Functional analysis and operator theory

Contents.

Special chapters

Introduction.

Basic facts about normed vector spaces and bounded linear operators appear in § 1-6 under the headline Special Chapters. The less experienced reader may prefer to start directly with § 1-6 and postpone the reading of earlier text until later. However, a survey appears under the headline *Some basic facts in functional analysis* and a glimpse upon this introductory section is helpful as a guidance for the material in the special chapters. In addition to § 1-6, section 7 contains material which is needed in § 8-9 where we treat the spectral theorem for both bounded and unbounded self-adjoint operators in Hilbert spaces. Basic facts about commutative Banach algebras are established in § 10 and § 11 treats scatterred results which foremost deal with operators on Hilbert spaces.

Differential operators. In addition to general theory devoted to topological vector spaces and their linear operators, Chapter xx is devoted to partial differential operators. An instructive lesson appears in § xx where we construct fundamental solutions to second order elliptic operatars with variable coefficients which in general need not be symmetric. As we shall see this is achieved via Newton's classical formula for fundamental solutions of second order elliptic operators with constant coefficients, together with extensive use of Neumann series which are used to solve integral equations. The reward is that fundamental solutions and Greens' functions, are found in a canonical fashion with accurate estimates. Another example is the solution to first order differential equations of hyperbolic type in § xx where the crucial step to get solutions employs unbounded operators on Hilbert spaces of the Sobolev type. A veritable challenge appears during the study of the time-dependent Schrödinger equation

$$i \cdot \frac{\partial f}{\partial t} = -\Delta(f) + V(x)f$$

where Δ is the Laplace operaror in $\mathbf{R}^n$ for some positive integer n and $V(x)$ a real-valued potential function. One seks solutions $f(x,t)$ where x belongs to some domain Ω which may be bounded or unbounded, whike $t \geq 0$. In addition one has some intial conditions and eventual boundary conditions to ensure that the solution is unique, It goes without saying that this leads to quite delicate problems. In § xx we discuss certain situations where the equation above occurs. Let us mention that already when $n = 1$ leads to a quite rich theory.

Remark. It is a matter of taste if one should regard all this as a topic within functional analysis or not. Personally I think the general theory only becomes interesting when constructions of this nature are performed. So in that perspective § xx is an essential and "rewarding" section for the whole text.

The spectrum of a linear operator. The passage to infinite dimension leads to phenomena which do not occur in the finite dimensional case. Consider as an example the Hilbert space ℓ^2 whose vectors are sequences of complex numbers $x_0, x_1, x_2, \ldots$ where $\sum |x_k|^2 < \infty$. Now there exists the forward shift operator S defined by $S(x_0, x_1, x_2, \ldots) = (x_1, x_2, x_3, \ldots)$ It is clear that S is surjective, i.e. its range is ℓ^2 while the kernel is the 1-dimensional subspace generated by $e_0 = (1, 0, 0, \ldots)$. On the other hand, linear algebra teaches that if $A \colon V \to V$ is a linear and operator on a finite dimensional vector space whose range is V, then A is injective. The spectrum of the linear operator S is consists of the set of complex numbers λ for which $\lambda \cdot E - S$ is not invertible, where E is the identity on ℓ^2. It turns out that this is a compact set in the complex λ-plane which we denote by $\sigma(S)$. In the infinite dimensional case there does not exist a direct substitute for determinants which are used to descibe spectra of linear operators acting on finite dimensional vector spaces. To find $\sigma(S)$ one must therefore perform a special study. if $\lambda| < 1$ there exists a non-zero vector x such that $S(x) = \lambda \cdot x$. In fact, this follows when we take $x = (1, -\lambda, -\lambda^2, \ldots$. Hence $\lambda \cdot E - S$ has a non-zero kernel and we conclude that the open unit disc D is contained in $\sigma(S)$. If $\lambda = e^{i\theta}$ the situation is different. Here one checks that no eigenvectors exist, i.e. the operator $e^{i\theta} \cdot E - S$ is injective. But it is not invertible. The reason is that the range of $e^{i\theta} \cdot E - T$ is a dense subspace of ℓ^2, but not equal to the whole space. To see this one can use another model of ℓ^2 given by the Hardy space $H^2(T)$ whose vectors are square integrable functions on the unit circle T. which are boundsty values of analytic functions in the unit disc. Every such function has a Fourier series expansion

$$f(e^{i\theta}) = \sum_{n=0}^{\infty} a_n \cdot e^{in\theta}$$

which yield the analytic extension via the series $f(z) = \sum_{n=0}^{\infty} a_n \cdot z^n$. Then S can be expressed by

$$S(f) = \frac{f(z) - f(0)}{z}$$

With $\theta = 0$ and some $q \in H^2(T)$ we seek f so that

$$f - S(f) = q(z)$$

which means thar

$$(z+1)f = zq(z) + f(0)$$

Set $f = f(0) + zg(z)$ and now g must satisfy

$$(z+1)zg + zf(0) = zq(z) \implies (z+1)g = q(z) - f(0)$$

Above we are free to chose $f(0)$. If $q(-1) = a$ we try $f(0) = a$, and if q is continuous on T one may expect that g exists. However, even with this choice of $f(0)$ one can constrcut continuuos q-functions where no $g \in H^2(T)$ solves (xx). See § for further details.

Remark. The example above is a first example of a bounded linear operator which spectrum fills up a show disc and for spectral poinits which belong to the boundary T there do not exist non-zero eigenvectors.

Adjoint operators. In general, let $A \colon X \to X$ be a bounded inear operstor from a Banach space X into itself. We shall learn how to construct the dual space X^* and the adjoint operator $A^* \colon X^* \to X^*$. A basic result asserts that A is invertible if and only if A^* is so which entails that $\sigma(A) = \sigma(A^*$. We prove this in § xx.

Unbounded linear operators. They arise from pairs $(A, \mathcal{D}(A)$ where $\mathcal{D}(A)$ is a dense subspace of a Banach space X and $A \colon \mathcal{D}(A) \to X$ a linear map. The graph

(i) $$\Gamma(A) = \{(x, A(x) : x \in \mathcal{D}(A)\}$$

is a subspace of $X \times X$ and one says that A is closed when (i) is a closed subspace. The study of densely defined and closed linear operators is a natural subject because one encounters many examples in applications. A crucial point is that one can associate a well-defined spectrum $\sigma(A)$ of every densely defined and closed linear operator. But in contrast to bounded operators, the spectrum is in general not a compact set of the complex λ-plane.

Example. We have the Hilbert space $L^2(\mathbf{R})$ of complex-valued and square integrable functions on the real x-line. Consider a differential operator P with polynomial coefficients:

$$P(x, \frac{d}{dx}) = p_0(x) + p_1(x)\frac{d}{dx} + \dots p_m(x)\frac{d^m}{dx^m} +$$

So above $m \geq 1$ and $\{p_k(x)\}$ are polynomials with complex coefficients. Since test-functions belong to $\mathcal{D}(P)$, it follows that P is densely defined and in § xx we shall learn that it has a closed graph. The determination of $\sigma(P)$ is not at all clear, and it is only in the case when P has constant coefficients that one can perform a precise analysis. Namely, using Fourier's inversin formula and Parseval's equality this amounts to study linear operators on square integrable functions on the real ξ-line defined by $f(\xi) \mapsto p(\xi) \cdot f(\xi)$ where $P(\xi)$ is a polynomial We shall discuss this in § xx and here we only mention that if

$$P(x, \frac{d}{dx}) = \frac{d^2}{dx^2}$$

then the spectrum of P is the unbounded real interval $[0, +\infty)$.

Schödinger operators on the real line With $X = L^2(\mathbf{R})$ we there exists a densely defined operator

$$A(f)(x) = -\frac{d^2 f}{dx^2} + V(x)f(x)$$

when $V(x)$ is some complex-valued function which at least is locally square integrable on the real x-line. Here the spectral analysis becomes quite involved. An account appears in § xx where we treat the case when $V(x)$ is integrable, i.e. the L^1-norm

$$||V||_1 = \int_{-\infty}^{\infty} |V(x)|\, dx < \infty$$

Here one can search eventual eigenvalues λ outside $[0, +\infty)$ for which there exists a non-zero $f \in L^2(\mathbf{R})$ such that

$$-\frac{d^2 f}{dx^2} + V(x)f(x) = \lambda \cdot f(x)$$

In § xx we show that the absolute value of every such eigenvalue λ satisfies

$$|\lambda| \leq \frac{1}{4} \cdot ||V||_1^2$$

The proof relies upon Neumann's equations for resolvents and a pertubation argument using properties of the function

$$P_V(s) = \max_x \int e^{-s|x-y|}\, v(y)|\, dy$$

Remark. We could continue by listing further examples and results about spectral properties of unbounded operators of the Schrödinger type, i.e. arising via second order elliptic operators. Already the 1-dimensional case contains a wealth of results and many problems for current research remain open. The interested reader should consult the final chapters in the text-book *Linear Operators and their spectra* by E. Brian Davies which in addition to its excellent presentation of quite recent discoveries contains an extensive list of literature covering work up to 2006.

Fredholm operators. Let $A\colon X \to Y$ be a bounded linear operator from one Banach space into another. It is called Fredholm if both the kernel $\mathrm{Ker}(A) = \{x : A(x) = 0\}$ and $\mathrm{Coker}(A) = \frac{Y}{A(X)}$ are finite dimensional vector spaces. The index of a Fredholm operator is defined by

$$\mathrm{ind}(A) = \dim(\mathrm{Ker}(A) - \dim(\mathrm{Coker}(A))$$

In § xx we shall learn that if $T\colon X \to X$ is a compact operator on a Banach space X, then $A = E + T$ is Fredholm and has index zero. Next, to each linear operator $A\colon X \to X$ on a complex Banach space its *essential spectrum* is defined as the set of complex numbers λ for which $\lambda \cdot E - A$ is not a Fredholm operator. It is denoted by $\sigma_{\text{ess}}(A)$ and we shall learn that it is a compact subset of $\mathbf{C}$. The full spectrum $\sigma(A)$ is the disjoint union of its essential spectrum and a (possibly empty) set of spectral values $\{\lambda_k\}$ where $\lambda_k \cdot E - A$ are Fredholm. If $T\colon X \to X$ is compact, then a result which goes back to Fredholm, and for operators on general Banach spaces was put forward by F. Riesz around 1910, asserts that $\sigma_{\text{ess}}(T)$ is reduced to $\{0\}$.

Toeplitz' operators on $H^2(T)$ We consider the Hardy space from § xx. Analytic function theory shows that an L^2-function g on T belong to $H^2(T)$ if and only if

$$\int_0^{2\pi} e^{in\theta} \cdot g(e^{i\theta})\, d\theta = 0 \quad : n = 1, 2, \ldots$$

There exists the bounded projection operator $\Pi : L^2(T) \to H^2(T)$. More precisely, each $f \in L2(T)$ has a unique Fourier expansion

$$f(e^{i\theta}) = \sum a_n \cdot e^{in\theta}$$

where the sum extends over all integers and $\sum |a_n|^2 < \infty$. Now

$$\Pi(f)(e^{i\theta}) = \sum_{n=0}^{\infty} a_n \cdot e^{in\theta}$$

Next, let $a(\theta)$ belong to $L^\infty(T)$, i.e. it is a bounded measurable function in the sense of Lebesgue. The Toeplitz operator T_a on $H^2(T)$ is defined by

$$T_a(f) = \Pi(af)$$

Thus one multiplies functions $f \in H^2(T)$ with a and after they are projected to the Hardy space. The determination of the essential spectrum of T_a is quite delicate and is an active area for research up to the present date. Only the case hen a is a continuous function is easy. In § x we show that $\text{ess}(\sigma(T)$ is equal to the range $a(T)$. But already the case when a is piecewise continuous leads to situations where no such simple formula exists.

The Dirichlet problem. Let Ω be a bounded and connected open set in $\mathbf{R}^n$ for some $n \geq 3$. We assume that the boundary $\partial\Omega$ is a finite union of closed surfaces of class C^1 which implies that a well-defined area measure $d\omega$. By a result which goes back to Newton each continuuos boundary function $g \in C^0(\partial\Omega)$ yields a harmonic function G in Ω defined by

$$G(p) = \int_{\partial\Omega} |p - q|^{-n+2} \cdot g(q)\, d\omega(q)$$

Moreover, the C^1-hypothesis and elementary calculus entail that G extends to a continuous function on the closure of Ω. In particular there exists the boundary function $g^* = G|\partial\Omega$ which by Newton's construction has G as its harmonic extension. Now

$$T\colon g \mapsto g^*$$

is a linear operator from $C^0(\partial\Omega)$ into itself. In § xx we explain how the maximum principle for harmonic functions implies that the Dirichlet problem has a solution for every continuous boundary function if the range of the linear operator T is dense. At this stage one profits upon general functional analysis. Namely, the dual of $C^0(\partial\Omega)$ consists of Riesz measures on the boundary, so if the range is not dense there exists a non-zero Riesz measure μ on $\partial\Omega$ with the property that

$$\int_{\partial\Omega} \big[\, |p - q|^{-n+2} \cdot g(q)\, d\omega(q) \,\big]\, d\mu(p) = 0$$

for every $g \in C^0(\partial\Omega)$. In § xx we explain why this vanishing entials that the harmonic function defined in the complement of $\partial\Omega$ by

$$U_\mu(x) = \int_{\partial\Omega} \big[\, |p-x|^{-n+2}\, d\mu(p)$$

vanishes identically and from this one can deduce that $\mu = 0$ which gives a contradiction and proves that the Dirchlet problem always can be solved for every $g \in C^0(\partial\Omega)$.

The Riemann hypothesis. One does not need "sophisticated theoretical concepts" to arrive at deep problems. An outstanding challenge is to prove (or disprove) the famous Riemann hypothesis about zeros of the ζ-function. Beurling has proved that this hypothesis is equivalent to a certain closure result which at first sight appears to be "elementary" and goes as follows: To each positive real number x we denote by $\{x\}$ the unique largest integer such that $0 \le x - \{x\} < 1$. If M is a positive integer we define functions $\rho_1(x), \ldots, \rho_M(x)$ on the open unit interval $(0,1)$ by

$$\rho_j(x) = \frac{j}{Mx} - \{\frac{j}{Mx}\}$$

Set

$$\beta(M) = \min \int_0^1 (c_1 \cdot \rho_1(x) + \ldots c_M \cdot \rho_M(x) - 1)^2\, dx$$

with the minimum taken over M-tuples of real numbers $\{c_j\}$ which satisfy

(i) $$\sum_{j=1}^{j=M} jc_j = 0$$

Beurling proved that Riemann's hypothesis is true if and only if

(*) $$\lim_{M \to \infty} \beta(M) = 0$$

In spite of the appearantly simple assertion one is left with an outstanding open question. Notice that each $\beta(M)$ is given as the minimum of a quadratic form of the M-tuple $c_1, \ldots, c_M$ satisfying the linear equation (i).The difficulty to check if (*) holds or not is that the ρ-functions become quite irregular as M increases, and they hide subtle arithmetic properties which is the reason why the Riemann conjecture remains open.

A result by Hardy and Littlewood. For students who enter the study of functional analysis, a device is to first perform some serious studies about additive series and learn about such beautiful results as the Tauberian theorem by Hardy and Littlewood which asserts the following: Let $\{a_n\}$ be a sequence of non-negative real numbers such that there exists the limit

$$A = \lim_{r \to 1} (1-r) \cdot a_n r^n$$

Then it follows that

$$\lim_{N \to \infty} \frac{a_1 + \ldots + a_N}{N}$$

also exist and its limit value is equal to A. The proof offers an instructive lesson about how to handle inequalities which is the core in functional analysis. It is given in Chapter 1 from my notes in analytic function theory. Let me remark that the result above cannot be deduced from "soft methods" which rely upon Baire's category theorem. So the Hardy-Littelwood theorem is not covered by the material in this chapter.

A result by Beurling. We shall learn about the construction of dual spaces and the notion of weak convergence. An example of a Banach space consists of bounded continuous functions $f(x)$ on the real line equipped with the maximum norm. A subspace arises via the inverse Fourier transform, i.e. for every Riesz measure γ on the real ξ-line with a finite total variation we get the bounded continuous function

$$f(x) = \int e^{ix\xi}\, d\gamma(\xi)$$

Let $\mathcal{A}$ be its closure taken in the Banch space $C_b^0(\mathbf{R})$. Next, a sequence $\{\mu_n\}$ of Riesz measures on the real x-line is said to converge weakly in Beurling's sence if there exists a constant M such that the total variations $||\mu_n|| \leq M$ for each n, and

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

hold for each fixed ξ. It is easily seen that if $f \in \mathcal{A}$ then

$$\lim_{n\to\infty} \int f(x)\, d\mu_n(\xi) = 0$$

for every sequence $\{\mu_n\}$ which converges to zero in Beurling's sense. In § xx we prove thst the converse holds, i.e. the convergence property above for a given bounded continuous function f entails that $f \in \mathcal{A}$. In contrast to proofs of general nature which are derived via Baire's theorem, the result by Beurling requires a quite involved proof. So the reader should be aware of the fact that the general theory is helpful but not always sufficent to attain more refined results.

Historic comments.

The need for functional analysis in a general set-up was recognized after Weierstrass had found cases where "obscure phenomena" appear in the calculus of variation, starting from the study of minimal surfaces with fixed boundaries in $\mathbf{R}^3$ which leads to some peculiar situations which were discovered by Hermann Schwarz in 1866 while he considered surfaces with vanishing mean curvature. To overcome these obstacles new methods were invented by Carl Neumann in his pioneering work *Untersuchungen über das logarithmische und Newtonsche Potential*[Teuber. Leipzig. 1877]. Here elliptic boundary value problems were solved by operator methods which at that time was a revolutionary idea. Neumann's constructions of bounded resolvent operators is fundamental. In § 0.0 we expose Neumann's spectral theory for unbounded densely defined operators which leads to an operational calculus with a wide range of applications.

Results which settled Neumann's boundary value problem in non-convex domains were established by Poincaré in the article *La méthode de Neumann et le problème de Dirichlet* from 1897. His analysis is based upon solutions to certain variational problems where a crucial step was to use a symmetrization of the Neumann kernel $K(p, q)$ which is defined on the boundary S of a domain in $\mathbf{R}^3$ of class C^2. In § xx below we discuss two famous value problems whose solutions have inspired later work.

More extensive classes of integral operators than those defined by the Newton potential were introduced by Fredholm whose article *Sur une classe d'équationes fonctionnelles* [Acta math. 1903] has been very influential during the early development of functional analysis. Fredholm's studies of integral equations inspired Hilbert to create an abstract theory which led to the notion of infinite dimensional Hilbert spaces. His book *Integralgleichungen* from 1904 gave the first general account of linear operators and contains the spectral theorem for bounded self-adjoint operators on Hilbert spaces. After the contributions by Fredholm and Hilbert, spectral properties of operators on Hilbert spaces were investigated by many authors. An example is a result due to Jentsch from 1912 concerned with symmetric integral operators with positive kernels. His theorem asserts that for every such operator there exists a positive real eigenvalue λ_* whose eigenspace is one-dimensional and generated by a positive function, while the remaining eigenvalues all have absolute value $\geq \lambda_*$. This should be compared with earlier results due to Perron and Frobenius about positive matrices where we recall that if $A = \{a_pq\}$ is an $N \times N$-matrix whose elements are real and positive numbers, then the characteristic polynomial $P_A(\lambda)$ has a simple real root $s^* > 0$ and $|\alpha| < s^*$ hold for every other root.

The need for a spectral theory of unbounded linear operators was put forward by Birkhoff and Weyl around 1910 while they studied differential operators. Consider for example a second order ODE-equations

$$P(u) = \frac{d^2u}{dx^2} + q(x) \cdot u$$

where $q(x)$ is a real-valued continuous function on the real x-line. Given a bounded interval $[0, \ell]$ one seeks solutions u to the equation

$$(*) \qquad P(u) + \lambda \cdot u = 0$$

which satisfy the boundary conditions $u(0) = u(\ell) = 0$. Here it is easy to show that the set of real numbers λ for which (*) has a non-trivial solution form an infinite discrete sequence of real numbers. Passing to an infinite interval $[0 \leq x < +\infty)$ the situation becomes more involved. Here one seeks solutions u of (*) such that $u(0) = 0$ and

$$\int_0^\infty u(x)^2 \, dx < \infty$$

In contrast to the case of bounded intervals non-real λ can occur in (*). Following Weyl one says that the "Grenzpunktfall" occurs if (*) has no non-trivial solutions $u \in L^2\mathbf{R})$ when λ is non-real. The significance of this condition can be understood after one has constructed spectral resolutions of unbounded self-adjoint operators. A crucial fact is that when Weyl's "Grenzpunktfall" occurs, then there exists a unique spectral function of the given ODE-operator P. Extension to elliptic PDE-equations was undertaken by Carleman and we give further comments and examples in the subsection below about self-adjoint operators on Hilbert spaces.

A personal remark. The reader will recognize that many results are attributed to Torsten Carleman (1892-1949). This is motivated by his extensive contributions to the general theory during the years 1916-1923. Starting from pioneering work by C. Neumann, Poincaré, Hadamard, Fredholm and Hilbert, followed by some further contributions by Schmidt, Weyl and Toeplitz whose articles at an early stage treated various unbounded operators, Carleman established the spectral theory for unbounded self-adjoint operators on separable Hilbert spaces in 1923. He also proved several fundamental inequalites which lead to quite precise estimates for resolvent operators and can be used to obtain asymptotic formulas for eigenvalues of elliptic PDE-operators.

Next follow some comments which point out some of the major issues and expose some basic facts which are used frequently in functional analysis.

Matrices and determinants

.

A separate section is devoted to finite dimensional constructions which are essential to study linear operators on normed vector spaces of infinite dimension. We expose the Cayley-Hamilton calculus for matrices which is helpful as a start before Neumann's resolvents are studied for operators on infinite dimensional spaces. In § xx we describe some results from Hilbert's text-book *Integralgleichungen* about the construction of Fredholm's determinants which are used to express the resolvents of integral operators. Here is an important outcome of the Fredholm-Hilbert theory. Let Ω an open set in $\mathbf{R}^n$ and $K(x,y)$ a real-valued and square integrable function in $\Omega\times\Omega$. In general $K(x,y)\neq K(y,x)$ so the integral operator on the Hilbert space $L^2(\Omega)$ defined by

$$\mathcal{K}(f)(x)=\int_\Omega K(x,y)\cdot f(y)\,dy$$

is not self-adjoint and its spectrum needs not be real. A major fact in the Fredholm-Hilbert theory asserts that the kernels $\Gamma(x,y,\lambda)$ of the resolvent operators $R_K(\lambda)=(\lambda\cdot E-\mathcal{K})^{-1}$ are operator-valued meromorphic functions of the complex variable λ whose poles are controlled via Fredholm's resolvent determinant $\mathcal{D}_K(\lambda)$ which makes the operator valued function $\lambda\mapsto\mathcal{D}(\lambda)\cdot R_K(\lambda)$ entire. In §§ we prove a result due to Carleman which asserts that if $\{\lambda_n\}$ is the sequence of complex numbers where $\{\frac{1}{\lambda_n}\}$ give the non-zero spectral values of the compact operator $\mathcal{K}$, then

$$\mathcal{D}_K(\lambda)=\prod\,(1-\frac{\lambda}{\lambda_n})\cdot e^{\frac{\lambda}{\lambda_n}}$$

where the right hand side is the convergent Hadamard product of an entire function of exponential type. We have announced this to underline that more delicate inequalities in operator theory often rely heavily upon analytic function theory or Fourier analysis. So the reader should be aware of the fact that functional analysis is not "an independent subject", i.e. the results in the general theory serve as a guidance while one often needs specific computations from other disciplines to attain sharp inequalities. Another example where analytic function theory is used appears in Thorin's theorem. The proof is given in § xx using Hadamard's convexity theorem for maximum norms of analytic functions in strip domains.

Hadamard's radius formula. A veritable "high-point" appears in §§ XX. One starts with a local Taylor series $\sum c_nz^n$ which converges and gives an analytic function in some disc centered at the origin. In his thesis from 1894, Hadamard found necessary and sufficient conditions in order that f extends to a meromorphic function to larger discs, expressed via properties of the recursive Hankel determinants defined by the Taylor sequence $\{c_n\}$. Moreover, Hadamard's precise analysis of determinants, give insight about absolute values of the poles which occur while f extends to a meromorphic function.

2. Measure theory.

We assume that the reader is already familiar with Lebesgue theory, as well as Borel-Stieltjes integrals and the description of the dual space of continuous functions on a compact space via the representation theorem by F. Riesz. Relevant background appears in the appendix of my notes devoted to measure theory. An exception is the construction of integrals in the sense of Hellinger where quadratic differences appear while the construction of Stieltjes' integrals is extended. In § xx we expose results from his work *Die Orthogonalinvarianten quadratischer Formen von unendlichen vielen Variabeln* [Dissertation. Göttingen 1907]. Let us remark that the construction of Borel-Stieltjes integrals of complex-valued functions extends verbatim to vector-valued functions. For example, let X be a Banach space and consider a function

$$\text{(i)} \qquad s \mapsto x(s)$$

defined on a real interval $\{a \leq s \leq b\}$. Introducing the norm on X we consider variations with respect to partitions of $[a, b]$, i.e. when $a = s_0 < s_1 < \ldots < s_M = b$ we set

$$V = \sum_{\nu=0}^{M-1} ||x(s_{\nu+1}) - x(s_\nu)||$$

If there exists a constant C such that $V \leq C$ for all partitions as above one says that the vector-valued function (i) has a bounded total variation. The same construction as for complex-valued functions give Borel-Stieltjes integrals

$$\int_a^b g(s) \cdot \frac{dx}{ds}$$

for every $g \in C^0([a, b])$ with values in X. More precisely, the integral is a limit of Borel-Stieltjes sums:

$$\sum_{\nu=0}^{M-1} g(s_\nu) \cdot (x(s_{\nu+1}) - x(s_\nu))$$

taken over partitions where $\{\max s_{\nu+1} - s_\nu\}$ tend to zero. The function in (i) is absolutely continuous if there to each $\epsilon > 0$ exists δ such that

$$\sum ||x(s_\nu - x(t_\nu)|| < \epsilon$$

whenever $\{[t_\nu, s_\nu]\}$ is a finite family of of subintervals of $[a, b]$ arranged so that $s_\nu \leq t_{\nu+1} < s_{\nu+1}$ and $\sum (s_\nu - t_\nu) < \delta$. So many constructions which appear in ordinary calculus and measure theory extend while one regards vector-valued integrals.

2.1 Weak limits of L^2-functions. Let us recall some fundamental results established by F. Riesz in the article *Memoire über Systeme integrierabarer Funktionen* [Math. Ann. 1910]. Without essential loss of generality we expose his results for functions in the Hilbert space $L^2[0, 1]$ where $0 \leq x \leq 1$ is the unit interval. Following Riesz, a sequence $\{f_n\}$ of Lebesgue measurable and square integrable functions converges weakly to an L^2-function f if there exists a constant M such that the L^2-norms $\{||f_n||_2 \leq M\}$ for every n and

$$\lim_{n\to\infty} \int_0^x f_n(s)\, ds = \int_0^x f(s)\, ds$$

hold for every x. A first result is that this weak convergence entails that

$$\lim_{n\to\infty} \int_0^1 f_n(s) \cdot \phi(s)\, ds = \int_0^1 f(s) \cdot \phi(s)\, ds$$

hold for every L^2-function ϕ. Next, let $\{g_n\}$ be another sequence of L^2-functions which under dominated converge to a limit function g. It means that there exists some non-negative L^2-function $\gamma(x)$ such that

$$|g_n(x)| \leq \gamma(x)$$

for all n, and $\lim_{n\to\infty} g_n(x) = g(x)$ holds almost everywhere on $[0,1]$. Under this assumptions, Riesz proved that g is an L^2-function and one has the limit formula

$$(*) \qquad \lim_{n\to\infty} \int_0^1 f_n(s)\cdot g_n(s)\,ds = \int_0^1 f(s)\cdot g(s)\,ds$$

This ressembles the Lebesgue theorem for dominated convergence in $L^1[0,1]$. A proof of Riesz' result in L^2 is given in § xx. Another useful fact by Riesz goes as follows: Let $\{f_n\}$ be a sequence in L^2 such that

$$\lim_{n\to\infty} \int_0^1 f_n(s)\cdot \phi(s)\,ds = 0$$

hold for every L^2-function ϕ. Then there exists a constant M such that

$$||f_n||_2 \leq M$$

for all n, i.e. the sequence is automatically bounded in the L^2-norm., Moreover, there exists a sequence of convex combinations

$$s_N = \sum_{k=0}^{M} c_0^{(N)} f_N + \ldots + c_M^{(N)} f_{N+M}$$

where each sum is taken over a finite set of integers and the c-coefficients are non-negative while $c_0^N + \ldots c_0^N = 1$ hold for all N and one has the strong limit

$$\lim_{N\to\infty} ||S_N||_2 = 0$$

Thus, by taking convex combinations from $\{f_n\}$ we obtain a sequence of s-functions which tend to zero in the L^2-norm.

3. Weak limits in separable Hilbert spaces.

Every separable Hilbert space is isomorphic to ℓ^2 whose vectors a sequenes of complex numbers $x = \{x_p\}$ with a finite norm

$$||x|| = \sqrt{\sum_{p=1}^{\infty} |x_p|^2}$$

In spite of this elementary definition, the geometry in ℓ^2 is very rich. A crucial fact is that if Π is a closed subspace in ℓ^2 then its complement

$$\text{(i)} \qquad \Pi^\perp = \{y \in \ell^2 : \langle x, y\rangle = 0\}$$

is a new subspace and

$$\ell^2 = \Pi \oplus \Pi^\perp$$

Next, consider a sequence of closed subspaces $\{\Pi_1, \Pi_2, \ldots\}$. Each vector x in ℓ^2 is decomposed via (i) to give

$$\text{(ii)} \qquad x = \pi_n(x) + \pi_n^\perp(x)$$

One says that the sequence $\{\Pi_n\}$ converges weakly to a closed subspace Π if the following hold:

$$(1) \qquad \lim_{n\to\infty} ||\pi_n(x) - x|| = 0 \quad : x \in \Pi$$

$$(2) \qquad \lim_{n\to\infty} ||\pi_n^\perp(x) - x|| = 0 \quad : x \in \Pi^\perp$$

It turns out that every sequence $\{\Pi_n\}$ of closed subspaces contains at least one convergent subsequence which tends weakly to a limit space Π. We prove this in § xx. The result is remarkable since the Hilbert space ℓ^2 has infinite dimension. while the assertion above gives a "weak compactness property" for the family of closed subspaces.

3.1. Kernel functions. Let us consider take $L^2[0,1]$ as a model for a separable Hilbert space. Thus, the vectors are complex-valued and square integrable functions in the sense of Lebesgue.

If Π is a closed subspace there exists an orthonormal basis $\phi_n(x)\}$, where each ϕ-function has L^2-norm equal to one and

$$\int_0^1 \phi_n(x) \cdot \overline{\phi_k}(x)\, dx = 0 \quad : k \neq n$$

We associate the kernel function

$$\phi(x,y) = \sum \phi_n(x) \cdot \overline{\phi_n}(y)$$

Now there exists the linear operator which sends each $h \in L^2[0,1]$ to

$$\Phi(h)(x) = \int_0^1 \phi(x,y) \cdot h(y)\, dy$$

Then one easily verifes that Φ is the orthogonal projection onto the subspace Π.

3.2.The general Bessel inequality. It asserts that for every pair h, g in $L^2[0,1]$ one has:

$$|\iint g(x)\phi(x,y)h(y)\, dxdy\,| \leq ||g||_2 \cdot ||h||_2 \tag{2.1}$$

with equality if and only if both h and g belong to Π. The proof of (2.1) is easy but we shall supply details in § xx.

3.3. Weak limits of operators Let $[a,b]$ be a bounded real interval where $0 < a < b$. Given a subspace Π and an orthonormal famly $\{\phi_n\}$ as above we consider a kernel function of the form

$$\rho(x,y) = \sum \mu_n \cdot \phi_n(x) \cdot \overline{\phi_n}(y)$$

where $0 \leq \mu_n \leq b$ hold for each n. In this case one has the Bessel inequality

$$|\iint g(x)\rho(x,y)h(y)\, dxdy\,| \leq b \cdot ||g||_2 \cdot ||h||_2$$

Again a weak compactness principle holds. More precisely, let $\{\rho_1, \rho_2, \ldots\}$ be a sequence of kernel functions constructed as above via pairs of an orthonormal famliy of ϕ-functions and a μ-sequence. Then there exists at least one subsequence $\{\rho_{n_k}\}$ which converges weakly to a bounded linear operator A which means that:

$$\lim_{k\to\infty} \int_0^1 \rho_{n_k}(x,y) \cdot h(y)\, dy = A(h)(x) \tag{3.1}$$

hold for every $h \in L^2[0,1]$.

4. Symmetric operators

Much attention will be given to the spectral theorem for unbounded self-adjoint operators on a separable Hilbert space. So let us already in this introduction give a brief exposition to illustrates our major theme. Consider a doubly indexed and symmetric sequence of real numbers $\{a_{pq}\}$, i.e. $a_{pq} = a_{qp}$ hold for all pairs of positive integers. The sequence is said to be of the Hilbert-Schmidt type if

$$\sum\sum a_{pq}^2 < \infty \tag{*}$$

When (*) holds we get a linear operator A on ℓ^2 which sends a vector x to a vector $y = \{y_p\}$ where

$$y_p = \sum_{q=1}^{\infty} a_{pq}x_q$$

The spectral theorem for symmetric matrices in the finite dimensional case was extended by Hilbert to bounded and symmetric linear operators on ℓ^2. This fundamental result is exposed in § 8 from special chapters. Applied to a Hilbert-Schmidt operator it gives a sequece $\{\phi_n\}$ of pairwise orthogonal eigenvectors of unit norm in ℓ^2 with non-zero real eigenvalues $\{\lambda_n\}$, i.e.

$$A(\phi_n) = \lambda_n \cdot \phi_n$$

hold for each n. Let $V = \oplus\, \mathbf{C} \cdot \phi_n$ be the closed subspace of ℓ^2 generated by these eigenvectors. Then the kernel of A is the orthogonal complement

$$V^\perp = \{y : \langle y, V\rangle = 0\}$$

and one has the direct sum decomposition

$$\ell^2 = V \oplus V^\perp$$

If $\{\psi_k\}$ is an orthonormal basis in $V^\perp$, then $\{\phi_n\}$ and $\{\psi_k\}$ give an ON-basis in ℓ^2 and there exists a unitary matrix U which sends the standard e-basis to the ON-basis given by the pair $\{\phi_n\}$ and $\{\psi_k\}$. It follows that

$$A = U^* \circ B \circ U$$

where B is the linear operator expressed by a diagonal matrix where the eigenvalues $\{\lambda_n\}$ appear.

4.1 Unbounded symmetric matrices. Consider a real and symmetric doubly indexed sequence $\{\alpha_{pq}\}$ such that

$$\sum_{q=1}^{\infty} \alpha_{pq}^2 < \infty \tag{1}$$

hold for each p. This yields a densely defined linear operator A on ℓ^2 whose domain of definition $\mathcal{D}(A)$ consists of ℓ^2-vectors x such that

$$\sum_{p=1}^{\infty} \Big|\sum_{q=1}^{\infty} \alpha_{pq} \cdot x_q\Big|^2 < \infty \tag{2}$$

A sequence of Hilbert-Schmidt operators $\{A_n\}$ defined by the double indexed sequences $\{a_{pq}^{(n)}\}$ of real numbers converges weakly to A if

$$\lim_{n\to\infty} \sum_{q=1}^{\infty} (a_{pq}^{(n)} - \alpha_{pq})^2 = 0 \tag{3}$$

hold for every positive integer p, and in addition there exists a constant M such that

$$\sum_{q=1}^{\infty} (a_{pq}^{(n)})^2 \leq \sum_{q=1}^{\infty} \alpha_{pq}^2 \tag{4}$$

for each p. The question arises if (3-4) entails that

$$\lim_{n\to\infty} A_n(x) = A(x) \tag{5}$$

hold for each vector $x \in \mathcal{D}(A)$, where the limit means that the ℓ^2-norms of $\{A_n(x) - A(x)\}$ tend to zero. It turns out that in general (5) fails. This leads us to the following: Denote by $\mathcal{S}$ the family of symmetric sequences for which (1) hold.

4.2 Definition. *An operator in $\mathcal{S}$ is of Class I if (5) hold for every weak approximation by Hilbert-Schmidt operators.*

4.3 Carleman's criterion. We shall learn that a necessary and sufficient condition for an operator A in $\mathcal{S}$ to be of Class I is that there does not exist a non-zero vector $\phi \in \mathcal{D}(A)$ such that

$$A(\phi) = i \cdot \phi$$

where i is the imaginary unit. If $A \in \mathcal{S}$ is not of Class I it is said to be of Class II. We will also give examples of Class II-operators A and mention that they include Stieltjes' symmetric matrices for which

$$a_{p,p+1} = a_{p+1,p} = b_p \quad : p = 1, 2, \ldots \tag{i}$$

where $\{b_p\}$ is a sequence of positive real numbers, while $a_{p,q} = 0$ for all other pairs. The condition that a Stieltjes matrix A is of class I is expressed by Stieltjes' criterion for determined moment problems which amounts to check if the infinite continued fractions associated to a given b-sequence converges or not. Let us remark that growth conditions on $\{b_p\}$ alone do not give criteria for the convergence of its associated continued fractions. This illustrates that in general one cannot expect "easy criteria " in order that a matrix in $\mathcal{S}$ is of Class I or not. However, the following "positive result" is proved in § XX: If the series

$$\sum_{p=1}^{\infty} \frac{1}{b_p} = +\infty$$

then the Stieltjes matrix in (i) is of Class I. However, this sufficient condition is not necessary, i.e. there exist Case I-operators for which the series above is convergent.

4.4 Weyl's example. In most applications one is given a separable complex Hilbert space $\mathcal{H}$ and a densely defined symmetric operator A where a denumerable orthonormal basis in $\mathcal{H}$ which represents A by an inifinite matrix is not given in advance. Here one often needs extra work in order to decide if A is a Case I operator or not. Of special interest are densely defined operators given by symmetric PDE-operators acting on Hibert spaces $L^2(\Omega)$ where Ω is an open set in some euclidian space $\mathbf{R}^n$ with $n \geq 1$. Already the case $n = 1$ leads to involved situations which were treated by Weyl in an article from 1908 while he studied second order differential operators acting on the Hilbert space $\mathcal{H}$ of square integrable functions on the non-negative real line $[0, +\infty)$. If $q(x)$ is a real-valued continuous function on $[0, +\infty)$ there exists the symmetric and densely defined operator

$$P = \frac{d^2}{dx^2} + q(x)$$

By Carleman's criterion in (4.3) P yields a self-adjoint operator if and only if there does not exist any non-zero L^2-function $\phi(x)$ such that

$$\text{(i)} \qquad \phi''(x) + q(x) \cdot \phi(x) = i \cdot \phi(x)$$

Weyl gave exampes of q-functions for which (i) has non-trivial solutions which means that P fails to be a Case I-operator. We discuss this further in § xx.

5. Neumann's resolvent operators

.

