \oplus E - E_T(\lambda_0)$$

Here $V = E - E_T(\lambda_0)$ is a closed subspace of X which is invariant under T and there exists some $c > 0$ such that

$$||\lambda_0 - Tx|| \geq ||x|| \quad x \in V \cap \mathcal{D}(T)$$

Remark. In applications it is often an important issue to decide when $E_T(\lambda_0)$ has a finite dimensional range for an isolated point in $\sigma(T)$. The Kakutani-Yosida theorem in § 11.9 is an example where this finite dimensionality will be established for certain operators T.

§ 0.0.7 Semi-groups and infinitesmal generators.

We give details of proofs of a result due to Hille, Phillips and Yosida. It has been inserted at this early stage to illustrate Neumann's theory about resolvents. The less expericenced reader may consult the subsequent proof after studies of basic material in the special chapters. Let X be a Banach space. A family of bounded operators $\{T_t\}$ indexed by non-negative real numbers is a semi-group if $T_0 = E$ is the identity and

$$T_{t+s} = T_s \circ T_t$$

for all pairs of non-negative real numbers. In particular the T-operators commute. The semi-group is said to be strongly continuous if the vector-valued functions

$$x \mapsto T_t(x)$$

are continuous with respect to the norm in X for each $x \in X$.

1. Proposition. *Let $\{T_t\}$ be a strongly continuous semi-group and set*

$$\omega = \log |||T_1||$$

Then the operator norms satisfy

$$||T_t|| \leq e^{\omega t} \quad : t \geq 0$$

Exercise. Prove this using calculus applied to sub-multiplicative functions.

With ω as above we consider the open half-plane

$$U = \{\mathfrak{Re}(\lambda) > \omega\}$$

Let $\lambda \in U$ and x is a vector in X. The Borel-Stieltjes construction of integrals with values in a Banach space gives the X-valued integral

$$\int_0^\infty e^{-\lambda t} \cdot T_t(x)\, dt \tag{1.1}$$

whose value is denoted by $\mathcal{T}(\lambda)(x)$. It is clear that

$$x \mapsto \mathcal{T}(\lambda)(x)$$

is linear and the triangle inequality gives

$$||\mathcal{T}(\lambda)(x)|| \leq \int_0^\infty e^{-\mathfrak{Re}\lambda t} \cdot e^{\omega t}\, dt \cdot ||x|| = \frac{1}{\mathfrak{Re}\,\lambda - \omega} \cdot ||x|| \tag{1.2}$$

2. Infinitesmal generators. Let $\{T_t\}$ be a a strongly continuous semi-group. Then there exists a densely defined and closed operator which is called its infinitesmal generator and obtained as follows:

2.1 Theorem. *There exists a dense subspace $\mathcal{D}$ in X such that*

$$\lim_{h \to 0} \frac{T_h(x) - x}{h} \quad : x \in \mathcal{D}$$

exists. If $A(x)$ is the limit value then A is a densely defined operator and

$$\sigma(A \subset \{\mathfrak{Re}\,\lambda \leq \omega\} \tag{i}$$

Moreover, in the open half-space U one has the equality

$$R_A(\lambda) = \mathcal{T}(\lambda) \tag{ii}$$

3. The Hille-Phillips-Yosida theorem. Theorem 2.1 produces infinitesmal generators of strongly continuous semi-groups. This specific class of densely defined and closed operators can be described via properties of their spectra and behavious of the resolvent operators. Denote by

$\mathcal{HPY}$ the family of densely defined and closed linear operators A with the property that $\sigma(A)$ is contained in a half-space $\{\mathfrak{Re}\,\lambda \leq a\}$ for some real number a, and if $a^* > a$ there exists a constant M such that

$$||R_A(\lambda)|| \leq M \cdot \frac{1}{\mathfrak{Re}\,\lambda - \omega} \quad : \mathfrak{Re}\,\lambda \geq a^*$$

3.1 Theorem. *Each $A \in \mathcal{HPY}$ is the infinitesmal generator of a uniquely determined strongly continuous semi-group.*

Remark. Notice that (iii) in Theorem 2.1 together with (1.2) imply that the infinitesmal generator of a strongly continuous semi-group belongs to $\mathcal{HPY}$. Hence Theorem 3.1 this gives a 1-1 correspondence between $\mathcal{HPY}$ and the family of strongly contionuous semi-groups.

