Appendix J

Functional Analysis

J.1 Introduction

introducing heavy mathematical tools, often unfamiliar to the average physicist

a lot of underlying structure is assumed, and taken granted for, even if the physicist is unaware of this. The physicist will not worry so much about the possibilities in the underlying structure if he obtains experimental results confirming all the assumptions made. Field theory used to make predictions with QED which agree with experiment with astounding accuracy - and this is even though the quantum field theory may not exist in a mathematically rigorous way.

in the absence of any experimental observation (at least for the moment) by having a consistent theory

to achieve certainty is to work at a high level of mathematical precision.

we search for a mathematical precision is that in quantum gravity

The study of the spaces of functions. applied mathematics, and in physics, engineering and statistics.

Our reason for studying functions is as a model of the physical world. In selecting a certain class of functions... need to find a balance

general theory of topological spaces. The need to study of ... comes from that the class of .. is too restrictive for our purposes

investigate infinite dimensional representations arise when we consider eigenvalues of operators with continuous spectra. These are already present in undergraduate quantum mechanics - position and momentum operators.

However, now the vector space is infinitely large, the analysis is much more involved and the intuitive expectations coming from the study of finite dimensional vector spaces have to be suitably modified.

A blow-by-blow account of much of the material presented in the *Part II Mathematical tools* [32] T. Thiemann, Introduction to Modern Canonical Quantum General Relativity, [gr-qc/0110034].

J.2 Finite Dimensional Vector Spaces

• Linear Operators transform one vector into another: $\hat{T}f = g$

$$\hat{T}(c_1 f_1 + c_2 f_2) = c_1(\hat{T}f_1) + c_2(\hat{T}f_2)$$

$\hat{T}$ is a linear transformation in the Hilbert space Examples:

1. multiplication
$$\hat{T}_1 = x: \quad \hat{T}_1 f(x) = x f(x)$$
2. differentiation
$$\hat{T}_2 = \frac{d}{dx}: \quad \hat{T}_2 f(x) = \frac{d}{dx} f(x)$$
3. combined multiplication and differentiation
$$\hat{T}_1\hat{T}_2 = x\frac{d}{dx}: \quad \hat{T}_1\hat{T}_2 \neq \hat{T}_2\hat{T}_1$$

J.2.1 Linear Operator and Functionals on Finite Dimensional Spaces

As we saw in section ?? linear operators can be represented in terms of matrices.

Linear functionals where dim$V = n$ and $\{e_1, \ldots, e_n\}$ is a basis for V, as before. These functionals constitute teh algebraic dual space V^* of V. For every such functional f and every $x = \sum \zeta_j e_j \in V$ we have

$$f(x) = f\left(\sum_{j=1}^{n} \xi_j e_j\right) = \sum_{j=1}^{n} \xi_j f(e_j) = \sum_{j=1}^{n} \xi_j \alpha_j \tag{J.1}$$

where

$$\alpha_j = f(e_j) \qquad j = 1, \ldots, n \tag{J.2}$$

and f is uniquely determined by its values α_j at the n basis vectors of V.

Conversely, every $n-$tuple of scalars $\alpha_1, \ldots, \alpha_n$ determines a linear functional on V.

Theorem J.2.1 *Let V be an $n-$dimensional vector space and $\{e_1, \ldots, e_n\}$ a basis for V. Then $\{f_1, \ldots, f_n\}$ given by*

$$f_k(e_j) = \delta_{jk} \tag{J.3}$$

a basis for the algebraic dual V^* of V, and $dimV^* = dimV = n$.

Proof: $\{f_1, \ldots, f_n\}$ is a linear independent set since

$$\sum_{k=1}^{n} \beta_k f_k(x) = 0 \tag{J.4}$$

as we can see by setting $x = e_j$,

$$\sum_{k=1}^{n} \beta_k f_k(e_j) = \sum_{k=1}^{n} \beta_k \delta_{jk} = \beta_j = 0.$$

We show that every $f \in V^*$ can be represented as a linear combination of elements of $\{f_1, \ldots, f_n\}$ in a unique way. We write $f(e_j) = \alpha_j$. We know that

$$f(x) = \sum_{j=1}^{n} \xi_j \alpha_j$$

for every $x \in V$. On the other hand

$$f_j(x) = f_j(\xi_1 e_1 + \cdots + \xi_n e_n) = \xi_j.$$

Together,

$$f(x) = \sum_{j=1}^{n} \alpha_j f_j(x).$$

Hence the unique representaion of the arbitrary linear functional f on V in terms of the functionals $f_1, \ldots, f_n$ is

$$f = \alpha_1 f_1 + \cdots + \alpha_n f_n.$$

□

J.3 Finite Hilbert Space.