The reader should at an early stage become familiar with the construction of spectra of densely defined unbounded operators, and series expansions of their resolvents outside the spectrum. The construction and properties of Neumann's resolvent operators which can be defined for every densely defined operator on a Banach space is presented in § 0.0. The power of the Neumann calculus is illustrated in § xx where we construct fundamental solutions to elliptic second order differential operators in a canonical way, starting from Newton's explicit solutions for PDE-operators with constant coefficients which already appear in his famous text-books from 1666. Another example where Neumann's resolvent operators occur appears in the study of strongly continuous semi-groups of bounded operators on a Banach space. In § xx we prove a result due to Hille, Phillips and Yosida which gives a criterion for a closed and densely defined linear operator T on a Banach space to be the infinitesmal generator of a strongly continuous semi-group $\{S(t)\}$. The Hille-Phillips-Yosida theorem asserts that the necessary and sufficient condition for T to be an infinitesmal generator is that the spectrum $\sigma(T)$ is contained in a half-plane $\mathfrak{Re}(\lambda) \leq a$ for some real number a, and there exists a constant K and some positive real number $a^* > a$ such that the operator norms of the resolvents $R_T(\lambda)$ for real values of λ satisfy

$$||R_T(\lambda)|| \leq \frac{K}{\lambda} \quad : \lambda \geq a^*$$

where the left hand side means that we take operator norms of Neumann's resolvents.

6. Geometry of normed vector spaces.

This topic is treated rather scarcely. For example, we shall not discuss in detail more advanced facts about extreme points of a convex sets and barycentric representations which includes the notion of the Choquet boundary. But we expose some results which occur rather frequently in applications. The notion of convex subsets in a vector space X over the real field is fundamental. A convex set K which contains the origin is absorbing if every sufficiently dilated vector belongs to K. It means that for every $x \in X$ there exists some $a > 0$ such that x belongs to $aK = \{a\xi : \xi \in K\}$. When it holds the support function $\mathfrak{k}_K$ is defined for every $x \in X$ by

$$\text{(i)} \qquad \mathfrak{k}_K(x) = \inf_{a>0} x \in aK$$

The tangent functional of K assigns a number for each pair of vectors x, y in X by:

$$\text{(ii)} \qquad \mathfrak{T}_K(x,y) = \lim_{a \to 0^+} \frac{1}{a}(\mathfrak{k}_K(x+ay) - \mathfrak{k}_K(x)$$

The proof that (ii) exists is easy. See § xx for details. Above no topology has been introduced, i.e. one deals with pure geometry related to the convex set K. When X is a normed vector space further notions are introduced. The support function of a convex set K is strongly differentiable at a point $x \in X$ if

$$\text{(iii)} \qquad \lim_{||y|| \to 0} \frac{1}{||y||}\big(\mathfrak{k}(x+y) - \mathfrak{k}(x) - \mathfrak{T}_K(x,y)\big) = 0$$

where $||y||$ denotes norms during the passage to the limit. This definition, as well as several notions of differentiable with respect to other topologies than that defined by the norm, were investigated in a series of articles by Smulian around 1938. In § xx we prove one of his results which gives criteria for stong differentiability. Let X be a normed vector space and B the closed unit ball, i.e. vectors of norm ≤ 1. Then the following hold where X^* is the dual of X and B^* its unit ball.

6.1 Smulian's criterion. *A necessary and sufficient condition that (iii) holds with $K = B$ for a non-zero vector x, is that there exists a unique $x^* \in X^*$ with norm one such that $x^*(x) = ||x||$ with the additional property that whenever $\{x_n^*\}$ is a sequence in B^* such that $x_n^*(x) \to 1$, it follows that $||x_n^* - x^*|| \to 0$.*

6.2 Further results. In § xx we prove a result due to Mazur which asserts that when $X = L^p(\mu)$ for some $1 < p < \infty$ with respect to a measure space, then strong differentiability with respect to the norm hold for all $x \in X$. In §§ we prove a theorem by Hörmander in § xx which describes support functions on convex sets in locally convex topological vector spaces, and a non-linear functional appear in a theorem due to Beurling and Lorch in § xx.

7. Fixed point theorems.

A compact topological space S has the fixed point property if every continuous map $T : S \to S$ has at least one fixed point. An example is the closed unit ball in $\mathbf{R}^n$ whose fixed-point property is an immediate consequenc of Stokes Theorem. See § xxx for details. More generally, consider a locally convex vector space X. In § xx 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.

7.1 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 X we mean a family of bijective linear maps $g : X \to X$ such that composed maps $g_2 \circ g_1$ again belong to the group as well as the inverse of every g.

7.2 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 $g \in \mathbf{G}$.

7.3 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.

8. The approximation property.

Normed vector spaces can be quite involved. This is best illustrated by Enflo's example of a separable Banach space on which there exists a compact operator which cannot be approximated by operators with finite dimensional range. Enflo's ingenious construction was presented in a seminar at Stockholm University in 1972 which after led to an intense study where one seeks to determine when a given separable Banach space X has the approximation property in the sense that every compact operator on X can be approximated in the strong operator norm by operators with finite dimensional range. Fortunately most Banach spaces encountered in applications enjoy this property. For example, if S is a compact metric space which contains a denumerable dense subset, then the Arzela-Ascoli theorem for equi-continuous functions entails that the Banach space $C^0(S)$ has the approximation property. See § xxx for the proof. Let us also remark that if the approximation property hold for the family of compact operators in a separable Banach space then there exists a Schauder basis. See § xx for further comments.

Preliminary examples.

While entering the study of a general theory it is useful to consider special situations. So we describe some results which from a historic perspective inspired later studies of operators on Hilbert spaces. We begin with an example which in a histriic perspective inspired the theory about linear operstors.

Neumann's boundary value problem.

Consider a domain Ω in $\mathbf{R}^2$ whose boundary is the union of a finite family of closed Jordan curves of class C^1. Given a pair of real-valued continuous functions a and f on $\partial\Omega$ one seeks a harmonic function U in Ω which extends continuously to the boundary. In addition the inner normal derivatives along the boundary exist where they satisfy the equation

$$\text{(1)} \qquad \frac{\partial U}{\partial n_*}(p) = a(p) \cdot U(p) + f(p) \quad : p \in \partial\Omega$$

Above the left hand side refers to inner normal derivatives. Here the situation is more involved since neither uniqueness nor existence are automatic. However, if a is a positive function then it turns out that the equation has a unique solution U. To prove this one regards a positive number M which exceeds the diameter of $\partial\Omega$, i.e. $M > |p-q|$ for every pair of points on the boundary. Now there exist the linear operator on $C^0(\partial\Omega)$ defined by

$$\text{(2)} \qquad \mathcal{L}(u)(p) = \int_{\partial\Omega} \log \frac{M}{|p-q|} \cdot u(q)\, ds(q)$$

where ds is the arc-length measure on the boundary. Since the function $(x,y) \mapsto \log \sqrt{x^2+y^2}$ is harmonic outside the origin, it follows that (2) defines a harmonic function U in Ω when p varies in this domain. To solve (1) is therefore equivalent to find $u \in C^0(\partial\Omega)$ such that

$$\text{(3)} \qquad \frac{\partial \mathcal{L}(u)}{\partial n_*}(p) = a(p) \cdot \mathcal{L}(u)(p) + f(p) \quad : p \in \partial\Omega$$

This is the gateway to settle Neumann's boundary value problem. The crucial point that the composed operator

$$u \mapsto \frac{\partial}{\partial n_*} \circ \mathcal{L}(u)$$

acting on $C^0(\partial\Omega)$ is expressed by the linear operator

$$\text{(iv)} \qquad \mathcal{S}(u)(p) = -\pi \cdot u(p) + \int_{\partial\Omega} \frac{\langle n_*(p), q-p\rangle}{|q-p|^2} \cdot u(q)\, ds(q)$$

The verifiction of (iv) relies on calculus and given in § xx. Next, the assumption that $\partial\Omega$ is of class C^1 gives a constant C such that

$$\text{(v)} \qquad \max_{p,q} \Big| \frac{\langle n_*(p), q-p\rangle}{|q-p|^2} \Big| \leq C$$

with the maximum taken over all pairs p, q on the boundary, We shall learn that this entails that the linear operator on $C^0(\partial\Omega)$ defined by

$$\mathcal{K}(u)(p) = \int_{\partial\Omega} \frac{\langle n_*(p), q-p\rangle}{|q-p|^2} \cdot u(q)\, ds(q)$$

is compact. Neumann's operator is defined by

$$\text{(vi)} \qquad \mathcal{N}(u)(p) = a(p) \cdot \mathcal{L}(u)(p) - \mathcal{K}(u)(p)$$

As explained in § xx, Lebesgue theory and the Arzela-Ascoli theorem entail that the operator

$$u \mapsto a(p) \cdot \mathcal{L}(u)(p)$$

also is compact. Hence Neumann's operator is compact and we shall learn that it has a discrete spectrum. Next, for each M chosen as above we set

$$\rho(M) = \min_{p,q} a(p) \cdot \log \frac{M}{|p-q|} - C$$

Let E be the identity operator. If $\rho(M) > 0$ one easily verifies that the operator

$$u \mapsto s \cdot u + \mathcal{N}(u)$$

is invertible for all real numbers s such that

$$s + \rho(M) > 0$$

Next, notice that (1) can be written as

$$\pi \cdot u + \mathcal{N}(u) = -f$$

When $\rho(M) > -\pi$ it follows from the above that there exists the bounded inverse operator

$$(\pi \cdot E + \mathcal{N})^{-1}$$

and (1) is solved by

$$u = -(\pi \cdot E + \mathcal{N})^{-1}(f)$$

The case when the boundary is not smooth. Let Ω be a bounded domain in $\mathbf{R}^2$ such that $\partial\Omega$ is sufficiently smooth in order Stokes formula is valid as explained in Chapter 2 from my notes in analytic function theory. Let $\text{reg}(\partial\Omega)$ denote the regular part of the boundary. Given a pair a, f in $C^0(\partial\Omega$ one now asks for a harmonic function U in Ω which together with its inner normal derivatives extend to be continuous on $\text{reg}(\partial\Omega)$ and

$$\frac{\partial U}{\partial n_*}(p) = a(p) \cdot U(p) + f(p) \quad : p \in \text{reg}(\partial\Omega) \tag{1}$$

Again one tries to solve this via suitable linear operators. Since our assumption in particular means that the 2-dimensional area of $\text{reg}(\partial\Omega)$ is finite, it follows $\mathcal{L}$-operator is still defined, where the integral in (1) is taken over the regular part of $\partial\Omega$. Next, let us impose the condition that there exists a constant C such that

$$\max_{p \in \text{reg}(\partial\Omega)} \int_{\text{reg}(\partial\Omega)} \Big| \frac{\langle n_*(p), p-q \rangle}{|p-q|^2} \Big| \, ds(q) \leq C \tag{2}$$

Then each $u \in C^0(\partial\Omega)$ yields a function on the regular part of the boundary defined by

$$\mathcal{K}(u)(p) = \int_{\text{reg}(\partial\Omega)} \frac{\langle n_*(p), p-q \rangle}{|p-q|^2} \cdot u(q) \, ds(q) \tag{3}$$

where $\mathcal{K}$ becomes a linear operator from the Banach space $C^0(\partial\Omega)$ into the Banach space of bounded and continuous functions on $\text{reg}(\partial\Omega)$ whose operator norm is majorized by C. Hence there also exists the Neumann operator $\mathcal{N}$ given as the sum of the $\mathcal{L}$-operator minus the bounded operator $\mathcal{K}$. The question rises if we can produce bounded inverse operators similar to the smooth case for sufficiently large M. To achieve this we suppose in addition to (2) that there exists a constant C^* such that

$$\max_{p \in \text{reg}(\partial\Omega)} \int_{\text{reg}(\partial\Omega)} \frac{M}{\log|p-q|} \, ds(q) \leq C^* \tag{4}$$

Basic measure theory entails that $\mathcal{L}$ now becomes a compact operator whose spectrum depends on M. The whole discussion has shown that the essential point is to analyze the spectrum of Neumann's operator defined as in (vi) frome the smooth case while p stays in $\text{reg}(\partial\Omega)$. We shall treat this in § xx and settle Neumann's boundary value problem for domains where (2) and (4) hold.

A non-linear boundary value problem. In his lecture *L'avenir des mathématiques* at the international congress at Rome in 1908, Poincaré pointed out the significance of regarding infinite systems of linear equations which can be used to treat non-linear equations. His proposed strategy was used by Carleman in an article from 1921 to establish the result below. Let Ω be as above and fix a point $q_* \in \Omega$. Consider functions of the form

$$u(p) = \frac{1}{4\pi} \cdot \frac{1}{|p - q_*|} + H(p) \tag{i}$$

where H is harmonic in Ω. Let $F(s,p)$ be a real-valued continuous function defined on $\{s \geq 0\} \times \partial\Omega$ such that $s \mapsto F(s,p)$ is strictly decreasing on $[0, +\infty)$ with $F(0,p) = 0$ for each $p \in \partial\Omega$.

Theorem. *There exists a unique function u from (i) such that its inner normal derivative along $\partial\Omega$ satisfies the equation*

$$\frac{\partial u}{\partial n}(p) = F(u(p), p)$$

We shall prove this result in § xx where an essential ingredient is the systematic use of Neumann series of linear operators.

a. Toeplitz' operators.

Let $\{a_p\}$ be a Hermitian sequence of complex numbers indexed over all integers where

$$(*) \qquad \sum_{-\infty}^{\infty} |a_p|^2 < \infty$$

The hermitian condition means that

$$a_{-p} = \bar{a}_p \quad : p = 1, 2, \ldots$$

In an article [Göttingen Nachrichten: 1907], Toeplitz studied quadratic forms of infinitely many variables:

$$J(x) = \sum_{p=1}^{\infty} \sum_{q=1}^{\infty} a_{p-q} x_q x_p$$

which arise from sequences satisfying (*). If λ is a complex number with $\mathfrak{Im}(\lambda) \neq 0$ and $\{c_p\}$ is a vector such that $\sum_{-\infty}^{\infty} |c_p|^2 < \infty$, then Toeplitz proved that there exists a unique vector $\{x_p\}$ which again satisfies (*) and

$$(1) \qquad x_p = \lambda \cdot \sum_{q=-\infty}^{\infty} a_{p-q} x_q = c_p \quad : p \in \mathbf{Z}$$

In § xx we prove an integral version which gives Toeplitz' result as a special case and includes an inversion formula which represents the solution vector $\{x_p\}$ for each given c-vector. It goes without saying that the discoveries above by Toeplitz inspired later studies of linear operators. An example occurs in the moment problem of Stieltjes and its extended version by Hamburger. Here the starting point is the close interplay between expansions in continued fractions of a sequence of real numbers $\{c_1, c_2, \ldots\}$ with the property that the quadratic form

$$\mathcal{C}(x) = \sum \sum c_{p+q} x_p x_q$$

is positive, i.e. the value is > 0 for every x-vector where only finitely many $x_p \neq 0$. The article *Zur Einordnung der Kettenbruchtheorie in die Theorie der quadratischen Formen von unendlichvielen veränderlichen* [Crelle J. für Math. 1914] by Hellinger and Toeplitz led to a new perspective of Stieltjes' original solutions to the moment problem. Conclusive results which cover both determined and undetermined moment problems were attained by Carleman in the article *Sur le problème des moments* [C.R. Acad. Sci. Paris 1922] which we expose in § xx. A first acount about this is given in the next subsection.

b. The moment problem.

We shall briefly resume some results and refer to § xx for further details. Consider a pair of sequences of real numbers $\{a_p\}$ and $\{b_p\}$, indexed by integers $p \geq 1$. We assume that every $b_p > 0$ while no conditions are imposed on the a-sequence. It can for example occur that $a_p = 0$ for every p. To each $n \geq 2$ we get the symmetric $n \times n$-matrix

$$A_n = \begin{pmatrix} a_1 & -b_1 & 0 & \ldots & 0 & 0 \\ -b_1 & a_2 & -b_2 & \ldots & 0 & 0 \\ 0 & -b_2 & a_3 & \ldots & \ldots & \\ \ldots & \ldots & \ldots & \ldots & \ldots & \\ 0 & 0 & \ldots & a_{n-1} & -b_{n-1} & \\ 0 & 0 & \ldots & -b_{n-1} & a_n & \end{pmatrix}$$

Let μ be a complex number with $\mathfrak{Im}(\mu) > 0$. Then there exists a unique complex n-vector $\Phi(\mu)$ which solves the equation

$$\text{(i)} \qquad A_n(\Phi(\mu)) = \mu \cdot \Phi(\mu) + e_1$$

where $e_1 = (1, 0, \ldots 0)$. Indeed, we see that (i) holds if and only if

$$a_1\phi_1(\mu) - b_1\phi_2(\mu) = \mu\phi_1(\mu) + 1$$

$$-b_{k-1}\phi_{k-1}(\mu) + a_k\phi_k(\mu) - b_k\phi_{k+1}(\mu) = 0 \quad : 2 \leq k \leq n-1$$

$$b_{n-1}\phi_{n-1}(\mu) = a_n\phi_n(\mu)$$

From this it is clear that $\phi_1(\mu)$ determines the remaining components and treating μ as an indeterminate. In § xx we show that

$$\phi_1(\mu) = \frac{P_n(\mu)}{Q_n(\mu)}$$

where $Q_n(\mu)$ is a polynomial of degree n while P_{n-1} has degree $(n-1)$ Moreover, since the solution exists for every non-real μ it follows that the zeros of Q_n are real and since $\phi_1(\mu) \neq 0$ for non-real μ, the zeros of P_{n-1} are also real. Finally one shows that the polynomials satisfy the recursive equations

$$\frac{P_{n+1}(\mu)}{Q_{n+1}(\mu)} = \frac{P_n(\mu)}{Q_n(\mu)} + \frac{b_1^2 \cdots b_n^2}{Q_n(\mu) \cdot Q_{n+1}(\mu)}$$

Since the degree of the polynomial $Q_n(\mu) \cdot Q_{n+1}(\mu)$ is $\geq 2n+1$ for every n, it follows that when one takes the Laurent series expansions of the rational g-functions at $\mu = +\infty$, then the coefficients in negative μ-powers up to degree $2n$ are the same for $g_n(\mu)$ and $g_{n+1}(\mu)$. This gives a unique asymptotic series

(1) $$-\sum_{k=1}^{\infty} \frac{c_k}{\mu^{k+1}}$$

whose series expansion up to order $2n$ is matching that of $g_n(\mu)$ for every positive integer n. As explained in § xx, the series (1) corresponds to the expansion in continued fractions attached to the two sequences $\{a_k\}$ and $\{b_k\}$. In §§ we also show that there exists a probability measure ρ_n on the real t-line such that

(2) $$\frac{P_n(\mu)}{Q_n(\mu)} = \int_{-\infty}^{\infty} \frac{d\rho_n(t)}{t-\mu}$$

b.1 Passage to a limit. From (2) it follows that the functions $\{g_n(\mu) = \frac{P_n(\mu)}{Q_n(\mu)}\}$ are analytic in the upper half-plane and satisfy

$$|g_n(\mu)| \leq \frac{1}{\mathfrak{Im}(\mu)}$$

Montel's theorem entails that $\{g_n\}$ is a normal family of analytic functions in the upper half-plane. Denote by $\mathcal{G}$ the familly of all analytic functions $g(\mu)$ in the upper half-plane which are limits from a subsequence $g_{n_1}, g_{n_2}, \ldots$, i.e.

$$\lim_{k \to \infty} g_{n_k}(\mu) = g(\mu)$$

where the convergence holds uniformly over compact subsets of the upper half-plane. By weak-star compactenss of Riesz measures and the compatibility between the Laurent series of the g-functions, it follows that for each $g \in \mathcal{G}$, there exists a probability measure ρ such that

$$g(\mu) = \int_{-\infty}^{\infty} \frac{d\rho(t)}{t-\mu}$$

and the moments of the ρ-measure satisfy

(3) $$\int t^{\nu} \cdot d\rho(t) = c_{\nu} \quad : \nu = 1, 2, \ldots$$

Above the c-sequence is quite special since Cauchy-Schwarz inequality entails that the quadratic form in an infinite number of variables $x_1, x_2, \ldots$ defined by

(4) $$J(x) = \sum\sum c_{p+q}x_px_q$$

is positive definite. One refers to $\{c_n\}$ as a Hankel sequence. This termionology stems from the fact that J is positive definite if and only if the recursive Hankel determinants of the c-sequence are all positive. See the section about matrices for details.

b.2 The determined case. If $\mathcal{G}$ is reduced to a single function one says that the moment problem is determined which by the above means that there exists a unique probability measure ρ for which (3) holds. Necessary and sufficient conditions on the sequence $\{c_n\}$ in order that the moment problem is determined is not known. It appears as a very hard (and perhaps even unolvable) problem However, the following sufficiency result by Carleman holds:

B.2.1 Theorem. *A Hankel sequence $\{c_n\}$ yields a derermined moment problem if*

$$\int_1^\infty \log\big[\sum_{n=0}^\infty \frac{r^{2n}}{C_{2n}^2}\big]\,\frac{dr}{r^2} = +\infty$$

The proof of this result goes beyond the scope of the present chapter since it is based upon quite involved analysis where estimates for harmonic measures and constructions of suitable subharmonc functions are needed. In the case of a non-determined moment problem the following is proved in § xx: For each complex number μ in the open upper half-plane we set

$$\mathcal{R}(\mu) = \{g(\mu) :\ g \in \mathcal{G}\}$$

B.2.2 Theorem. *In a non-determined moment problem each set $\mathcal{R}(\mu)$ is a compact disc with positive radius placed in the upper half-plane.*

c. Picard's operator.

Consider the Hilbert space $L^2(\mathbf{R})$ whose vectors are square integrable complex valued functions on the real line. For each $f \in L^2\mathbf{R})$ we get a function

$$x \mapsto P_f(x) = \int_0^\infty e^{-[x-y|}\cdot f(y)\,dy$$

The Cauchy-Schwarz inequality shows that when x is fixed, then the integral in the right hand side is absolutely convergent which gives a well-defined function $P_f(x)$ on the real x-line. In § xx we use Fourier's inversion formula to calculate the spectral function $\theta(x,y;\lambda)$ of Picard's operator. More precisely, it is zero on the interval $(-\infty,1/2]$, and when $\lambda > 1/2$ one has

$$\theta(x,y;\lambda) = \int_{\frac{1}{2}} \frac{\sin\left(\sqrt{2s-1}\cdot(x-y)\right)}{x-y}\,ds$$

From this explicit equation for the spectral function we shall demonstrate in § xx how classic formulas in Fourier analysis can be interpretated via Carleman 's general results about their spectral functions. In our present situation one recaptures the Fourier-Parseval equation

$$(*)\qquad \int_{-\infty}^\infty f(x)g(x)\,dx = \frac{1}{\pi}\cdot\int_{-\infty}^\infty \frac{d}{d\lambda}\Big(\iint \frac{\sin\left(\sqrt{2\lambda-1}\cdot(x-y)\right)}{x-y}\cdot f(x)g(y)\,dxdy\Big)\,d\lambda$$

for each pair of L^2-functions f and g. More generally one considers operators on $L^2(\mathbf{R})$ defined by

$$\mathcal{H}(f)(x) = \int_0^\infty H_1(|x-y|)\cdot f(y)\,dy + \int_0^\infty H_2(x+y)\cdot f(y)\,dy$$

where the H-functions satisfy

$$\int_0^\infty H_\nu(s)^2\,ds < \infty$$

Again one computes the spectral function of this self-adjoint operator which is used to obtain equations similar to (*)

d. Inequalities via spectra.

An inequality which goes back to early work by Hilbert and Schur asserts that for every sequence $x_1, x_2, \ldots$ of real numbers one has the inequality

$$|\sum_{p,q} \frac{x_p x_q}{p+q}| \leq \pi \cdot \sum_{p=1}^{\infty} x_p^2$$

where the double sum is extended over all pairs of non-negative integers. The integral version asserts that if $f(x)$ is a square-integrable function on $\{x \geq 0\}$ then

$$\text{(d.1)} \qquad \iint \frac{1}{x+y} f(x) f(y)\, dx dy \leq \pi \cdot \int_0^{\infty} f(x)^2\, dx$$

where the double integral is taken over $\{x, y \geq 0\}$. To prove this one considers the linear operator T defined by

$$\text{(i)} \qquad T(f)(x) = \int_0^{\infty} \frac{1}{x+y} \cdot f(y)\, dy$$

It is clear that (d.1) follows if the linear operator T on the Hilbert space $L^2(\mathbf{R}^+$ has norm $\leq \pi$. To prove this we shall use variable substitutions. For each square integrable function g on the non-negative y-line we set Set $g_*(s) = g(e^s)$, and notice that

$$\text{(ii)} \qquad \int_{-\infty}^{\infty} g_*(s)^2 \cdot e^s\, ds = \int_0^{\infty} g(y)^2\, dy$$

Using a similar transformation $t \mapsto e^t = x$ we define the linear operator T_* on $L^2(\mathbf{R})$ by

$$\text{(iii)} \qquad T_*(f_*)(t) = \int_{-\infty}^{\infty} \frac{e^{\frac{t}{2}+\frac{s}{2}}}{e^t + e^s} \cdot f_*(s)\, ds$$

From (ii) the reader can check that the operator norms of T_* and T are the same. Define the function

$$\text{(iv)} \qquad K(u) = \frac{1}{e^{\frac{u}{2}} + e^{-\frac{u}{2}}}$$

Then it is clear that

$$\text{(v)} \qquad T_*(f_*)(t) = \int_{-\infty}^{\infty} K(t-s) \cdot f_*(s)\, ds$$

Hence T_* is defined via a convolution and Fourier's inversion formula entails that

$$T_*(f_*)(t) = \frac{1}{2\pi} \cdot \int e^{-it\xi} \cdot \widehat{K}(\xi) \cdot \widehat{f}(\xi)\, d\xi$$

Next, the Fourier transform of K is given by

$$\int_{-\infty}^{\infty} e^{-iu\xi} \cdot \frac{1}{e^{\frac{u}{2} - \frac{u}{2}}}\, du$$

Exercise. Use residue calculus to show that the Fourier transform

$$\widehat{K}(\xi) = \int_{-\infty}^{\infty} e^{-iu\xi} \cdot \frac{1}{e^{\frac{u}{2} - \frac{u}{2}}}\, du = \frac{2\pi}{e^{\xi} + e^{-\xi}}$$

Finally, Parseval's equality and the observation that

$$\min_{\xi} e^{\xi} + e^{-\xi} = 2$$

imply that the operator norm of T_* is at most π. Moreover, one has equality since we can take L^2-function $f(t)$ where the Fourier transform is the characteristic function of a small interval $[-\delta, \delta]$ on the real ξ-line in which case

$$||T_*(f)||^2 \geq \pi^2 \cdot \frac{4}{(e^\delta + e^{-\delta})^2} \cdot ||f||^2$$

Since $\frac{4}{(e^\delta+e^{-\delta})^2} \simeq 1$ when δ is small we conclude that the operator norm of T is π. Hence (d.1) is sharp and notice that equality cannot be attained for a single f-function, i.e. one must pass to the infimum while the operator norm of T is determined.

e. Infinite systems of linear equations.

In functional analysis one usually imposes constraints on infinite vectors. One may also study linear systems of equations under the sole condition that solutions are expressed by convergent additive series. Thus, let $\mathcal{C} = \{c_{pq}\}$ be an infinite matrix where the doubly-indexed elements are compex numbers. Given a complex sequence $\{a_p\}$ one studies the system:

$$\sum_{q=1}^{\infty} c_{pq} \cdot x_q = a_p \quad : p = 1, 2, \ldots$$

The infinite x-vector is a solution when the additive series in the left hand side converges for each fixed p with a sum equal to a_p. A necessary and sufficient condition for the existence of at least one x-solution is described in Carleman's lecture from the IMU-congress at Zürich in 1932. As expected the criterion is involved and relies upon solutions with bounds to a recursively constructed family of variational problems. A special case occurs when $\{c_{pq}\}$ are given via evaluations at pairs of integers of rational funtions of two variables. Here follows an example.

The Dagerholm series. Consider the homogeneous system

$$\sum_{q \neq p} \frac{x_q}{p-q} = 0 \quad : p = 1, 2, \ldots$$

It means that we seek vectors $x = (x_1, x_2, \ldots$ such that the left hand side above converges for each fixed positive integer p and the series sum is zero. Without loss of generality we only ask for real x-vectors. It turns out that the system has a unique non-trivial solution up to a multiple with a constant. This result is due to Dagerholm and a proof is exposed in a section from *Special Topics* from my notes devoted to analytic function theory. It goes without saying that the proof is considerably more demanding compared to situations in normed vector spaces where existence and uniqueness properties are controlled via imposed norms.

f. Maxwell's equations in electrodynamics

.

Consider the following triple of 3×3-matrices

$$A_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix} \quad : A_2 = \begin{pmatrix} 0 & 0 & i \\ 0 & 0 & 0 \\ -i & 0 & 0 \end{pmatrix} \quad : A_3 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

One seeks a vector valued function $f = (f_1(x_1, x_2, x_3, t), f_2(x_1, x_2, x_3, t), (f_3(x_1, x_2, x_3, t)$ which satisfies the differential system

$$(*) \qquad \frac{\partial}{\partial t} \begin{pmatrix} f_1 \\ f_2 \\ f_3 \end{pmatrix} = \sum_{j=1}^{j=3} A_j \circ \frac{\partial}{\partial x_j} \begin{pmatrix} f_1 \\ f_2 \\ f_3 \end{pmatrix}$$

and the initial conditions $\{f_\nu(x, 0) = g_\nu(x)\}$ where g_1, g_2, g_3 are given real-valued C^∞-functions in $\mathbf{R}^3$. This turns out to be a well-posed Cauchy problem, i.e. for every triple of g-functions there exists a unique solution f to (*) with the boundary conditions as above.

A veritable propagande for operator theory appears during the proof of this result which actuslly extends to symmetric hyperbolic systems. They arise as follows. Let $x = (x_1, \ldots, x_n)$ be space variables for some positive integer n. Next, let m be another positive integer and

$$(x,t) \mapsto A_\nu(x,t) \quad : 1 \leq \nu \leq n$$

is an n-tuple of maps from $\mathbf{R}^n \times \{t \geq 0\}$ into the space $m \times m$-matrices where the elements of these A-matrices are complex-valued C^∞-functions defined on $\mathbf{R}^n \times \{t \geq 0\}$. In addition we consider another matrix-valued function $B(x,t)$ whose m^2-many elements again are C^∞-functions. Now one seeks vector-valued functions $f = (f_1(x,t), \ldots, f_m(x,t)$ which satisfy the differential system

$$(**) \qquad \frac{\partial}{\partial t}\begin{pmatrix} f_1 \\ \ldots \\ f_m \end{pmatrix} = \sum_{j=1}^{j=n} A_j(x,t) \cdot \frac{\partial}{\partial x_j}\begin{pmatrix} f_1 \\ \ldots \\ f_m \end{pmatrix} + B(x,t)\begin{pmatrix} f_1 \\ \ldots \\ f_m \end{pmatrix}$$

and the boundary value conditions

$$f_\nu(x,0) = g_\nu(x) \quad : 1 \leq \nu \leq m$$

where $\{g_\nu(x)\}$ is an m-tuple in $C^\infty(\mathbf{R}^n)$.

The symmetric case. If the $n \times n$-matrices $\{A_1(x,s), \ldots, A_m(x,t)\}$ are Hermitian for each fixed pair (x,t) one says that the system (**) is symmetric. No special condition is imposed on B, except that its elements are complex-valued C^∞-functions. Notice that the A-matrices in Maxwell's equations are Hermitian so (*) is an example of a symmetric system. The existence and uniqueness of solution to Maxwell's equations is a special case of the result below proved by Friedrichs in the article *Symmetric hyperbolic differential systems* [Comm. pure.appl.math. Vol 7. 1954].

0.1 Theorem. *Assume that the system (**) is symmetric and that the matrix elements of $A_1, \ldots, A_n$ and of B are bounded functions in $\mathbf{R}^{n+1}$. Then Cauchy's boundary value problem has a unique C^∞ solution $f = (f_1, \ldots, f_m)$ for every m tuple $g = (g_1, \ldots, g_m)$ of C^∞-functions in the n-dimensional x-space.*

Remark. The proof is given in § xx. It involves a number of steps where unbounded operators on suitable Hilbert spaces which emerge from Sobolev spaces appear. So the proof relies heavily upon operator theory and gives therefore a good motivation for learning material to be exposed later on.

g. Positive definite quadratic forms

Inequalities play a major role in functional analysis. They are often established via geometric considerations. Let us illustrate this by some remarkable results presented by Beurling during seminars at Uppsala university in 1935. See *Collected work* [vol. xx p. xx-xx]. The theorem below can for example be used to establish the existence of certain spectral measures which arise in ergodic processes. Let f and g be a pair of square integrable functions on the real line. For each real number t we set

$$\psi(t) = \int_{-\infty}^{\infty} [f(t+s) - g(s)|^2 \cdot ds$$

g.1 Theorem. *There exists a measure μ on the real line whose total variation $\leq 2\sqrt{||f||_2 \cdot ||g||_2}$ such that*

$$\psi(t) = ||f||_2 + ||g||_2 + \int_{-\infty}^{\infty} e^{ixt} \cdot d\mu(x)$$

The proof of this theorem relies upon a study of certain bi-linear forms in euclidian spaces which we describe in detail to illiustrate the usefulness of geometric considerations while objects such as quadratic forms are studied. Let $m \geq 2$ and denote m-vectors in $\mathbf{R}^m$ with capital letters, i.e. $X = (x_1, \dots, x_m)$. Let $N \geq 2$ be some positive integer and $X_1, \dots, X_N$ an N-tuple of real m-vectors. To each pair $j \neq k$ we set

$$b_{ij} = ||X_j|| + X_k|| - ||X_j - X_k||$$

where $||\cdot||$ is the usual euclidian length in $\mathbf{R}^m$ and consider the quadratic form

$$H(\xi_1, \dots, \xi_N) = \sum\sum b_{ij} \cdot \xi_i \cdot \xi_j$$

g.2. Theorem. *If the X-vectors are all different then H is positive definite.*

The proof relies upon an integral formula which determines the length of a vector in $\mathbf{R}^m$.

g.3. Lemma *There exists a constant C_m such that*

$$(*) \qquad ||X|| = C_m \cdot \int_{\mathbf{R}^m} \frac{1 - \cos\langle X, Y\rangle}{||Y||^{m+1}} \cdot dY$$

hold for every m-vector X.

Proof. We use polar coordinates. Denote by dA the area measure on the unit sphere S^{m-1} and $\omega = (\omega_1, \dots, \omega_m)$ are points on the unit sphere S^{m-1}. Notice that the integrals

$$\int_{S^{m-1}} (1 - \cos\langle X, \omega\rangle) \cdot dA$$

only depend upon $||X||$. Hence it suffices to prove Lemma 2.xx when $X = (R, \dots, 0)$ where $R = ||X||$. It is clear that

$$\text{(i)} \qquad \int_{\mathbf{R}^m} \frac{1 - \cos\langle X, Y\rangle}{||Y||^{m+1}} \cdot dY = \int_0^\infty \big[\int_{S^{m-2}} (1 - \cos Rr\omega_1) \cdot dA_{m-1} \big] \cdot \frac{dr}{r^2}$$

where dA_{m-1} is the area measure on S^{m-2}. Set

$$B(R, \omega_1) = \int_0^\infty (1 - \cos Rr\omega_1) \cdot \frac{dr}{r^2}$$

for each $-1 < \omega_1 < 1$. The variable substitution $r \to s/R$ gives

$$B(R, \omega_1) = R \cdot \int_0^\infty \frac{1 - \cos s\omega_1}{s^2} \cdot ds = R \cdot B_*(\omega_1)$$

With these notations the right hand side in (i) is equal to

$$(1) \qquad R \cdot \int_{S^{m-2}} B_*(\omega_1) \cdot dA_{m-2}$$

Above $||X|| = R$ and Lemma 2.xx follows with

$$C_m = \big(\int_0^\infty \frac{1 - \cos s\omega_1}{s^2} \cdot ds \big)^{-1}$$

Proof of Theorem g.2. For a given pair $1 \leq i, j \leq N$ the addition formula for the cosine-function gives:

$$1 - \cos\langle X_i, Y\rangle + 1 - \cos\langle X_j, Y\rangle + \cos\langle (X_i - X_j), Y\rangle =$$

$$(1) \qquad (1 - \cos\langle X_i, Y\rangle) \cdot (1 - \cos\langle X_i, Y\rangle) + \sin\langle X_i, Y\rangle \cdot \sin\langle X_j, Y\rangle$$

It follows that the matrix element b_{ij} is given by

$$(3) \qquad C_m \cdot \int_{\mathbf{R}^m} \frac{(1 - \cos\langle X_i, Y\rangle) \cdot (1 - \cos\langle X_j, Y\rangle) + \sin\langle X_i, Y\rangle \cdot \sin\langle X_j, Y\rangle}{||Y||^{m+1}} \cdot dY$$

From this we see that

$$(3) \qquad H(\xi) = C_m \cdot \int_{\mathbf{R}^m} \big([\sum (\xi_k \cdot (1 - \cos\langle X_k, Y\rangle)]^2 + [\sum (\xi_k \cdot (\sin\langle X_k, Y\rangle)]^2 \big) \cdot \frac{dY}{||Y||^{m+1}}$$

Since the vectors $\{X_k\}$ are different it is clear that

$$\big([\sum(\xi_k\cdot(1-\cos\langle X_k,Y\rangle)\,]^2+[\sum(\xi_k\cdot(\sin\langle X_k,Y\rangle)\,]^2\big)\neq 0 \tag{4}$$

for all non-zero m-vectors Y. of norm ≥ 1. By continuity this gives for example a positive coinstant ρ such that (4) is $\geq \rho^*\cdot||\xi||^2$ for all vectors Y with norm between 2^{-1} and 2 and from this the reader can conclude that H is positive definite.

g.4. Exercise. Prove more generally that for every $1<p<2$ that the quadratic form

$$H_p(\xi)=\big(||X_j||^p+X_k||^p-||X_j-X_k||^p\big)\cdot\xi_j\cdot\xi_k$$

again is positive when the X-vectors are all different. Show also the following: Let $z_1,\ldots,z_N$ be an N-tuple of distinct and non-zero zero complex numbers and put

$$b_{ij}=\{\frac{z_i}{z_j}\}$$

Show that the Hermitian matrix $B=\{b_{ij}\}$ is positive definite.

Self-adjoint operators on Hilbert spaces.

The spectral theory for unbounded self-adjoint operators was developed by Carleman in the monograph *Sur les équations singulières à noyau reèl et symmétrique* [Uppsala University. 1923]. An account of this work appears in Carleman's lecture from [Comptes rendus du VI^e Congrés des Mathématiques Scandinavies. Copenhagen 1925]. Let us remark that [ibid] was preceeded by his earlier studies of singular integral operators. For example, kernels associated to the Neumann's boundary value problem whose boundary has corner points, produce operators whose spectra contain closed sets with interor points, a situation which in 1916 was quite new.