4. The case of bounded operators. Before we give the proofs of the two theorems above we consider bounded operators. If B is a bounded linear operator on X there exists the strongly continuous semi-group where

$$T_t = e^{tB} = 1 + \sum_{n=1}^{\infty} \frac{t^n}{n!} \cdot B^n$$

If $h > 0$ and $x \in X$ we have

$$|| \lim_{h \to 0} \frac{T_h(x) - x}{h} - B(x)|| = || \sum_{n=2}^{\infty} | \frac{h^n}{n!} \cdot B^n(x)||$$

The triangle inequality entails that the last term is majorised by

$$\frac{h^2}{2} \cdot e^{||B(x)||}$$

We conclude that B is the infinitesmal generator of the semi-group. $\mathcal{T} = \{T_t\}$. Next, if $\omega = ||B||$ then $\sigma(B)$ is contained in the disc of radius ω and hence in the half-plane $U = \{\mathfrak{Re}\,\lambda \leq \omega\}$. If $\lambda \in U$ the operator-valued integral

$$\int_0^{\infty} e^{-\lambda t} \cdot e^{Bt}\, dt = \frac{1}{\lambda} \cdot E + \sum_{n=1}^{\infty} \big(\int_0^{\infty} e^{-\lambda t} \cdot t^n\, dt \big) \cdot \frac{B^n}{n!}$$

Evaluating the integrals the right hand side becomes

$$\frac{1}{\lambda} \cdot E + \sum_{n=1}^{\infty} \frac{B^n}{\lambda^{n+1}}$$

The last sum is equal to the Neumann series for $R_B(\lambda)$ from § xx. which gives the equation

$$R_B(\lambda) = \mathcal{T}(\lambda)$$

Hence Theorem 3.1 is confirmed for bounded operators.

4.1 Uniformly continuous semi-groups. Let B be a bounded operator. The semi-group $\{T_t = e^{tB}\}$ has the additional property that

(i) $$\lim_{t \to 0} |||T_t - E|| = 0$$

In fact, the triangle inequality gives

$$||T_t - E|| \leq \sum_{n=1}^{\infty} \frac{t^n}{n!} \cdot ||B||^n$$

and it is clear that the right hand side tends to zero as $t \to 0$.

4.2 Definition. *A semi-group $\{T_t\}$ is called uniformly continuous if*

(4.2.1) $$\lim_{t \to 0} |||T_t - E|| = 0$$

When (4.2.1) holds we find $h_* > 0$ such that

$$||T_t - E|| \leq 1/2 \quad : \; 0 \leq t \leq h_*$$

As explained in § xx this gives bounded operators $\{S_t : \; 0 \leq t \leq h_*\}$ such that

$$T_t = e^{S_t}$$

This entails that if $0 < h \leq h_*$ then

$$\frac{T_h - E}{h} = \frac{S_h}{h} + \sum_{n=2}^{\infty} \frac{h^{n-1}}{n!} \cdot S_h^n$$

4.3 Exercise. Verify that the semi-group equations entail that if N is a positive integer and h so small that $Nh \leq h_*$ gives the equation

$$\frac{S_{Nh}}{N} = S_h \tag{4.3.1}$$

Next, for each small and positive $h < h_*$ we choose the largest positive integer N_h such that

$$\frac{h_*}{N_h + 1} < h \leq \frac{h_*}{N_h}$$

With this choice of N_h we apply (4.3.1) and get

$$\frac{S_h}{h} = \frac{1}{N_h \cdot h} \cdot S_{N_h \cdot h}$$

Passing to the limit as $h \to 0$ the reader can check the limit equation

$$\lim_{h \to 0} \frac{S_h}{h} = h_*^{-1} \cdot S(h_*)$$

4.4 Exercise. Show from the above that

$$\lim_{h \to 0} \frac{T_h - E}{h} = h_*^{-1} \cdot S(h_*)$$

which means that the infinitesmal generator of $\{T_t\}$ is given by the bounded operator $h_*^{-1} \cdot S(h_*)$.

4.5 Conclusion. There exists a 1-1 correspondence between uniformly continuous semi-groups and bounded linear operators on X.