For a finite-dimensional space every Hermitian or unitary operator can be described completely by its eigenvalues and eigenvectors.

For a finite-dimensional space the eigenvectors span the whole space.

This comprises the so-called spectral theorem.

J.3.1 The Hamilton-Cayley Theorem.

Let $\mathbf{A}$ be a square $N \times N$ matrix representing in some basis the operator A , and let λ be a parameter. The equation

$$\varphi(\lambda) := \det(\lambda \mathbf{E} - \mathbf{A}) = 0 \tag{J.5}$$

is called the characteristic equation of the operator A(or of the matrix $\mathbf{A}$). It is evident that $\psi(\lambda)$ is a polynomial of the Nth degree in λ with numerical coefficients the leading (that of λ^N being equal to 1

$$\varphi(\lambda) = \varphi_0 + \varphi_1 \lambda + \varphi_2 \lambda^2 + \cdots + \varphi_{N-1} \lambda^{N-1} + \varphi_N \lambda^N \tag{J.6}$$

The form of the characteristic (and thus the numerical values of the coefficients, ψ_i) does not depend on the choice of basis, since the determinant of a matrix, in our case, the matrix $\det(\lambda \mathbf{E} - \mathbf{A})$, is a scalar.

Replacing λ by the operator A we get the operator

$$A(\lambda) = A_0 + A_1 \lambda + A_2 \lambda^2 + \cdots + A_{N-1} \lambda^{N-1} + A_N \lambda^N \tag{J.7}$$

J.3.2 Projection Operators

In chapter 3 we saw that a projection operator is equivalent to operators that are self-adjoint and satify $P^2 = P$.

We say that a subspace $\mathcal{M}$ **reduces** a linear operator T if $T\psi$ is in $\mathcal{M}$ for every $\psi \in \mathcal{M}$ and $T\phi$ is in $\mathcal{M}^\perp$ for every ϕ in $\mathcal{M}^\perp$. Let P be the projection operator onto $\mathcal{M}$. We say that a subspace $\mathcal{M}$ is **invariant** under an operator if $T\psi$ is in $\mathcal{M}$ for every vector in $\mathcal{M}$.

Theorem J.3.1 *Let P be the projection operator onto the subspace $\mathcal{M}$. The following statements are equivalent:*

i) $\mathcal{M}$ reduces T;

ii) $PT = TP$

Proof: First we show i) implies ii). For any vector

$$\psi = \psi_{\mathcal{M}} + \psi_{\mathcal{M}^\perp}$$

we have

$$T\psi = T\psi_{\mathcal{M}} + T\psi_{\mathcal{M}^\perp}.$$

If $\mathcal{M}$ reduces T, then $T\psi_{\mathcal{M}}$ is in $\mathcal{M}$ and $T\psi_{\mathcal{M}^\perp}$ is in $\mathcal{M}^\perp$. Therefore

$$\begin{aligned} PT\psi &= PT(\psi_{\mathcal{M}} + \psi_{\mathcal{M}^\perp}) = PT\psi_{\mathcal{M}} \\ &= T\psi_{\mathcal{M}} = TP\psi_{\mathcal{M}} \\ &= TP\psi. \end{aligned} \tag{J.8}$$

Thus i) implies ii). We now show the converse. If ψ is in $\mathcal{M}$ and $PT = TP$, then $P\psi = \psi$ and

$$T\psi = TP\psi = PT\psi$$

so $T\psi$ is in $\mathcal{M}$. It is easy to see that $PT = TP$ and $(1-P)T = T(1-P)$ are equivalent. If ϕ is in $\mathcal{M}^\perp$ and $(1-P)T = T(1-P)$, then

$$T\phi = T(1-P)\phi = (1-P)T\phi$$

so $T\phi$ is in $\mathcal{M}^\perp$.