Hundreds of text-books and thousands of articles treat spectral theory for unbounded self-adjoint operators. But the basic material is covered in Carleman's original work and to this I would like to add a personal comment. Even though "abstract methods" in mathematics often are useful, one only becomes truly familiar with the "source" of a subject after executing explicit computations. From this point of view [ibid] appears as an outstanding text about unbounded linear operators on Hilbert spaces which in addition to theoretic results contain several concrete applications exposed with detailed and very elegant proofs. At the same time specific situations illustrate how to perform the pasage to limits which are taken in various weak topologies. Let us brielfy describe the major results from [ibid] which will be exposed in more detail in § xx.

1. Densely defined hermitian operators. Let A be a linear operator on a separable Hilbert space $\mathcal{H}$ whose domain of definition $\mathcal{D}(A)$ is dense and satisfies the hermitian equation:

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

for each pair x, y in $\mathcal{D}(A)$. A first result in [ibid] asserts that the inhomogenous equation

$$(*) \qquad x = \lambda \cdot A(x) + \xi$$

has at least one solution $x \in \mathcal{D}(A)$ for every vector $\xi \in \mathcal{H}$ and non-real complex number λ. To prove this Carleman used weak approximations of A by hermitian Hilbert-Schmidt operators. Using the material from § xx it follows that every solution in (*) satisfies

$$(**) \qquad ||x|| \leq |\mathfrak{Im}(\lambda)|^{-1} \cdot ||\xi||$$

If (*) has a unique solution for some λ_* whose imaginary part is > 0, then it is proved in [ibid] that (*) has unique solutions for all λ in the upper half-plane which gives an operator-valued function $\lambda \mapsto T(\lambda)$ where

$$T(\lambda)(\xi) = x(\lambda)$$

when $x(\lambda)$ is the unique solution to (*) for the pair ξ and λ. In the same way one can consider (*) when λ varies in the lower half-plane. If uniqueness again holds, then the operator-valued function above is defined in the complement of the real axis. We shall learn that the uniquness above hold for non-real λ if and only if A is self-adjoint which means that A is equal to its adjoint A^*.

2. A criterion for existence of self-adjoint extensions. Suppose that the solution space in (*) has a common finite dimension N^* for each non-real λ_*. Then there exist a distinguished family of self-adjoint operators whose graphs contain that of A. More precisely, this family is parametrized by points on an N^*-dimensional torus. Details appear in § xx.

3. The case of integral operators. Here one starts with a real-valued function $K(x, y)$ defined in a product $\Omega \times \Omega$ where Ω is an open set in $\mathbf{R}^n$ for some $n \geq 1$. Assume the symmetry $K(x, y) = K(y, x)$ and that there exists a nullset $\mathcal{N}$ in Ω such that

$$\int_\Omega K(x, y)^2 \, dy < \infty \quad : x \in \Omega \setminus \mathcal{N}$$

Let $f(x)$ is a continuous and square integrable function in Ω. Now there exist the non-empty family of L^2-functions ϕ which satisfy

$$(3.1) \qquad \phi(x) = \lambda \cdot \int_\Omega K(x, y) \cdot \phi(y) \, dy + f(x)$$

To begin with this equality hold for almost every $x \in \Omega$, i.e. outside a null-set which in general contains $\mathcal{N}$. Suppose that E is a compact subset of $\Omega \setminus \mathcal{N}$ such that

$$\max_{|x_1|-x_2|\leq\delta} \int_\Omega (K(x_1,y)-K(x_2,y))^2\,dy$$

tends to zero as $\delta \to 0$ while x_1, x_2 stay in E. Then it is easily seen that each ϕ-function is continuous on E and one is led to introduce the sets

$$\mathcal{D}(x:\lambda) = \bigcup \phi(x)$$

with the union taken over all solutions to (3.1). Properties of these sets are studied in [ibid] and in the section devoted to the monent problem we shall give some precise results.

4. The spectral function of a self-adjoint operator.

A major result in Carleman's cited monograph asserts the following: Let A be a densely defined self-adjoint operator on a separable Hilbert space $\mathcal{H}$. Then there exists an operator valued function

$$\lambda \mapsto \Theta(\lambda) \tag{*}$$

of finite total variation with respect to the operator norms on the Banach space of bounded linear operators on $\mathcal{H}$ with the following properties. First, a vector $x \in \mathcal{H}$ belongs to $\mathcal{D}(A)$ if and only if

$$\lim_{\delta\to 0}\Big[\int_{-1/\delta}^{-\delta} |\lambda|^{-1}\cdot ||\frac{d}{d\lambda}(\Theta(\lambda)(x))|| + \int_\delta^{1/\delta} |\lambda|^{-1}\cdot ||\frac{d}{d\lambda}(\Theta(\lambda)(x))||\Big] < \infty \tag{**}$$

and when this holds

$$A(x) = \lim_{\delta\to 0}\Big[\int_{-1/\delta}^{-\delta} \lambda^{-1}\cdot \frac{d}{d\lambda}(\Theta(\lambda)(x)) + \int_\delta^{1/\delta} \lambda^{-1}\cdot \frac{d}{d\lambda}(\Theta(\lambda)(x))\Big] \tag{***}$$

Above the integrals are taken in the sense of Stieltjes. For example, the finiteness of the last integral in (**) means that if $\delta = \lambda_0 < \ldots < \lambda_1 < \lambda_N = 1/\delta$ is a partition of the interval $[\delta, 1/\delta]$, then

$$\sum_{\nu=1}^{\nu=N-1} \lambda_\nu^{-1}\cdot ||\Theta(\lambda_\nu)(x)) - ||\Theta(\lambda_{\nu-1})(x))|| \leq C$$

for a constant which is independent of δ and the chosen partitions of the corresponding interval. One refers to (*) as the spectral operator-valued function attached to A. In addition to the above the operators which arise via (*) have the property that whenver $[a,b]$ and $[c,d]$ are pairwise disjoint intervals on the real λ-line then the composed opertor

$$(\Theta(b)-\Theta(a))\circ(\Theta(d)-\Theta(c)) = 0$$

and $E_{a,b} = (\Theta(b)-\Theta(a))$ is a self-adjoint projection for every interval $[a,b]$ which are zero when the interval $[a,b]$ is disjoint from $\sigma(A)$. Finally, for every vector $x \in \mathcal{H}$ one has the Bessel inequality

$$\int_{-\infty}^{\infty} ||\frac{d}{d\lambda}(\Theta(\lambda)(x)|| \leq ||x||$$

4.1 The construction of Θ. It will be given in § xx and here we ony remark that the spectral resolution of A is found via a passage to the limit where one approximates A by bounded self-adjoint operators. By this procedure one gets insigh about the Θ-function in (*).

4.2 The closure property. When A is a self-adjoint operator one can impose the condition that its kernel is zero, i.e. $A\colon \mathcal{D}(A) \to \mathcal{H}$ is injective. In [Carleman: Chapire 4] it is proved that when this holds, then every every vector $x \in \mathcal{H}$ satisfies the equation:

$$x = \lim_{\delta\to 0}\Big[\int_{-1/\delta}^{-\delta} \frac{d}{d\lambda}(\Theta(x)) + \int_\delta^{1/\delta} \frac{d}{d\lambda}(\Theta(x))\Big] \tag{****}$$

Following [ibid. page 136] one says that the self-adjoint operator A has the closure property when (****) holds. In the case where A is presented as an integral operator a more refined sufficient condition for A to have the closure property is established in [Carleman: page 138-142] which is expressed by the existence of certain weak approximations of A.

5. Unitary groups.

Starting from (*) above one constructs a unitary group $\{U_t\}$ indexed by real numbers where

$$U_t(x) = i \cdot \int_{-\infty}^{\infty} e^{i\frac{t}{\lambda}} \cdot \frac{d}{d\lambda}(\Theta(\lambda)(x)) \quad : \ x \in \mathcal{H}$$

Taylor's limit formula gives

$$i \cdot \lim_{t \to 0} t^{-1}(e^{i\frac{t}{\lambda}} - 1) = \lambda^{-1} \quad : \ \lambda \neq 0$$

Then (***) entails that A is the infinitesmal generator of $\{U_t\}$ to be defined in § xx. The converse was proved by Stone and exposed in the section devoted to the Hille-Phillips-Yosida theorem. Together, the results by Carleman and Stone give a 1-1 corespondence between the family of strongly continuous unitary groups and the family of self-adjoint operators.

5.1 Example. In $\mathbf{R}^3$ the Laplace operator Δ yields a densely defined linear operstor on the Hilbert space $\mathcal{H} = L^2(\mathbf{R}^3)$. Introducing the Fourier transform, Parseval's equality gives the unitary group $\{U_t\}$ where

$$U_t(f)(x) = i \cdot (2\pi)^{-3} \cdot \int_{\mathbf{R}^3} e^{i\langle x,\xi\rangle} \cdot e^{\frac{it}{|\xi|^2}} \cdot \widehat{f}(\xi)\, d\xi$$

When f is such that $\widehat{f}(\xi) \in C_0^\infty(\mathbf{R}^3)$ we have

$$\lim_{t \to 0} \frac{U_t(f) - f}{t} = -(2\pi)^{-3} \cdot \int_{\mathbf{R}^3} |\xi|^{-2} \cdot e^{i\langle x,\xi\rangle} \cdot \widehat{f}(\xi)\, d\xi = \Delta(f)(x)$$

It follows that the infinitesmal generator of $\{U_t\}$ is the self-adjoint operator A where $\mathcal{D}(A)$ consists of L^2-functions f for which $\Delta(f)$ taken in the sense of distribution theory is square integrable. To find $\Theta(\lambda)$ one employs Fourier's inversion formula to Gaussian L^2-functions of the form $e^{-a|x|^2}$ where a are real and positive numbers. The spectrum of A is $(-\infty, 0]$. Indeed, this follows since Fourier's inversion formula gives

$$R_A(\lambda)(f)(x) = -(2\pi)^{-3} \cdot \int_{\mathbf{R}^3} \frac{e^{i\langle x,\xi\rangle}}{\lambda + |\xi|^2} \cdot \widehat{f}(\xi)\, d\xi$$

Parseval's equality entails that this gives a bounded linear operator when λ is outside $\mathbf{C} \setminus (-\infty, 0]$.

6. The case of non-separable Hilbert spaces.

Carleman's studies were restricted to separable Hilbert spaces which means that densely defined and symmetric operators are represented by infinite hermitian matrices indexed by pairs of non-negative integers. Using the Cayley transform and Hilbert's spectral theorem for bounded normal operators one can establish the spectral theorem for densely defined self-adjoint operators on non-separable Hilbert spaces. This was carried out in J. von Neumann's article *Allgemeine Eigenwerttheorie Hermitscher Funktionaloperatoren* [Math. Annalen, vol. 102 (1929)]. See the next section for a brief account and § 9 in Special Chapters for details. During the search of self-adjoint extensions of densely defined symmetric operators we remark that the notion of hypermaximality introduced by Schmidt is useful. It stems from Schmidt's early contributions in the article *Auflösing der allgemeinen linearen Integralgleichung* [Math. Annalen, vol. 64 (1907].

Non-separable Hilbert spaces arise in many situations. An example is the Hilbert space of square integrable functions on the compact Bohr group which by definition is the dual of the discrete abelian group of real numbers. The theory of almost periodic functions was created by Harald Bohr (brother to the phycisist Niels Bohr) and presented in the articles *Zur Theorie der fastperiodischen Funktionen I-III* [Acta. Math. vol 45-47 (1925-26]. This led in a natural way to consider of operators on non-separable Hilbert spaces and motivated a more "abstract account" where unbounded operators are not described via infinite matrices indexed by pairs of integers. Other examples emerge from quantum mechanics. The uncertainty principle led to questions concerned with actions by pairs of non-commutating densely defined operators on a Hilbert space. In 1928 this was reformulated by Weyl to specific problems about non-commuting families of unitary groups. See § for further details where we present von Neumann's proof of a result which settled original problems posed by de Broigle, Heisenberg and Schrödinger.

6.1 The Cayley transform.

Recall a classic result about matrices which goes back to work by Cayley and Hamilton. If $N \geq 2$ is an integer we have the family $\mathfrak{h}(N)$ of Hermitian $N \times N$-matrices. Next, a matrix R is normal if it commutes with the adjoint matrix R^*. The Cayley transform sends a Hermitian matrix A to the normal matrix

$$R_A = (iE_N - A)^{-1}$$

where E_N is the identity matrix in $M_N(\mathbf{C})$. If $\{\alpha_\nu\}$ is the real spectrum of A it is readily seen that

$$\sigma(R_A) = \{\frac{1}{i - \alpha_\nu}\}$$

Next, when a is a real number we notice that

$$\frac{1}{i-a} + \frac{i}{2} = \frac{1}{2(i-a)}(2 + i^2 - ia) = \frac{1 - ai}{2i(1+ai)}$$

The last quotient is a complex number with absolute value 1/2. Hence $\sigma(R_A)$ is contained in the circle

$$\mathcal{C} = \{|\lambda + \frac{i}{2}| = \frac{1}{2}\}$$

Conversely, let R be a normal operator such that $\sigma(R)$ is contained in $\mathcal{C}$ and does not contain the origin. Now there exists the inverse operator R^{-1} and

(i) $$A = iE_N - R^{-1}$$

is a normal operator whose spectrum $\sigma(A)$ is real, and since A also is normal it follows that it is Hermitian. Notice also that $R = R_A$ in (i). *Summing up*, the Cayley transform gives a bijective map between Hermitian matrices and the family of invertible normal operators whose spectra are contained in $\mathcal{C}$.

The passage to Hilbert spaces. Consider a complex Hilbert space $\mathcal{H}$ and denote by $\mathcal{N}$ the family of bounded normal operators R with the property that $\sigma(R)$ is contained in the circle $\mathcal{C}$, and in addition R is injective and has a dense range. To each such R we find a densely defined operator S_R on $\mathcal{H}$ which for each $y = R(x)$ in the dense range assigns the vector x, i.e. the composed operator $S_R \circ R = E$, where E is the identity on $\mathcal{H}$. Now there also exists the densely defined operator

(*) $$A = iE - S_R$$

We shall learn that the densely operator A is self-adjoint, and conversely every densely defined and self-adjoint operator A is of the form (*) for a unique R in $\mathcal{C}$. So this gives 1-1 correspondence which extends the previous result for matrices. Let us remark that bounded self-adjoint operators appear in (*) when the normal operator R is invertible.

7. An ugly example.

Following [Carleman - page 62-66] we give examples which show that the existence of self-adjoint extensions of densely defined symmetric operator is not automatic. Every separable Hilbert space can be identified with $L^2[0,1]$ whose vectors are complex-valued square integrable functions. Let us then consider a real-valued Lebegue measurable function $K(x,y)$ defined on the square

$\{0 \leq x \leq 1\} \times \{0 \leq y \leq 1\}$ which is symmetric, i.e. $K(y,x) = K(x,y)$. Assume that there exists a null set $\mathcal{N}$ such that

$$(1) \qquad \int_0^1 K(x,y)^2 \, dy < \infty \quad : x \in [0,1] \setminus \mathcal{N}$$

If $f \in L^2[0,1]$ the Cauchy-Schwarz inequality entails that $y \mapsto K(x,y) \cdot f(y)$ is absolutely integrable when x is outside $\mathcal{N}$. Hence there exists the almost everywhere defined function

$$(2) \qquad \mathcal{K}(f)(x) = \int_0^1 K(x,y) \cdot f(y) \, dy$$

However, (1) does not imply that $\mathcal{K}(f)$ is square integrable. So we consider the subspace $\mathcal{D}$ of $L^2[0.1]$ where $f \in \mathcal{D}$ gives $\mathcal{K}(f)$ in $L^2[0.1]$. To analyze this subspace we consider positive integers N and denote by E_N the subset of the unit x-interval such that

$$(3) \qquad \int_0^1 K(x,y)^2 \, dy \leq N \quad : x \in E_N$$

If $f \in L^2[0,1]$ is supported by E_N and g is another L^2-function, the symmetry of K and Fubini's theorem give

$$\int_0^1 g(x)\mathcal{K}(f)(x) \, dx = \int_0^1 g(x) \cdot (\int K(x,y) f(y) \, dy) \, dx = \int_0^1 f(x) \cdot (\int K(x,y) g(y) \, dy) \, dx$$

The Cauchy-Schwartz inequality and (3) entail that the absolute value of the last term is majorised by

$$\sqrt{N} \cdot ||g||_2 \cdot \int_{E_N} |f(x)| \, dx \leq \sqrt{N} \cdot ||g||_2 \cdot ||f||_2$$

Since this hold for every L^2-function g, a wellknown fact from integration theory implies that $\mathcal{K}(f)$ is square integrable and its L^2-norm is majorised by $\sqrt{N} \cdot ||f||_2$. Hence $\mathcal{D}$ contains every L^2-function supported by E_N. Since $\cup E_N = [0,1] \setminus \mathcal{N}$ it follows that $\mathcal{D}$ is a dense subspace. The question arises if there exists some complex-valued function $\phi \in \mathcal{D}$ which satisfies the eigenvalue equation

$$(*) \qquad \phi = i \cdot \mathcal{K}(\phi)$$

When $\mathcal{K}$ is a bounded linear operator on the Hilbert space $L^2[0,1]$ such non-zero solutions do not exist. In fact, this follows from Hilbert's theorem for bounded symmetric operators. So if a non-trivial solution to (*) exists, then $\mathcal{K}$ must be unbounded. To exhibit examples of such "ugly operators", Carleman introduced the an orthonormal basis for $L^2[0,1]$ given by a sequence $\{\psi_n\}$ where $\psi_0(x) = 1$ and $\psi_1(x) = -1$ on $(0,1/2)$ and $+1$ on $(1/2,1)$. Finally, for each $n \geq 2$ we set

$$\psi_n(x) = -2^{\frac{n-1}{[}2} : \ 1 - 2^{-n+1} \leq x < 1 - 2^{-n} \quad \text{and} \quad \psi_n(x) = 2^{\frac{n-1}{2}} : \ 1 - 2^{-n} < x < 1$$

while

$$\psi_n(x) = 0 \quad : \ 0 < x < 1 - 2^{-n+1}$$

It is easily seen that this gives an orthonormal basis. Next, for every sequece $\{a_p\}$ of real numbers we define the kernel function on $[0,1) \times [0,1)$ by

$$(\text{i}) \qquad K(x,y) = \sum a_p \cdot \psi_p(x)\psi_q(x)$$

To this symetric function we associate the operator

$$\mathcal{K}(u)(x) = \int_0^1 K(x,y)u(y) \, dy$$

The construction of the ψ-functions show that $\mathcal{D}$ contains L^2-functions u supported by $0 \leq x \leq x_*$ for every $x_* < 1$. In § xx we shall prove the following:

Theorem. *The equation (*) has a non-trivial L^2-solution if and only if*

$$\sum_{p=0}^{\infty} \frac{2^p}{1+a_p^2} < \infty$$

8. Two examples from PDE-theory.

The study of unbounded self-adjoint operators is especially relevant while one regards PDE-operators. We present two examples below.

1. Propagation of sound. With (x, y, z) as space variables in $\mathbf{R}^3$ and a time variable t, the propagation of sound in the infinite open complement $U = \mathbf{R}^3 \setminus \overline{\Omega}$ of a bounded open subset Ω is governed by solutions $u(x, y, z, t)$ to the wave equation

$$\frac{\partial^2 u}{\partial t^2} = \Delta u \tag{1}$$

where Δ is the Laplace operator in x, y, z. So here (1) holds when $p = (x, y, z) \in U$ and $t \geq 0$. We assume that $\partial\Omega$ is of class C^1, i.e. given as a union of pairwise disjoint closed surfaces of class C^1 along which normal vectors are defined. A well-posed Cauchy problem arises when we seek solutions such that $p \mapsto u(p, t)$ belong to $L^2(U)$ for each t, and the outer normal derivatives taken along $\partial\Omega$ are zero, i.e. for every t

$$\frac{\partial u}{\partial n}(p, t) = 0 \quad : p \in \partial\Omega \tag{2}$$

Initial conditions are expressed by a pair of C^2-functions $f_1 x, y, z)$ and $f_2(x, y, z)$ defined in U such that f_1, f_2 together with $\Delta(f_1$ and $\Delta(f_2)$ belong to $L^2(U)$, and their outer normal derivatives along $\partial\Omega$ are zero. Then there exists a unique solution u which satisfies (1-2) and the initial conditions $u(p, 0) = f_0(p)$ and $\frac{\partial u}{\partial t}(p, 0) = f_1(p)$. A proof of existence and uniqueness relies upon the construction of a time-independent self-adjoint operator A acting on $L^2(U)$ whose kernel is expressed via a Green's function which is found by Neumann's standard elliptic boundary value problem in the domain U. After this has been done one gets an integral representation of the solution u expressed via the pair f_0, f_1 and the spectral function associated to A. For physical reaons one expects that if u is a solution, then the first order partial derivatives of $u(p, t)$ with respect to the space variables tend to zero as $t \to +\infty$. In [Carleman 1923] this is proved by analyzing the spectral function associated with A where the crucial point is that its associated spectral function is absolutely continuous with respect to the λ-parameter. In §§ we show that this absolute continuity is derived from a general result which goes as follows:

Let $\{a \leq s \leq b\}$ be a compact interval on the real s-line and $s \mapsto G_s$ is a function with values in the Hilbert space $L^2(U)$ which is continuous in the sense that

$$\lim_{s \to s_0} ||G_s - G_{s_0}||_2 = 0$$

hold for each s_0, where we introduced the L^2-norms. The function has a finite total variation if there exists a constant M such that

$$\sum ||G_{s_{\nu+1}} - G_{s_\nu}||_2 \leq M$$

hold for every partition $a = s_0 < s_1 < \ldots < s_M = b$. When this holds one construct Stieltjes integrals and for every subinterval $[\alpha, \beta]$ there exists the L^2-function in U

$$\Phi_{[\alpha,\beta]} = \int_\alpha^\beta s \cdot \frac{dG_s}{ds}$$

We impose the extra conditions that the normal deriviatives $\frac{\partial G_s}{\partial n}$ exist and vanish $\partial\Omega$ for every $a \leq s \leq b$ and the following differential equation holds for every sub-interval $[\alpha, \beta]$ of $[a, b]$:

$$\Delta(G_\beta - G_\alpha) + \Phi_{[\alpha,\beta]} = 0$$

Theorem. *The equations above imply that $s \mapsto G_s$ is absolutely continuous.*

Remark. The absolute continuity of $s \mapsto G_s$ means that whenever $\{\ell_1, \ldots, \ell_M\}$ a finite family of disjoint intervals in $[a, b]$ where the sum of their lengths is $< \delta$, then the sum of the total variations over these intervals is bounded by $\rho(\delta)$ where ρ is a function of δ which tends to zero as $\delta \to 0$. We refer to § xx for an account of the proof of this result.

2. The Bohr-Schrödinger equation.

In 1923 quantum mechanics had not yet appeared so the studies in [Carleman: 1923] were concerned with singular integral equations, foremost inspired from previous work by Fredholm, Hilbert, Weyl and Volterra. The creation of quantum mechanics gave new challenges for the mathematical community,. The interested reader should consult the lecture held by Niels Bohr at the Scandianavian congress in mathematics in Copenhagen 1925 where he speaks about the interplay between the new physics and mathematics. Bohr's lecture presumably inspired Carleman when he some years later resumed work from [Car 1923]. Recall that the fundamental point in Schrödinger's theory is the hypothesis on energy levels which correspond to orbits in Bohr's theory of atoms. For an account about the physical background the reader may consult Bohr's plenary talk when he received the Nobel Prize in physics 1923. Mathematically the Bohr-Schrödinger theory leads to the equation

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

Here Δ is the Laplace operator in the 3-dimensional (x, y, z)-space, m the mass of a particle and h Planck's constant while $U(x, y, z)$ is a potential function. Finally E is a parameter and one seeks values on E such that (*) has a solution ϕ which belongs to $L^2(\mathbf{R}^3)$. Let us cite an excerpt from Carlemans lectures in Paris at Institut Henri Poincaré held in 1930:

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

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

The theory from [Carleman:1923] applies to the following PDE-equations attached to a second order differential operator

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

where $c(x, y, z)$ is a real-valued function. The L-operator is densely defined and symmetric on the subspace of test-functions in $\mathbf{R}^3$. The problem is to find conditions on the c-function in order that L is self-adjoint on the Hilbert space $L^2(\mathbf{R}^3)$ with a real spectrum confined to $[a, +\infty)$ for some real number a. The following sufficiency result was presented by Carleman's during his lectures at Sorbonne in the spring 1930:

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

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

Then $\Delta + c(x, y, z)$ is self-adjoint.

Example. The result applies when c is given by a potential function:

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

where $\{q_k\}$ is a finite subset of $\mathbf{R}^3$ and the α-numbers and β are real and positive. Here the requested self-adjointness is easy to prove and solutions to the Bohr-Schrödinger equation are found via robust limit formulas. See §§ xx for details.

Further comments. The literature about the Schrödinger equation and other equations which emerge from quantum mechanics is extensive. For the source of quantum mechanics the reader should first of all consult the plenary talks by Heisenberg, Dirac and Shrödinger when they received the Nobel prize in physics. Apart from physical considerations the reader will find expositions where explanations are given in a mathematical framework. Actually Heisenberg was sole winner 1931 while Dirac and Schrödinger shared the prize in 1932. But they visited Stockholm together in December 1932.

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

Schrödinger's non-linear Cauchy problem

The mathematical foundations in non-relativistic quantum mechanics rely upon studies of wave-equations. This was put forward in Schrödinger's article *Quantizierung als Eigenwertproblem* from 1926. In the article *Théorie relativiste de l'electron et l'interprétation de la mécanique quantique* [xxx 1932], Schrödinger raised a new and unorthodox question concerned with Brownian motions leading to new mathematical problems of considerable interest which up to the present date remain open. Consider a Brownian motion which takes place in a bounded region Ω of some euclidian space $\mathbf{R}^d$ for some $d \geq 2$. At time $t = 0$ the densities of particles under observation is given by some non-negative function $f_0(x)$ defined on Ω. Classically the density at a later time $t > 0$ is equal to a function $x \mapsto u(x,t)$ which satisfies the heat equation

$$\frac{\partial u}{\partial t} = \Delta(u)$$

with boundary conditions

$$(1) \qquad u(x,0) = f_0(x) \quad \text{and} \quad \frac{\partial u}{\partial \mathbf{n}}(x,t) = 0 \quad \text{when} \quad x \in \partial\Omega \quad \text{and} \quad t > 0$$

Schrödinger took into the account the reality of quantum physics which means that in an actual experiment the observed density of particles at a time $t_1 > 0$ does not coincide with $u(x,t_1)$. He posed the problem to find the most probable development during the time interval $[0, t_1)$ which leads to the state at time t_1, and concluded that the requested density function which substitutes the heat-solution $u(x,t)$ belongs to a non-linear class of functions formed by products

$$(*) \qquad w(x,t) = u_0(x,t) \cdot u_1(x,t)$$

where u_0 is a solution to (1) while $u_1(x,t)$ is a solution to an adjoint equation

$$(2) \qquad \frac{\partial u_1}{\partial t} = -\Delta(u) \quad : \quad \frac{\partial u_1}{\partial \mathbf{n}}(x,t) = 0 \quad \text{on} \quad \partial\Omega$$

defined when $t < t_1$. This leads to a new type of Cauchy problems where one asks if there exists a w-function in (*) satisfying

$$w(x,0) = f_0(x) \quad : \quad w(x,t_1) = f_1(x)$$

whenever f_0, f_1 are non-negative functions such that

$$\int_\Omega f_0 \cdot dx = \int_\Omega f_1 \cdot dx$$

The solvability of this non-linear boundary value problem was left open by Schrödinger. When Ω is a bounded set with a smooth boundary one can use the Poisson-Greens function for the classical equation (*) and rewrite Schrödinger's equation to a system of non-linear integral equations. The reader should consult the talk by I.N. Bernstein a the IMU-congress at Zürich 1932 for a first account about eventual solutions to the Schrödinger equation above. Non-trivial examples occur already on the product of two copies of the real line where solutions to Schrödinger's equation is equivalent to find solutions of non-linear equationa expressed via measures. More precisely, consider the Gaussian density function

$$g(x)=\frac{1}{\sqrt{2\pi}}\cdot e^{-x^2/2}$$

on the real x-line. Denote by the family $\mathcal{S}_g^*$ of all non-negative product measures $\gamma_1\times\gamma_2$ for which

$$\text{(i)}\qquad \iint g(x_1-x_2)\cdot d\gamma_1(x_1)\cdot d\gamma_2(x_2)=1$$

Next, $\gamma_1\times\gamma_2$ gives a new product measure

$$\mathcal{T}_g(\gamma_1\times\gamma_2)=\mu_1\times\mu_2$$

such that

$$\mu_1(E_1)\cdot\mu_2(E_2)=\iint_{E_1\times E_2} g(x_1-x_2)\cdot d\gamma_1(x_1)\cdot d\gamma_2(x_2)$$

hold for all pairs of bounded Borel sets. Notice that $\mu_1\times\mu_2$ becomes a probability measure since (i) above holds. With these notations the following result was proved by Beuling in the article *xxx*[xxx]:

Theorem. *For every product measure $\mu_1\times\mu_2$ with total mass one there exists a unique $\gamma_1\times\gamma_2$ in S_g^* such that*

$$\mathcal{T}_g(\gamma_1\times\gamma_2)=\mu_1\times\mu_2$$

Remark. Beurling's proof relies upon an involved variational problem and goes beyond the scope of the present chapter. But see § xxx in my notes *Special Topics* for an account of Beurling's studies related to Schrödinger's "unorthodox" Cauchy problem.

Some special examples.

It is valuable to become familiar with various "concrete examples" in order to appreciate the general theory about normed spaces and their operators. At the same time we remark that many problems which from the start are put in the context of "pure functional analysis" can only be solved by analytic function theory or other disciplines. Here follows some examples which illustrates this. The less experienced reader may neglect the present section for a while since the subsequent material is independent of the general theory in the chapters.

1. A linear system in infinitely many unknowns

Let $\{b_q\}$ be an infinite sequence of complex numbers indexed by positive integers. We assume it is separated which means that there exists $c > 0$ such that

(i)
$$|b_q - b_p| \geq c \quad : p \neq q$$

Let $\{a_q\}$ be another sequence of complex numbers where $a_q \neq b_p$ for all pairs p, q. We do not assume that $\{a_q\}$ is separated but it is a discrete subset of $\mathbf{C}$. Denote by ℓ^2 the linear space of sequences of complex numbers $\{x_n\}$ for which $\sum |x_n|^2 < \infty$. To each vector $\{x_n\} \in \ell^2$ we get the sequence

(ii)
$$\xi_p = \sum_{q=1}^{\infty} \frac{x_q}{a_p - b_q}$$

It is easily seen that (i) entails that

$$\sum_{q=1}^{\infty} \frac{1}{|a_p - b_q|^2} < \infty \quad : p = 1, 2, \ldots$$

The Cauchy-Schwarz inequality gives $\{\xi_p\} \in \ell^2$ and hence (ii) becomes a linear operator from ℓ^2 into itself. The question arises when this operator is injective, i.e. if the system

$$\sum_{q=1}^{\infty} \frac{x_q}{a_p - b_q} = 0 \quad : p = 1, 2, \ldots$$

only has the trivial null solution. Put in this generality one cannot give an answer. However, certain conclusions are available if the two sequences have angular distributions. More precisely, suppose that none of the sequences contain zero. To each p we get the poinit $\beta_p = |b_p|^{-1} b_p$ on the unit circle T. If $r \geq 1$ we get the disctete measure $\mu_b(r)$ on T defined by

$$\mu_b(r) = \frac{1}{r} \cdot \sum \delta(\beta_p)$$

where the sum extends over integers p such that $|b_p| \leq r$ and $\delta(\beta_p)$ assign unit point masses at the β-points. One says that $\{b_p\}$ has an angular distribution if there exists a non-negative Riess measure μ_b of finite mass such that

$$\lim_{r \to \infty} \mu_b(r) = \mu_b$$

holds in the weak-star topology of Riesz measures on T. In the same way we impose the condition that $\{a_p\}$ has an angular distribution and let μ_a be the resulting limit measure. Now we consider the support functions

$$h_b(\theta) = \int_0^{2\pi} \cos^+(\theta - \phi)\, d\mu_b(\phi) \quad : \ h_a(\theta) = \int_0^{2\pi} \cos^+(\theta - \phi)\, d\mu_a(\phi)$$

They define closed convex sets

$$C_b = \{re^{i\theta} : \ 0 \leq r \leq h_b(\theta)\} \quad : \ C_a = \{re^{i\theta} : \ 0 \leq r \leq h_a(\theta)\}$$

1.1 Theorem. *If $C_b \setminus C_a \neq \emptyset$ the system (*) only has the trivial null solution. Conversely, if C_b is contained in the interior of C_a there exists non-trivial solutions to (*).*

Remark. The proof of this result relies upon analytic function theory and cannot be derived via general considerations in functional analysis, even if the question is posed in the Hibert space ℓ^2. So the result above illustrates that in many "concrete situations" the general theory serves as a guidance only, while the solutions require a more delicate analysis. We refer to § xx in special Topics for an account of the proof of the theorem above and remark that equations of thid nature were studied in the thesis *Sur une classe de systèmes d'équations lineaires à une infinité d'innconnues* by K. Dagerholm at Uppsala in 1938.

2. A singular integral.

Given a real-valued and continuous function $a(x)$ defined on $[-1,1]$ we consider the operator

$$\mathcal{K}_u(x) = a(x)u(x) - \int_{-1}^{1} \frac{u(x)-u(y)}{x-y}\,dx$$

Here some care must be taken, i.e. the u-functions should be sufficiently regular in order that the integral above exists. A sufficient condition is that u is Hölder continuous of some order > 0, i.e.

$$|u(x)-u(y)| \leq C \cdot |x-y|^{\delta}$$

hold for a constant C and some $\delta > 0$. Consider the inhomogeneous equation

$$(*) \qquad \mathcal{K}(u) = f$$

where f is a given function. When a is identically the following uniqueness result holds:

Theorem. *If $u(x)$ is Hölder continuous of some order > 0 on the unit interval $[0,1]$ and satisfies*

$$\int_0^1 \frac{u(x)-u(y)}{x-y}\,dx = 0 \quad : 0 < x < 1$$

then u is a constant.

Let us now turn to the equation (*). If the function $a(x)$ is real-analytic it turns out that (*) has a solution u for every real-analytic function f. Moreover, u is found by the inversion formula:

$$u(x) = \frac{a(x)}{a(x)^2+\pi^2} \cdot f(x) + \frac{e^{\omega(x)}}{\sqrt{a(x)^2+\pi^2}} \cdot \int_{-1}^{1} \frac{e^{-\omega(s)}f(s)}{\sqrt{a(s)^2+\pi^2}} \cdot \frac{1}{s-x}\,ds$$

where the ω-function is defined by

$$\omega(x) = \int_{-1}^{1} \frac{\theta(s)}{s-x}\,ds \quad : \theta(s) = \frac{1}{2\pi i} \cdot \log \frac{a(x)+\pi i}{a(x)-\pi i}$$

and the argument of the complex log-function has been chosen above so that $0 < \theta(x) < 1$ hold for each $-1 < x < 1$.

Renark. The results above appear in Carleman's article *Sur la résolution de certaines équations intègrales* [Arkiv för matematik, fysik och astronomi: Vol 16]. It goes without saying that the proof of the inversion formula relies heavily upon analytic function theory. So the here the solutions are not found by "abstract reasoning" or via calculus of variation. Instead one employs residue calculus and integral formulas taken from analytic function theory where mutli-valued branches of the complex log-function play an essential role.

3. Carleson-Kronecker sets.

Let us begin with a result due to F. Riesz. Consider a doubly indexed sequence $\{c_{\nu.n}\}$ of complex numbers where (ν, n) are pairs of non-negative integers. Suppose there exists a strictly increasing sequence of positive numbers $\{A_n\}$ which tends to $+\infty$ such that

$$|c_{\nu.n}| \leq \frac{1}{A_n} \quad \text{hold for all pairs} \quad \nu, n$$

Under this assumption we study the inhomogeneous system of linear equations

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

where $\{y_\nu\}$ is a bounded sequence of complex numbers. The equation is said to be solvable in ℓ^1 if there exists a sequence $\{x_n\}$ such that $\sum |x_n| < \infty$. The following result was proved by Riesz:

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

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

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

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

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

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

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

With these notations the result below is proved in § XX from *Special Topics.*

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

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

4. L^p-inequalities.

Many results which originally were constrained to specific situations can be extended when notions in functional analysis are adopted. An example is a theorem from the article [Hörmander-1960] where singular integrals are treated in a general context. Here is the set up: Let B_1 and B_2 be Banach spaces and n a positive integer. Let $x \mapsto K(x)$ be a function which for every $x \in \mathbf{R}^n$ assigns a vector $K(x)$ in the Banach space $L(B_1, B_2)$ of continuous linear operators from B_1 to B_2. One only assumes that K is a continuous function, i.e. it need not be linear. If $f(x)$ is a continuous B_1-valued function defined in $\mathbf{R}^n$ with compact support there exists the convolution integral

$$(1) \qquad \mathcal{K}f(x) = \int_{\mathbf{R}^n} K(x-y)(f(y))\, dy$$

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

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

whose Borel-Stieltjes integral with respect to y exists. The resulting value of this integral yields the B_2-vector in the left hand side in (1). When x varies, $x \mapsto \mathcal{K}f(x)$ becomes a B_2-valued function.

The Hörmander conditions. Following [ibid] we impose two conditions on K where norms on the three Banach spaces $L(B_1, B_2), B_1, B_2$ appear. The first is that there exist a constant C_K such that the following hold for every real number $t > 0$:

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

where the left hand side employs the operator norms of K defined for pairs of points tx and $t(x-y)$ in $\mathbf{R}^n$. The second is the following L^2-inequality for every function f where we employ the same constant C_K as above:

$$\int_{\mathbf{R}^n} ||\mathcal{K}f(x)||^2\, dx \leq C_K^2 \cdot \int_{\mathbf{R}^n} ||f(x)||^2\, dx \tag{2}$$

where one has used norms on B_1 respectively B_2 during the integration.

Using an extension of Vitali's covering lemma to a similar covering principle on arbitrary normed spaces, the following result is proved in [ibid]:

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

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

Remark. This result has a wide range of applications when it is combined with interpolation theorems by Markinkiewicz and Thorin. For example, classical results which involve L^p-inequalties during the passage to Fourier transforms or other kernel functions can be put into a general frame where one employs vector-valued functions rather than scalar-valued functions. Typical examples occur when K sends real-valued functions to vectors in a Hilbert space, i.e. B_1 is the 1-dimensional real line and $B_2 = \ell^2$. So Hörmander's theorem constitutes a veritable propagande for learning general notions in functional analysis. The idea to regard vector-valued functions was of course considered at an early stage. An example is a theorem due to Littlewood and Paley which goes as follows: Let $f(x)$ be a function on the real line such that both f and f^2 are integrable. To each integer $n \geq 0$ we set

$$f_n(x) = \big|\int_{2^{n-1}}^{2^n} e^{ix\xi} \cdot \widehat{f}(\xi)\, d\xi\,\big| + \big|\int_{-2^{-n}}^{-2^{n-1}} e^{ix\xi} \cdot \widehat{f}(\xi)\, d\xi\,\big|$$

Then it is proved in [L-P] that for each $1 < p \leq 2$ there exists a constant $C_p \geq 1$ such that

$$C_p^{-1}||f||_p \leq \cdot \int_{-\infty}^{\infty} \big(\sum_{n=0}^{\infty} f_n(x)^2\big)^{\frac{2}{p}}\, dx \leq C_p \cdot ||f||_p \tag{*}$$

where $||f||_p = \big(\int |f(x)|^p\, dx\big)^{\frac{1}{p}}$ is the L^p-norm of f. Hörmander's theorem can applied to get (*) which has the merit that the use a suitable vector-valued operator during the proof becomes transparent.