5. The unbounded case.

First we prove Theorem 2.1. If $x \in X$ and $\delta > 0$ we put

$$x_\delta = \int_0^\delta T_t(x)\, dt$$

For every $h > 0$ the semi-group equation gives

$$\frac{T_h(x_\delta) - x_\delta}{h} = \frac{1}{h} \cdot \int_\delta^{\delta + h} T_t(x)\, dt$$

The strong continuity entails that the limit in the right hand side exists as $h \to 0$ and gives a vector $T_\delta(x)$. Hence the space $\mathcal{D}$ contains x_δ. The continuity of $t \mapsto T_t(x)$ at $t = 0$ implies that $||x_\delta - x|| \to 0$ which proves that $\mathcal{D}$ is dense and the construction of the infinitesmal generator A gives

$$A(x_\delta) = T_\delta(x) \tag{5.1}$$

for every $\delta > 0$.

Next, consider some vector $x \in \mathcal{D}(A)$. Now there exists the integral which defines $\mathcal{T}(\lambda)(Ax)$ when λ belongs to the half-plane U. If λ is real and $h > 0$ a variable substitution gives

$$\mathcal{T}(\lambda)(T_h(x)) = \int_0^\infty e^{-\lambda t} \cdot T_{t+h}(x)\, dt = e^{\lambda h} \cdot \int_h^\infty e^{-\lambda s} T_s(x)\, ds$$

It follows that

$$\mathcal{T}(\lambda)(\frac{T_h(x)) - x}{h}) = \frac{e^{\lambda h} - 1}{h} \cdot \int_h^\infty e^{-\lambda s} T_s(x)\, ds - \frac{1}{h} \cdot \int_0^h e^{-\lambda s} T_s(x)\, ds$$

Passing to the limit as $h \to 0$ the reader can check that the right hand side becomes

$$\lambda \cdot \mathcal{T}(\lambda)(x) - x$$

So with $x \in \mathcal{D}(A)$ we have the equation

$$\mathcal{T}(\lambda)(A(x)) = \lambda \cdot \mathcal{T}(\lambda)(x) - x$$

which can be written as

$$\mathcal{T}(\lambda)(\lambda \cdot E - A)(x) = x$$

Exercise. Conclude from the above that

$$\mathcal{T}(\lambda) = R_A(\lambda)$$

and deduce Theorem 2.1.

5.2 proof of Theorem 3.1

Let A belong to $\mathcal{HPY}$. So here $\sigma(A)$ is contained in $\{\mathfrak{Re}(\lambda) \leq a\}$ for some real number a, and when λ varies in the open half-plane $U = \{\mathfrak{Re}(\lambda) \leq a\}$ there exist the resolvents $R(\lambda)$ where the subscript A is deleted while we consider the operator A. By assumption there exists a constant K and some $a^* \geq a$ such that

(5.2.0) $$||R(\lambda)|| \leq \frac{K}{\lambda}$$

when $\mathfrak{Re}(\lambda \geq a^*$.

The operators B_λ. For each $\lambda \in U$ we set

(ii) $$B_\lambda = \lambda^2 \cdot R(\lambda) - \lambda \cdot E$$

Notice that (5.2.0) gives

(iv) $$||B_\lambda|| \leq (K+1)|\lambda|$$

Consider a vector $x \in \mathcal{D}(A)$. Now

$$B_\lambda(x) = \lambda \cdot (\lambda R(\lambda)(x) - x) = \lambda \cdot R(\lambda)(Ax)$$

We have also

$$\lambda \cdot R(\lambda)(Ax) - R(\lambda(A(x)) = A(x)$$

Hence

$$B_\lambda(x) - A(x) = R(\lambda)(A(x))$$

Hence (5.2.0) gives

$$||B_\lambda(x) - A(x)|| \leq \frac{K}{\lambda} \cdot ||A(x)||$$

Hence we have the limit formula

$$\lim_{\lambda \to \infty} ||B_\lambda(x) - A(x)|| = 0 \quad : x \in \mathcal{D}(A)$$

The semi-groups $\mathcal{S}_\lambda = \{e^{tB_\lambda} : t \geq 0\}$. To each $t \geq 0$ and $\lambda \in U$ we set

(5.2.1) $$S_\lambda(t) = e^{tB_\lambda} = E + \sum_{n=1}^\infty \frac{t^n}{n!} \cdot B_\lambda^n$$

By (iv) the series converges and with t fixed termwise differentiation with respect to λ gives

$$\frac{d}{d\lambda}(S_\lambda(t)) = t \cdot \frac{d}{d\lambda}(B_\lambda) \cdot S_t(\lambda) = -t \cdot S_\lambda(t) + t \cdot (\lambda \cdot R(\lambda)^2 - R(\lambda)) \cdot S_\lambda(t)$$

Keeping t fixed we get after an integration

$$S_\mu(t) - S_\lambda(t) = t \cdot \int_\lambda^\mu (\xi \cdot R(\xi)^2 - R(\xi))S_\xi(t)\, ds \quad : \mu > \lambda \geq a^* \tag{5.2.2}$$

Next, we notice that (5.2.0) gives

$$\lim_{\xi \to \infty} ||(\xi \cdot R(\xi)^2 - R(\xi)|| = 0 \tag{5.2.3}$$

Together with the general differential inequality from (xx) we conclude that if t stays in a bounded interval $[0, T)$ and $\epsilon > 0$, then there exists some large ξ^* such that for every $0 \leq t \leq T$ the operator norms satisfy

$$||S_\mu(t) - S_\lambda(t)|| < \epsilon \quad : \mu \geq \lambda \geq \xi^* \tag{5.2.4}$$

5.2.5 The semi-group $\{S(t)\}$. Since the bounded operators on X is a Banach space, it follows from the above that each t gives a bounded operator $S(t)$ given by

$$\lim_{\lambda \to +\infty} S_\lambda(t) = S(t)$$

where the limit is taken in the operator norm and the convergence holds uniformly when t stays in a bounded interval. Since $t \mapsto S_\lambda(t)$ is a semi-group for each large λ, the same holds for the family $\{S(t)\}$. Moreover, the convergence properties entail that the semi-group $\{S(t)\}$ is strongly continuous. Neumann's differential equation from (§ 0.x) gives

$$\frac{d}{d\lambda}(B_\lambda) = R(\lambda) - E - \lambda \cdot R(\lambda)^2 \tag{iii}$$

The infinitesmal generator of $\{S(t)\}$. Returning to the series (5.2.1) it is clear that

$$\frac{d}{dt}(S_\lambda(t) = S_\lambda(t) \cdot B_\lambda \tag{5.2.6}$$

An integration and the equality $S_\lambda(0) = E$ give for every $t > 0$ and each vector $x \in X$:

$$S_\lambda(t)(x) - x = \int_0^t S_\lambda(\xi) \circ B_\lambda(\xi)(x)\, d\xi \tag{5.2.7}$$

When $x \in \mathcal{D}(A)$ we have the limit formula (xx) and together with the limit which produces the semi-group $\{S(t)\}$ we conclude that

$$S(t)(x) - x = \int_0^t S(\xi) \circ A(x)\, d\xi \quad : x \in \mathcal{D}(A)$$

We can take a limit as $t \to 0$ where the strong continuity of the semi-group $\{S(t)\}$ applies to vectors $A(x) : x \in \mathcal{D}(A)\}$. Hence

$$\lim_{t \to 0} \frac{S(t)(x) - x}{t} = A(x) \quad : x \in \mathcal{D}(A)$$

So if $\widehat{A}$ is the infinitesmal generator of the semi-group $\{S(t)\}$ then its graph contains that of A, i.e. $\widehat{A}$ is an extension of the densely defined and closed operator A. However, we have equality because $\widehat{A}$ being an infinitesmal generator of a strongly continuous semi-group has its spectrum confined to a half-space $\{\mathfrak{Re}(\lambda) \leq b\}$ for some real number b, i.e. here we used Theorem 2.1. In particular there exist points outside the union of $\sigma(A)$ and $\sigma(\widehat{A})$ and then the equality $A = \widehat{A}$ follows from the general result in § xx. Hence A is an infinitesmal generator of a semi-group which finishes the proof of Theorem 3.1.

§ 10. Hilbert spaces.