The class BMO. The space of functions with bounded mean oscillation was introduced by F. John and Nirenberg. It can be used to establish various estimates for singular integrals. A fundamental discovery due to C. Fefferman and E.M. Stein is the duality between BMO($\mathbf{R}^n$) and the Hardy space $H^1(\mathbf{R}^n)$. See their joint article *[*xxxx [Acta mathematica. xx 19xx] and also Fefferman's article *xxxx* [ibid. xxx 19xx]. In addition we refer to the excellent text-books by E.M Stein. A merit to use BMO($\mathbf{R}^n$) is that one often can establish requested inequalities within this space, and in this way ignore the study of so called weak-type estimates. Here is an example of a result due to the present author in the article *xxx*[Math. Scand. xxx 1973]: Let $n \geq 1$ and for a given positive number M we denote by $\mathfrak{Db}_M(\mathbf{R}^n)$ the family of all distributions in $\mathbf{R}^n$ with compact support in the ball $\{|x|| \leq M\}$ whose Fourier transform satisfy the decay condition

$$|\widehat{\mu}(\xi)| \leq (1 + |\xi|)^{-n} \tag{*}$$

Theorem. *For each n there exists a constant C_n such that every $\mu \in \mathfrak{Db}_M(\mathbf{R}^n)$ has bounded mean oscillation and its BMO-norm is bounded by C_n·??.*

Improvements of this result were later given by P. Sjölin. For example, he showed that (*) can be relaxed to the condition

$$\sum xxxx < \infty$$

A final remark. For results where translation invariant convolution operators are replaced by kernels of pseudo-differential operators we refer to volume 3 in Hörmander's eminent text-book series on linear PDE theory.

4.2 Sobolev inequalities

Sobolev's article *On a theorem of functional analysis* [Mat. Sbornik. Vol 4 (1938)] established inequalities where integrability assumptions on derivative of a function lead to various regularity properties of the function itself. For functions of one or two variables such results were established earlier by Landau, Hardy and Littlewoood. The results in [Sob] rely upon integral formulas which recover f from first order partial derivatives when f is defined in some half-space of $\mathbf{R}^n$, or more generally in a domain with smooth boundary. We present this in § xx. Let us remark that Sobolev's averaged Taylor formula was inspired by the methods used by Hardy and Littlewood in the proof of Theorem xx below. For L^2-norms Sobolev inequalites were a well understood by Hermann Weyl at an early stage who used them to solve the Dirichlet problemand more general elliptic boundary value problems. Thanks to the general interpolation results by Markinsiewiz and Thorin one arrives at quite sharp estimates of the Sobolev type where one often profits upon weak-type estimates. Finally, using the Fefferman-Stein duality between functions of bounded mean oscillation in $\mathbf{R}^n$ and the Hardy space $H^1(\mathbf{R}^n)$ one gets inequalities which go beyond those in Sobolev's original work. Since the literature dealing with inequalities of the Sobolev type is very extensive we shall refrain fgrom a more detailed discussion. But we shal, expose some crucial results, foreost restricted to functions of two variables since already here the general methods to attain Sobolev inequalities appear.

The Hardy-Littelwood theorem. Let f be a real-valued function defined in the open half-plane $\{x > 0\}$ in $\mathbf{R}^2$ with a bounded support, i.e. say that $f = 0$ when $x^2 + y^2 \geq 1$. Assume that f has first order dervatives in the open half-disc D^+ where $x > 0$ and $x^2 + y^2 < 1$ and impose the condition that

$$\iint_{D^+} \big[\, |\frac{\partial f}{\partial x}(x,y)|^p + |\frac{\partial f}{\partial y}(x,y)|^p \,\big] \cdot dxdy < \infty$$

holds for some $p \geq 1$, i.e. the partial derivatives belongs to $L^p(D^+)$.

4.2.1 Theorem. *If the first order derivatives of f belong to $L^p(D^+)$ for some $1 < p < 2$, then*

$$\iint_{D_+} |f(x,y)|^{p_*} < \infty \quad \text{where} \quad \frac{1}{p_*} = \frac{1}{p} - \frac{1}{2}$$

4.2.2 The case $p = 2$. Before we give the proof we notice that $p = 2$ means that $p_* = +\infty$. The question arises if square integrability of the partial derivatives imply that f is a bounded function. The answer is negative. To see this we apply Riemann's Mapping Theorem. Namely, consider a simply connected unbounded open set Ω in $\mathbf{C}$ which has finite area while $\mathfrak{Re}(z)$ can become arbitrary large in the unbounded domain. Let $f \colon D \to \Omega$ be the conformal mapping. Now one has the equality

$$\text{area}(\Omega) = \iint |f'(z)|^2 \cdot dxdy$$

So with $u = \mathfrak{Re}\, f$ the first order derivatives of u are square integrable while Ω while the choice of Ω entails that the harmonic function u is unbounded in D.

Proof of Theorem 4.2.1

Let $b(r)$ be a smooth function defined on $r \geq 0$ where $b(r) = 1$ if $0 \leq r \leq 2$ and $b(r) = 0$ if $r \geq 3$. Next, let $a(\theta)$ be a smooth and non-negative function on the periodic θ-interval where

$$\int_0^{2\pi} a(\theta)\, d\theta = 1$$

We also assume that the support of a stays close to the point 1 on the unit circle, say that $a(\theta) =$ if $|e^{i\theta} - 1| \geq 1/10$. Set:

$$F(x,y) = \iint \frac{\partial}{\partial_r}(f(x + r\cos\theta, y + r\sin\theta) \cdot b(r) \cdot a(\theta)\, dr d\theta$$

where the double integral is over $\mathbf{R}^+ \times T$. Partial integration with respect to r gives:

$$F(x,y) = f(x,y) - \iint f(x + r\cos\theta, y + r\sin\theta) \cdot b'(r) \cdot a(\theta)\, dr d\theta$$

Here $b'(r) = 0$ when r is outside $[2,3]$ so the constraints on the supports of f and a imply that the last integral vanishes when $x^2 + y^2 \leq 1$. In the half-disc D_+ we therefore have $F = f$. At the same time rules of differentation give

$$\frac{\partial}{\partial_r}(f(x + r\cos\theta, y + r\sin\theta) =$$

$$\frac{\partial}{\partial_x} f(x + r\cos\theta, y + r\sin\theta)) \cdot \cos\theta + \frac{\partial}{\partial_y} f(x + r\cos\theta, y + r\sin\theta)) \cdot \sin\theta$$

Define the functions

$$h(r,\theta) = \frac{\cos\theta \cdot b(r) \cdot a(\theta)}{r} \quad : \quad g(r,\theta) = \frac{\sin\theta \cdot b(r) \cdot a(\theta)}{r}$$

From the above we get

$$F(x,y) = \iint_{\mathbf{R}^2} \big[\frac{\partial}{\partial_x} f(x+t, y+s) \cdot h(s,t) + \frac{\partial}{\partial_y} f(x+t, y+s) \cdot g(s,t)\big] \cdot ds dt$$

Hence F is the sum of two convolution integrals. The hypothesis in Theorem 0.1 means that each term is the convolution given by a compactly supported L^p-function and h, respectively g. Ths reduces the proof to an inequality for convolution integrals. More precisely, Theorem 4.2.1 follows we have proved:

4.2.3 Lemma. *To each $1 < p < 2$ there exists a constant C_p such that*

$$||\phi * h||_{p_*} \leq C_p \cdot ||\phi||_p$$

hold for each L^p-function $\phi(x,y)$ supported by the compact half-disc $\overline{D^+}$, and with a similar result when h is replaced by g.

4.2.4 How to prove Lemma 4.2.3

.

In their original proof, Hardy and Littlewood used weak-type estimates to analyze the two extreme cases when $p = 1$ or $p = 2$, and prior to the Fefferman-Stein theory text-books have exposed the Hardy-Littelwood theorem via interpolation results due to Markinzievich and Thorin. A more precise study when $p = 2$ is available which gives a proof of Lemma 4.2.3 via the general interpolation results from —Fef-Stein]. Namely, the constructions of h and g entail that both have Fourier transforms which decay like $|\xi|^{-1}$ with $\xi = (\xi_1, \xi_2)$. This leads us to consider the following set-up:

To each $R \geq 1$ we denote by D_R the closed disc of radius R centered at the origin in the (x, y)-space. Next, $\mathfrak{Db}_*(R)$ is the family of distributions μ with compact support in in D_R whose Fourier transforms satisfy

(i) $$|\widehat{\mu}(\xi)| \leq (1 + |\xi|)^{-1}$$

for all ξ. Each such distribution gives the convolution operator T_μ which to begin with is defined on test-functions $f(x, y)$ by

$$T_\mu(f) = f * \mu$$

4.2.5 Theorem. *To each $R \geq 1$ there exists a constant C_R such that the inequality below hold when f is a test-function with support in D_R and $\mu \in \mathfrak{Db}_*(R)$*

$$||f * \mu||_{\text{BMO}} \leq C_R \cdot ||f||_2$$

where the right hand side refers to the norm in the John-Nirenberg space of functions havnig bounded mean oscillation.

Remark. This result appears nowadays as a standard result from the theory of singular integrals. The reader may consult my cited article [Björk] for the methods which give the theorem above, or text-books by E.M. Stein et. al.

4.2.6 Use of interpolation. Theorem 4.2.5 can be spplied with μ equal to h or g. Notice that these functions also are integrable and recall that convolution of L^1-functions again belongs to L^1, where

$$||f * \mu||_1| \leq ||\mu||_1 \cdot ||f||_1$$

hold for all test-funtions and no constraint on their support is needed. Together with the general interpolation result from the work by Fefferman and Stein, it follows that Theorem 4.2.5 gives the result below where $\mathfrak{Db}_{**}(R)$ denotes the class if distributions $\mu \in \mathfrak{Db}_{**}(R)$ which in addition are L^1-densities with norms $||\mu||_1 \leq 1$.

Theorem. *If $\mu \in \mathfrak{Db}_{**}(R)$ and f has support in D_R, the following hold for each $1 < p < 2$:*

$$||f * \mu||_{p_*} \leq$$

4.2.7 Regularity up to the boundary.

Consider the open rectangle

$$\square = \{0 < x < 1\} \times \{-1 < y < 1\}$$

Let $F(x, y)$ be a function in $\square$ with the additional property that the first order derivatives belong to $L^2(\square)$. Now one may ask if there exists the limit function

(i) $$f(y) = \lim_{x \to 0} F(x, y)$$

and if it yields a continous function of y. This turns out that to be a subtle question. Consider the special case when f is a harmonic function in $\square$. Then the limit exist for all y outside a set whose outer logarithmic capacity is zero. This result (which is sharp) was proved by Beurling in the article *xxxx* [Acta math. 1940]. The proof goes beyond the scope of the present chapter since it would require a rather extensive detour into the theory about subharmonic functions and properties of logarithmic potential functions in the complex plane. For open cubes in $\mathbf{R}^3$ where $y = (y_1, y_2)$ the situation becomes more involved and it appears that no definite answers are known about the properties of the limit function f under the condition that the first order derivatives of F belong to $L^2(\square)$. So apart from cases where standard estimates from the theory about singular integrals are available, questions of a more refined nature related to Sobolev inequalities remain to be investigated.

4.2.8 Easier cases. Avoiding the case $p = 2$ we impose the condition that the first order derivatives of F belong to L^p for some $p > 2$. In that case the proof of Hardy and Littewood entails that a continuous limit function f exists in (i). In fact, this follows from the observation that the functions h and g belong to L^r for each $r < 2$. With $p > 2$ we can take r such that

$$r^{-1} + p^{-1} = 1$$

and Hölder's inequality entails that

$$||F * h||_\infty \leq ||F||_p \cdot ||h||_r$$

Approximating F by a sequence of test-functions whose L^p-norms are uniformly bounded by that of F the reader should verify that F extends to a continuous function along $x = 0$ in (i). Moreover, the maximum norm of $f(y)$ is bounded by a constant times the sum of the L^p-norms of the first order derivatives of F.

4.2.9 Sobolev's integral formula.

Here follows a crucial construction from Sobolev's original work. Let $n \geq 2$ and consider a real-valued function $F(x) = F(x_1, \ldots, x_n)$ in $\mathbf{R}^n$ which is zero outside a ball of radius $K/2$ centered at the origin for some positive number K. We shall perform some constructions which lead to a certain integral formula for F in (*) below.

Choose a C^∞-function $b(r)$ on the real line where $b(r) = 1$ if $0 \leq r \leq 2K$ and zero if $r > 3K$. Let ω denote points on the unit sphere S^{n-1}. For each fixed ω we set

$$F_\omega(x) = \int_0^\infty \frac{\partial}{\partial r}(F(x - r\omega) \cdot b(r)\, dr$$

A partial integration shows that the right hand side becomes

$$F(x) + \int_0^\infty F(x - r\omega) \cdot b'(r)\, dr$$

Now $b'(r) \neq 0$ only occurs if $2K < r < 3K$ so if $|x| \leq K/2$ it is clear that the last integral is zero because F vanishes when $|x| > K/2$. Hence the following hold for each $\omega \in S^{n-1}$:

(1) $$F_\omega(x) = F(x) \quad : |x| \leq K/2$$

Notice that

(i) $$\frac{\partial}{\partial r}(F(x - r\omega) = \sum_{j=1}^{j=n} \omega_j \cdot \frac{\partial}{\partial x_j}(F(x - rw)$$

Hence (1) entails that

$$F(x) = \frac{1}{|S^{n-1}|} \cdot \sum_{j=1}^{j=n} \int_{S^{n-1}} \int_0^\infty \omega_j \cdot \frac{\partial}{\partial x_j}(F(x - rw) \cdot b(r)\, dr d\omega \quad : |x| \leq K/2$$

where the volume of the unit sphere was introduced. Define the functions $h_1, \ldots, h_n$ in $\mathbf{R}^n$ expressed in polar coordinates r, ω by

$$h_j(r\omega) = \frac{1}{|S^{n-1}|} \cdot \frac{b(r) \cdot \omega_j}{r^{n-1}}$$

With $y = r\omega$ as a variable in $\mathbf{R}^n$ it follows that

(*) $$F(x) = \sum_{j=1}^{j=n} \int \frac{\partial}{\partial x_j}(F(x - y) \cdot h_j(y)\, dy \quad : |x| \leq K/2$$

Remark. Above (*) which can be regarded as an "averaged Taylor formula", and since the right hand side is a sum of convolution integrals one can apply results about singular integrals. Notice that the right hand side in (xx) entails that that the restrictions of the h-functions to every sphere $\{y| = \delta\}$ have mean-value zero. The h-functions have by construction compact support and are C^∞ outside the origin. The denominator r^{n-1} which appears in (x) causes singularities. However, the h-functions are integrable and behave better than Riesz kernel where r^n appears in the denominator. Let us finally remark that it is a "matter of taste" how to describe regularity of F when the partial derivatives belong to L^p for some $p > 1$. Hölder continuity of some order $\alpha > 0$ is one possibility. Another is to express regularity via properties of the Fourier transform of F which often is used in PDE-theory. More precsely, let $\{\phi_j(\xi)\}$ be the Fourier transforms of the partial derivatives. Fourier's inversion formula entails that the convolution integrals which appear in (*) are given by

$$x \mapsto (\frac{1}{(2\pi)^n} \cdot \int e^{\{i\langle x,\xi\rangle} \cdot \phi_j(\xi) \cdot \widehat{h}_j(\xi)\, d\xi$$

4.2.10 Passage to higher order derivatives. Passing to second order derivatives of F one gets similar averaged Taylor formulas as in (*). For example, one finds doubly indexed functions $\{h_{jk}\}$ such that

$$F(x) = \sum\sum \int \frac{\partial^2}{\partial x_j \partial x_k}(F(x-y) \cdot h_{jk}(y)\, dy$$

This time the compactly supported h-functions are of the form

$$\frac{\rho(x)}{|x|^{n-2}}$$

where ρ are test-functions. Moreover, the Fourier transforms of the doubly indexed h-functions decay like $\xi|^{-2}$ as $|\xi| \to +\infty$. Passing to the case of n:th order derivatives it follows that F is represented as a sum of convolution integrals of n:th order partial derivtives times compactly supported distributions whose Fourier transforms decay like $|\xi|^{-n}$. Applying the cited result in (§ 4.1. xx) one arrives at the following:

4.2.11 Theorem. *Let $\square$ be a fixed bounded cube in $\mathbf{R}^n$. Then there exists a constant C whuch only depends on the size of $\square$ and n such that the following hold for n-times differentiable functions supported by $\square$:*

$$||F||_{\text{BMO}} \leq \sum ||\frac{\partial^\alpha(F)}{\partial x^\alpha}||_1$$

with the sum extened over multi-indices of length $\leq n$.

4.2.11 Smooth extensions to boundaries. Consider a bounded open set Ω in $\mathbf{R}^n$ with a smooth boundary, i.e. of class C^∞. Let $p \geq 1$ and k is a positive integer which yields the largest integer m such that

$$\frac{1}{m} < k - \frac{n}{p}$$

CHECK

4.2.3 Theorem. *Let $F(x)$ be a function in Ω whose partial derivatives up to order k belong to $L^p(\Omega)$. Then every derivative of order $\leq m$ exists and extends to a continuous function on the closure of Ω.*

5. The complex Hilbert transform

.

Identify the complex z-plane with the 2-dimensional real (x, y)-space. Now we have the Hilbert space $L^2(\mathbf{C})$ whose vectors are complex-valued and square integrable functions in the (x, y)-plane. If g is a continuously differentiable function with compact support there exists a limit

$$G(z) = \lim_{\epsilon \to 0} \iint_{|z-\zeta|>\epsilon} \frac{g(\zeta)}{(z-\zeta)^2}\, d\xi d\eta \tag{5.1}$$

for every $z \in \mathbf{C}$. In fact, using Stokes theorem one has

$$G(z) = \iint \frac{\frac{\partial g}{\partial \zeta}(\zeta)}{z - \zeta} d\xi d\eta$$

where the last integral exists since $\frac{\partial g}{\partial \zeta}$ is a continuous function with compact support and $(z-\zeta)^{-1}$ is locally integrable in the (ξ,η)-space for each fixed z. The complex Hilbert transform is defined by the linear operator:

$$T_g(z) = \frac{1}{\pi}\, G(z)$$

It turns out that T is an isometry with respect to the L^2-norm, i.e.

$$\iint |T_g(z)|^2\, dxdy = \iint |g(z)|^2\, dxdy \quad : g \in C_0^1(\mathbf{C}) \tag{5.2}$$

There are several proofs of (5.2). One relies upon the Fourier transform and Parseval's theorem which is exposed in § xx. Another proof employs an abstract reasoning and goes as follows: Let $\mathcal{H}$ be a Hilbert space and Φ us a subset of vectors with the property that the subspace of vectors in $\mathcal{H}$ of the form

$$c_1\phi_1 + \ldots + c_k\phi_k \quad : \langle \phi_j, \phi_i\rangle = 0 \colon j \neq i$$

is dense in $\mathcal{H}$. Thus, we assume that finite linear combinations of pairwise orthogonal Φ-vectors is dense. Let us now consider a linear operator T on $\mathcal{H}$ such that the following hold:

$$||T(\phi)|| = ||\phi|| \quad : \phi \in \Phi \tag{*}$$

$$\langle \phi_1, \phi_2\rangle = 0 \implies \langle T(\phi_1), T(\phi_2)\rangle = 0 \quad : \phi_1, \phi_2 \in \Phi \tag{**}$$

Exercise Show that (*)-(**) imply that T is an isometry, i.e. $||T(g)|| = ||g||$ for every $g \in \mathcal{H}$.

Apply the result above to the complex Hilbert transform where Φ is the family of vectors in $L^2(\mathbf{C})$ given by characteristic functions of dics in $\mathbf{C}$. It is clear that this family satisfies the density condition above. Now we study the action by the Hilbert transform on such functions. Let $z_0 \in \mathbf{C}$ and $r_0 > 0$ be given and ϕ is the characteristic function of the disc of radius r_0 centered at z_0.

5.1 Proposition. *One has $T_\phi(z) = 0$ when $|z - z_0| < r_0$ while*

$$T_\phi(z) = \frac{r_0^2}{(z-z_0)^2} \quad : |z - z_0| > r_0$$

Proof. After a translation we can take $z_0 = 0$ and via a dilation of the scale reduce the proof to the case $r_0 = 1$. If $|z| > 1$ we have

$$T_\phi(z) = \frac{1}{\pi}\iint_D \frac{d\xi d\eta}{(z-\zeta)^2} = \frac{1}{\pi z^2}\iint_D \frac{d\xi d\eta}{(1-\zeta/z)^2}$$

Expand $(1-\zeta/z)^{-2}$ in a power series of ζ and notice that

$$\iint_D \zeta^m\, d\xi d\eta = 0 \quad : m = 1, 2, \ldots$$

Then we see that $T_\phi(z) = z^{-2}$. There remains to show that $|z| < 1$ gives $T_\phi(z) = 0$. To prove this we use the differential 2-form $d\zeta \wedge d\bar{\zeta}$ and there remains to show that

$$\lim_{\epsilon \to 0} \iint_{D_\epsilon} \frac{d\zeta \wedge d\bar{\zeta}}{(z-\zeta)^2} = 0 \tag{i}$$

where $D_\epsilon = D \setminus \{|\zeta - z| \leq \epsilon\}$. Now $\partial_\zeta((z-\zeta)^{-1}) = (z-\zeta)^{-2}$ and Stokes theorem gives (i) if

$$\int_{|\zeta|=1} \frac{d\bar{\zeta}}{z-\zeta} - \lim_{\epsilon \to 0} \int_{|\zeta - z| = \epsilon} \frac{d\bar{\zeta}}{z-\zeta} = 0$$

In the first integral we use the series expansion $(z-\zeta)^{-1} = -\sum_{n=0}^\infty \zeta^{-n-1}\cdot z^n$ and get a vanishing since

$$\int_{|\zeta|=1} \zeta^{-n-1}\, d\bar{\zeta} = 0 \quad : n = 0, 1, \ldots$$

The verification that the limit as $\epsilon \to 0$ in the second integral vanishes is left to the reader.

The Proposition entails that

$$||T(\phi)||^2 = r_0^4 \cdot \iint_{|z|>r_0} \frac{dxdy}{|z|^4} = r_0^4 \cdot 2\pi \int_{r_0}^{\infty} \frac{dr}{r^3} = \pi \cdot r_0^2$$

Since $||\phi||^2 = \pi \cdot r_0^2$ we conclude that $||T(\phi)|| = ||\phi||$. The verification that (**) holds when ϕ_1, ϕ_2 are characteristic functions of disjoint discs is left as an exercise.

6. The Kakutani-Yosida theorem.

Let $(\Omega, \mathcal{B}, \mu)$ be a probability space defined in [Appendix: Measure theory]. Consider a real-valued function P defined on the product set $\Omega \times \mathcal{B}$ with the following two properties:

(*) $\quad t \mapsto P(t, E)$ is a bounded measurable function for each $E \in \mathcal{B}$

(**) $\quad E \mapsto P(t, E)$ is a probability measure for each fixed $t \in \Omega$

When (*-**) hold one refers to P as a transition function. Inductively we get the sequence $\{P^{(n)}\}$:

$$P^{(n+1)}(t, E) = \int_\Omega P^{(n)}(s) \cdot dP(s, E)$$

It is clear that $\{P^{(n)}\}$ yield new transition functions. The probabilistic interpretation is that one has a Markov chain with independent increments. More precisely, if E and S are two sets in $\mathcal{B}$ and $n \geq 1$. then

$$\int_S P^{(n)}(t; E) \cdot d\mu(t)$$

is the probability that the random walk which starts at some point in E has arrived to some point in S after n steps. One says that the transition function P yields a stationary Markov process if there exists a finite family of disjoint subets $E_1, \ldots, E_m$ in $\mathcal{B}$ and some $\alpha < 1$ and a constant M such that the following hold: First, for each $1 \leq i \leq m$ one has:

(1) $$P(t, E_i) = 1 \quad : t \in E_i$$

Next, if $\Delta = \Omega \setminus E_1 \cup \ldots \cup E_m$ then

(2) $$\sup_{t \in S} P^{(n)}(t, \Delta) \leq M \cdot \alpha^n \quad : \ n = 1, 2, \ldots$$

Remark. One refers to Δ as the disippative part of Ω and $\{E_i\}$ are the ergodic kernels of the process. Since $\alpha < 1$ in (2) the probabilistic interpretation of (2) is that as n increase then the disippative part is evacuated with high probability while the Markov process stays inside every ergodic kernel. A sufficient condition for a Markov process to be stationary is as follows: Denote by $\mathfrak{B}$ the Banach space of complex-valued and bounded $\mathcal{B}$-measurable functions on the real s-line. Now P gives a linear operator T which sends $f \in \mathfrak{B}$ to the function

$$T(f)(x) = \int_\Omega f(s) \cdot dPx, ds)$$

Theorem. *The Markov process is stationary if there exists a triple α, n, K where K is a compact operator on $\mathfrak{B}$, $0 < \alpha < 1$ and n some positive integer such that the operator norm*

(*) $$||T^n + K|| \leq \alpha$$

Remark. The proof relies upon general facts about linear operators on Banach spaces. First $\mathfrak{B}$ is a commutative Banach algebra with a maximal ideal space denoted by S and the Gelfand transform enable us to identify $\mathfrak{B}$ with the space of continuous complex-valued functions on the compact Hausdorff space S, Now T can be regarded as a positive linear operator on $C^0(S)$. The crucial step in the proof is to show that (*) entails that the spectrum of T consists of a finite set of points on the unit circle together with a compact subset in a disc of radius < 1. Moreover, for each isolated point $e^{i\theta} \in \sigma(T)$ the corresponding eigenspace is finite dimensional. Each such eigenvalue corresponds to an ergodic kernel and when the eigenspace has dimension $m \geq 2$, the corresponding ergodic kernel, say E_1, has a further decomposition into pairwise disjoint subsets $e_1, \ldots, e_m$ where the process moves in a cyclic manner between these sets, i.e.

$$\int_{e_{i+1}} P(s, e_i) = 1 \quad : 1 \leq i \leq m \quad \text{where we put} \quad e_{m+1} = e_1$$

Above we desfribed the probabilistic consequences of the theorem. The proof is given in § xx after we have introduced Neumann's resolvent operators and establish a result about positive operators on Banach spaces of continuous functions on compact topological spaces which gives the theorem above as a special case.

Basic facts in functional analysis.

Baire's category theorem for complete metric spaces is used in several proofs below. The detailed proofs of the results below are given in § xx.

A. Three basic principles.

A.1 The Banach-Steinhaus Theorem. *Let X be a vector space equipped with a complete norm $||\cdot||^*$. Then, every stronger norm is equivalent to the complete norm, i.e. if $||\cdot||$ is a norm for which there is a constant C such that*

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

hold for all vectors $x \in X$, then there is a constant C_1 such that

$$||x|| \leq C_1 \cdot ||x||^*$$

A.2 The Open Mapping theorem. Let $T\colon X \to Y$ be a bounded linear operator where Y is a Banach space and X some normed space. Assume that T is surjective, i.e. $T(X) = Y$. If $y \in Y$ we set

(i)
$$||y|| = \min_{x \in X} ||x|| : y = Tx$$

It is easily seen that this gives a norm on Y. Let $||\cdot||^*$ be the complete norm on Y. The construction in (i) gives

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

where $||T||$ is the operator norm of T. Now the Banach-Steinhaus theoren gives a constant C_1 such that every vector y in Y norm < 1 is the image of a vector x in X with norm $\leq C_1$. In particular T maps the open unit ball in X to an open set in Y which is convex and contains the origin,

A.3 The Hahn-Banach theorem. Let X be a real vector space. A function $\rho\colon X \to \mathbf{R}^+$ 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:

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

Using the axiom of choice one verifies that if X_0 is a subspace and $\lambda_0\colon X_0 \to \mathbf{R}$ is a linear functional such that

$$\lambda_0(x_0) \leq \rho(x_0) \quad : x_0 \in X_0$$

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

$$\lambda(x) \leq \rho(x) \quad : x \in X$$

A.4 Example. One is often confronted with severe difficulties in specific cases. 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.5 Separating hyperplanes. The Hahn-Banach theorem employs the axiom of choice and is therefore not very constructive. So one often asks for more explicit proofs. An example is the separation of a pair of disjoint closed and convex sets K and K_1 in a normed vector space. So here there exists some $\delta>0$ such that

(i) $$\min_{x,y}\,||x-y||=\delta$$

where the minimum is taken over pairs $x\in K$ and $y\in K_1$. Using the Hahn-Banach theorem one verifies that there exists a continuous linear functional x^* of norm one such that

(ii) $$\max_{x\in K}\,x^*(x)+\delta\leq\min_{x\in K_1}\,x^*(x)$$

In specific situations one often wants to find x^* in some constructive manner. When $X=\mathbf{R}^n$ is an euclidian space we can get x^* as follows. By a translation we may assume that the origin belongs to K and let us for a moment also assume that K contains a small open ball of radius δ centered at $x=0$. In this case the convexity of K easily entails that the interior is a dense open set. In fact, if $p\in\partial K$ and $0<s<1$ the convexity implies that K contains the open ball of radius $(1-s)\delta$ centered at sx. Since $||x-sx||=(1-s)||x||$ tends to zero as $s\to 1$ we conclude that the interior is dense. Next, let $p\in\partial K$ and denote by $\mathcal{H}(p)$ the family of hyperplanes π in X such that K is placed in one of the closed half-spaces separated by the affine hyperplane $\{p\}+\pi$.

Exercise. Show with the aid of a figure that the hyperplane

$$\{p\}^\perp=\{x\in X:\langle x,p\rangle=0\}$$

belongs to $\mathcal{H}(p)$.

Next, let K_1 be another closed convex set which gives the positve δ in (i). Since bounded sets in X are relatively compact, there exists a point $p\in K$ such that

$$\delta=\min_{y\in K_1}\,||p-y||$$

Exercise. With p as above we define x^* by

$$x^*(x)=-\frac{1}{||p||}\cdot\langle x,p\rangle$$

Then x^* has unit norm and we leave it to the reader to check that (ii) holds for this linear functional which by construction has unit norm. Finally, show also how to get x^* in (ii) when the two convex sets are not fat.

B. Dual spaces and weak topologies.

Let X be a normed complex vector space. The dual X^* consists of $\mathbf{C}$-linear maps $x^* \colon X \to \mathbf{C}$ which satisfy

$$|x^*(x)| \leq C \cdot ||x||$$

for a constant C which depends on the given linear functional x^*. The smallest C is the norm of x^* and in this way X^* becomes a normed vector space. The Hahn-Banach theorem entails that

$$||x|| = \max_{x^* \in S^*} |x^*(x)|$$

where S^* is the family of vectors in X^* with unit norm. Now we introduce other topologies on X. To begin with a Banach space X is equipped with the weak topology. Here a sequence $\{x_n\}$ converges weakly to a limit vector x if

$$\lim_{n \to \infty} x^*(x_n) = x^*(x) \quad : x^* \in X^*$$

To indicate weak convergnce we use w as a subscript and write

$$\lim_{n \to \infty} x_n \overset{w}{\to} x$$

B.0 The weak-star topology on X^*. It is defined by the family of semi-norms $\{\rho_X \colon x \in X\}$ where

$$\rho_x(x^*) = |x^*(x)|$$

for each pair $\in X$ and $x^* \in X^*$. Equipped with the weak-star topology, X^* becomes a locally convex vector space. Recall Tychonoff's theoremt in topology which asserts that the product of compact spaces is compact. It follows that if S^* is the unit sphere in X^*, then the weak-star topology restricted to S^* gives a compact topological space. Next, starting from the normed space X^* one constructs the bi-dual X^{**} which by definiton is the dual of X^*. If $x \in X$ we get the vector $\mathfrak{i}_X(x) \in X^{**}$ defined by

$$\mathfrak{i}_X(x)(x^*) = x^*(x) \quad : x^* \in X^*$$

The Hahn-Banach theorem entails that the map $x \to \mathfrak{i}_X(x)$ is an isometry, i.e.

$$||x|| = ||\mathfrak{i}_X(x)||$$

One refers to $\mathfrak{i}_X \colon X \to X^{**}$ as the bi-dual embedding. A Banach space X is reflexive if $\mathfrak{i}_X$ is surjective In § xx we prove that X is reflexive if and only if the unit sphere S in X is compact with respect to the weak topology which was defined above. 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. An example of a non-reflexive Banach space is $\mathbf{c}_0$ whose vectors are complex sequences $\{x_n\}$ for which $x_n \to 0$, and the norm is defined by

$$||x|| = \max_n |x_n|$$

Here the dual space is ℓ^1 and the bi-dual $\mathbf{c}_0^{**} = \ell^\infty$ where ℓ^∞ is the Banach space whose vectors consist of bounded sequences of complex numbers.

B.1 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. Each $x \in X$ gives a linear function $\widehat{x}$ on X^* defined by

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

By definition the weak-star topology on X^* is the weakest topology so that $\widehat{x}$ becomes a continuous function for every vector $x \in X$. It means that a fundamental system of open neighborhoods of the origin in X^* consists of sets

$$U(x_1, \ldots, x_N; \delta) = \{x^* \colon |x^*(x_\nu)| < \delta\}$$

Next we construct the *bounded weak-star topology* on X^*. Here one takes the open unit ball S^* 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 locally convex topology on X^* whose corresponding topological vector space is denoted by X^*_{bw}, while X^*_w denotes the locally convex 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:

B.1.1. Theorem. *The dual of X^*_{bw} is equal to $\widehat{X}$.*

We prove this in § x. For the moment we only remark that an essential ingredient in the proof employs a result due to Stefan Banach which goes as follows: Consider the Banach space $\mathbf{c}_0$ whose vectors consist of sequences $\{x_1, x_2, \ldots\}$ of complex numbers where $x_n \to 0$ as $n \to \infty$ and the norm

$$||x|| = \max_{n \geq 1} |x_n|$$

Banach proved that the dual space $\mathbf{c}_0^*$ is ℓ^1 whose vectors are complex sequences $\{y_n\}$ where $\sum |y_n| < \infty$.

B.2 Bounded operators with closed range. Consider a bounded linear operator $T \colon X \to Y$ where X and Y are Banach spaces. Then there exists the adjoint linear operator

$$T^* \colon Y^* \to X^*$$

where a continuous linear functional y^* on Y^* is mapped into the vector

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

B.2.1. Theorem. *T has closed range if and only if T^* has a closed range.*

Remark. Thus, the subspace $T(X)$ is closed in Y if and only if $T^*(Y^*)$ is closed in X^*. We prove this result in §§§ using the Krein-Smulian theorem.

B.3 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:

B.3.1 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.

C. 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.

C.1 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

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

C.2 Exercise. Use (C. 1) to 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.

C.3 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.

D. Locally convex topologies and 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

(D.1) $$\rho(\alpha \cdot x) = |\alpha| \cdot \rho(x) \quad : \rho(x_1 + x_2) \leq \rho(x_1) + \rho(x_2)$$

where x, x_1, x_2 are vectors in X and $\alpha \in \mathbf{C}$. If X is a real vector space the left hand side equality is only required for real scalars α. Notice that (0.0.1) entails that the kernel of ρ is a subspace of X denoted by $\mathrm{Ker}(\rho)$. If it is reduced to the zero vector one says that ρ is a norm. Next, a subset K in X is convex if x_1, x_2 in K implies that the line segment

$$[x_1, x_2] = \{sx_1 + (1-s)x - 2 : 0 \leq s \leq 1\} \subset K$$

If ρ is a pseudo-norm we put

$$\mathcal{B}_\rho = \{x : \ \rho(x) \leq 1\}$$

This is a convex set with the property that

(i) $$x \in \mathcal{B}_\rho \implies e^{i\theta} \cdot x \in \mathcal{B}_\rho \quad : \ 0 \leq \theta \leq 2\pi$$

Moreover $\mathcal{B}_\rho$ is absorbing which means that for every vector $x \in X$ there exists some real $s > 0$ such that $s \cdot x \in \mathcal{B}_\rho$.

D.2 A converse result. Denote by $\mathcal{K}_X$ the family of convex, symmetric and absorbing subsets K with the property that

(ii) $$K = e^{i\theta} \cdot K \quad : 0 \leq \theta \leq 2\pi$$

For a real vector space one only imposes symmetry with respect to the origin, i.e. that $K = -K$.

Exercise. Show that (ii) gives a pseudo-norm ρ defined by

(iii) $$\rho(x) = \inf_{s>0} x \in s \cdot K$$

and verify the inclusions

$$s \cdot \mathcal{B}_\rho \subset K \subset B_\rho \quad : \forall \ 0 < s < 1$$

D.3 Locally convex topologies. Let $\{\rho_\alpha\}$ be a family of semi-norms on X such that

(i) $$\bigcap \mathrm{Ker}(\rho_\alpha) \ = \{0\}$$

One defines a topology whose open neighborhoods of the origin are of the form:

(ii) $$U = \cap \ \{\rho_\alpha < \epsilon\}$$

where $\epsilon > 0$ are arbitrary small and the intersection is taken over a finite set α-indices. Finally, a set Ω in X is open if there for every $x \in \Omega$ exists some finite ρ-family and $\epsilon > 0$ such that $x + U \subset \Omega$ with U as in (ii). One refers to a locally convex topology defined by the ρ-family. Notice that (i) implies that the topology is Haussdorf.

D.4 Frechet spaces. Let $\{\rho_n\}$ be a denumerable sequence of pseudo-norms where (i) holds above. Put

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

It is clear d is a metric which defines the $\{\rho_n\}$-topology. If every Cauchy-sequence with respect to d converges to a vector in X one refers to X as a Frechet space

D.5 Theorem. *Let X be a Frechet space with the Frechet metric d. Then every d-closed convex set K in the family $\mathcal{K}_X$ 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.

D.6 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 Theorem 0.1.5.

D.7 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.

D.8 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.

Dual spaces.

Let X be a locally convex real vector space where $\{\rho_\alpha\}$ is a family of pseudo-norms defining the topology on X. The dual space X^* consists of continuous linear forms on X. The continuity for a vector $x^* \in X^*$ means that there exist a finite family of pseudo-norms $\{\rho_\alpha\}$ and a constant C such that

$$|x^*(x)| \leq C \cdot \sum \rho_\alpha(x) \quad : x \in X$$

If $x^* \in X^*$ and a is a real number one gets the closed half-space

$$H(x^*; a) = \{x : \; x^*(x) \leq a\}$$

The result below appears in the article *La dualité dans les espaces vectorielles* [Ann. Ecoles Sup. 1942] by Dieudonné.