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

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

(0.1) $$(x, y) \mapsto (x + y, y)$$

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

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

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

(ii) $$\langle Bu, v\rangle = \langle Au, Av\rangle$$

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 $n = 3$. After general matrices and their determinants were introduced, the spectral theorem for symmetric matrices was established by A. Cauchy in 1810 under the assumption that the eigenvalues are different. Later Weierstrass found the proof in the general case, and independently Gram and Weierstrass found a method to produce an orthonormal basis of eigenvectors for a given symmetric $n \times n$-matrix B. To find an eigenvector with largest eigenvalue one studies the extremal problem

(1) $$\max_x \langle Bx, x\rangle \quad : \quad ||x|| = 1$$

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

$$Bx_* = a_1 x_*$$

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

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

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

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

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

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

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

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

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

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 studied systems of linear equations in an infinite number of variables with certain bounds. In Fredholm's investiagations one starts with a sequence of matrices $A_1, A_2, \ldots$ where A_n is an $n \times n$-matrix and an infinite dimensional vector space

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

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

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

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

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

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

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. By the Gram-Schmidt construction there exists an orthonormal sequence $\{e_n\}$ in V. However, in order to be sure that it suffices to take a *denumerable* orthonormal basis it is necessary and sufficient that the normed space V is *separable*. Assuming this every $v \in V$ has a unique representation

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

The existence of this orthonormal family means that every separable Hilbert space is isomorphic to the standard space ℓ^2 whose vectors are infinite sequences $\{c_n\}$ where the square sum $\sum c_n^2 < \infty$. So in order to prove general results about separable Hilbert spaces it is sufficient to regard ℓ^2. However, the abstract notion of a Hilbert space is useful since inner products on specific linear spaces appear in many different situations. For example, in complex analysis 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 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 set

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

For each fixed p we get

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

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

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

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

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

It follows that

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

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

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

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

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

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

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

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

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

4.11 General results on Hilbert spaces.

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

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

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

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

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

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

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

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

By the identity (*) the right hand side is

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

It follows from (i-iii) that

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

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

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

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

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

then we have

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

from which the uniquness of k_* follows.

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 study eigenvalues of an $n \times n$-matrices A with complex elements. The spectrum $\sigma(A)$ is the n-tuple of roots $\lambda_1, \ldots, \lambda_n$ of the characteristic polynomial $P_A(\lambda) = \det(\lambda \cdot E_n - A)$, where eventual multiple eigenvalues are repeated.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

4.B.3 An inequality by Pick.

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

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

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

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

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

Now one has

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

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

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

Then (1) and the triangle inequality give

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

Hence there only remains to show that

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

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

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

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

4.B.4 Results by A. Brauer.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Results by Perron and Frobenius

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

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

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

4:B.8 Theorem. *Let A as above be a positive matrix which gives the eigenvalue s^*. For every complex $n \times n$-matrix $B = \{b_{pq}\}$ such that $|b_{pq}| \leq a_{pq}$ 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 $B = A$.*

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

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

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

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

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

§ 11. Commutative Banch algebras

Contents

Introduction Let B be a complex Banach space equipped with a commutative product whose norm satisfies the multiplicative inequality

$$(*) \qquad ||xy|| \leq ||x|| \cdot ||y|| \quad : x, y \in B$$

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 commutative Banach algebra with a multiplicative unit. A $\mathbf{C}$-linear form λ on B is multiplicative if:

$$(**) \qquad \lambda(xy) = \lambda(x) \cdot \lambda(y) \quad \text{for all pairs} \quad x, y \in B$$

When λ satisfies (**) and is not identically zero it is clear that $\lambda(e) = 1$.

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 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 the B-valued power series

$$(1) \qquad (e - x)^{-1} = e + x + x^2 + \ldots$$

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

$$\rho(x) = \lim_{n \to \infty} \rho_n(x)$$

One refers to the limit $\rho(x)$ as the spectral radius of x which reflects the inequality below.

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. By definition this is the weakest topology on $\mathcal{M}(B)$ for which every Gelfand transform $\widehat{x}$ becomes a continuous function. Hence there exists an algebra homomorphism from B into the commutative algebra $C^0(\mathcal{M}(B))$:

$$(*) \qquad x \mapsto \widehat{x}$$

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 transforms. 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 with the identity element 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

(**) $$S_*(x) = e + \sum_{k=1}^{\infty} x^k$$

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

$$(***) \qquad (\lambda \cdot e - x)^{-1} = \lambda^{-1} \cdot \big[\, e + \sum_{k=1}^{\infty} \lambda^{-k} \cdot x^k \,\big]$$

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

$$(1) \qquad \delta = ||(\lambda_0 \cdot e - x)^{-1}||$$

To each complex number λ we set

$$(2) \qquad y(\lambda) = (\lambda_0 - \lambda) \cdot (\lambda_0 \cdot e - x)^{-1}$$

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:

$$(3) \qquad (e - y(\lambda))^{-1} = e + \sum_{\nu=1}^{\infty} (\lambda_0 - \lambda)^{\nu} \cdot (\lambda_0 \cdot e - x)^{-\nu}$$

Next, for each complex number λ we notice that

$$(\lambda \cdot e - x) \cdot (\lambda_0 \cdot e - x)^{-1} =$$

$$\big[(\lambda_0 \cdot e - x) + (\lambda - \lambda_0) \cdot e\big](\lambda_0 \cdot e - x)^{-1} = e - y(\lambda) \implies$$

$$(4) \qquad (\lambda \cdot e - x) = (\lambda_0 \cdot e - x)^{-1} \cdot (e - y(\lambda))$$

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:

$$(\text{A.2.1}) \qquad (\lambda \cdot e - x)^{-1} = (\lambda_0 \cdot e - x) \cdot [\, e + \sum_{\nu=1}^{\infty} (\lambda_0 - \lambda)^{\nu} \cdot (\lambda_0 \cdot e - x)^{-\nu} \,]$$

We refer to (A.2.1) as a local Neumann series. The triangle inequality gives the norm inequality:

$$||(\lambda \cdot e - x)^{-1}|| \leq ||(\lambda_0 \cdot e - x)|| \cdot \big[1 + \sum_{\nu=1}^{\infty} |\lambda - \lambda_0|^{\nu} \cdot \delta^{\nu} \,\big] =$$

$$(\text{A.2.2}) \qquad ||(\lambda_0 \cdot e - x)|| \cdot \frac{1}{1 - |\lambda - \lambda_0| \cdot \delta}$$

A.3. The analytic function $R_x(\lambda)$. From the above $\gamma(x)$ is an open subset of $\mathbf{C}$. Put:

$$R_x(\lambda) = (\lambda \cdot e - x)^{-1} \quad : \; \lambda \in \gamma(x)$$

The local Neumann series (A.2.1) shows that $\lambda \mapsto R_x(\lambda)$ 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.6 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

$$\mathfrak{r}(x) = \max_{\lambda \in \sigma(x)} |\lambda| \qquad (*)$$

We refer to $\mathfrak{r}(x)$ as the spectral radius of x. Notice that it gives 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. *For each* $x \in B$ *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

$$\lim_{n\to\infty} \xi(n) = \xi_* \qquad \text{(i)}$$

There remains to prove the equality

$$\xi_* = \mathfrak{r}(x)\,. \qquad \text{(ii)}$$

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

$$\xi_* \leq \mathfrak{r}(x) \qquad \text{(iii)}$$

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 of radius $> \frac{1}{\mathfrak{r}(x)}$. Hence Hadamard's limit formula for power series with values on a Banach space 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 it follows that $\mathcal{M}_B$ is 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

(*) $$\sigma(x) = \hat{x}(\mathcal{M}_B)$$

Hence Theorem 4.1 gives the equality:

(**) $$\lim_{n \to \infty} ||x^n||^{\frac{1}{n}} = \max_{\mathfrak{m}} \hat{x}(\mathfrak{m}) = |\hat{x}|_{\mathcal{M}_B}$$

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

(*) $$|\widehat{x}|_{\mathcal{M}_B} \leq ||x||$$

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. 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:

(1) $$\sigma_{B^*}(x) \subset \sigma_B(x)$$

The example above shows that this inclusion in general is strict. However, one has the opposite inclusion

(2) $$\partial(\sigma_B(x)) \subset \sigma_{B^*}(x)$$

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 $\mathbf{R}^n$ semi-analytic strata 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 each measure in $L^1(S)$ is identified with a Riesz measure in $\mathbf{R}^n$. One can easily prove that every $\mu \in L^1(S)$ has a power which belongs to the Varopolulos algebra. From this it follows that the closed subalgebra of $M(\mathbf{R}^n)$ generated by the family $\{L^1(S)\}$ and $V(\mathbf{R}^n))$ yields a Wiener algebra.

3.6 Olofsson's example. Above real analytic strata were used to obtain $\mathcal{V}^*$. That real-analyticity is essential was demonstrated 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 similar examples on curves defined by C^∞-functions outside the Carleman-Denjoy class of quasi-anaytic functions.