D.9 Theorem. *Every closed and convex subset K in X which is not equal to X is the intersection of closed half-spaces.*

Proof. Clearly it suffieces to prove the result when K contains the origin. If $\xi \in X \setminus K$ the closedness gives an open set U and some $0 < s < 1$ such that $K \cap \{s \cdot \xi\} + U = \emptyset$. By translations this entails that the convex set $\widehat{K} = K + U$ does not contain $s \cdot \xi$. To this convex set we construct the associated function ρ as in (ii) from (D.2) and (iii) entials that

$$\rho(s \cdot \xi) \geq 1 \implies \rho(\xi) = s^{-1} > 1$$

Next, the Hahn-Banach theorem gives a linear funtional λ on X such that $\lambda(\xi) = s^{-1}$ while $|\lambda(x) \leq \rho(x)$ for each $x \in \widehat{K}$. Since the open set u is contained in $\widehat{K}$ it follows that the functional λ is continuous and from the above K is contained in the half-space $H(\lambda : 1)$ while ξ is outside this half-space. Since $\xi \in X \setminus K$ was arbitrary the proof is finished.

D.10 Support functions of convex sets.

In §§ we prove a theorem due to Hörmander which describes support functions of closed and convex subsets in a locally convex vector space. Hörmander's result goes as follows. Let X be a real vector space, i.e. the elements are vectors and each $x \in X$ yields vectors $s \cdot x$ for every real number s. A pseudo-norm on X is a sub-additive real-valued map ρ such that $\rho(s \cdot x) = |s| \cdot \rho(x))$ hold for every real number s and every $x \in X$. Let $\mathcal{F} = \{\rho_\alpha\}$ be a family of pseudo-norms indexed by a set which may be finite or infinite. One associates the space $X^*_{\mathcal{F}}$ of linear functionals λ on X for which there exists some finite family $\{\rho_{\alpha_1}, \ldots, \rho_{\alpha_n}\}$ in $\mathcal{F}$ and a constant C such that

$$|\lambda(x)| \leq C \cdot \sum_{\nu=1}^{\nu=n} \rho_{\alpha_\nu}(x)$$

for every $x \in X$. To each pseudo-norm ρ_α in the family and $\epsilon > 0$ we put

$$U(\alpha; \epsilon) = \{\lambda \in X^*_{\mathcal{F}} : |\lambda(x)| < \epsilon \cdot \rho_\alpha(x) \, x \in X\}$$

The weak-star topology on $X^*_{\mathcal{F}}$ arises as follows: A subset Ω in $X^*_{\mathcal{F}}$ is weak-star open if there for each $\lambda_0 \in \Omega$ exists a finite family of pseudo-norms $\{\rho_{\alpha_\nu}\}$ and some $\epsilon > 0$ such that

$$\lambda_0 + \cap U(\alpha_\nu; \epsilon) \subset \Omega$$

Next, denote by $\mathcal{S}(X^*_{\mathcal{F}})$ the family of functions g on $X^*_{\mathcal{F}}$ with values in $(-\infty, +\infty]$ such that the following hold for each pair λ_1, λ_2 in $X^*_{\mathcal{F}}$ and every real $s > 0$:

$$g(\lambda_1 + \lambda_2) \leq g(\lambda_1) + g(\lambda_2) \quad : g(s\lambda) = s \cdot g(\lambda)$$

where the rule is that $+\infty + a = +\infty$ for every real number a. It is easily seen that every convex subset K in X gives rise to such a g-function defined by

$$\mathcal{H}_K(\lambda) = \max_{x \in K} \lambda(x)$$

With these notations one has:

D.11 Theorem. *Let g be a function in $\mathcal{S}(X^*_{\mathcal{F}})$ which is lower semi-contoinuous in the weak-star topology, i.e. the sets $\{\lambda : g(\lambda) > a\}$ are weak-star open for every real s. Then there exists a convex set K such that $g = \mathcal{H}_K$.*

E. Non-linear calculus on Banach spaces.

Ordinary differential calculus in $\mathbf{R}^n$ can be extended to Banach spaces. We present some results due to Clarkson, Beurling and Lorch. Let X be a Banach space X and to each $\epsilon > 0$ we denote by $\mathcal{F}(\epsilon)$ the set of of pairs of unit vectors x, y in X such that

$$||x + y|| \geq 2 - \epsilon$$

Following Clarkson's article *Uniformly convex spaces* in [Trans. Amer.Soc. 40: 1936] we give

E.1 Definition. *A Banach space X is uniformly convex if*

$$\lim_{\epsilon \to 0} \max_{x,y \in \mathcal{F}(\epsilon)} ||x - y|| = 0$$

Next, a uniformly convex Banach space X is differentiable if

$$\lim_{a \to 0} \frac{||p + ax|| - ||p||}{a}$$

exists for evey pair of non-zero vectors x, p in X. During the passage to the limit a can be both > 0 or < 0. From now on X is uniformly convex and differentiable. Let S be the unit sphere in X, i.e. the set of vectors of unit norm. Similarly, S^* is the unit sphere in X^*. With $p \in X$ kept fixed the differentiablity gives a map from X into $\mathbf{R}$ defined by

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

With these notations the following result was proved by Clarkson in the cited article:

E.2 Theorem. *For each $p \in S$ the function $x \mapsto D_p(x)$ is a linear functional of unit norm and $p \mapsto D_p$ is a bijective map between S and S^*. Finally, $D_p(p) = 1$ and for every $x^* \in S^* \setminus \{D_p\}$ one has*

$$|x^*(p)| < 1$$

E.3 Non-linear duality. Clarkson's result is proved in § xx and can be used to construct non-linear duality maps. Consider a continuous and strictly increasing function $\phi(r)$ defined on $\{r \geq 0\}$ where $\phi(0) = 0$ and $\phi(r) \to +\infty$ as $r \to \infty$. Each $p \in S$ gives the Clarkson vector $D_p \in S^*$. For each non-zero vector $x \in X$ we put

$$\mathcal{D}_\phi(x) = \phi(||x||) \cdot D_{||x||^{-1} \cdot x}$$

Next, if $V \subset X$ is a closed subspace we set

$$V^\perp = \{x^* \in X^* : x^*(V) = 0\}$$

E.4 The Beurling-Lorch Theorem. *For each closed subspace V of X which is not equal to the whole space X, and every pair of vectors $x \in X$ and $y \in X^*$ the intersection*

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

consists of a single vector in X^*.

The result above is proved in § xx.

F. The logarithmic potential.

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. A favourable case is when $T(X)$ is a closed subspace of Y. But in general one encounters situations where the closedness fails to hold. 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

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

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. The Arzela-Ascoli theorem implies that L is a compact operator. It turns out that L is injective and has a dense range. See § xx for the proof which includes an inversion formula for L which therefore describes the range of L. But as we shall see this description is rather involved. It turns out that the range contains all C^2-functions and is therefore dense in $C^0[0,1]$ but its precise description is not easy to give. Next, general results about compact operators imply that the spectrum of L is discrete. Here $\sigma(L)$ contains the origin and outside the origin the spectrum is a discrete subset of $\mathbf{C}$ with a sole cluster point at zero. Moreover. general facts from Fredholm's theory sbout integral equations show that the spectrum is real and since the kernel function $\frac{1}{|x-t|}$ is everywhere postive, each eigenvalue μ for which there exists a non-trivial eigenfunction g satisfying

$$L_g(x)=\mu\cdot g(x)$$

is >0. One often prefers to replace the μ-numbers by μ^{-1} which gives a discrete non-decreasing sequence of eigenvalues $0<\lambda_1\leq\lambda_2\leq\ldots$ and associated eigenfunctions $\{g_n\}$ such that

$$L_{g_n}(x)=\lambda_n^{-1}\cdot g_n(x)$$

G. A result about spectral values.

Consider a linear operator $\mathcal{K}$ defined by

$$\mathcal{K}_g(x)=\int_0^1\frac{K(x,y)}{|x-y|^\alpha}\cdot g(y)\,dy$$

where $0<\alpha<1$ and $K(x,y)$ is continuous and symmetric, i.e. $K(x,y)=K(y,x)$ hold when $\{0\leq x,y\leq 1\}$. We regard $\mathcal{K}$ as a linear operator on the Hilbert space $L^2[0,1]$. The Cauchy-Schwartz inequality shows that it is a bounded operator whose operator norm is estimated above by:

$$||\mathcal{K}||\leq\max_{0\leq x\leq 1}\int_0^1\frac{|K(x,y)|}{|x-y|^\alpha}\,dy$$

If $0<\alpha<1/2$ the kernel function $K(x,y)=\frac{K(x,y)}{|x-y|^\alpha}$ is square integrable over the unit square $\{0\leq x,y\leq 1\}$ which means that $\mathcal{K}$ is a Hilbert-Schmidt operator. In this case Hilbert proved in his text-book from 1904 that the eigenvalues are real, and the eigenfunctions normalised with L^2-integrals of unit norm give an orthonormal basis for $L^2(0,1]$. The extension to the case $1/2\leq\alpha<1$ was treated by Carleman in his thesis from 1916, and again the spectrum consists real numbers with a sole cluster point at zero. Moreover, the eigenfunctions give an orthonormal basis in $L^2[0,1]$. It turns out that the set of eigenvalues, which are repeated when the eigenspaces have dimension ≥ 2 is a quite ample set. One has for example:

G.1 Theorem. *If there exists some interval (a,b) in $[0.1]$ such that $K(x,x)>0$ when $a<x<b$, then*

$$\sum\bigl(\frac{1}{\lambda_n^+}\bigr)^{\frac{1}{1-\alpha}}=+\infty$$

where the sum is taken over positive eigenvalues.

Remark. This theorem appears in Carleman's article *Sur la distribution des valeurs singuliérs d'une classe des noyaux infinis* [Arkiv för matematik, astronomi och fysik. Vol. 12] and a proof is given in § xx.

§ 0. Neumann's resolvent operators

Introduction. The less expericenced reader may prefer to first study basic material in the special chapters. But is it essential to become familiar with unbounded densely defined linear operators and their bounded resolvent operators. So we expose the crucial facts about this and remark that the major result appears in Theorem 0.xxx where analytic function theory is used to perform an operational calculus using resolvent operators.

0.1 The class $\mathcal{I}(X)$. Let X be a Banach space. Denote by $\mathcal{I}(X)$ the classs 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 whose domain of defintion $\mathcal{D}(T) = R(X)$. Namely, 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. It is also clear that

$$R \circ T(x) = x \quad : x \in \mathcal{D}(T)$$

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 by definition 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 os closed in $X \times X$, i.e. $\Gamma(T)$ is closed which we express by saying 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 (xx) 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 in § xx.

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 injective 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 $\rho(T)$. Put

$$\sigma(T) = \mathbf{C} \setminus \rho(T)$$

which by definition is the spectrum of T. Next, each $\lambda \in \rho(T)$ the construction in (0.1) gives a unique bounded operator $R_T(\lambda) \in \mathcal{I}(X)$. So here

$$(\lambda \cdot E - T) \circ R_T(\lambda)(x) = x$$

and since $\mathcal{D}(T) = \mathcal{D}(\lambda \cdot E - T)$ it follows that the range of $R_T(\lambda)$ is equal to $\mathcal{D}(T)$ for every resolvent value.

0.2.1 Definition. *Tbe bounded operators $\{R_T(\lambda) : \lambda \in \rho(T)\}$ are called Neumann's resolvent operators 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.

Lwet T be closed and densely defined. Assume that $\rho(T) \neq \emptyset$. The fundamental 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 (*) and at the same time this proves that the resolvent operators commute.

0.4 Neumann series.

If $\lambda_0 \in \rho(T)$ we construct the operator valued series

$$(1) \qquad S(\zeta) = R_T(\lambda_0) + \sum_{n=1}^{\infty} (-1)^n \cdot \zeta^n \cdot R_T(\lambda_0)^{n+1}$$

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. Thjs oroves 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 therimit as $\mu \to \lambda$ in Neumann's equation and conclude that the complex derivative is given by

(**) $$\frac{d}{d\lambda}(R_T(\lambda) = -R_T^2(\lambda)$$

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 oair of resolvent values of T we can write Neumann's equation in the form

(1) $$R_T(\lambda)(E + (\lambda - \mu)R_T(\mu)) = R_T(\mu)$$

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) $$\sigma(T) = \{\lambda : \frac{1}{\mu - \lambda} \in \sigma(R_T(\mu))\}$$

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 (xxx) 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 Exercise. 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. Apply this in Neumann's equation (xx) to conclude that if one resolvent operator $R_T(\lambda_0)$ is compact, then all resolvent operators 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 its 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) $$|y(Tx)| \leq C(y) \cdot ||x|| \quad : x \in \mathcal{D}(T)$$

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) $$y(Tx) = T^*(y)(x) \quad : x \in \mathcal{D}(T)$$

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. Let T be densely defined. There may exist 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 densely defined and closed and 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

(*) $$\rho(T) \subset \rho(T^*)$$

Proof. By translations it suffices to show that if the origin beongs 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

(i) $$||x|| \leq c^{-1} \cdot ||Tx|| \quad : x \in \mathcal{D}(T)$$

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

(ii) $$|y(Tx)| \geq 1/2$$

Now

(iii) $$y(Tx) = T^*(y)(x)$$

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.

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 euality

$$\sigma(T) \subset \sigma(T^*)$$

We refer to § xx for the proof.

0..6 Operational calculus.

Let T be a densely defined and closed operator on a Banach apace 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 triagle inequality entials 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

(0.6.2) $$T_f = \int_{\partial\Omega} f(z) R_T(z)\, dz$$

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

$$\text{(i)} \qquad T_g = \int_{\partial\Omega_*} g(z) R_T(z_*)\, dz_*$$

where we use z_* to indicate that integration takes place along $\partial\Omega_*$. Now

$$\text{(ii)} \qquad T_f \circ T_g = \iint_{\partial\Omega_* \times \partial\Omega} f(z) g(z_*) R_T(z) \circ R_T(z_*)\, dz_* dz$$

Neumann's equation (*) from (0.0.3) entails that the right hand side in (ii) becomes

$$\text{(iii)} \qquad \iint_{\partial\Omega_* \times \partial\Omega} \frac{f(z) g(z_*) R_T(z_*)}{z - z_*}\, dz_* dz + \iint_{\partial\Omega_* \times \partial\Omega} \frac{f(z) g(z_*) R_T(z)}{z - z_*}\, dz_* dz = A + B$$

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 **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

$$T_K = T E_K$$

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

$$S(\lambda_0E_K-T)=E_K$$

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

$$(\lambda_0E-T)E_T(\lambda_0)=\lim_{\epsilon\to 0}\frac{1}{2\pi i}\cdot\int_{|\lambda-\lambda_0|=\epsilon}(\lambda_0-\lambda)R(\lambda)\,d\lambda=0$$

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

$$N_k(\lambda_0)=\{x\colon(\lambda_0E-T)^k(x)=0\}\quad:1\leq k\leq M$$

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 expericence dreader may consult the subsequent proof after studies of basic material in the apecial 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 there exist*

$$A(x) = \lim_{h \to 0} \frac{T_h(x) - x}{h} \quad : x \in \mathcal{D}$$

Here the densely defined operator A is closed 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. The semi-group $\{e^{tB}\}$ has the additional property that

$$t \mapsto e^{tB}$$

is continuous with respect to operator norms. In fact, by v the triangle inequality the reader can check that each pair $0 \leq t < s$ gives

$$(4.2) \qquad ||e^{sB} - e^{tB}|| = ||e^{tB}|| \cdot ||e^{(s-t)B} - E|| \leq ||e^{tB}|| \cdot ||B|| \cdot ||e^{(s-t)||B||} \cdot (s-t)$$

4.3 A converse result. A semi-group $\{T_t\}$ is uniformly continuous if

$$\text{(i)} \qquad \lim_{t \to 0} |||T_t - E|| = 0$$

When (i) holds the operators T_t are invertible when $t \simeq 0$ and if h_* is chosen so that

$$||T_t - E|| \leq 1/2 \quad : 0 \leq t \leq h_*$$

there exist 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

$$\frac{T_h - E}{h} = \frac{S_h}{h} + \sum_{n=2}^{\infty} \frac{h^{n-1}}{n!} \cdot S_h^n$$

4.4 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

$$\frac{S_{Nh}}{N} = S_h \tag{4.4.1}$$

Next, $h \simeq 0$ we choose the largest positive integer N_h such that $N_h \cdot h \leq h_*$. So here

$$\frac{h_*}{N_h + 1} < h \leq \frac{h_*}{N_h}$$

With this choice of N_h we apply (4.4.1) and get

$$\frac{S_h}{h} = \frac{1}{N_h \cdot h} \cdot S_{N_h \cdot h}$$

Passing to the imit as $h \to 0$ the reader can check the limit equation

$$\lim_{h \to 0} \frac{S_h}{h} = h_* \cdot S(h_*)$$

4.5 Exercise. Show from the above that

$$\lim_{h \to 0} \frac{T_h - E}{h} = h_* \cdot S(h_*)$$

which means that the infinitesmal generator of $\{T_t\}$ is given by the bounded operator $h_* \cdot S(h_*)$.

4.6 Conclusion. There exists a 1-1 correspondence between uniformly continuous semi-groups and bounded linear operators on X.

5. Proofs in 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

$$\frac{T_h(x_\delta) - x_\delta}{h} = \frac{1}{h} \cdot \int_\delta^{\delta + h} T_t(x)\, dt$$

By strong contoinuity 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 a fixed 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

(5.2.2) $$S_\mu(t) - S_\lambda(t) = t \cdot \int_\lambda^\mu (\xi \cdot R(\xi)^2 - R(\xi)) S_\xi(t)\, ds \quad : \mu > \lambda \geq a^*$$

Next, we notice that (5.2.0) gives

(5.2.3) $$\lim_{\xi \to \infty} ||(\xi \cdot R(\xi)^2 - R(\xi)|| = 0$$

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

(5.2.4) $$||S_\mu(t) - S_\lambda(t)|| < \epsilon \quad : \mu \geq \lambda \geq \xi^*$$

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

(iii) $$\frac{d}{d\lambda}(B_\lambda) = R(\lambda) - E - \lambda \cdot R(\lambda)^2$$

The infinitesmal generator of $\{S(t)\}$. Returning to the series (5.2.1) it is clear that

(5.2.6) $$\frac{d}{dt}(S_\lambda(t) = S_\lambda(t) \cdot B_\lambda$$

An integration and the equality $S_\lambda(0) = E$ give for every $t > 0$ and each vector $x \in X$:

(5.2.7) $$S_\lambda(t)(x) - x = \int_0^t S_\lambda(\xi) \circ B_\lambda(\xi)(x)\, d\xi$$

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 inifinitesmal 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 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.

Special Chapters

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 is surjective so that $X = X^{**}$ 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 in X is relatively compact in Y. In applications one often encounters compact operators. The case when the target space $Y = C^0(S)$ for a compact metric space is of special interest. Let B denote the unit ball in X. A classic result due to Arzela and Vitali asserts that $T(B)$ is relatively compact if and only if this family is equi-continuous, i.e. to each $\epsilon > 0$ there exists $\delta > 0$ such that

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

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 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. With $Y = C^0(S)$ we recall that Riesz' representation theorem gives the equality $Y^* = \mathfrak{M}(S)$ where $\mathfrak{M}(S)$ is the space of Riesz measures on S with a finite total mass. Let us then consider a sequence $\{\mu_n\}$ in the unit ball of $\mathfrak{M}(S)$ which converges in the weak star topology to a limit measure μ. When T^* is compact there exists at least one subsequence $\{\mu_{n_k}\}$ such that $\{T(\mu_{n_k})\}$ converges in the norm topology to some $x^* \in X^*$. We leave as an exercise to the reader to show that x^* is independentof the chosen subsequence. In other words, one has the following:

Theorem. *Let $T\colon X \to C^0(S)$ be a compact operator and $\{\mu_n\}$ a sequence in the unit ball of $\mathfrak{M}(S)$ which converges in the weak star topology to a limit measure μ. Then it follows that*

$$\lim_{n\to\infty} ||T^*(\mu_n) - T^*(\mu)|| = 0$$

The result as above illustrates why general theory has a great merit since we can apply the theorem in many different situations.

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

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

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:

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

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

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

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.

1. Normed spaces.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Proof. The verification of the triangle inequality:

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

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

1.5 A converse. Let $||\cdot||$ be a norm on Y which gives 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 (i-ii) above are equivalent if they define the same norm. Starting from this norm we get K_* and K^* and then the reader may verify that

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

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

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

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

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

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

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

$$||y||_e = \sqrt{\sum_{\nu=1}^{\nu=n} |a_\nu|^2} \quad : \quad y = a_1e_1 + \ldots + a_ne_n \tag{i}$$

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

$$||y||^* = \sum_{\nu=1}^{\nu=n} |a_\nu| \quad : \quad y = a_1e_1 + \ldots + a_ne_n \tag{ii}$$

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

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

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

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

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

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

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

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

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

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

This inclusion gives

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

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

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

1.8 The complex case. If X is a complex vector space we obtain complex norms 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.

1.9 Non-linear convexity.

Let $f(x)$ be a real-valued function in $\mathbf{R}^n$ of class C^2. To every point x we assign the Hessian $H_f(x)$ which is the symmetric matrix 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) $$\lim_{|x|\to+\infty} f(x) = +\infty$$

Under these conditions one has the results below:

1.10 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 which is due to Lagrange and Legendre.

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

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

For each fixed ω the reader should verify that this yields a strictly increasing function of s and we leave it to the reader to plot level sets $\{f = a\}$ when $a > 0$ for some different f-functions as above.

1.11 On 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.

1.12 Exercise. Show that a cone Γ is proper if and only if $\widehat{\Gamma}$ is fat.

1.13 Dual cones. The *dual cone* is defined by

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

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

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

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

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

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

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.

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

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

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

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

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

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

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

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

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

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

2.3 Exercise. Prove Theorem 2.1 using Baire's result.

2.4 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:

2.5 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*

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

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

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

Above X and Y are not necessarily Banach spaces. But one verifies easily that if $\widehat{X}$ 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 $\widehat{Y}$. 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 addition, $T(X)$is equipped with the norm defined by

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

It is clear that

$$||y||_* \leq ||\bar{T}|| \cdot ||y|| = ||T||| \cdot ||y||$$

Ths, up to a constant the complete norm majorises then norm in (xx). The Banach-Steinhaus theorem applies and gives a constant C such that

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

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

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

Remark. The inequality (*) implies that if $\epsilon > 0$ and $B_X(\epsilon) = \{||x|| < \epsilon\}$, then $T(B_X(\epsilon)$ is an open neighborhood of the origin in the Banach space $T(X)$. Thus, (*) gives the so called *Open Mapping Theorem.*

3.4 The closed graph theorem Let X and Y be Banach spaces and let T be 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

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

3.5 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. Notice that

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

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

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

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

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

Theorem 3.3 XX applies to $\mathfrak{i}$ and proves that the inverse map is continuos. This gives a constant C such that

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

By the observation in XX this implies that T is bounded with an operator norm $\leq C$.

3.6 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 also is closed.

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

3.8 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.

3.9 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

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

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 and was used by Weyl in pioneering work which for example gave existence theorems for elliptic boundary value problems. In later work the systematic use of "abstract functional analysis" has been used with great succes. A result of this nature is *Gårding's inequality* established by Lars Gårding in [Gå] and later extended to the so called sharp Gårding inequality by L. Hörmander in [Hö]. This illustrates the usefulness of functional analysis, though one must not forget that delicate parts in the proofs rely upon "hard analysis". See also § xx for further comments related to PDE-theory.

4. Hilbert spaces.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

$$UAU^* = \Lambda$$

where Λ is a diagonal matrix whose elements are real.

4.7 The passage to infinite dimension.

Around 1900 the need for a spectral theorem in infinite dimensions became urgent. In his article *Sur une nouvelle méthode pour la resolution du problème de Dirichlet* from 1900, Ivar Fredholm extended earlier construction by Volterra and 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

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

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

Remark. If V is a "abstract" Hilbert space the restriction of the norm to any finite dimensional subspace W is determined by an ellipsoid and exactly as in linear algebra one constructs an orthonormal basis on W. 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

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

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

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

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

4.11 General results on Hilbert spaces.

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

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

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

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

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

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

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

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

By the identity (*) the right hand side is

$$\text{(iii)} \qquad 2||\xi - k_n||^2 + 2||\xi - k_m||^2 - ||k_n - k_m||^2$$

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

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

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

(i)
$$|\lambda_1| \leq \sqrt{\mu_1}$$

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

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

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

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

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*

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

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

Now one has

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

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

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

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

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:*

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

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

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

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.

5. Dual vector spaces

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

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

The set of these continuous linear forms is denoted by X^*. It is obvious that X^* is a vector space and the smallest constant C above is the norm of γ. In this way X^* is a normed space. Since Cauchy-sequences of complex numbers converge it follows easily that X^* is a Banach space. Notice that this is true even if X from the start is not complete. 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 tht one has a canonical isomorphism

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

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.

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

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

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

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

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

5.3 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^*.

5.3.1 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

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

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

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

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 Weak Convergence.

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

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

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

Let k be the dimension of the vector space generated by $x_1, \ldots, x_n$. By Linear Algebra the kernel of $\mathfrak{res}_Y$) has codimension k in X. So the U-set in (*) contains a subspace of X^* with finite codimension, i.e. the open U-sets in (*) are quite large when X has inifinite dimension. One refers to the *weak-star* topology on X^*.

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

$$(5.5.1) \qquad d(\gamma_1, \gamma_2) = \sum_{n=1}^\infty 2^{-n} \cdot 2^{-n} \cdot \frac{|\gamma_1(x_n) - \gamma_2(x_n)|}{1 + |\gamma_1(x_n) - \gamma_2(x_n)|}$$

Exercise. Verify that the weak-start topology induces a topology $B(X^*)$ which is equal to the topology defined by the metric above. So in particular the topology defined by the metric in (5.5.1) does not depend upon the chosen dense subsequence in X.

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

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

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

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

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

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

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

This proves that the given γ-sequence contains a convergent subsequence. So by the definition of compact metric spaces Theorem 5.6 follows.

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

5.8 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.

5.9 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 $x \in X$.

5.10 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

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

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.

5.11 The weak topology on X

.

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

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

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

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

5.12 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.

5.13 Remark. If X is an arbitrary infinite dimensional Banach space 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 it is clear that 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.

5.14 The bidual X^{**} and reflexive spaces. A *reflexive* Banach space X is reflexive if the canonical map from X into its bi-dual X^{**} is surjective. For such spaces one often profits upon the existence of weak convergence. In general, let X be a Banach space and A is some bounded subsets. One says that A is *weakly sequentially compact* if every sequence $x_1, x-2, \ldots$ of points from A contains at least one subsequence $\{x_{n_k}\}$ such that $\{\xi_k\}$ converges weakly to a limit x_*, i.e.

$$\lim_{k \to \infty} \gamma(x_{n_k}) = \gamma(x_*)$$

hold for every $\gamma \in X^*$. In a reflexive space every bounded subset is weakly sequentially compact. This is often used to get certain existence results when the reflexive condition holds. A remarkable result due to Eberlein and Smulian goes as follows:

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

Remark. For the proof we refer to [D-S. Vol 1: page 490-492]. The theorem above is non-trivial since general toplogy gives examples of spaces Y which are non-compact while every denumerable sequence in Y has a convergent subsequence. The main difficulty in the proof of Theorem 5.13 appears when the weak topology on X is not metrisable which occurs when X^* is non-separable. The Krein-Smulian result is for example very useful when we take the Banach space $L^1(\mu)$ of integrable functions with respect to some σ-additive measure μ.

5.14 A theorem by Pietsch

.

We prove a result which will be used in § 11.xx in relation to ergodic theorems. Let T be a bounded linear operator on a Banach space X and $\{p_n(\lambda)\}$ a sequence of polynomial with complex coefficients which give the linear operators

$$A_n = p_n(T)$$

Assume that $p_n(1) = 1$ for every n and impose the following two extra conditions for each vector $x \in X$ where E denotes the identity operator on X:

(i) $$A_n(E - T)(x) \xrightarrow{w} 0$$

(ii) $$\{A_n(x)\} \quad \text{is a weakly compact sequence}$$

In X we define the closed subspaces:

(iii) $$X_1 = \overline{(E - T)(X)} \quad \text{and} \quad X_2 = \text{Ker}(E - T)$$

Theorem. *When (i-ii) hold it follows that $\{A_n\}$ converge weakly to a bounded operator B where is a projection whose kernel is X_1 while B restricts to the identity on X_2 and one has the direct sum decomposition*

$$X = X_1 \oplus X_2$$

Proof. To begin with the result in § XX and the hypothesis (ii) give a constant M such that the operator norms

(1) $$||A_n|| \leq M$$

for all n. Next, if $\xi = (E - T)(x) = x - T(x)$ for some $x \in X$, then (i) entails that

(2) $$\lim_{n\to\infty} A_n(\xi) \xrightarrow{w} 0$$

Together with (1) (2) also holds for each vector in the closure of $(E-T)(X)$, i.e. for each $\xi \in X_1$. Next, we prove that there exists a bounded operator B such that

(3) $$\lim_{n\to\infty} A_n(x) \xrightarrow{w} Bx$$

hold for every $x \in X$. To begin with (1) implies that if the sequenes $\{A_n(x)\}$ converge weakly for every x, then we get a bounded linear operator B in (3). To prove that these weak limits exist we argue by a contradiction, i.e. let $x \in X$ and suppose there are two subsequences of integers $\{n_k\}$ and $\{m_k\}$ such that

$$\lim_{k\to\infty} A_{n_k}(x) = \xi \quad : \lim_{k\to\infty} A_{m_k}(x) = \eta$$

where $\xi \neq \eta$. At the same time (2) entails that

$$0 = \lim_{k\to\infty} (E - T)A_{n_k}(x) = \xi - T(\xi)$$

and in the same way we get $T(\eta) = \eta$. In particular

$$T(\xi - \eta) = \xi - \eta$$

Since $p_n(1) = 1$ for every n this entails that

$$A_n(\xi - \eta) = \xi - \eta \quad : n = 1, 2, \ldots$$

In particular the non-zero vector $\xi - \eta$ does not belong to the closed subspace X_1. The Hahn-Banach theorem gives some $x^* \in X^*$ such that

(4) $$x^* \in X_1^\perp \quad : x^*(\xi - \eta) \neq 0$$

Consider the adjoint operator T^*. The inclusion $x^* \in X_1 \perp$ means that

(5) $$x^*(x - Tx) = x^*(x) - T^*x^*(x) = 0$$

hold for every $x \in X$ and hence $T^*(x^*) = x^*$. Next, passing the adjoints of $\{A_n\}$ it follows that

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

where we again used that $p_n(1) = 1$ for every n, This gives

$$x^*(\xi) = \lim x^*(A_{n_k}(x)) = \lim A_{n_k}^*(x^*)(x) = x^*(x) \tag{6}$$

in the same way we find that $x^*(\eta) = x^*(x)$ which would give $x^*(\xi - \eta) = 0$ and contradict (4) above.

We have proved (2) which give the bounded operator B. If $\xi = (E - T)(x)$ is vector in the range of $E - T$, then (i) entails that $A_n(\xi) \xrightarrow{w} 0$ which gives $B(\xi) = 0$. Passing to the closure we conclude that

$$X_1 \subset \mathrm{Ker}(B) \tag{7}$$

Next, if $x^* \in X_1^\perp$ we have $T^*(x^*) = x^*$ which entails that $A_n^*(x^*) = x^*$ for every n. It follows that

$$B^*(x^*) = x^*$$

i.e. the adjoint operator B^* restricts to the identity on the subspace $X_1^\perp$ in the dual space X^*. Together with (7) this gives a direct sum decomposition

$$X = X_1 \oplus X_2 \tag{8}$$

where B restricts to the identity on X_2. Thus, if $x_2 \in X_2$ we have $B(x_2) = x_2$. At the same time

$$(E - T)(A_n(x_2)) \xrightarrow{w} 0$$

which entals that

$$(E - T)(Bx_2) = x_2 - T(x_2) = 0$$

Hence X_2 is the kernel of $E - T$ which together with the direct sum in (8) finishes the proof of the theorem.

5.B Locally convex spaces

Let E be a real vector space. Denote by $\mathcal{C}_E$ the family of convex and symmetric subsets U which contain the origin and are absorbing in the sense that for every $x \in E$ there exists some real $s > 0$ such that $s \cdot \in U$. The symmetry means that $U = -U$. If $U \in \mathcal{C}_E$ and x is a vector in E it may occur that the whole line $\mathbf{R}x$ is contained in U in which case we say that x is fully absorbed by U. The convexity of U entails that the set of absorbed vectors is a linear subspace of E which we denote by $\mathcal{L}_U$.

5.B.1 The pseudo-norm ρ_U. For each non-absorbed vector x we get a positive 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

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

We leave it to the reader to verify that the convexity of U entails that ρ_U satisfies the triangle inequlity

(i) $$\rho_U(x_1 + x_2) \leq \rho_U(x_1) + \rho_U(x_2)$$

for every pair in X. Moreover, $\rho_U(x) = 0$ if and only if x beongs to $\mathcal{L}_U$ and ρ_U is positively homogeneous, i.e. the equality below holds when a is real and positive:

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

Finally, the symmetry of U entails that $\rho_U(x) = \rho_U(-x)$ for every x.

Exercise. Conversely, let $\rho \colon X \to \mathbf{R}^+$ satisfiy (i-ii) and the symmetry $\rho(x) = \rho(-x)$. Set

$$U = \{x \colon\ \rho(x) < 1\}$$

Show that U belongs to $\mathcal{C}_E$ and that $\rho_U = \rho$.

5.B.2 Locally convex topologies. A topology on the real vector space X is said to be locally convex and separated if there exists a family $\mathfrak{U} = \{U_\alpha\}$ in $\mathcal{C}_E$ such that the intersection $\mathcal{L}_{U_\alpha}\}$ is reduced to the zero vector. A bais for open neighborhoods of the origin in X consists of sets defined by

$$\{x \colon \rho_{U_i}(x) < \epsilon\}$$

where $\epsilon > 0$ and $\{\alpha_1, \ldots, \alpha_k\}$ is a finite set of indices defining the U-family.

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:

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

5.B.3 The dual space E^*. Let E be equipped with a locally convex $\mathfrak{U}$-topology. A linear functional ϕ is $\mathfrak{U}$-continuous if and only if there exists some $U \in \mathfrak{U}$ such that ϕ is continuous with respect to the ρ_U-norm, i.e there is a constant C such that

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

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

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

The next result is due to Dieudonné from the article *La dualité dans les espaces vectorielles* (Ann.Ecol. Norm.Sup. 1942).

5.B.4. Theorem. *Each closed convex set K in E is the intersection of closed half-spaces.*

Tbe proof is left as an exercise. A hint is that K to begin with is closed with respect to the ρ_U-toplogy for every U which this gives a real number a_U such that

$$K + \mathcal{L}_U = \{\rho_U \leq a_U\}$$

5.C Support functions of convex sets.

From now on E is a locally convex space. Denote by $\mathcal{S}(E)$ the family of real-valued functions G on E^* which are subadditive and positively homogeneous, i.e. for each real $t > 0$ and all pairs y_1, y_2 in E^* one has:

$$(*) \qquad G(y_1 + y_2) \leq G(y_1) + G(y_2) \quad : \quad G(ty) = tG(y)$$

To each closed and convex subset K we define a function $\mathcal{H}_K$ on the dual E^* by:

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

Notice that these $\mathcal{H}$-functions take values in $(-\infty, +\infty]$, i.e. it may be $+\infty$ for some vectors $y \in E^*$.

5.C.1 Exercise. Show that $\mathcal{H}_K \in \mathcal{S}(E)$ for every closed convex set K.

Next, in the weak-star topology on E^* a basis for open neighborhoods of the origin consists of open sets

$$U(x_1, \ldots, x_N; \epsilon) = \{y \in E^* : \, |y(x_k)| < \epsilon : \, 1 \leq k \leq N\}$$

where $x_1, \ldots, x_N$ are finite subsets of E and $\epsilon > 0$. Dieudonné's theorem shows that if K is a closed convex set then H_K is an upper semi-continuous function with respect to the weak-star topology. Indeed, this is so because the $y \mapsto y(x)$ is weak-star continuous for every fixed $x \in E$ and the supremum of an arbitrary family of weak-star continuoes functions is upper semi-continuous. Less obvious is the converse which 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].

5.C.2. Theorem. *Each $G \in \mathcal{S}(E)$ which is upper semi-continuous with respect to the weak-star topology is of the form $\mathcal{H}_K$ for a unique closed convex subset K in E.*

Remark. As pointed out by Hörmander in [ibid] Theorem 5.C.2 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 5.C.2 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 5.C.2 Put $F = E \oplus \mathbf{R}$ which is a new vector space where the 1-dimensional real line is added. Here $F^* = E^* \oplus \mathbf{R}^*$ and for each vector (y, η) in F^* we declare that

$$(1) \qquad (y \oplus t)(x, t) = y(x) - \eta t$$

Notice the minus sign in the right hand side. Next, each closed and convex set K in E gives the convex cone in F defined by

$$(1) \qquad C_K = \{(tx, t) : \, t \geq 0\}$$

It is in general not closed but is closure is again a convex cone denoted by $\overline{C_K}$ and the reader may notice that

$$\overline{C_K} \setminus C_K \subset E \oplus \{0\}$$

Exercise. Show that $K \mapsto \overline{C_K}$ yields 1-1 correpondence between closed convex sets K in E which contain at least one point outside the origin and the family of closed cones in F contained in $\{t \geq 0\}$ but not entirely in the hyperplane $\{t = 0\}$.

Next, when $K \subset E$ is a closed convex set we get a cone in F^* defined by:

$$(2) \qquad C_K^* = \{(y, t) : \, y(x) - t \geq 0 : \, (x, t) \in \bar{C}_K \quad \text{and} \quad s \geq 0\}$$

This means that $(y,s) \in C_K^*$ if and only if $y(x) \leq t$ for each $(x,t) \in \bar{C}_K$, or equivalently that $\mathcal{H}_K(y) \leq t$.

After these preliminary observations we begin to study a function G on E^* which satisfies the conditions in Theorem 5.B.2 Put

(i) $$G_* = \{(y,\eta) \in F^* \quad : , G(y) \leq \eta\}$$

Condition in (*) entails that G_H 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^*. Since $G(0) = 0$ we have

(ii) $$(0,\eta) \in G_* \implies \eta \geq 0$$

Next, in F we define the cone

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

Since $\eta \geq 0$ for every vector $(y,\eta) \in G_*$ we see that G_{**} is contained in the half-space $t \geq 0$. Moreover it is obviously closed in F Next, let $\widehat{C}$ be the cone in F^* defined by

(iii) $$\widehat{C} = \{(y,\eta) :\ y(x) - \eta t \leq 0 :\ (x,t) \in G_{**}\}$$

It is clear that $G_* \subset \widehat{C}$ and if equality holds the construction G_* in (i) and the remarks at the beginning of the proof entail that $G = \mathcal{H}_K$ for closed convex cone K in E. Hence there only remains to prove that inclusion below is not strict:

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

To prove this we use Theorem 5.B.0. Namely, since the two sets above 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 must have $\alpha \leq 0$. and since it also is a cone the implication above gives $(x_*, t_*) \in G_{**}$ and then 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 the proof of Theorem 5.C.2 is finished.

5.C.3 The case of normed spaces. If X is a normed vector space Theorem 5.C.2 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

$$U_\epsilon(K_0) = \{K \in \mathcal{K} :\ \text{dist}(K, K_0) < \epsilon\}$$

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.

5.C.4 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.*

5.C.5 Exercise. Deduce this result from Theorem C.5.2.

5.D The Krein-Smulian theorem.

Let X be a Banach space and X^* its dual. The weak-star topology on X^* was defined in § 5.4. We have also the bounded weak-star topology desrcibed in § xx from the introduction. So now we have a pair of locally convex spaces X_w^* and X_{bw}^*. Open sets in the weak-star topology are by defintion also open in the bounded weak-star topology, i.e. the latter topology contains more open sets and is therefore strongethan the ordinary weak-star topology. Hence there is a natural inclusion of dual spaces:

(*) $$(X_w^*)^* \subset (X_{bw}^*)^*$$

The Krein-Smulian theoren assers that eqaulity holds in (*). To prove this we first study the bounded weak-star topology on X^*. In general, if A is a finte set in X we put

$$A^0 = \{x^* : \max_{x \in A} |x^*(x)| \leq 1\}$$

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

(i) $$S^* \cap A_1^0 \subset U$$

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

(ii) $$nS^* \cap A_n^0 \subset U$$

To each finite set B of vectors in X with norm $\leq n^{-1}$ we notice that $(A_n \cup B)^0$ contains A_n^0. Let $U^c = X^* \setminus U$ and put

$$F(B) = (n+1)S^* \cap (A_n \cup B)^0 \cap U^c$$

Each such set if weak-star closed. If they are non.empty for every finite set B as above, it follows by the weak-star compactness of $(n+1)S^*$ that the whole intersection is non-empty. So we find x^* in this intersection. In particular $x^* \in B^0$ for every finite set of vectors in X of norm $\leq n^{-1}$. It means that $|x^*(x)| \leq 1$ when $||x| \leq 1$ and hence the norm $||x^*|| \leq n$. This gives the inclusion

(ii) $$x^* \in nS^* \cap A : n^0 \cap U^c$$

But this contradicts (ii) and hence $F(B) = \emptyset$ for some finite set B if X-vectors of norm $\leq n^1$.

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

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

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

Proof 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 iof finite sets by $x_1, x_2, \ldots$, i.e. start with the finite string of vectors in A_1, and so on. In particular $||x_n|| \to 0$ as $n \to \infty$. If x^* is a vector in X^* we associate the complex sequence

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

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

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

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

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

we have by defintion $x^* \in A_n^0$ for each n. Choose n so large that $||x^*|| \leq n$. Thus entails that

$$x^* \in nS^* \cap A_n^0$$

and 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 ρ with norm one at most. The Hahn-Banach theoren gives $\lambda \in \mathbf{c}_0^*$ of norm one at most such that

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

Next, a wellknown result due to Banach gives 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.E 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:

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

$$\text{(i)} \qquad 1 = |x + a(x - \phi(x)|^2 = |x|^2 + 2a(1 - \langle x, \phi(x)\rangle) + a^2|x - \phi(x)|^2$$

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

$$\text{(E.1)} \qquad a(x) = 0 \quad : x \in S$$

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

$$\text{(ii)} \qquad \frac{\partial f_i}{\partial x_k} = e_{ik} + t\frac{\partial g_i}{\partial x_k}$$

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

$$\text{(iii)} \qquad J(t) = \int_B D(x;t)\,dx$$

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

$$\text{(E.2)} \qquad \frac{d}{dt}(D(x;t) = \sum\sum (-1)^{j+k} \cdot \frac{\partial g_i}{\partial x_k}$$

where the double sum extends over all pairs $\leq j, k \leq n$.

Next, for all pairs $i.k$, Stokes theorem gives

$$\text{(iv)} \qquad \int_B \frac{\partial g_i}{\partial x_k}\,dx = \int_S g_i \cdot \mathbf{n}_k\,d\omega$$

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.

The Hilbert cube $\mathcal{C}$. In the Hilbert space ℓ^2 we have the closed subset $\mathcal{C}$ which consists of vectors $x = (x_1, x_2, \ldots)$ such that $|x_k| \leq 1/k$ for each k.

2. Proposition. *Every closed and convex subset of $\mathcal{C}$ has the fixed point property.*

Exercise. Deduce this result from Theorem 1.

Next, let X be a locally convex vector space and X^* its dual. Consider the family $\mathcal{K}(X)$ of convex subsets which are compact with respect to the weak topology. Let $K \in \mathcal{K}(X)$ and $T \colon K \to K$ a continuous map in 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:

3. 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.*

4. 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 3 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-deterministic 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. 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.

6. 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:

6.1 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 $T(K_0) \subset K_0$.*

7. 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.F 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

(i) $$(K_* - K_*) \setminus \overline{V} \neq \emptyset$$

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

(ii) $$\mathbf{G}(k_1 - k_2) \subset V$$

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

(iii) $$\mathbf{G}\overline{U^*}) = \overline{U^*}$$

We find the unique positive number δ such that the following hold for every $\epsilon > 0$:

(iv) $$K_* - K_* \subset (1+\epsilon) \cdot U^* \quad : \ (K_* - K_*) \setminus (1-\epsilon) \cdot \delta \cdot \overline{U^*} \neq \emptyset$$

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

(v) $$K_* \subset \ \cup (k_\nu + \frac{\delta}{2} \cdot U)$$

Put

(vi) $$K_{**} = K_* \cap \bigcup_{k \in K_*} (k + (1 - 1/4n)\delta \cdot \overline{U})$$

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

(v) $$k_1 \in (k_2 + (1-1/4n)\delta \cdot \overline{U}) \implies k_1 - k_2 \in (1-1/4n)\delta \cdot \overline{U}$$

which cannot hold since $1 - 1/4n < 1 - \epsilon$. The proof of Kakutani's theorem is therefore finished if we have shown that

(vi) $$p \in K_{**}$$

To see this we take an arbitrary $k \in K_*$. From (v) we find some $1 \leq i \leq n$ such that

(vii) $$k_i - k \in \frac{\delta}{2} \cdot U$$

Without loss of generality we can assume that $i = 1$ and get a vector $u \in U$ such that

(viii) $$k_1 = k + \frac{\delta}{2} \cdot u$$

It follows that

(ix) $$p = \frac{k_1 + \ldots + k_n}{n} = k + \frac{\delta}{2n} \cdot u + \sum_{i=2}^{i=n} \frac{1}{n}(k_i - k)$$

Next, for each $\epsilon > 0$ the left hand inclusion in (iv) and the convexity of U give

(x) $$\sum_{i=2}^{i=n} \frac{1}{n}(k_i - k) \subset \frac{n-1}{n}(1+\epsilon) \cdot \delta \cdot U$$

It follows that

(xi) $$\frac{\delta}{2n} \cdot u + \sum_{i=2}^{i=n} \frac{1}{n}(k_i - k) \in (\frac{n-1}{n}(1+\epsilon)\delta + \frac{\delta}{2n}) \cdot U$$

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.

6. Fredholm theory.

Introduction. A major result in this section is Theorem 6.6.1 which is due to Fredholm in the article [Fredholm: 1901]. The original result was stated for integral operators but the proof adapted to bounded linear operators on Banach spaces is verbatim the same. 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 $\bar{u}$. Let $u\colon X \to Y$ be a bounded linear operator which gives the closed null space N_u in X. On the Banach space $\frac{X}{N_u}$ there exists 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 If $x \in N_u$. 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 hence assume that $u\colon X \to Y$ from the start is injective. It means that $N_u^\perp = Y^*$ so there remains to show that u^* is surjective, i.e. that

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

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

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

If 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.

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:

6.4.2 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 the image

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

Let us first notice that $V \neq X^*$. To see this 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 $V \cap nS^*$ where $\xi_n \xrightarrow{w} x^*$ 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 $\{y\}^\perp_w$ to a limit vector y^*. If $x \in X$ we get

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

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 a non-zero vector $x \in X$ such that $\widehat{x}(V) = 0$ which gives the implication

(i) $$y^*(y) = 0 \implies \widehat{x}(T^*(y^*)) = y^*(Tx) = 0 \quad : \quad y^* \in Y^*$$

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

6.4.3 Operators with finite dimensional range. The range of u is the image space $u(X)$. Suppose the range 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 $u(X)$. Consider the N-dimensional subspace of X :

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

The u-kernel is given by:

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

It is clear that there exists a direct sum:

$$X = N_u \oplus V$$

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

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

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

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

and conclude that the range of u^* is N-dimensional.

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 postive 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

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

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

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

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:

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

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

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

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

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

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

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

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

6.6 Compact pertubations.

The main result goes as follows:

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.*

Before we enter the proof we recal from § xx that the hypothesis on u via the Open Mppaing Theorem gives $c > 0$ such that

$$\text{(i)} \qquad ||u(x)|| \geq c \cdot ||x||$$

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

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

Let $\{x_n\}$ be a sequence in B. 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 then (i) entails that $\{\xi_j\}$ is a Cauchy sequence and hence has a limit vector. This proves that B 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.* Next, let y be in the closure of $\mathrm{Im}(u+T)$. It means that there is a sequence ξ_n in X such that

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

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

(ii) $$\lim\, u(x_n^*) + T(x_n^*) = 0$$

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

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

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

Hence y belongs to $\mathrm{Im}(u+T)$ 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

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

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

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

At the same time the separation gives

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

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. Ths 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:

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

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 $T_*|colonX_* \to Y$ 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

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

Remark. There are more "sophisticated proofs" which employ suitable diagram chasings and Theorem xx can be extended to compact pertubations of Fredholm operators between Frechet spaces. For elegant proofs and conclusive results the reader may consult the lecture by Serre from the Cartan seminars. See [Cartan-xxxx].

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

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

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

$$y^* \mapsto \chi(y^*) \tag{2}$$

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

$$||\chi(y^*)|| \leq C \cdot ||y^*|| \tag{3}$$

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

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

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

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

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

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

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

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

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

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

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

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

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

Let $f(z)$ be an analytic function in Ω which extends to a continuous function on $\bar{\Omega}$. 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. Bounded self-adjoint operators.

Introduction. Let $\mathcal{H}$ be a complex Hilbert space. A bounded linear operator S on $\mathcal{H}$ is self-adjoint if $S = S^*$, or equivalently

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

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

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

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

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

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

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

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

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

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

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

Next, we consider the spectrum of self-adjoint operators.

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

Proof. Let λ be a complex number and for a given x we set $y = \lambda x - Sx$. It follows that

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

Since S is self-adjoint we get

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

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

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

With $\lambda = a + ib$ the right hand side becomes

$$b^2||x||^2 + a^2||x||^2 + ||Sx||^2 - 2a \cdot ||Sx|| \cdot ||x|| \geq b^2||x||^2$$

Hence we have proved that

(ii) $$||\lambda x - Sx||^2 \geq (\mathfrak{Im}\,\lambda)^2 \cdot ||x||^2$$

This implies that $\lambda E - S||$ is invertible for every non-real λ which proves Theorem 8.1.

Theorem 8.1 together with the general result from 7.XX gives the following: Denote by $\mathbf{S}$ the closed subalgebra of $L(\mathcal{H}, \mathcal{H})$ generated by S and the identity operator. Then $\mathbf{S}$ is a sup-norm algebra which is isomorphic to the sup-norm algebra $C^0(\sigma(S))$.

8.2 Exercise. Let T be an arbitary bounded opertor on $\mathcal{H}$. Show that the operator $A = T^*T$ is self-adjoint and that $\sigma(A)$ is a compact subset of $[0, +\infty)$, i.e. every point in its spectrum os real and non-negative. A hint is to use the biduality formula $T = T^{**}$ and if s is real the reader should verify that

$$||sx + T^*Tx||^2 = s^2||x||^2 + 2s \cdot ||Tx||^2 + ||T^*Tx||^2$$

8.3 Normal operators.

A bounded linear operator A is normal if it commutes with its adjoint A^*. Let A be normal and put $S = A^*A$ which yields a self-adjoint by Exercise 8.2. Here (8.0.1) gives

(1) $$||S||^2 = ||S^2|| = ||A^2 \cdot A(^*)^2|| \leq ||A^2||\ddot{|}|(A^*)^2||$$

where we used the multiplicative inequality for operator norms. Now $(A^*)^2$ is the adjoint of A^2 and we recall from § xx that thr norms of an operator and its adjoint are equal. Hence the right hand side in (1) is equal to $||A^2||^2$. At the same time

$$||S|| = ||A^*A|| = ||A||^2$$

and we conclude that (1) gives

(3) $$||A||^2 \leq ||A^2||$$

Exactly as in the self-adjoint case we can take higher powers and obtain the equality

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

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

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

Remark. The spectrum $\sigma(A)$ is some compact subset of $\mathbf{C}$ and in general analytic polynomials restricted to $\sigma(A)$ do not generate a dense subalgebra of $C^0(\sigma(A))$. To get a more extensive algebra we consider the closed subalgebra $\mathcal{B}$ of $L(\mathcal{H}, \mathcal{H})$ which is generated by A and A^*. Since every polynomial in A and A^* again is a normal operator it follows that $\mathcal{B}$ is a sup-norm algebra and here the following holds:

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

Proof. If $S \in \mathcal{B}$ is self-adjoint then we know from the previous section that its Gelfand transform is real-valued. Next, let $Q \in \mathcal{B}$ be arbitrary. Now $S = Q + Q^*$ is self-adjoint. So if $p \in \mathfrak{M}_\mathcal{B}$ it first follows that the gelfand transform $\widehat{Q}(p) + \hat{Q}^*(p)$ is real, i.e. with $\widehat{Q}(p) = a + ib$ we must have $\hat{Q}^* = a_1 - ib$ for some real number a_1. Next, QQ^* is also self-adjoint and hence $(a + ib)(a_1 - ib)$ is real. This gives $a = a_1$ and hence we have proved that the Gelfand transform of Q^* is the complex conjugate function of $\widehat{Q}$. Hence the Gelfand transforms of $\mathcal{B}$-elements is a self-adjoint algebra and the Stone-Weierstrass theorem implies that the Glefand transforms of $\mathcal{B}$-elements is equal to the whole algebra $C^0(\mathfrak{M}_\mathcal{B}$. Finally, since $\widehat{A}^*$ is the complex conjugate function of $\widehat{A}$ it follows that the Gelfand transform $\widehat{A}$ separates points on $\mathfrak{M}_\mathcal{B}$ which means that this maximal ideal space can be identified with $\sigma(A)$.

8.6 Spectral measures.

Given $\mathcal{B}$ and $\sigma(A)$ as above we construct certain Riesz measures on $\sigma(A)$. Namely, each pair of vectors x, y in $\mathcal{H}$ yields a linear functional on $\mathcal{B}$ defined by

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

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

(8.6 *) $$\langle Tx, y \rangle = \int_{\sigma(A)} \hat{T}(z) \cdot d\mu_{x,y}(z)$$

hold for every $T \in \mathcal{B}$. Since $\widehat{A}(z) = z$ we have

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

Similarly one has

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

8.7 The operators $E(\delta)$. Notice that (8.6 *) implies that the map from $\mathcal{H} \times \mathcal{H}$ into the space of Riesz measures on $\sigma(A)$ is bi-linear. We have for example:

$$\mu_{x_1+x_2,y} = \mu_{x_1,y} + \mu_{x_2,y}$$

Moreover, since $\mathcal{B}$ is the sup-norm algebra $C^0(\sigma(A))$ the total variations of the μ-measures satisfy the equations:

$$||\mu_{x,y}|| \leq \max_{T \in \mathcal{B}_*} |\langle Tx, y\rangle| \tag{8.7.1}$$

where $\mathcal{B}_*$ is the unit ball in $\mathcal{B}$. From this we obtain

$$||\mu_{x,y}|| \leq ||x|| \cdot ||y|| \tag{8.7.2}$$

Next, let δ be a Borel subset of $\sigma(A)$. Keeping y fixed in $\mathcal{H}$ we obtain a linear functional on $\mathcal{H}$ defined by

$$x \mapsto \int_\delta d\mu_{x,y}(z) = \mu_{x,y}(\delta)$$

By (8.7.2) it has norm $\leq ||y||$ and is represented by a vector $E(\delta)x$ in $\mathcal{H}$. More precisely

$$\langle E(\delta)x, y\rangle = \int_\delta d\mu_{x,y}(z) = \mu_{x,y}(\delta) \tag{8.7.3}$$

8.8 Exercise. Show that $x \mapsto E(\delta)x$ is linear and that the resulting linear operator $E(\delta)$ commutes with all operators in $\mathcal{B}$. Moreover, show that it is a self-adjoint projection, i.e.

$$E(\delta)^2 = E\delta) \quad \text{and} \quad E(\delta)^* = E(\delta)$$

Finally, show that

$$E(\delta_1 \cap \delta_2) = E(\delta_1)E(\delta_2)$$

holds for every pair of Borel subsets and with $\delta = \sigma(A)$ one gets the identity operator.

8.9 Resolution of the identity. If $\delta_1, \ldots, \delta_N$ is any finite family of disjoint Borel sets whose union is $\sigma(A)$ then

$$1 = E(\delta_1) + \ldots + E(\delta_N)$$

At the same tine we get a decomposition of the operator A:

$$A = A_1 + \ldots + A_N \quad \text{where} \quad A_k = E(\delta_k) \cdot A$$

For each k the spectrum $\sigma(A_k)$ is equal to the closure of δ_k. So the normal operator is represented by a sum of normal operators where the individual operators have small spectra when the δ-partition is fine.

9. Unbounded operators on Hilbert spaces

Let T be a densely defined linear operator on a complex Hilbert space $\mathcal{H}$. We suppose that T is unbounded so that:

$$\max_{x\in\mathcal{D}_*(T)} ||Tx|| = +\infty \quad \mathcal{D}_*(T) = \text{ the set of unit vectors in } \quad \mathcal{D}(T)$$

9.1 The adjoint T^*. If $y \in \mathcal{H}$ we get a linear functional on $\mathcal{D}(T)$ defined by

(i) $$x \mapsto \langle Tx, y\rangle$$

If there exists a constant $C(y)$ such that the absolute value of (i) is $\leq C(y)\cdot||x||$ for every $x \in \mathcal{D}(T)$, then (i) extends to a continuous linear functional on $\mathcal{H}$. The extension is unique because $\mathcal{D}(T)$ is dense and since $\mathcal{H}$ is self-dual there exists a unique vector T^*y such that

(9.1.1) $$\langle Tx, y\rangle = \langle x, T^*y\rangle \quad : \quad x \in \mathcal{D}(T)$$

This gives a linear operator T^* where $\mathcal{D}(T^*)$ is characterised as above, To describe the graph of T^* we consider the Hilbert space $\mathcal{H}\times\mathcal{H}$ equipped with the inner product

$$\langle (x, y), (x_1, y_1)\rangle = \langle x, x_1\rangle + \langle y, y_1\rangle$$

On $\mathcal{H}\times\mathcal{H}$ we define the linear operator

$$J(x, y) = (-y, x)$$

9.2 Proposition. *For every densely defined operator T one has the equality*

$$\Gamma(T^*) = J(\Gamma(T))^\perp$$

Proof. Let (y, T^*y) be a vector in $\Gamma(T^*)$. If $x \in \mathcal{D}(T)$ the equality (9.1.1) and the construction of J give

$$\langle (y, -Tx\rangle + \langle T^*y, x\rangle = 0$$

This proves that $\Gamma(T^)* \perp J(\Gamma(T))$. Conversely, if $(y, z) \perp J(\Gamma(T))$ we have

(i) $$\langle y, -Tx\rangle + \langle z, x\rangle = 0 \quad : x \in \mathcal{D}(T)$$

This shows that $y \in \mathcal{D}(T^*)$ and $z = T^*y$ which proves Proposition 9.2.

9.3 Consequences. The orthogonal complement of a subspace in a Hilbert space is always closed. Hence Proposition 9.2 entails that T^* has a closed graph. Passing to the closure of $\Gamma(T)$ the decompostion of a Hilbert space into a direct sum of a closed subpace and its orthogonal complement gives

(9.3.1) $$\mathcal{H}\times\mathcal{H} = \overline{J(\Gamma(T))} \oplus \Gamma(T^*)$$

Notice also that

(9.3.2) $$\Gamma(T^*)^\perp = \overline{J(\Gamma(T))}$$

9.4 Closed extensions of operators. A closed operator S is called a closed extension of T if

$$\Gamma(T) \subset \Gamma(S)$$

9.4.1 Exercise. Show that if S is a closed extension of T then

$$S^* = T^*$$

9.4.2 Theorem. *A densely defined operator T has a closed extension if and only if $\mathcal{D}(T^*)$ is dense. Moreover, if T is closed one has the biduality formula $T = T^{**}$.*

Proof. Suppose first that T has a closed extension. If $\mathcal{D}(T^*)$ is not dense there exists a non-zero vector $0 \neq h \perp \mathcal{D}(T^*)$ and (9.3.2) gives

(ii) $$(h, 0) \in \Gamma(T^*)^\perp = J(\Gamma(T))$$

By the construction of J this would give $x \in \mathcal{D}(T)$ such that $(h, 0) = (-Tx, x)$ which cannot hold since this equation first gives $x = 0$ and then $h = T(0) = 0$. Hence closedness of T implies that $\mathcal{D}(T^*)$ is dense. Conversely, assume that $\mathcal{D}(T^*)$ is dense. Starting from T^* we construct its adjoint T^{**} and Proposition 9.3.2 applied with T^* gives

(i) $$\Gamma(T^{**}) = J(\Gamma(T^*)^\perp$$

At the same time $J(\Gamma(T^*)^\perp$ is equal to the closure of $\Gamma(T)$ so (i) gives

(ii) $$\overline{\Gamma(T)} = \Gamma(T^{**})$$

which proves that T^{**} is a closed extension of T.

9.4.3 The biduality formula. Assume that the densely defined operator T is closed. We constrict T^* which by the above is densely defined and therefored its dual exists. It is denoted by T^{**} and called the bi-dual of T. With these notations one has:

(*) $$T = T^{**}$$

9.4.4 Exercise. Prove the equality (*).

9.5 Inverse operators.

.

Denote by $\mathfrak{I}(\mathcal{H})$ the set of closed and densely defined operators T such that T is injective on $\mathcal{D}(T)$ and the range $T(\mathcal{D}(T))$ is dense in $\mathcal{H}$. If $T \in \mathfrak{I}(\mathcal{H})$ there exists the densely defined operator S where $\mathcal{D}(S)$ is the range of T and

$$S(Tx) = x \quad : \quad x \in \mathcal{D}(T)$$

By this construction the range of S is equal to $\mathcal{D}(T)$. Next, on $\mathcal{H} \times \mathcal{H}$ we have the isometry defined by $I(x, y) = (y, x)$, i.e we interchange the pair of vectors. The construction of S gives

(i) $$\Gamma(S) = I(\Gamma(T))$$

Since $\Gamma(T)$ by hypothesis is closed it follows that S has a closed graph and we conclude that $S \in \mathfrak{I}(\mathcal{H})$. Moreover, since I^2 is the identity on $\mathcal{H} \times \mathcal{H}$ we have

(ii) $$\Gamma(T) = I(\Gamma(S))$$

We refer to S is as the inverse of T. It is denoted by T^{-1} and (ii) entails that T is the inverse of T^{-1}, i.e one has

$$T = (T^{-1})^{-1}$$

9.5.1 Exercise. Let T belong to $\mathfrak{I}(\mathcal{H})$. Use the description of $\Gamma(T^*)$ in Proposition 9.3 to show that T^* belongs to $\mathfrak{I}(\mathcal{H})$ and the equality

(*) $$(T^{-1})^* = (T^*)^{-1}$$

9.6 The operator T^*T. Each $h \in \mathcal{H}$ gives the vector $(h, 0)$ in $\mathcal{H} \times \mathcal{H}$ and (9.3.1) gives a pair $x \in \mathcal{D}(T)$ and $y \in \mathcal{D}(T^*)$. such that

$$(h, 0) = (x, Tx) + (-T^*y, y) = (x - T^*y, Tx + y)$$

With $u = -y$ we get $Tx = u$ and obtain

(9.6.1) $$h = x + T^*(Tx)$$

9.6.2 Proposition. *The vector x in (9.6.1) is uniquely determined by h.*

Proof. Uniqueness follows if we show that

$$x + T^*(Tx) \implies x = 0$$

But this is clear since the construction of T^* gives

$$0 = \langle x, x)\rangle + \langle x, T^*(Tx)\rangle = \langle x, x)\rangle + \langle Tx, Tx)\rangle \implies x = 0$$

9.7 The density of $\mathcal{D}(T^*T)$. This is the subspace of $\mathcal{D}(T)$ where the extra condition for a vector $x \in \mathcal{D}(T)$ is that $Tx \in \mathcal{D}(T^*)$. To prove that $\mathcal{D}(T^*T)$ is dense we consider some orthogonal vector h. Proposition 9.6 gives some $x \in \mathcal{D}(T)$ such that $h = x + T^*(Tx)$ and for every $g \in \mathcal{D}(T^*T)$ we have

(i) $$0 = \langle x, g\rangle + \langle T^*Tx, g\rangle = \langle x, g\rangle + \langle Tx, Tg\rangle = \langle x, g\rangle + \langle x, T^*Tg\rangle$$

Here (i) hold for every $g \in \mathcal{D}(T^*T)$ and by another application of Proposition 9.6 we find g so that $x = g + T^*Tg$ and then (i) gives $\langle x, x\rangle = 0$ so that $x = 0$. But then we also have $h = 0$ and the requested density follows.

Conclusion. Set $A = T^*T$. From the above it is densely defined and (9.6.1) entails that the densely defined operator $E + A$ is injective. Moreover, its range is equal to $\mathcal{H}$. Notice that

$$\langle x + Ax, x + Ax\rangle = c + \langle x, Ax\rangle + \langle Ax, x\rangle$$

Here

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

and from this the reader can conclude that

$$||x + Ax||^2 = ||x||^2 + ||Ax||^2 + 2 \cdot ||Tx||^2 \quad : x \in \mathcal{D}(A)$$

The right hand side is $\geq ||x||^2$ which implies that $E + A$ is invertible in Neumann's sense.

9.8 The equality $A^* = A$. Recall the biduality formula $T = T^{**}$ and apply Proposition 9.6 starting with T^*. It follows that $\mathcal{D}(TT^*)$ also is dense and exactly as in (9.6.1) every $h \in \mathcal{H}$ has a unique representation

$$h = y + T(T^*y)$$

9.9. Exercise. Verify from the above that A is self-adjoint, i.e one has the equality $A = A^*$.

9.10 Unbounded self-adjoint operators.

A densely defined operator A on the Hilbert space $\mathcal{H}$ for which $A = A^*$ is called self-adjoint.

9.11 Proposition *The spectrum of a self-adjoint operator A is contained in the real line, and if λ is non-real the resolvent satisfies the norm inequality*

$$||R_A(\lambda)|| \leq \frac{1}{|\mathfrak{Im}\,\lambda|}$$

Proof. Set $\lambda = a + ib$ where $b \neq 0$. If $x \in \mathcal{D}(A)$ and $y = \lambda x - Ax$ we have

$$||y||^2 = |\lambda|^2 \cdot ||x||^2 + ||Ax||^2 - 2 \cdot \mathfrak{Re}(\lambda) \cdot \langle x, Ax\rangle$$

The Cauchy-Schwarz inequality gives

(i) $$||y||^2 \geq b^2||x||^2 + a^2||x||^2 + ||Ax||^2 - 2|a| \cdot ||Ax|| \cdot ||x|| \geq b^2||x||^2$$

This proves that $x \to \lambda x - Ax$ is injective and since A is closed the range of $\lambda \cdot E - A$ is closed. Next, if y is $\perp$ to this range we have

$$0 = \lambda\langle x, y\rangle - \langle Ax, y\rangle \quad : x \in \mathcal{D}(A)$$

From this we see that y belongs to $\mathcal{D}(A^*)$ and since A is self-adjoint we get

$$0 = \lambda\langle x, y\rangle - \langle x, Ay\rangle$$

This holds for all x in the dense subspace $\mathcal{D}(A)$ which gives $\lambda \cdot y = Ay$ Since λ is non-real we have already seen that this entails that $y = 0$. Hence the range of $\lambda \cdot E - A$ is equal to $\mathcal{H}$ and the inequality (i) entails $R_A(\lambda)$ has norm $\leq \frac{1}{|\mathfrak{Im}\lambda|}$.

9.12 An conjugation formula. Let A be self-adjoint. For each complex number λ the hermitiain inner product on $\mathcal{H}$ gives the equation

$$\bar{\lambda} - A = (\lambda \cdot E - A)^*$$

So when we take the complex conjugate of λ it follows that § 9.5 that

$$R_A(\lambda)^* = R_A(\bar{\lambda}) \tag{9.12.1}$$

9.13 Properties of resolvents. Let A be self-adjoint. By Neumann's resolvent calculus the family $\{(R_A(\lambda)\}$ consists of pairwise commuting bounded operators outside the spectrum of A. Since $\sigma(SA)$ is real there exist operator-valued analytic functions $\lambda \mapsto R_A(\lambda)$ in the upper- respectively the lower half-plane. Moreover, since Neumann's resolvents commute, it follows from (9.12.1) that $R_A(\lambda)$ commutes with its adjoint. Hence every resolvent is a bounded normal operator.

9.14 A special resolvent operator. Take $\lambda = i$ and set $R = R_A(i)$. So here

$$R(iE - A)(x) = x \quad : \quad x \in \mathcal{D}(A)$$

9.15 Theorem. *The spectrum $\sigma(R)$ is contained in the circle*

$$C_* = \{|\lambda + i/2| = 1/2\}$$

Proof. Since $\sigma(A)$ is confined to the real line, it follows from § 0.0.5.1 that points in $\sigma(R)$ have the form

$$\lambda = \frac{1}{i - a} \quad : a \in \mathbf{R}$$

This gives

$$\lambda + i/2 = \frac{1}{i - a} + i/2 = \frac{1}{2(i - a)}(2 + i^2 - ia) = \frac{1 - ia}{2i(1 + ia)}$$

and the last term has absolute value $1/2$ for every real a.

9.B. The spectral theorem for unbounded self-adjoint operators.

The operational calculus in § 8.3-8.6 applies to the bounded normal operator R in § 9.14. If N is a positive integer we set

$$C_*(N) = \{\lambda \in C_* : \mathfrak{Im}(\lambda) \leq -\frac{1}{N}\} \quad \text{and} \quad \Gamma_N = C_*(N) \cap \sigma(R)$$

Let χ_{Γ_N} be the characterstic function of Γ_N. Now

$$g_N(\lambda) = \frac{1 - i\lambda}{\lambda} \cdot \chi_{\Gamma_N}$$

is Borel functionon $\sigma(R)$ which by operational calculus in § 8.xx gives a bounded and normal linear operator denoted by G_N. On Γ_N we have $\lambda = -i/2 + \zeta$ where $|\zeta| = 1/2$. This gives

$$\frac{1 - i\lambda}{\lambda} = \frac{1/2 - i\zeta}{-i/2 + \zeta} = \frac{(1/2 - i\zeta)(i/2 + \bar{\zeta})}{|\zeta - i/2|^2} = \frac{\mathfrak{Re}\,\zeta}{|\zeta - i/2|^2} \tag{1}$$

By § 8.x the spectrum of G_N is the range of the g-function on Γ_N and (1) entails that $\sigma(G_N)$ is real. Since G_N also is normal it follows that it is self-adjoint. Next, notice that

$$\lambda \cdot (\frac{1 - i\lambda}{\lambda} + i) = 1 \tag{2}$$

holds on Γ_N. Hence operational calculus gives the equation

$$R(G_N + i) = E(\Gamma_N) \tag{3}$$

where $E(\Gamma_N)$ is a self-adjoint projection. Notice also that

$$R \cdot G_N = (E - iR) \cdot E(\Gamma_N) \tag{4}$$

Hence (3-4) entail that

$$(5) \qquad E(\Gamma_N) - iRE(\Gamma_N) = (E - iR) \cdot E(\Gamma_N)$$

Next, the equation $RA = E - iR$ gives

$$(*) \qquad RAE(\Gamma_N) = (E - iR)E(\Gamma_N) = R \cdot G_N$$

9.B.1 Exercise. Conclude from the above that

$$(*) \qquad AE(\Gamma_N) = G_N$$

Show also that:

$$(**) \qquad \lim_{N\to\infty} AE(\Gamma_N)(x) = A(x) \quad \text{for each} \quad x \in \mathcal{D}(A)$$

9.B.2 A general construction. For each bounded Borel set e on the real line we get a Borel set $e_* \subset \sigma(R)$ given by

$$e_* = \sigma(R) \cap \{\frac{1}{i-a} a \in e\}$$

The operational calculus gives the self-adjoint operator G_e constructed via $g \cdot \chi_{e_*}$. We have also the operstor $E(e)$ given by χ_{e_*} and exactly as above we get

$$AE(e) = G_e$$

The bounded self-adjoint operators $E(e)$ and G_e commute with A and $\sigma(G_e)$ is contained in the closure of the bounded Borel set e. Moreover each $E(e)$ is a self-adjoint projection and for each pair of bounded Borel sets we have

$$E(e_1)E(e_2) = E(e_1 \cap E(e_1)$$

In particular the composed operators

$$E(e_1) \circ E(e_2) = 0$$

when the Borel sets are disjoint.

9.C The spectral measure. Exactly as for bounded self-adjoint operators the results above give rise to a map from $\mathcal{H} \times \mathcal{H}$ into the space of Riesz measures:

$$(x, y) \mapsto \mu_{x,y}$$

For each real-valued and bounded Borel function $\phi(t)$ on the real line with compact support there exists a bounded self-adjoint operator ϕ such that

$$\langle \Phi(x), y \rangle = \int g(t) \cdot d\mu_{x,y}(t)$$

All these Φ operators commute with A. If $x \in \mathcal{D}(A)$ and y is a vector in $\mathcal{H}$ one has

$$\langle A(x), y \rangle = \lim_{M\to\infty} \int_{-M}^{M} t \cdot d\mu_{x,y}(t)$$

9.D Remark. To digest the results above one should regard specific examples. For exasmpöe, consider A-opertors which arise via integral kernels on L^2-spaces. Then μ-measures above are related to the spectral functions constructed in § xx by approximating of the unbounded self-adjoint operator AS by a sequence of bounded self-adjoint operators.

10. 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 **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:

$$(A.2.1) \qquad (\lambda \cdot e - x)^{-1} = (\lambda_0 \cdot e - x) \cdot [\, e + \sum_{\nu=1}^{\infty} (\lambda_0 - \lambda)^{\nu} \cdot (\lambda_0 \cdot e - x)^{-\nu} \,]$$

We refer to (A.2.1) as a local Neumann series. The triangle inequality gives the norm inequality:

$$||(\lambda \cdot e - x)^{-1}|| \leq ||(\lambda_0 \cdot e - x)|| \cdot \big[1 + \sum_{\nu=1}^{\infty} |\lambda - \lambda_0|^{\nu} \cdot \delta^{\nu} \,\big] =$$

$$(A.2.2) \qquad ||(\lambda_0 \cdot e - x)|| \cdot \frac{1}{1 - |\lambda - \lambda_0| \cdot \delta}$$

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

$$(*) \qquad \mathfrak{r}(x) = \max_{\lambda \in \sigma(x)} |\lambda|$$

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

$$\text{(i)} \qquad \lim_{n\to\infty} \xi(n) = \xi_*$$

There remains to prove the equality

$$\text{(ii)} \qquad \xi_* = \mathfrak{r}(x)\,.$$

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

$$\text{(iii)} \qquad \xi_* \leq \mathfrak{r}(x)$$

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

$$(*) \qquad |\widehat{x}|_{\mathcal{M}_B} \leq ||x||$$

It is in general strict. When equality holds one says that B is a *uniform algebra*. In this case the Gelfand transform identifies B with a closed subalgebra of $C^0(\mathcal{M}_B)$. For an extensive study of uniform algebras we refer to the books [Gamelin] and [Wermer].

C. Examples of Banach algebras.

Below we illustrate the general theory by some examples. 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:

$$(1) \qquad \sigma_{B^*}(x) \subset \sigma_B(x)$$

The example above shows that this inclusion in general is strict. However, one has the opposite inclusion

$$(2) \qquad \partial(\sigma_B(x)) \subset \sigma_{B^*}(x)$$

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

(*) $$\mu = e^\nu * \mu_*^k$$

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

(*) $$\mathcal{M}(B) = \mathbf{R}^n_\xi \cup \{\lambda^*\}$$

was originally put forward by Wiener prior to the general theory about Banach algebras. Another Banach algebra is $M_d(\mathbf{R})^{\mathbf{n}}$ whose elements are discrete measures with a finite total variation. Thus, the elements are measures

$$\mu = \sum c_\nu \cdot \delta(p_\nu)$$

where $\{p_\nu\}$ is a sequence of points in $\mathbf{R}^n$ and $\{c_\nu\}$ a sequence of complex numbers such that $\sum |c_\nu| < \infty$. Here the Gelfand space is more involved. To begin with the Fourier transform identifies $\mathbf{R}^n_\xi$ with a subset of $\mathcal{M}(B)$. But the compact space $\mathcal{M}(B)$ is considerably and given by a compact abelian group which is called the Bohr group after Harald Bohr whose studies of almost periodic functions led to the description of $\mathcal{M}(B)$. However one has the following result:

3.2 Bohr's Theorem. *The subset $\mathbf{R}^n_\xi$ is dense in $\mathcal{M}(B)$.*

Remark. See XX for an account about almost periodic functions which proves Bohr's theorem in the case $n = 1$.

3.3 Beurling's density theorem. Consider the Banach algebra B generated by $M_d(\mathbf{R}^n)$ and $L^1(\mathbf{R}^n)$. So its elements are measures of the form

$$\mu = \mu_d + f$$

where μ_d is discrete and f is absolutely continuous. Here the Fourier transform identifies $\mathbf{R}^n_\xi$ with an open subset of $\mathcal{M}(B)$. More precisely, a multiplicative functional λ on B belongs to the open set $\mathbf{R}^n_\xi$ if and only if $\lambda(f) \neq 0$ for at least some $f \in L^1(\mathbf{R}$. The remaining part $\mathcal{M}(B) \setminus \mathbf{R}^n_\xi$ is equal to the Bohr group above.. It means that when λ is an arbitrary multiplicative functional on B then there exists $\lambda_* \in \mathcal{M}(B)$ such that λ_* vanishes on $L^1(\mathbf{R}^n$ while $\lambda_*(\mu) = \lambda(\mu)$ for every discrete measure. The density of $\mathbf{R}^n_\xi$ follows via Bohr's theorem and the fact that Fourier transforms of L^1-functions tend to zero as $|\xi| \to +\infty$.

3.4 Varopoulos' density theorem. For each linear subspace Π of arbitrary dimension $1 \leq d \leq n$ we get the space $L^1(\Pi)$ of absolutely continuous measures supported by Π and of finite total mass. Thus, we identify $L^1(\Pi)$ with a subspace of $M(\mathbf{R}^n)$. We get the closed subalgebra of $M(\mathbf{R}^n)$ generated by all these L^1-spaces and the discrete measures. It is denoted by $\mathcal{V}(\mathbf{R}^n)$ and called the Varopoulos measure algebra in $\mathbf{R}^n$. In [Var] it is proved that the Fourier transform identifies $\mathbf{R}^n_\xi$ with a dense subset of $\mathcal{M}(\mathcal{V}(\mathbf{R}^n))$.

3.5 The extended $\mathcal{V}$-algebra. In $\mathbf{R}^n$ we can consider semi-analytic strata which consist of locally closed real-analytic submanifolds S whose closure $\bar{S}$ is compact and the relative boundary $\partial S = \bar{S} \setminus S$ is equal to the zero set of a real analytic function. On each such stratum we construct measures which are absolutely continuous with respect to the area measure of S. Here the dimension of S is between 1 and $n-1$ and now each measure in $L^1(S)$ is identified with a Riesz measure in $\mathbf{R}^n$ which happens to be supported by S, One can easily prove that every $\mu \in L^1(S)$ has a power which belongs to the Varopolulos algebra and from this deduce that if $\mathcal{V}^*$ is the closed subalgebra of $M(\mathbf{R}^n)$ generated by the family $\{L^1(S)\}$ and $V(\mathbf{R}^n))$ then one gets a new Wiener algebra.

3.6 Olofsson's example. Above real analytic strata were used to obtain $\mathcal{V}^*$. That real-analyticity is 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 further examples as above on curves defined by C^∞-functions outside the Carleman class of quasi-anaytic functions.

11. Miscellaneous results.

Introduction. The subsequent sections treat topics related to the material in § 8 and 9.

11.1. Symmetric operators.

A densely defined and closed operator T on a Hilbert space $\mathcal{H}$ is symmetric if

$$(*) \qquad \langle Tx, y\rangle = \langle x, Ty\rangle \quad \text{hold for all pairs} \quad x, y \in \mathcal{D}(T)$$

The symmetry means that the adjoint T^* extends T, i.e.

$$\Gamma(T) \subset \Gamma(T^*)$$

Recall that adjoints always are closed operators. Hence $\Gamma(T^*)$ is a closed subspace of $\mathcal{H} \times \mathcal{H}$ and becomes a Hilbert space equipped with the inner product

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

Moreover, since T also is closed. it follows that $\Gamma(T)$ appears as a closed subspace of this Hilbert space. Consider the eigenspaces:

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

11.1.1 Proposition. *The following orthogonal decomposition exists in the Hilbert space* $\Gamma(T^*)$*:*

$$(*) \qquad \Gamma(T^*) = \Gamma(T) \oplus \mathcal{D}_+ \oplus \mathcal{D}_-$$

Proof. The verification that the three subspaces are pairwise orthogonal is left to the reader. To show that the direct sum above is equal to $\Gamma(T^*)$ we use duality and there remains only to prove that

$$(1) \qquad \Gamma(T)^\perp = \mathcal{D}_+ \oplus \mathcal{D}_-$$

To show (1) we pick a vector $y \in \Gamma(T)^\perp$. Here $(y, T^*y) \in \Gamma(T^*)$ and the definition of orthogonal complements gives:

$$\langle x, y\rangle + \langle Tx, T^*y\rangle = 0 \quad : \quad x \in \mathcal{D}(T)$$

From this we see that $T^*y \in \mathcal{D}(T^*)$ and obtain

$$\langle x, y\rangle + \langle x, T^*T^*y\rangle = 0$$

The density of $\mathcal{D}(T)$ entails that

$$0 = y + T^*T^*y = (T^* + iE)(T^* - iE)(y) \implies$$

$$\xi = T^*y - iy \in \mathcal{D}_- \quad \text{and} \quad \eta = T^*y + iy \in \mathcal{D}_+ \implies$$

$$y = \frac{1}{2i}(\eta - \xi) \in \mathcal{D}_- \oplus \mathcal{D}_+$$

which proves (1).

11.1.2 The case $\dim(\mathcal{D}_+) = \dim(\mathcal{D}_-)$**.** Suppose they are finite dimensional with equal dimension $n \geq 1$. Then self-adjoint extensions of T are found as follows: Let $e_1, \ldots, e_n$ be an orthonormal basis in $\mathcal{D}_+$ and $f_1, \ldots, f_n$ a similar basis in $\mathcal{D}_-$. For each n-tuple $e^{i\theta_1}, \ldots, e^{i\theta_n}$ of complex numbers with absolute value one we have the subspace of $\mathcal{H}$ generated by $\mathcal{D}(T)$ and the vectors

$$\xi_k = e_k + e^{i\theta_k} \cdot f_k \quad : \quad 1 \leq k \leq n$$

On this subspace we define a linear operator A_θ where $A_\theta = T$ on $\mathcal{D}(T)$ while

$$A_\theta(\xi_k) = ie_k - ie^{i\theta_k} \cdot f_k$$

11.1.3 Exercise. Verify that A_θ is self-adjoint and prove the converse, i.e. if A is an arbitrary self-adjoint operator such that

$$\Gamma(T) \subset \Gamma(A) \subset \Gamma(T^*)$$

then there exists a unique n-tuple $\{e^{i\theta_\nu}\}$ such that

$$A = A_\theta$$

11.1.4 Example. Let $\mathcal{H}$ be the Hilbert space $L^2[0,1]$ of square-integrable functions on the unit interval $[0,1]$ with the coordinate t. A dense subspace $\mathcal{H}_*$ consists of functions $f(t) \in C^1[0,1]$ such that $f(0) = f(1) = 0$. On $\mathcal{H}_*$ we define the operator T by

$$T(f) = if'(t)$$

A partial integration gives

$$\langle T(f), g\rangle = i\int_0^1 f'(t) \cdot \bar{g}(t) \cdot dt = \int_0^1 \bar{g}'(t) \cdot f(t)dt = \langle f, T(g)\rangle$$

Hence T is symmetric. Next, an L^2-function h belongs to $\mathcal{D}(T^*)$ if and only if there exists a constant $C(h)$ such that

$$\Big|\int_0^1 if'(t) \cdot \bar{h}(t)dt\Big| \leq C(h) \cdot ||f||_2 \quad : f \in \mathcal{H}_*$$

This means that $\mathcal{D}(T^*)$ consists of all L^2-functions h such that the distribution derivative $\frac{dh}{dt}$ again belongs to L^2.

Exercise. Show that

$$\mathfrak{D}_+ = \{h \in L^2 \quad : \frac{dh}{dt} = ih\}$$

is a 1-dimensional vector space generated by the L^2-function e^{ix}. Similarly, $\mathcal{D}_-$ is 1-dimensional and generated by e^{-ix}.

Self-adjoint extensions of T. For each complex number $e^{i\theta}$ we get the linear space $\mathcal{D}_\theta$ of functions $f(t) \in \mathcal{D}(T^*)$ such that

$$f(1) = e^{i\theta} f(0)$$

Exercise. Verify that one gets a self-adjoint operator T_θ which extends T where is $\mathcal{D}(T_\theta) = \mathcal{D}_\theta$. Conversely, show every self-adjoint extension of T is equal to T_θ for some θ. Hence the family $\{T_\theta\}$ give all self-adjoint extensions of T with their graphs contained in $\Gamma(T^*)$.

11.2 Semi-bounded symmetric operators.

Let T be closed, densely defined and symmetric. It is said to be bounded below if there exists some positive constant k such that

(*) $$\langle Tx, x\rangle \geq k \cdot ||x||^2 \quad : \quad x \in \mathcal{D}(T)$$

On $\mathcal{D}(T)$ we have the Hermitian bilinear form:

(1) $$\{x, y\} = \langle Tx, y\rangle \quad \text{where (*) entails that} \quad \{x, x\} \geq k \cdot ||x||^2$$

In particular a Cauchy sequence with respect to this inner product is a Caichy sequence in the given Hilbert space $\mathcal{H}$. So if $\mathcal{D}_*$ is the completion of $\mathcal{D}(T)$ with respect to the inner product above, then it appears as a subspace of $\mathcal{H}$. Put

$$\mathcal{D}_0 = \mathcal{D}(T^*) \cap \mathcal{D}_*$$

11.2.1 Proposition. *One has the equality*

(*) $$T^*(\mathcal{D}_0) = \mathcal{H}$$

Proof. A vector $x \in \mathcal{H}$ gives a linear functional on $\mathcal{D}_*$ defined by

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

We have

(i) $$|\langle y, x\rangle| \leq ||x| \cdot ||y|| \leq ||x||| \cdot \frac{1}{\sqrt{k}} \cdot \sqrt{\{y, y\}}$$

where we used (1) above. The Hilbert space $\mathcal{D}_*$ is self-dual. This gives a vector $z \in \mathcal{D}_*$ such that

(iii) $$\langle y, x\rangle = \{y, z\} = \langle Ty, z\rangle$$

Since $\mathcal{D}(T) \subset \mathcal{D}_*$ we have (iii) for every vector $y \in \mathcal{D}(T)$, and the construction of T^* entails that $z \in \mathcal{D}(T^*)$ so that (iii) gives

(iv) $$\langle y, x\rangle = \langle y, T^*(z)\rangle$$

The density of $\mathcal{D}_*$ in $\mathcal{H}$ implies that $x = T^*(z)$ and since $x \in \mathcal{H}$ was arbitrary we get (*) in the proposition.

11.2.2 A self-adjoint extension. Let T_1 be the restriction of T^* to $\mathcal{D}_0$. We leqve it to the reader to check that T_1 is symmetric and has a closed graph. Moroever, since $\mathcal{D}(T) \subset \mathcal{D}_0$ and T^* is an extension of T we have

$$\Gamma(T) \subset \Gamma(T_1)$$

Next, Proposition 11.2.1 gives

$$T_1(\mathcal{D}(T_1) = \mathcal{H}$$

i.e. the T_1 is surjective. But then T_1 is self-adjoint by the general result below.

11.2.3 Theorem . *Let S be a densely defined, closed and symmetric operator such that*

(*) $$S(\mathcal{D}(S)) = \mathcal{H}$$

Then S is self-adjoint.

Proof. Let S^* be the adjoint of S. When $y \in \mathcal{D}(S^*)$ we have by definition

$$\langle Sx, y\rangle = \langle x, S^*y\rangle \quad : \quad x \in \mathcal{D}(S)$$

If $S^*y = 0$ this entails that $\langle Sx, y\rangle = 0$ for all $x \in \mathcal{D}(S)$ so the assumption that $S(\mathcal{D}(S)) = \mathcal{H}$ gives $y = 0$ and hence S^* is injective. Finally, if $x \in \mathcal{D}(S^*)$ the hypothesis (*) gives $\xi \in \mathcal{D}(S)$ such that

(i) $$S(\xi) = S^*(x)$$

Since S is symmetric, S^* extends S so that (i) gives $S^*(x - \xi) = 0$. Since we already proved that S^* is injective we have $x = \xi$. This proves that $\mathcal{D}(S) = \mathcal{D}(S^*)$ which means that S is self-adjoint.

11.3 Contractions

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

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

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

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

From the result in § 8.xx it follows that this non-negative self-adjoint operator has a square root:

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

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

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

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

11.3.1 Propostion. *One has the equations*

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

Proof. If n is a positive integer we notice that

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

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

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

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

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

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

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

In particular we can take $f(t) = \sqrt{1-t}$ and Proposition 11.3.1 follows.

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

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

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

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

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

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

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

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

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

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

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

Using (iii) we have

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

Similarly, (ii) gives

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

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

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

To get (v) we use Proposition 11.5.1 which gives

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

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

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

and (v) follows.

11.3.4 The Nagy-Szegö theorem.

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

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

$$A^n = \mathcal{P} \cdot U_A^n \quad : \quad n = 1, 2, \ldots$$

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

Proof. On the product $\mathcal{H}_1 = \mathcal{H} \times \mathcal{H}$ we have the unitary operator U_A from (*) in 11.3.2. Let $\mathcal{P}(x, y) = x$ be the projection onto the first factor. Then (*) in (11.3.2) gives $A = \mathcal{P}U_A$ and the constructions from the proof of Propostion 11.3.4 imply that $A^n = \mathcal{P} \cdot U^n$ hold for every $n \geq 1$ which finishes the proof.

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

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

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

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

Proof. Theorem 11.3.5 gives $p(A) = \mathcal{P} \cdot p(U_A)$. Since the orthogonal $\mathcal{P}$-projection is norm decreasing we get

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

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

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

The right hand side is majorized by $|p|_D^2$ and Theorem 11.3.7 follows.

11.3.8 An application. Let $A(D)$ be the disc algebra. Since each $f \in A(D)$ can be uniformly approximated by analytic polynomials, Theorem 11.3.7 entails that if A is a contraction then each $f \in A(D)$ gives a bounded linear operator $f(A)$, i.e. we have a map from $A(D)$ into the space of bounded linear operators on $\mathcal{H}$.

11.4 A product formula

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

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

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

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

where the self-adjointness gives the equality

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

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

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

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

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

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

Hilbert's the spectral theorem for bounded and self-adjoint operators on a Hilbert space $\mathcal{H}$ gives a similar equation as in (*) above using the non-negative and self-adjoint operator $\sqrt{T^*T}$. More generally, let T be densely defined and closed. From § XX there exists the densely defined self-adjoint operator T^*T and we can also take its square root.

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

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

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

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

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

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

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

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

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

$$Bx = j(Ax_*)$$

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

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

We have also

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

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

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

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

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

Keeping y fixed the linear functiional

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

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

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

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

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

Conclusion. From the above we have the inclusion

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

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

FINISH ..

11.5 A result about positive operators.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

$$f(s_0) = 1$$

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

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

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

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

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

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

Now

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

Taking real parts we get

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

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

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

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

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

A point s_* in this non-empty intersection gives

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Moreover, by (*) we also have

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

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

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

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

Compact pertubations to finish Kakutani-Yosida !!!

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Together with (ii) we obtain

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

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

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

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

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

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

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

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

$$\frac{A_{-1}}{z} + A_0 + A_1 z + \ldots = B(z) + B(z)(E + B(z) \cdot K)^{-1} B(z)$$

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

$$(E + B(z) \cdot K)^{-1} = \frac{A^*_{-1}}{z} + A^*_0 + A^*_1 z + \ldots$$

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

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

Next, from (xx) one has

$$E = (E + B(z) \cdot K)(\frac{A^*_{-1}}{z} + A^*_0 + A^*_1 z + \ldots) \implies (E + B(0) \cdot K) A^*_{-1} = 0$$

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

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

Proof. We have the algebraic equation

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

It follows that

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

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

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

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

11.6 Factorizations of non-symmetric kernels.

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

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

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

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

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

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

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

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

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

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

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

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

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

11.6.1 Spectral properties of non-symmetric kernels.

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

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

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

(1) $$\mathcal{S}\phi_n = \kappa_n\phi_n$$

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

(2) $$u = \sum \alpha_n\cdot\phi_n$$

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

(3) $$u = \lambda\cdot\mathcal{K}u$$

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

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

Multiplying with $\phi_p(x)$ an integration gives

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

Next, using the expansion of u from (2) we get the equations:

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

Set

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

It follows that $\{\alpha_p\}$ satisfies the system

$$\kappa_p \alpha_p = \lambda \cdot \sum_{q=1}^{\infty} c_{qp}\alpha_q \tag{7}$$

Since $N(x,y) = N(y,x)$ the doubly indexed c-sequence is symmetric. Set

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

Next, put

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

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

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

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

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

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

Hence the symmetric and doubly indexed sequence

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

is of Hilbert-Schmidt type.

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

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

Uniqueness results for the exterior Laplace equation

Let Ω be a bounded open set in $\mathbf{R}^3$ whose boundary is a finite union of closed surfaces of class C^1 at least. Set $U = \mathbf{R}^3 \setminus \Omega$. Denote by $\mathcal{S}(U)$ the class of real-valued C^2-functions f in U which extend continuously to the boundary of U which of course is equal to $\partial\Omega$. Moreover, we assume that the the exterior normal derivaties $\frac{df}{dn}$ taken along the boundary exist and give a continuous function on ∂U. Consider large positive R-numbers so that the open ball $B(R)$ of radius R centered at the origin contains the closure of Ω. Set

$$D(f)^2(x,y,z) = (\frac{\partial f}{\partial x})^2 + (\frac{\partial f}{\partial y})^2 + (\frac{\partial f}{\partial z})^2$$

Green's formula applied to the domain $B(R) \setminus \overline{\Omega}$ gives

(1) $$\int_{B(R)\setminus\overline{\Omega}} D(f)^2\,dp + \int_{B(R)\setminus\overline{\Omega}} f\cdot\Delta(f)\,dp = \int_{\partial U}$$

where $dp = dxdxydz$. We shall use this to prove:

1. Theorem. *If f and $\Delta(f)$ both belong to $L^2(U)$, it follows that*

$$\int_U D(f)^2\,dp = \int_{\partial U} - \int_U f\cdot\Delta(f)\,dp$$

Proof Since f and $\Delta(f)$ are square integrable, the Cauchy-Schwarz inequality entails that $f\cdot\Delta(f)$ is absolutely integrable over U. Hence (1) gives the theorem if we prove the limit formula

(i) $$\liminf_{R\to\infty} \int_{\partial U} = 0$$

To prove (i) we consider the function

$$\psi(R) = \int_{B(R)\setminus\overline{\Omega}} u^2\,dp$$

The derivative becomes

$$\psi'(R) = \int_{S(R)} u^2\cdot d\omega$$

where $S(R)$ is the sphere of radius R and $d\omega$ its area measure. Passing to the second derivative the reader should verify the equation

$$\psi''(R) = \frac{2}{R}\cdot\psi'(R) + 2\cdot\int_{S(R)} f\cdot\frac{\partial f}{\partial n}\,d\omega$$

Hence (i) follows if

(ii) $$\liminf_{R\to\infty} \psi''(R) - \frac{2}{R}\cdot\psi'(R) = 0$$

To obtain (ii) we first notice that Ψ is non-decreasing and since f is square integrable it tends to a finite limit as $R\to+\infty$. Hence the first order derivative cannot stay above a positive constant for all large R. So for the derivative ψ' two cases can occur. Either it decreases in a monotone way to zero as $R\to+\infty$. In this case it is eveident that there exists a striuctly increasing sequebce $\{r_n\}$ such that $\psi''(r-n)\to 0$ and (ii) follows. In the second case the function $\psi'(R)$ attain a local minium at an infinite sequence $\{r_n\}$ which again tend to zero. Here $\psi''(r_n)=0$ and at the same time these local minimum values of the first order derivative tend to zero. So again (ii) holds and Theorem 1 is proved.

A vanishing result.

Let f satisfy the differential equation

(*) $$\Delta(f) + \lambda\cdot f = 0$$

in U for some real number λ. In addition we assume that

(**) $$\frac{df}{dn}(p)=0 \quad : p\in\partial U$$

2. Theorem. *If f satisfies (*-**) and belongs to $L^2(U)$, then f is identically zero.*

Proof. Notice that Theorem 1 gives the equality

$$\int_U D(f)^2\,dp=\lambda\cdot\int_U f^2\,dp$$

So if $\lambda\leq 0$ the vanishing of f is obvious. From now on $\lambda>0$. We shall work with polar coordinates, i.e. employ the Euler's angular variables ϕ and θ where

$$0<\theta \quad : 0<\phi<2\pi$$

The wellknown expression of Δ in the variables r,θ,ϕ shows that the equation (*) corresponds to

(i) $$xxxxx+xxxx=0$$

Let $n\geq 1$ and $Y_n(\theta,\phi)$ some spherical function of degree n with a normalised L^2-integral equal to one. For each r where $B(r)$ contains $\overline{\Omega}$ we set

(ii) $$Z(r)=\int_0^{2\pi}\int_0^{\pi} Y_n(\theta,\phi)\cdot f(r,\theta,\phi)\,\sin(\theta)\,d\theta d\phi$$

The Cauchy-Schwarz inequality gives

$$Z(r)^2\leq\int_0^{2\pi}\int_0^{\pi} Y_n^2\cdot\sin(\theta)\,d\theta d\phi\cdot\int_0^{2\pi}\int_0^{\pi} f^2(r,\theta,\phi)\cdot\sin(\theta)\,d\theta d\phi$$

Since the L^2-integral if Y_n is normalised the last product is reduced to

$$J(r)=\int_0^{2\pi}\int_0^{\pi} f^2(r,\theta,\phi)\cdot\sin(\theta)\,d\theta d\phi$$

Now

$$\int_{r_*}^{\infty} r^2\cdot J(r)\,dr$$

is equal to the finite L^2-integral of f in the exterior domain taken outside a ball $B(r_*)$. From (ii) we conclude that

(iii) $$\int_{r_*}^{\infty} r^2\cdot Z(r)^2\,dr<\infty$$

A differential equation. Recall that a spherical function of degree n satisfies

(iv) $$\frac{1}{\sin\theta}\frac{d}{d\theta}\big(\sin\theta\cdot\frac{\partial Y_n}{\partial\theta}\big)+\frac{1}{\sin^2\theta}\frac{\partial^2 Y_n}{\partial\phi^2}+(n+1)n\cdot Y_n=0$$

Exercise. Show via suitable partial integrations that (i) and (iv) imply that $Z(r)$ satisfies the differential equation

(v) $$\frac{1}{r^2}\cdot\frac{d}{dr}(r^2\cdot\frac{dZ}{dr})+(\lambda-\frac{n(n+1)}{r^2})Z=0$$

The second order differential equation has two linearly independent solutions of the form

$$Z_1=\cos(\sqrt{\lambda}r)\cdot\big[\frac{1}{r}+\frac{a_2}{r^2}+\ldots\big]$$

$$Z_2=\sin(\sqrt{\lambda}r)\cdot\big[\frac{1}{r}+\frac{b_2}{r^2}+\ldots\big]$$

It follows that there exist a pair of constants c_1,c_2 such that

(vi) $$Z = c_1 Z_1 + c_2 Z_2 = \frac{c_1 \cdot \cos(\sqrt{\lambda} r) + c_2 \cdot \sin(\sqrt{\lambda} r)}{r} + \frac{B(r)}{r^2}$$

where $r \mapsto B(r)$ stays bounded as $r \to +\infty$.

Exercise. Show that the finite integral in (iii) and (vi) give $c_1 = c_2 = 0$ and hence $Z(r)$ identically zero for large r. Since this hold for all spherical functions we conclude that f is identically zero outside the closed ball $B(r_*)$. Finally, by assumption U is connected and the elliptic equation (*) implies that f is a real-analytic function in U, So the vanishing outside a large ball entails that f is identically zero in U which finishes the proof of Theorem 2.

3. A result about absolute continuity.

We consider functions depending upon a real parameter μ which varies in an interval $[a, b]$. To each μ we are given a function $f(x, y, z; \mu)$ which is square integrable C^2-function in U and the normal derivative along ∂U is zero, i.e just as in the class $\mathcal{S}$ above. Moreover, the $L^2(U)$-valued function

(i) $$\mu \mapsto f(x, y, z; \mu)$$

has a finite total variation on $[a, b]$. Next, assume that for every sub-interval ℓ of $[a, b]$ one has the equality

$$\Delta\big(\int_\ell \frac{d}{d\mu} f(x, y, z; \mu) + \int_\ell \mu \cdot \frac{d}{d\mu} f(x, y, z; \mu) = 0$$

where the integrals as usual are taken in the sense of Borel-Stieltjes.

Theorem. *Every function from (i) which satisfies the conditions above is absolutely continues with respect to μ.*

About the proof. Using similar methods as in the proof of Theorem 1 one reduces the proof to study functions $g(r; \mu)$ where $r \mapsto g(r; \mu)$ is a C^2-function and square integrable on the interval $[r_*, +\infty)$ for a given $r_* > 0$ while μ as above varies in $[a, b]$. Moreover one has

$$\max_\mu \int_{r_*}^\infty g(r; \mu)\, dr < \infty$$

Next, for each sub-interval $\ell = [\alpha, \beta]$ we set

$$\delta_\ell(g(r, \mu) = g(r, \beta) - g(r, \alpha)$$

With these notations we say that $g(r; \mu)$ is absolutely continuous with respect to μ if there to each $\epsilon > 0$ exists $\delta > 0$ such that

$$\sum \int_{r_*} |\delta_{\ell_\nu}((g(r, \mu)\big|^2 \cdot r^2\, dr < \epsilon$$

for every finite family of sub-intervals $\{\ell_\nu\}$ when the sum of their lengths is $< \delta$.

Theorem. *Assume in addition to the above that the equation below holds for each sub-interval ℓ*

$$\frac{1}{r^2} \frac{d}{dr} \big[r^2 \frac{d}{dr}(\delta_\ell(g(r, \mu)\big] - \frac{n(n+1)}{r^2} \cdot \delta_\ell(g(r, \mu) + \int_\ell \mu \cdot \frac{d}{d\mu}(g(r, \mu) = 0$$

Then $g(r, \mu)$ is absolutely continuous with respect to μ.

A first example: Moment problems.

In the very impressive and highly original article *Recherches sur les fractions continues* [Ann.Fac. Sci. Toulouse. 1894], Stieltjes studied the moment problem on the non-negative real line. This amounts to find a non-negative Riesz measure μ on $\mathbf{R}^+$ with prescribed moments

$$(*) \qquad c_\nu = \int_0^\infty x^\nu d\mu(x) \quad : \nu = 0, 1, 2, \ldots$$

An obvious necessary condition for the existence of μ is that the quadratic form

$$J(x) = \sum_{p=1}^\infty \sum_{q=1}^\infty c_{pq} x_p x_q$$

is positive definite when it is restricted to vectors x-vectors where only finitely many $x_p \neq 0$. The moment problem is determined with respect to the given sequence $\{c_\nu\}$ if there exists a unique non-negative measure μ such that (*) holds. To prove existence and analyze uniqueness, Stieltjes employed the expansion in continued fractions of the c-sequence. Stieltjes' work was later reconsidered by Hamburger in the article *xxxx*[1922] where one allows solutions by Riesz measures μ on the whole real line, i.e. now integration takes place over $(-\infty, +\infty)$. In § xx we explain that this (extended) moment problem is closely related to the class of linear operators defined by inifinite and symmetric matrices of the form

$$xxx = xxx$$

where $\{a_n\}$ is an arbitrary real sequence while $\{b_n\}$ is a sequence of positive real numbers. A complete description of all solutions μ to a non-determined moment problem was established by Carleman in the article *Sur le probléme des moments* [C.R.Acad. Sci. Paris. 1922], and we shall expose this in § xx.

Returning to the case treated by Stieltjes the existence of non-determined moment problems, leads to a peculiar result for certain densely defined linear operators on the Hilbert space ℓ^2. Following Stieltjes we consider a positive sequence of real numbers $\{b_n\}$ and regard the operator S defined by

$$S(x) = (-b_1 x_2, -b_1 x_1, -b_2 x_4, -b_2 x_3, \ldots)$$

where $x = \{x_n\}$ is a vector in ℓ^2. It means that S is represented by the infinite symmetric matrix whose non-zero elements only appear in position $(j, j+1)$ or $(j+1, j)$ for every integer $j \geq 1$. Now there exists the dense subspace $\mathcal{D}(S)$ of vectors $x \in \ell^2$ such that $S(x)$ also belongs to ℓ^2. The question arises if the symmetry of the S-matrix persists in the following sense: Suppose that x and y both belong to $\mathcal{D}(S)$. Does it follow that

$$(**) \qquad \langle S(x), y\rangle = \langle x, S(y)\rangle$$

where we have used the Hermitian inner product on the complex Hilbert space ℓ^2. So above x and y can be complex vectors. In the cited article above, Carleman proved that (**) hold for every pair in $\mathcal{D}(S)$ if and only if the corresponding moment problem is determined. This means that in spite of the symmetry of the matrix which represents S, the resulting densely defined linear operator can fail to be self-adjoint. In § xx we show that whenever this peculair phenomenon occurs, the series

$$\sum_{n=1}^\infty \frac{1}{b_n} < \infty$$

However, the converse is not true, i.e. there exist determine moment problems where the series above is converges.

A final remark. The study of the Hamburger-Stieltjes moment problem offers an instructive lesson about unbounded but densely defined liear operators on Hilbert spaces. At the same time the material in §xx will teach that even if one starts with a problem expressed in terms of functional analysis, it is often necesssty to employ both results and methods from other disciplines such as analytic function theory and Fourier analysis.

Moment problems.

A great inspiration during the development of unbounded symmetric operators on Hibert spaces emerged from Stieltjes' pioneering article *Recherches sur les fractions continues* [Ann. Fac. Sci. Toulouse. 1894]. The moment problem asks for conditions on a sequence $\{c_0, c_1, \ldots\}$ of positive real numbers such that there exists a non-negative Riesz measure μ on the real line and

$$(*) \qquad c_\nu = \int_{-\infty}^{\infty} t^\nu \cdot d\mu(t)$$

hold for each $\nu \geq 0$. One easily verifies that if μ exists then the Hankel determinants

$$\begin{pmatrix} c_0 & c_1 & \ldots & c_n \\ c_1 & c_2 & \ldots & c_{n+1} \\ \ldots & \ldots & \ldots & \ldots \\ c_n & c_{n+1} & \ldots & c_{2n} \end{pmatrix}$$

are > 0 for every $n = 0, 1, \ldots$. For brevity we say that $\{c_n\}$ is a positive Hankel sequence when the determinants above are > 0. It turns out that this confition also is sufficient.

Theorem. *For each positive Hankel sequence there exists at least one non-negative Riesz measure such that (*) holds.*

Remark. The result above is due to Hamburger who extended Stieltjes' original results which treated the situation where one seeks non-negative measures supported by $\{\geq 0\}$. The moment problem for a positive Hankel sequence is said to be determined if (*) has a unique solution μ. Hamburger proved that the determined case holds if and only if the the associated continued fractions of $\{c_n\}$ is completely convergent. See § x below for an account about this convergence condition. A "drawback" in the Hamburger-Stieltjes theory is that it is often difficult to check when the associated continued fraction of a positive Hankel sequence is completely convergent. So one is led to seek sufficient conditions in order that (*) has a unique solution μ. Such a sufficient condition was established in Carleman's monograph [Carleman 1923. Page 189-220] and goes as follows:

1. Theorem. *The moment problem of a positive Hankel sequence $\{c_n\}$ is determined if*

$$\sum_{n=1}^{\infty} c_n^{-\frac{1}{n}} = +\infty$$

Another criterion for determined moment problems. Studies of continued fractions and their associated quadratic forms go back to work by Heine. See [Handbuch der theorie der Kugelfunktionen: Vol. 1. part 2]. An excellent account appears also in the article *Zur Einordnung der Kettenbruchentheorie in die theorie der quadratischen Formen von unendlichvielen Veränderlichen* [Crelle J. of math. 1914] by Hellinger and Toeplitz. A positive Hankel sequence $\{c_n\}$ corresponds to a quadratic form in an infinite number of variables:

$$J(x) = \sum_{p=1}^{\infty} a_p x_p^2 - 2 \cdot \sum_{p=1}^{\infty} b_p x_p x_{p+1}$$

where $b_p \neq 0$ for every p. More precisely, the sequences $\{a_\nu\}$ and $\{b_\nu\}$ arise when the series

$$-\left[\frac{c_0}{\mu} + \frac{c_1}{\mu^2} + \frac{c_2}{\mu^3} + \ldots\right]$$

is formally expanded into a continuous fraction

$$\cfrac{c_0}{a_1 - \mu - \cfrac{b_1^2}{a_1 - \mu - \cfrac{b_2^2}{a_2 - \mu}}}$$

Now there exists the infinite matrix A with diagonal elements $\alpha_{pp} = a_p$ while

$$\alpha_{p,p+1} = \alpha_{p+1,p} = -b_p \quad : p = 1, 2, \ldots$$

and all the other elements are zero. The symmetric A-matrix gives a densely defined linear operator on the complex Hilbert space ℓ^2. The following result is proved [Carleman:1923]:

2. Theorem. *The densely defined operator A on ℓ^2 is self-adjoint if and only if the moment problem for $\{c_n\}$ is determined.*

3. A study of A-operators.

Ignoring the "source" of the pair of real sequences $\{a_p\}$ and $\{b_p\}$ we consider a matrix A as above where the sole condition is that $b_p > 0$ for every p. If μ is a complex number one seeks infinite vectors $x = (x_1, x_2, \ldots)$ such that

$$\text{(i)} \qquad Ax = \mu \cdot x$$

It is clear that (i) holds if and only if the sequence $\{x_p\}$ satisfies the infinite system of linear equations

$$(a_1 - \mu)x_1 = b_1 x_2$$
$$(a_p - \mu)x_p = b_{p-1}x_{p-1} + b_p x_{p+1} \quad : p \geq 2$$

Since $b_p > 0$ for every p we see that x_1 determines the sequence and $x_2, x_3, \ldots$ depend on x_1 and the parameter μ. Keeping x_1 fixed while μ varies the reader can verify that

$$x_p = \psi_p(\mu) \quad : p \geq 2$$

where $\psi_p(\mu)$ is a polynomial of degree $p-1$ for each $p \geq 2$. These ψ-polynomials depend on the given pair of sequences $\{a_p\}$ and $\{b_p\}$ and with the following result is proved in [ibid]:

3.1 Theorem. *The densely defined operator A on ℓ^2 is self-adjoint if and only if*

$$\sum_{p=1}^{\infty} \psi_p(\mu) = +\infty$$

for every non-real complex number μ.

4. A general result. Let $\{c_\nu\}$ be a positive Hankel sequence, determined or not. Let ρ be a non-negative measure which solves the moment problem (*) and set

$$\widehat{\rho}(\mu) = \int \frac{d\rho(t)}{t - \mu}$$

This yields an analytic function in the upper half-plane. A major result in [Carleman] gives a sharp inclusion for the values which can be attained by these $\widehat{\rho}$ functions while ρ varies in the family of non-negative measures which solve the moment problem. It is expressed via constructions of discs in the upper half-plane which arise via a nested limit of discs constructed from a certain family of Möbius transformations. More precisely, to the given Hankel sequence we have the pair of sequences $\{a_\nu\}$ and $\{b_\nu\}$. For a fixed μ in the upper half-plane we consider the maps:

$$S_\nu(z) = \frac{b_\nu^2}{a_{\nu+1} - \mu - z} \quad : \nu = 0, 1, \ldots$$

Notice that

$$\mathfrak{Im}(S_\nu(z)) = \frac{b_\nu^2 \cdot \mathfrak{Im}(\mu + z)}{|a_{\nu+1} - \mu - z|^2} > 0$$

It follows that S_ν maps the upper half-plane U^+ conformally onto a disc placed in U^+.

Exercise. To each $n \geq 1$ we consider the composed map

$$\Gamma_n = S_0 \circ S_1 \circ \ldots \circ S_{n-1}$$

Show that the images $\{C_n(\nu) = \Gamma_n(U^+)\}$ form a decreasing sequence of discs and there exists a limit

$$C(\mu) = \cap C_n(\mu)$$

which either is reduced to a single point or is a closed disc in U^+.

4.1 Theorem. *For each $\mu \in U^+$ and every ρ-measure which solves the moment problem one has the inclusion*

$$\widehat{\rho}(\mu) \subset C(\mu)$$

Moreover, for each point z in this disc there exists ρ such that $\widehat{\rho}(\mu) = z$.

§ 12. Symmetric integral operators.

Consider the domain $\square = \{0 \leq x \leq 1\} \times \{0 \leq y \leq 1\}$ in $\mathbf{R}^2$. Let $K(x,y)$ be a real-valued and Lebesgue mesure function on $\square$ such that the integrals

$$\int_0^1 K(x,y)^2 \, dy < \infty$$

for all x outside a null-set on $[0,1]$. In addition we assume that K is symmetric, i.e.$K(x,y) = K(y,x)$. The K-kernel is bounded in Hilbert's sense if there exists a constant C such that

$$\iint_\square K(x,y)u(x)u(y) \, dxdy \leq C^2 \cdot \int_0^1 u(x)^2 \, dx$$

for each $u \in L^2[0,1]$. A special case occurs when $K(x,y)$ satisfies

$$\iint_\square |K(x,y)|^2 \, dxdy < \infty$$

Then $\mathcal{K}$ is called a Hilbert-Schmidt operator and a crucial fact is that it yields a compact operator on the Hilbert space $\mathcal{H} = L^2[0,1]$.

Exercise. Prove that Hilbert-.Schmidt operators are compact. Tne hint is that Lebesgue theory to begin with entails that when (zz) holds then there exists a seqience of symmwtric and continuous kernel functions $\{K_n(,y)\}$ such that

$$\iint_\square |K(x,y) - K(x,y)|^2 \, dxdy = 0$$

Next, the Cauchy-Schwarz inequality gives

$$||\mathcal{K} - \mathcal{K}_n|| \leq \iint_\square |K(x,y) - K(x,y)|^2 \, dxdy$$

for each n, where the right hand side refers to operator norms. Now one uses the general fact that a linear operator which can be approximated in the operator norm by compact operators is itself compact. Finally the reader should verify that if $K(x,y)$ is contoinous then $\mathcal{K}$ is compact.

Spectral functions. Let K be a symmetric Hilbert-Schmidt operator. A special case of Holbert's theorem from § 8 and the general facts about spectera of compact operators in § xx, entail that there exists a sequence $\{\phi_n\}$ of pairwise orthogonal functions in $\mathcal{H}$ with L^2-norms equal to one, a real eigenvalues $\{\mu_n\}$ so that

$$\mathcal{K}(\phi_n) = \mu_n \cdot \phi_n$$

The eigenvalues form a discrete set outside zero and arranged so that $\mu_1 \geq |\mu_2 \geq \ldots$ and repeated according to multiplicities, i.e. when the corresponding eigenspace has dimension ≥ 2. Set

$$\lambda_n = \mu_n^{-1}$$

For each $\lambda > 0$ we set

$$\rho(x,y;\lambda) = \sum_{0<\lambda_\nu<\lambda} \phi_\nu(x)\phi_\nu(y)$$

$$\rho(x,y;-\lambda) \sum_{-\lambda<\lambda_\nu<0} \phi_\nu(x)\phi_\nu(y)$$

Notice that the λ-numbers in (x) stay away from zero. Hence function ρ_N vanishes in a neighborhood of zero. The spectral theorem applied to symmetric Hilbert-Schmidt operators entails that

$$\mathcal{K}(h)(y) = \int_{-\infty}^{\infty} \frac{1}{\lambda} \cdot \frac{d}{d\lambda} \rho(x,y;\lambda) \cdot h(y) \, dy$$

hold for each $h \in \mathcal{H}$.

The Ω-kernel. If $h \in \mathcal{H}$ we have an expansion

$$h = \sum a_k \cdot \phi_k + h^*$$

where h^* is $|perp$ to the subspace of $\mathcal{H}$ generated by the eigenfunctions and the reader should check that

$$\sqrt{\sum |a_k|^2} \leq ||h||_2$$

Apply (x) to a pair h, g in $L^2[0,1]$. Then the reader should check that

$$|\iint_\square \rho(x,y;\lambda) \cdot h(x)g(y)\,dxdy| \leq ||h||_2 \cdot ||g||_2 \tag{xx}$$

hold for each λ which may be taken to be positive or negative. For each λ we set

$$\Omega(x,y;\lambda) = \int_a^x \int_a^y \rho(s,t;\lambda)\,dsdt$$

Exercise. Conclude from (xx) that the function

$$(x,y) \mapsto \Omega(x,y;\lambda)$$

is Hölder continuous of order 1/2. More precisely, for each λ and every pair $x_1 < x_2$ and $y_1 < y_2$ one has

$$|\Omega(x_2,y_2;\lambda) - \Omega(x_1,y_1;\lambda)| \leq \sqrt{(x_2-x_1)(y_2-y_1)}$$

A crucial point is that (e.1) holds for all Hilbert-Schmidt kernels, i.e. independent of the size of the Hilbert-Schmidt norms.

The ψ-functions. With λ and $0 < x \leq 1$ kept fixed we set

$$\psi(y;\lambda) = \int_0^x \rho(s,y;\lambda)\,ds$$

We have the characteristic function $\chi(x)$ defined as one on $[0,x_*]$ and zero if $x > x_*$. Considered as a vector in $\mathcal{H}$ one has an expansion

$$\chi(x) = \sum a_k\phi_k(x) + \chi_* \quad : \, a_k = \int_0^{x_*} \phi_k(s)\,ds$$

At the same time the construction of the ρ-function entails that

$$\psi(y;\lambda) = \sum_* \int_0^{x_*} \phi_k(s) \cdot \phi_k(y)$$

with the sum restricited over those k for which $0 < \lambda_k < \lambda$. Bessel's inequality gives

$$\int_0^1 \psi(y;\lambda)^2\,dy \leq \bigl(\sum_* \int_0^{x_*} \phi_k(s)\bigr)^2$$

In the last term we have taken a restricted sum which is majorsised by the sum over all k which again by Bessel's inequality is majorised by the L^2-integral of $|chi$, i.e by x^2. Hnece one has the inequality

$$\int_0^1 \psi(y;\lambda)^2\,dy \leq x \quad : 0 < x \leq 1$$

Exercise. More generally, let E be a sum of disjoint intervals on the x-interval and put

$$\psi_E(y;\lambda) = \int_E \rho(x,y;\lambda)\,dx$$

Show that

$$\int_0^1 \psi_E(y;\lambda)^2\,dy \leq |E|_1$$

where the last term is the Lebesgue measure of E.

Weak limits. Let us first notice that the contruction of the ψ-function means that

$$\frac{\partial\Omega(x,y;\lambda)}{\partial y} = \psi_x(y;\lambda)$$

So (xx) entails that the Hölder contionius Ω-function has a partial y-derivative in the sense of distributions which belongs to L^2 which moreover is absolutely continuu as a fuinction of x since (xx) is constucted as a primitive function in the sense of Lebesgue. It follows that one also has

$$\frac{\partial}{\partial x}(\frac{\partial\Omega(x,y;\lambda)}{\partial y}) = \rho(x,y;\lambda)$$

Let us then consider a sequence of Hilbert-Schmidt kernels $\{K_n\}$ and to each of them we get the spectral function $\rho_n(x,y;\lambda)$. Now $\{\Omega_n(x,y;\lambda)$ is an equi-contonuous family of functions on $\square$. So by the Arzela-Ascli theorem we find a subsequence which converges uniformly to a limit function $\Omega_*(x,y;\lambda)$ for each fixed λ. From (*) Ω is again uniformly Hölder continuous. Moreover, the uniform bound for the L^2-norms of $y \mapsto \psi_x(y;\lambda)$ entail that the partial y-erivatives of Ω_* yield L^2-functions which are absoutely continuous with respect to x. So there exists an almost everywhere defined limit function

$$\rho_*(x,y;\lambda) = \frac{\partial}{\partial x}(\frac{\partial\Omega_*(x,y;\lambda)}{\partial y})$$

for which the inequality (xx) holds.

and its first order partial derivatives are L^2-functions. Moreover, after the passage to the limit one still has Bessel's inequality which entails that

$$\int_{-\infty}^{\infty} \frac{d}{d\lambda} \int_0^1 h(x) \cdot \big(\frac{\partial}{\partial x} \int_0^1 \frac{\partial \Omega(x,y;\lambda)}{\partial y} \cdot h(y))\, dy\big)\, dx \leq \int_0^1 h(x)^2\, dx$$

for every $h \in L^2[0,1]$.

In the monograph [Carelan 1923] the condition (*) is imposed while (**) need not be valid. Then we encounter an unbounded operator. But notice that if $u|in\mathcal{H}$ then (*9 and the CuachySchwarz inequality entails that the functions

$$y \mapsto K(x,y)u(y)$$

are absolutely integrable in Lebegue's sense for all x outside the nullset $\mathcal{N}$ above. Hence it makes sense to refer to L^2-functions ϕ on $[0,1]$ which satisfy an eigenvalue equation

$$\phi(x) = \lambda \cdot \int_0^1 K(x,y)\phi(y)\,dy \tag{1}$$

where λ is a compkex number. one is also led to consider the integral equation

$$\phi(x) = \lambda \cdot \int_0^1 K(x,y)\phi(y)\,dy + f(x) \tag{2}$$

where $f \in \mathcal{H}$ is given and one seeks ϕ. It turns out that if $\mathfrak{Im}(\lambda) \neq 0$, then the inhomogenous equation (2) has at least one solution. Next, consider the equation (1). Let $\rho(\lambda)$ denote the number of linearly independent solution in (1). Carelan proived that the ρ-function is constant when λ varies in $\mathbf{C} \setminus \mathbf{R}$. Follwing Carelan one says that the kernel (x,y) gives an operator $\mathcal{K}$ of Class I if the ρ-function is zero. It means that (2) has unique solutions for every pair $\lambda, f)$ when λ are non-real.

A limit process. Let $\{G_n(x,y)\}$ be a sequece of symmetric kernel functions which are bounded in Hilbert's sense and approximate K in the sense that

$$\lim_{n\to\infty} \int_0^1 |K(x,y) - G_n(x,y)|^2\,dy = 0$$

for all x outside a null set. Fix some non-real λ. Hilbert's theory entials that if $f(x)$ is a continuous function, in general complex-valued, then we find unqiue continuous functions $\{\phi_n\}$ which satisfy the integral equations

$$\phi_n(x) = \lambda \cdot \int_0^1 G_n(x,y)\phi(y)\,dy \tag{1}$$

For a fixed x we consider the complex numbers $\{phi_n(x)\}$. Now there exists the set $Z(x)$ of all cluster points, i.e. a complex number z belongs to $Z(x)$ if there exists some sequence $1 \leq n_1 < n_2 < \ldots$ such that

$$z = \lim_{k\to\infty} \phi_{n_k}(x)$$

Theorem. *For every approximating sequence $\{G_n\}$ as above the sets $Z(x)$ are either reduced to points or circles in the complex plane. Moreover, each $Z(x)$ is reduced to a siungleton set when $\mathcal{K}$ is of class I.*

Remark. The result below was discovered by Weyl for some special unbounded operators which arise during the study of second order differential equations. See § for a comment-. If $\mathcal{K}$ is of Class I then the theorem sbove shows that each pair of a on-real λ and some $f \in \mathcal{H}$ gives a unique function $\phi(x)$ which satsfies (2) and it can be found via a pointwise limit of solutions $\{\phi_n\}$ to the equations (xx). In this sense the limit process is robust because one can emply an arbitrary approximating sequence $\{G_n\}$ under the sole condition that (xx) holds. This already indficartes that the Case I leads to a "consistent theory" even if $\mathcal{K}$ is unbounded. To make this precise Carleman constructed a unique spectral function when Case I holds. More preciely, if $K(x,y)$ is symmetric and Case I holds, then there exists a unique function $\rho(x,y;\lambda)$ defined for $(x,y) \in \square$ and every real λ such that

$$\mathcal{K}(h)(x) = \int_{-\infty}^{\infty} \frac{1}{\lambda} \cdot \frac{d}{d\lambda} \int_0^1 \theta(x,y;\lambda)\,h(y)\,dy$$

for all L^2-functions h. Concerning the spectral θ-function, Carleman proved that it enjoys the same properties as Hilbert's spectral function for bounded operators.

Remark. For each fixed $\lambda > 0$ one has a bounded self-adjoint operator on $\mathcal{H}$ defined by

$$\Theta_\lambda(h)(x) = \int_0^1 \theta(x, y; \lambda)\, h(y)\, dy$$

Moreover, the operator-valued function $\lambda \mapsto \Theta_\lambda$ has bounded variation over each interval $\{a \leq \lambda \leq b\}$ when $0 < a < b$. It means that there exists a constant $C = C(a, b]$ such that

$$\max \sum_{k=0}^{M} ||\Theta_{\lambda_{k+1}} - \Theta_{\lambda_k}|| \leq C$$

for all partitions $a = \lambda_0 < \lambda_1 < \ldots \lambda_{M+1} = b$ and we have taken operator norms of the differences $\Theta_{\lambda_{k+1}} - \Theta_{\lambda_k}$ in the sum above. These bounded variations entail that one can compute the integrals in the right hand side via the usual method by Borel and Stieltjes.

The construction of spectral functions. When case I holds one constructs the ρ-function via a robust limit process. Following [ibid:Chapitre 4] we expose this in § xx. The strategy is to regard an approximating sequence $\{G_n\}$ of Hilbert-Schmidt operator, i.e.

$$\iint_\square G_N(x, y)^2\, dxdy < \infty$$

hold for each N. In this case $\{G_N\}$ are compact opertors. With N fixed we get a discrete sequence of non-zero real numbers $\{\lambda\nu\}$ which are arragned with increasing avbsolute values and an orthonormal family of eigenfunctions $\{\phi_\nu^{(n)}\}$ where

$$G_N(\phi_\nu^{(N)}) = \lambda nu \cdot \phi_\nu^{(N)}$$

hold for each ν. Of course, the eigenvalues also depend on N. If $\lambda > 0$ we set

$$\rho_N(x, y; \lambda) = \sum_{0<\lambda_\nu<\lambda} \phi_\nu(x)\phi_\nu(y)$$

$$\rho_N(x, y; -\lambda) \sum_{-\lambda<\lambda_\nu<0} \phi_\nu(x)\phi_\nu(y)$$

let us notice that for each fixed n, the λ-numbers in (x) stay away from zero, i.e. there is a constant $c_N > 0$ such that $|\lambda_\nu \geq c_N$. So the function ρ_N vanishes in a neighborhood of zero. The spectral theorem applied to symmetric Hilbert-Schmidt operators entails that

$$\int_{-\infty}^{\infty} \frac{1}{\lambda} \cdot \frac{d}{d\lambda}\, \rho_N(x, y; \lambda) \cdot h(y)\, dy$$

The Ω-kernels. For each N we introduce the function

$$\Omega_N(x, y; \lambda) = \int_a^x \int_a^y \rho_N(s, t; -\lambda)\, dsdt$$

Let us notice that Bessel's inequality entails that

$$|\iint_\square \rho_N(x, y; \lambda) \cdot h(x)g(y)\, dxdy\big| \leq ||h||_2|cdot||g|7_2$$

for each pair h, g in $L^2[0, 1]$.

Exercise. Conclude from the above that the variation of Ω over $[0, x] \times [0, y]$ is bounfed above by

$$\sqrt{x \cdot y}$$

for every pair $0 \leq x, y \leq 1$. In partcular the functions

$$(x, y) \mapsto \Omega(x, y; \lambda)$$

are uniformly Hölder continuous of order $1/2$ in x and y respectively.

Using the inequalities above we elave it to the reader to check that there exists at least one subsequence $\{N_k\}$ such that the functions $\{\Omega_{N_k}(x, y; \lambda)$ converges uniformly with respect to x and y while λ stays in a bounded interval. When Case I holds one proves that the limit is independent of the subsequence, i.e. there exists a limit function

$$\Omega(x, y; \lambda) = \lim_{N \to \infty} \Omega_N(x, y; \lambda)$$

From the above the ω-function is again uniformly Hölder continuous and its first order partial derivatives are L^2-functions. Moreover, after the passage to the limit one still has Bessel's inequality which entails that

$$\int_{-\infty}^{\infty} \frac{d}{d\lambda} \int_0^1 h(x) \cdot \big(\frac{\partial}{\partial x} \int_0^1 \frac{\partial \Omega(x, y; \lambda)}{\partial y} \cdot h(y)) \, dy\big) \, dx \leq \int_0^1 h(x)^2 \, dx$$

for every $h \in L^2[0, 1]$.

Definition. *A Case 1 kernel $K(x, y)$ is closed if equality holds in (*).*

Theorem. *A Case I kernel $K(x, y)$ for which the equation*

$$\int_0^1 K(x, y) \cdot \phi(y) \, dy = 0$$

has no non-zero L^2-solution ϕ is closed.

A representation formula. When (*) hols we can apply to to $h + g$ for every pair of L^2-functiuons anbd since the Hilbert space $L^2[0, 1]$ is self-dual it follows that for each $f \in L^2[0, 1]$ one has the equality

$$f(x) = \frac{d}{dx} \int_{-\infty}^{\infty} \frac{d}{d\lambda} \int_0^1 \frac{\partial \Omega(x, y; \lambda)}{\partial y} \cdot f(y)) \, dy$$

almost everywhere with respect to x on $[0, 1]$.

Remark. The formula (**) shows that it often is important to decide when a Case I kernel is closed. Theorem xx gives such a sufficent condition. However, it can be extended to a quite geneeral result where one for can relax the passage to the limit via suitable linear operators. The reader may consult [ibid; page 139-143] for details. Here we are content to announce the conlsisive result ehich appear in [ibid;: page 142].

The $\mathcal{L}$-family. Let ξ dentoe a paramter which in general depends on several variables, or represents points in a manifold or vectors in a normed linear space. To each ξ we are given a linear operator

$$L(\xi) \colon f \mapsto L(\xi)(f)$$

from functions $f(x)$ on $[0, 1]$ to new functions on $[0, 1]$. The linear map is weakly continuous, i.e. if $\{f_\nu\}$ is a sequence in $L^2[0, 1]$ which converges weakly to a init function f in $L^2[0, 1]$ then

$$\lim L(\xi, f_\nu) \xrightarrow{w} L(\xi)(f)$$

Next, we are given $K(x, y)$ and the second condition for L to be in $\mathcal{L}$ is that for each pair ξ and $0 \leq y \leq 1$ there exists a constant $\gamma(\xi, y)$ which is independent of δ so that

$$|L(\xi)(K_\delta(\cdot, y))(x) \leq \gamma(\xi; y)$$

where $K_\delta(x, y)$ is the truncated kernel function from (xx) and in the left hand side we have applied $L(\xi$ to the function $x \mapsto K_\delta(x, y)$ for each fixed y. Moreover, we have

$$\lim_{\delta \to 0} L(\xi)(K(\cdot, y))(x) \xrightarrow{w} \lim_{\delta \to 0} L(\xi)(K_\delta(\cdot, y))(x)$$

where the convergence again holds weakly for $L2$-functions on $[0, 1]$. Funally the equality below holds for every L^2-function ϕ:

$$xxxx$$

Symmetric integral operators.

Consider the domain $\square = \{0 \leq x \leq 1\} \times \{0 \leq y \leq 1\}$ in $\mathbf{R}^2$. Let $K(x,y)$ be a real-vaued and Lebesgue mesure function on $\square$ such that the integrals

$$\int_0^1 K(x,y)^2\, dy < \infty$$

for all x outside a null-set on $[0,1]$. In addition K is symmetric, i.e.$K(x,y) = K(y,x)$. The K-kernel is bounded in Hilbert's sens if there exists a constant C such that

$$\iint_\square K(x,y)u(x)u(y)\, dxdy \leq C^2 \cdot \int_0^1 u(x)^2\, dx$$

for each $u \in L^2[0,1]$. In the text-book —emphIntyergralxxx [xxxx 1904], Hilbert proved that the linear operator on $L^2[0,1]$ defined by

$$\mathcal{K}(u)(x) = \int_0^1 K(x,y)u(y)dy$$

is bounded and its operator norm is majorsied by C-Moreover the spectrum is real and contained in $[-C,C]$ and just as for symmetric real matrices there exists a spectral resolution. More precisely, set $\mathcal{H} = L^2[0,1]$. Then there exists a map from $\mathcal{H} \times \mathcal{H}$ to the space of Riesz measures supported by $[-C,C]$ which to every pair u and v in $\mathcal{H}$ assigns a Riesz measure $\mu_{\{u,v\}}$ and $\mathcal{K}$ is recovered by the equation:

$$\langle \mathcal{K}(u), v\rangle = \int_{-C}^{C} t \cdot d\mu_{\{u,v\}}(t)$$

where the left hand side is the inner product on the complex Hilbert space defined by

$$\iint \mathcal{K}(u)(x) \cdot \overline{v}(x)\, dx$$

In the monograph [Carelan 1923] the condition (*) is imposed while (**) need not be valid. Then we encounter an unbounded operator. But notice that if $u|in\mathcal{H}$ then (*9 and the CuachySchwarz inequality entails that the functions

$$y \mapsto K(x,y)u(y)$$

are absolutely integrable in Lebegue's sense for all x outside the nullset $\mathcal{N}$ above. Hence it makes sense to refer to L^2-functions ϕ on $[0,1]$ which satisfy an eigenvalue equation

$$\phi(x) = \lambda \cdot \int_0^1 K(x,y)\phi(y)\, dy \tag{1}$$

where λ is a compkex number. one is also led to consider the integral equation

$$\phi(x) = \lambda \cdot \int_0^1 K(x,y)\phi(y)\, dy + f(x) \tag{2}$$

where $f \in \mathcal{H}$ is given and one seeks ϕ. It turns out that if $\mathfrak{Im}(\lambda) \neq 0$, then the inhomogenous equation (2) has at least one solution. Next, consider the equation (1). Let $\rho(\lambda)$ denote the number of linearly independent solution in (1). Carelan proived that the ρ-function is constant when λ varies in $\mathbf{C} \setminus \mathbf{R}$. Follwing Carelan one says that the kernel (x,y) gives an operator $\mathcal{K}$ of Class I if the ρ-function is zero. It means that (2) has unique solutions for every pair $\lambda, f)$ when λ are non-real.

A limit process. Let $\{G_n(x,y)\}$ be a sequece of symmetric kernel functions which are bounded in Hilbert's sense and approximate K in the sense that

$$\lim_{n\to\infty} \int_0^1 |K(x,y) - G_n(x,y)|^2\, dy = 0$$

for all x outside a null set. Fix some non-real λ. Hilbert's theory entials that if $f(x)$ is a continuous function, in general complex-valued, then we find unqiue continuous functions $\{\phi_n\}$ which satisfy the integral equations

$$\phi_n(x) = \lambda \cdot \int_0^1 G_n(x,y)\phi(y)\,dy \tag{1}$$

For a fixed x we consider the complex numbers $\{phi_n(x)\}$. Now there exists the set $Z(x)$ of all cluster points, i.e. a complex number z belongs to $Z(x)$ if there exists some sequence $1 \leq n_1 < n_2 < \ldots$ such that

$$z = \lim_{k\to\infty} \phi_{n_k}(x)$$

Theorem. *For every approximating sequence $\{G_n\}$ as above the sets $Z(x)$ are either reduced to points or circles in the complex plane. Moreover, each $Z(x)$ is reduced to a siungleton set when $\mathcal{K}$ is of class I.*

Remark. The result below was discovered by Weyl for some special unbounded operators which arise during the study of second order differential equations. See § for a comment-. If $\mathcal{K}$ is of Class I then the theorem sbove shows that each pair of a on-real λ and some $f \in \mathcal{H}$ gives a unique function $\phi(x)$ which satsfies (2) and it can be found via a pointwise limit of solutions $\{\phi_n\}$ to the equations (xx). In this sense the limit process is robust because one can emply an arbitrary approximating sequence $\{G_n\}$ under the sole condition that (xx) holds. This already indficartes that the Case I leads to a "consistent theory" even if $\mathcal{K}$ is unbounded. To make this precise Carleman constructed a unique spectral function when Case I holds. More preciely, if $K(x,y)$ is symmetric and Case I holds, then there exists a unique function $\rho(x,y;\lambda)$ defined for $(x,y) \in \square$ and every real λ such that

$$\mathcal{K}(h)(x) = \int_{-\infty}^{\infty} \frac{1}{\lambda} \cdot \frac{d}{d\lambda} \int_0^1 \theta(x,y;\lambda)\,h(y)\,dy$$

for all L^2-functions h. Concerning the spectral θ-function, Carleman proved that it enjoys the same properties as Hilbert's spectral function for bounded operators.

Remark. For each fixed $\lambda > 0$ one has a bounded self-adjoint operator on $\mathcal{H}$ defined by

$$\Theta_\lambda(h)(x) = \int_0^1 \theta(x,y;\lambda)\,h(y)\,dy$$

Moreover, the operator-valued function $\lambda \mapsto \Theta_\lambda$ has bounded variation over each interval $\{a \leq \lambda \leq b\}$ when $0 < a < b$. It means that there exists a constant $C = C(a,b]$ such that

$$\max \sum_{k=0}^{M} ||\Theta_{\lambda_{k+1}} - \Theta_{\lambda_k}|| \leq C$$

for all partitions $a = \lambda_0 < \lambda_1 < \ldots \lambda_{M+1} = b$ and we have taken operator norms of the differences $\Theta_{\lambda_{k+1}} - \Theta_{\lambda_k}$ in the sum above. These bounded variations entail that one can compute the integrals in the right hand side via the usual method by Borel and Stieltjes.

The construction of spectral functions. When case I holds one constructs the ρ-function via a robust limit process. Following [ibid:Chapitre 4] we expose this in § xx. The strategy is to regard an approximating sequence $\{G_n\}$ of Hilbert-Schmidt operator, i.e.

$$\iint_\square G_N(x,y)^2\,dxdy < \infty$$

hold for each N. In this case $\{G_N\}$ are compact opertors. With N fixed we get a discrete sequence of non-zero real numbers $\{\lambda\nu\}$ which are arragned with increasing avbsolute values and an orthonormal family of eigenfunctions $\{\phi_\nu^{(n)}\}$ where

$$G_N(\phi_\nu^{(N)}) = \lambda nu \cdot \phi_\nu^{(N)}$$

hold for each ν. Of course, the eigenvalues also depend on N. If $\lambda > 0$ we set

$$\rho_N(x,y;\lambda) = \sum_{0<\lambda_\nu<\lambda} \phi_\nu(x)\phi_\nu(y)$$

$$\rho_N(x,y;-\lambda) \sum_{-\lambda<\lambda_\nu<0} \phi_\nu(x)\phi_\nu(y)$$

let us notice that for each fixed n, the λ-numbers in (x) stay away from zero, i.e. there is a constant $c_N > 0$ such that $|\lambda_\nu \geq c_N$. So the function ρ_N vanishes in a neighborhood of zero. The spectral theorem applied to symmetric Hilbert-Schmidt operators entails that

$$\int_{-\infty}^{\infty} \frac{1}{\lambda} \cdot \frac{d}{d\lambda}\, \rho_N(x,y;\lambda) \cdot h(y)\, dy$$

The Ω-kernels. For each N we introduce the function

$$\Omega_N(x,y;\lambda) = \int_a^x \int_a^y \rho_N(s,t;-\lambda)\, dsdt$$

Let us notice that Bessel's inequality entails that

$$|\iint_\square \rho_N(x,y;\lambda) \cdot h(x)g(y)\, dxdy\big| \leq ||h||_2 |cdot||g|7_2$$

for each pair h, g in $L^2[0,1]$.

Exercise. Conclude from the above that the variation of Ω over $[0,x] \times [0,y]$ is bounfed above by

$$\sqrt{x \cdot y}$$

for every pair $0 \leq x, y \leq 1$. In partcular the functions

$$(x,y) \mapsto \Omega(x,y;\lambda)$$

are uniformly Hölder continuous of order $1/2$ in x and y respectively.

Using the inequalities above we elave it to the reader to check that there exists at least one subsequence $\{N_k\}$ such that the functions $\{\Omega_{N_k}(x,y;\lambda)$ converges uniformly with respect to x and y while λ stays in a bounded interval. When Case I holds one proves that the limit is independent of the subsequence, i.e. there exists a limit function

$$\Omega(x,y;\lambda) = \lim_{N\to\infty} \Omega_N(x,y;\lambda)$$

From the above the ω-function is again uniformly Hölder continuous and its first order partial derivatives are L^2-functions. Moreover, after the passage to the limit one still has Bessel's inequality which entails that

$$\int_{-\infty}^{\infty} \frac{d}{d\lambda} \int_0^1 h(x) \cdot \big(\frac{\partial}{\partial x} \int_0^1 \frac{\partial \Omega(x,y;\lambda)}{\partial y} \cdot h(y))\, dy\big)\, dx \leq \int_0^1 h(x)^2\, dx$$

for every $h \in L^2[0,1]$.

Definition. *A Case 1 kernel $K(x,y)$ is closed if equality holds in (*).*

Theorem. *A Case I kernel $K(x,y)$ for which the equation*

$$\int_0^1 K(x,y) \cdot \phi(y)\, dy = 0$$

has no non-zero L^2-solution ϕ is closed.

A representation formula. When (*) hols we can apply to to $h+g$ for every pair of L^2-functiuons anbd since the Hilbert space $L^2[0,1]$ is self-dual it follows that for each $f \in L^2[0,1]$ one has the equality

$$f(x) = \frac{d}{dx} \int_{-\infty}^{\infty} \frac{d}{d\lambda} \int_0^1 \frac{\partial \Omega(x,y;\lambda)}{\partial y} \cdot f(y))\, dy$$

almost everywhere with respect to x on $[0,1]$.

Remark. The formula (**) shows that it often is important to decide when a Case I kernel is closed. Theorem xx gives such a sufficent condition. However, it can be extended to a quite geneeral result where one for can relax the passage to the limit via suitable linear operators. The reader may consult [ibid; page 139-143] for details. Here we are content to announce the conlsisive result ehich appear in [ibid;: page 142].

The $\mathcal{L}$-family. Let ξ dentoe a paramter which in general depends on several variables, or represents points in a manifold or vectors in a normed linear space. To each ξ we are given a linear operator

$$L(\xi)\colon f \mapsto L(\xi)(f)$$

from functions $f(x)$ on $[0,1]$ to new functions on $[0,1]$. The linear map is weakly continuous, i.e. if $\{f_\nu\}$ is a sequence in $L^2[0,1]$ which converges weakly to a init function f in $L^2[0,1]$ then

$$\lim L(\xi, f_\nu) \xrightarrow{w} L(\xi)(f)$$

Next, we are given $K(x,y)$ and the second condition for L to be in $\mathcal{L}$ is that for each pair ξ and $0 \leq y \leq 1$ there exists a constant $\gamma(\xi,y)$ which is independent of δ so that

$$|L(\xi)(K_\delta(\cdot,y))(x) \leq \gamma(\xi;y)$$

where $K_\delta(x,y)$ is the truncated kernel function from (xx) and in the left hand side we have applied $L(\xi$ to the function $x \mapsto K_\delta(x,y)$ for each fixed y. Moreover, we have

$$\lim_{\delta\to 0} L(\xi)(K(\cdot,y))(x) \xrightarrow{w} \lim_{\delta\to 0} L(\xi)(K_\delta(\cdot,y))(x)$$

where the convergence again holds weakly for $L2$-functions on $[0,1]$. Funally the equality below holds for every L^2-function ϕ:

$$xxxx$$

11.5 Stones theorem.

11.3.2 Unitary semi-groups. Specialize the situation above to the case when B is a Hilbert space $\mathcal{H}$ and $\{U_t\}$ are unitary operators. Set $T = \xi_*$ so that

$$B(x) = \lim_{t \to 0} \frac{U_t x - x}{t} \quad : \quad x \in \mathcal{D}(T)$$

If x, y is a pair in $\mathcal{D}(B)$ we get

$$\langle Bx, y\rangle = \lim_{t \to 0} \langle \frac{U_t x - x}{t}, U_t y\rangle = \lim_{t \to 0} \frac{1}{t} \cdot \big(\langle U_t x, U_t y)\rangle - \langle x, U_t y)\rangle\,\big)$$

Since U_t are unitary we have $\langle U_t x, U_t y)\rangle = \langle x, y)\rangle$ for each t and conclude that the last term above is equal to

$$\lim_{t \to 0} \frac{\langle x, y - U_t y\rangle}{t} = -\langle x, By\rangle$$

Hence B is anti-symmetric, i.e.

$$\langle Bx, y\rangle = -\langle x, By\rangle$$

Set $A = i \cdot T$ which gives

$$\langle Ax, y\rangle = i \cdot \langle Tx, y\rangle = -i \cdot \langle x, Ty\rangle = -\langle x, i \cdot Ty\rangle = \langle x, Ay\rangle$$

where we used that the inner product is hermitian. Hence A is a densely defined and symmetric operator.

Theorem. *A is self-adjoint, i.e. one has the equality $\mathcal{D}(A) = \mathcal{D}(A^*)$.*

Proof. It suffices to prove that $\mathcal{D}(T) = |mathcalD(T^*)$. To btain this we take a vector y be a vector in $\mathcal{D}(T^*)$ which by definition gives a constant $C(y)$ such that

(i) $$|\langle Tx, y\rangle|| \leq C(y) \cdot ||x|| \quad : x \in \mathcal{D}(T)$$

Now $\langle Tx, y\rangle$ is equal to

(ii) $$\lim_{t \to 0} \langle \frac{U_t x - x}{t}, U_t y\rangle = -\lim_{t \to 0} \langle x, \frac{U_t y - y}{t}\rangle$$

So if

$$\eta_t = fracU_t y - yt$$

it follows that there exists

$$\lim_{t \to 0}, \langle x, \eta_t\rangle = -\langle Tx, y\rangle$$

for each $x \in \mathcal{D}(T)$.

The adjoint operators $\{U_t^*\}$ give another unitary semi-group with infinitesmal generator A_* where

$$A_*(x) = \lim_{t \to 0} \frac{U_t^* x - x}{t} \quad : \quad x \in \mathcal{D}(A_*)$$

Since U_t is the inverse operator of U_t^* for each t we get

(i) $$A_*(x) = \lim_{t \to 0} U_t(A_* x) = \lim_{t \to 0} \frac{x - U_t x}{t} \quad : \quad x \in \mathcal{D}(A_*)$$

From (i) we see that $\mathcal{D}(A) \subset \mathcal{D}(A_*)$ and one has the equation

(ii) $$A_* x) = -A(x) \quad : \quad x \in \mathcal{D}(A_*)$$

Reversing the role the reader can check the equality

(iii) $$\mathcal{D}(A) = \mathcal{D}(A_*)$$

Next, let x, y be a pair in $\mathcal{D}(A)$. Then

$$\langle Ax, y\rangle = \lim_{t \to 0} \langle \frac{U_t x - x}{t}, U_t y\rangle = \lim_{t \to 0} \frac{1}{t} \cdot \big(\langle U_t x, U_t y)\rangle - \langle x, U_t y)\rangle\,\big)$$

Since U_t are unitary we have $\langle U_t x, U_t y)\rangle = \langle x, y)\rangle$ for each t and conclude that the last term in (xx) is equal to

$$\lim_{t\to 0} \frac{\langle x, y - U_t y\rangle}{t} = -\langle x, Ay\rangle$$

Hence A is anti-symmetric. Set

$$B = iA$$

Exercise. Show that (i) gives the equality $\mathcal{D}(A_*) = \mathcal{D}(A)$ and that:

$$A_*(x) = -A(x) \quad : \quad x \in \mathcal{D}(A)$$

$$\langle Bx, y\rangle = -\langle x, By\rangle \quad : \quad x, y \in \mathcal{D}(B)$$

Exercise. Conclude from the above that the operator iB is self-adjoint.

Remark. The equations above constitute Stones theorem which was established in 1930. It has a wide range of applications. See for example von Neumann's article *Zur Operatorenmethode in der klassischen Mekanik* and Maeda's article *Unitary equivalence of self-adjoint operators and constant motion* from 1936.

11.3.3 A converse construction. Let A be a densely defined self-adjoint operator. From § 9.B A is approximated by a sequence of bounded self-adjoint operators $\{A_N\}$. With N kept fixed we get a semi-group of unitary operators where

$$U_t^{(N)} = e^{-itA_N}$$

The reader may verify that the inifintesmal generator becomes $-iA_N$. Next, for each $x \in \mathcal{D}(A)$ and every fixed t there exists the limit

$$\lim_{N\to\infty} U_t^{(N)}(x)$$

Remark. This gives a densely defined linear operator U_t whose operator norm is bounded by one which therefore extends uniquely to a bounded linear operator $\mathcal{H}$ and it is clear that this extension becomes a unitary operator. In this way we arrive at a semi-group $\{U_t\}$ and one verifies that its infinitesmal generator is equal to $-iA$. However, it is not clear that $\{U_t\}$ is strongly continuous and one may ask for conditions on the given self-adjoint operator A which ensures that $\{U_t\}$ is strongly continuous.

11.8 Transition probability functions.

Let $(\Omega, \mathcal{B}, \mu)$ be a probability space as defined in § XX. Consider a real-valued function P defined on the product set $\Omega \times \mathcal{B}$ with the following two properties:

$$(*) \qquad t \mapsto P(t, E) \quad \text{is a bounded measurable function for each} \quad E \in \mathcal{B}$$

$$(**) \qquad E \mapsto P(t, E) \quad \text{is a probability measure for each fixed} \quad t \in \Omega$$

When (*-**) hold one refers to P as a transition function. Given P we define inductively the sequence $\{P^{(n)}\}$ by:

$$P^{(n+1)}(t, E) = \int_\Omega P^{(n)}(s) \cdot dP(s, E)$$

It is clear that $\{P^{(n)}\}$ yield new transition functions. The probabilistic interpretation is that one has a Markov chain with independent increments. More precisely, if E and S are two sets in $\mathcal{B}$ and $n \geq 1$. then

$$\int_S P^{(n)}(t; E) \cdot d\mu(t)$$

is the probability that the random walk which starts at some point in E has arrived to some point in S after n steps. One says that the given transition function P yields a stationary Markov process if there exists a finite family of disjoint subets $E_1, \ldots, E_m$ in $\mathcal{B}$ and some $\alpha < 1$ and a constant M such that the following hold: First, for each $1 \leq i \leq m$ one has:

$$(1) \qquad P(t, E_i) = 1 \quad : t \in E_i$$

Next, if $\Delta = \Omega \setminus E_1 \cup \ldots \cup E_m$ then

$$(2) \qquad \sup_{t \in S} P^{(n)}(t, \Delta) \leq M \cdot \alpha^n \quad : \ n = 1, 2, \ldots$$

Remark. One refers to Δ as the disippative part of Ω and $\{E_i\}$ are the ergodic kernels of the process. Since $\alpha < 1$ in (2) the probabilistic interpretation of (2) is that as n increase then the disippative part is evacuated with high probability while the Markov process stays inside every ergodic kernel.

11.9 The Kakutani-Yosida theorem.

A sufficient condition for a Markov process to be stationary is as follows: Denote by X the Banach space of complex-valued and bounded $\mathcal{B}$-measurable functions on the real s-line. Now P gives a linear operator T which sends $f \in X$ to the function

$$T(f)(x) = \int_\Omega f(s) \cdot dPx, ds)$$

Kakutani and Yosida proved that the Markov process is stationary if there exists a triple α, n, K where K is a compact operator on X, $0 < \alpha < 1$ and n some positive integer such that the operator norm

$$(*) \qquad ||T + K|| \leq \alpha$$

The proof relies upon some general facts about linear operators on Banach spaces. First one identifies the Banach space X with the space of continuous complex-valued functions on the compact Hausdorff space S given by the maximal ideal space of the commutative Banach algebra X. Then T is a postive linear operator on $C^0(S)$ and in § 11.xx we shall prove that (*) implies that the spectrum of T consists of a finite set of points on the unit circle together with a compact subset in a disc of radius < 1. Moreover, for each isolated point $e^{i\theta} \in \sigma(T)$ the corresponding eigenspace is finite dimensional. Each such eigenvalue corresponds to an ergodic kernel and when the eigenspace has dimension $m \geq 2$, the corresponding ergodic kernel, say E_1, has a further decomposition into pairwise disjoint subsets $e_1,\ldots,e_m$ where the process moves in a cyclic manner between these sets, i.e.

$$\int_{e_{i+1}} P(s,e_i)=1 \quad : 1\leq i\leq m \quad \text{where we put} \quad e_{m+1}=e_1$$

11.9.1 Some results about linear operators.

The Kakutani-Yosida theorem follows from some results about linear operators which we begin to expose. Let X be a Banach space and denote by $\mathcal{F}$ the family of bounded liner operators T on X such that

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

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

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

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

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

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

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

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

3. Proposition. *If $T\in\mathcal{F}_*$ and $e^{i\theta}\in\sigma(T)$ for some θ, then $R_T(\lambda)$ has a simple pole at $e^{i\theta}$.*

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

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

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

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

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

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

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

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

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

$$\text{(i)} \qquad R_S(\lambda)(\lambda\cdot E-T)=E+R_S(\lambda)\cdot K$$

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

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

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

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

Together with (ii) we obtain

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

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

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

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

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

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

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

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

$$\frac{A_{-1}}{z} + A_0 + A_1 z + \ldots = B(z) + B(z)(E + B(z) \cdot K)^{-1} B(z)$$

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

$$(E + B(z) \cdot K)^{-1} = \frac{A^*_{-1}}{z} + A^*_0 + A^*_1 z + \ldots$$

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

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

Next, from (xx) one has

$$E = (E + B(z) \cdot K)(\frac{A^*_{-1}}{z} + A^*_0 + A^*_1 z + \ldots) \implies (E + B(0) \cdot K) A^*_{-1} = 0$$

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

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

Proof. We have the algebraic equation

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

It follows that

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

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

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

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

A result about positive operators.

Let S be a compact Hausdorff space and X is the Banach space of continuous and complex-valued functions on S. A linear operator T on X is positive if it sends every non-negative and real-valued function f to another real-valued and non-negative function.

6. Theorem. *If T is positive and belongs to $\mathcal{F}_*$ then each $e^{i\theta} \in \sigma(T)$ is a root of unity.*

Proof. The hypothesis gives $f \in C^0(S)$ which is not identically zero and

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

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

$$f(s_0) = 1$$

The dual space X^* consists of complex-valued Riesz measures on X. If $n \geq 1$ we get the measure μ_n such that the following hold for every $g \in C^0(S)$:

$$\int_S g \cdot d\mu_n = T^n g(s_0) \geq 0$$

Since T^n is positive the integrals in the left hand side are ≥ 0 when g are real-valued and non-negative. This entails that each μ_n is a real-valued and non-negative measure. Next, for each n we put

$$A_n = \{e^{-in\theta} \cdot f \neq 1\} = \{\mathfrak{Re}(e^{-in\theta} f) < 1\} \tag{6.0}$$

where the last equality follows since the sup-norm of f is one. Now

$$0 = f(s_0) - e^{-in\theta} \cdot T^n f(s_0) = \int_S [1 - e^{-in\theta} f(s)] \cdot d\mu_n(s) \tag{6.1}$$

Taking real parts we get

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

By (6.0) the integrand above is non-negative and since the whole integral is zero it follows that

$$\mu_n(A_n) = \mu_n(\{\mathfrak{Re}(e^{-in\theta} < 1\} = 0 \tag{6.2}$$

Hence (6.2) A_n is a null set with respect to μ_n. Suppose now that there exists a pair $n \neq m$ such that

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

A point s_* in this non-empty intersection gives

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

At this stage we use the hypothesis that T belongs to $\mathcal{F}_*$ which by Proposition XX entails that the resolvent $R_T(\lambda)$ has at most a simple pole at $\lambda = 1$. With $\epsilon > 0$ the Neumann series expansion gives

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

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

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

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

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

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

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

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

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

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

Moreover, by (*) we also have

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

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

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

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