$\square$

Theorem J.3.2 *If a subspace is invariant under T and $T^\dagger$ then it reduces T.*

Proof: Let $\mathcal{M}$ be a subspace that is invariant under T and $T^\dagger$ and let P be the projection operator onto $\mathcal{M}$. Then $TP\psi$ is in $\mathcal{M}$, so

$$TP\psi = PTP\psi$$

for every vector ψ, or

$$TP = PTP.$$

Because $\mathcal{M}$ is also invariant under $T^\dagger$, we have

$$T^\dagger P = PT^\dagger P.$$

Taking the adjoint, we have

$$PT = PTP.$$

Threfore

$$PT = TP.$$

This implies that $\mathcal{M}$ reduces T by theorem J.3.1.

$\square$

Theorem J.3.3 *If P and Q are the projections on closed linear subspaces M and N of H, then $M \perp N$ if and only if $PQ = 0$ or equivalently $QP = 0$.*

Proof: First note that $PQ = 0$ is equivalent to $QP = 0$ through taking adjoints. If $M \perp N$, so that $N \subseteq M^\perp$, as $Q\psi$ is in N for every x we have $PQ\psi = 0$, or $PQ = 0$. Conversely, if $PQ = 0$ then for every ψ in N we have $P\psi = PQ\psi = 0$, so $N \subseteq M^\perp$ and $M \perp N$.

$\square$

We say that two projections P and Q are orthogonal if $PQ = 0$.

Theorem J.3.4 *If $P_1, P_2, \ldots, P_n$ are the projections on closed linear subspaces $M_1, M_2, \ldots, M_n$ of H, then $P = P_1 + P_2 + \cdots + P_n$ is a projection if and only if the P_i's are pairwise orthognal ($P_iP_j = 0$ whenever $i \neq j$). In this case P is the projection onto*

$$M = M_1 + M_2 + \cdots + M_n.$$

Proof: Since P is self-adjoint, to prove it is a projection, we need only show it is idempotent, that is, $P^2 = P$. As the P_i's are pairwise orthogonal,

$$\begin{aligned} P^2 &= (P_1 + P_2 + \cdots + P_n)(P_1 + P_2 + \cdots + P_n) \\ &= P_1^2 + P_2^2 + \cdots + P_n^2 \\ &= P_1 + P_2 + \cdots + P_n = P. \end{aligned} \tag{J.9}$$

Therefore P is a projection. Conversely, assume that P is idempotent. Let ψ be a vector in the range of P_i then

$$\begin{aligned} \|\psi\|^2 &= \|P_i\psi\|^2 \leq \sum_{j=1}^{n} \|P_j\psi\|^2 \\ &= \sum_{j=1}^{n} (P_j\psi, \psi) = (P\psi, \psi) \\ &= \|P\psi\|^2 \leq \|\psi\|^2 \end{aligned}$$

Given we started with $\|\psi\|^2$ and ended with $\|\psi\|^2$, equality must hold throught, so

$$\sum_{j=1}^{n} \|P_j\psi\| = \|P_i\psi\|$$

and

$$\|P_j\psi\| = 0 \;\text{ for }\; j \neq i.$$

Therefore the range of P_i is contained in the null space of P_j, that is, $M_i \subseteq M_j^{\perp}$, for every $j \neq i$. So $M_i \perp M_j$ whenever $j \neq i$, and by the previous theorem we conclude that the P_i's are pairwise orthogonal. We now prove the final statement. Denote the range of P

by Ran(P). First note that since $\|P\psi\| = \|\psi\|$ for every ψ in M_i, each M_i is contained in the range of P, and therefore

$$M \subseteq Ran(P). \tag{J.10}$$

Next, if ψ is in the range of P, then

$$\begin{aligned} \psi &= P\psi \\ &= P_1\psi + P_2\psi + \cdots P_n\psi \end{aligned}$$

is obviously in M, and so ψ is in M, hence

$$Ran(P) \subseteq M. \tag{J.11}$$

Comparing (J.10) and (J.11) implies $M = Ran(P)$.

$\square$

J.3.3 Spectral Theorem for Finite Spaces

An operator on H is said to be **normal** if it commutes with its adjoint, that is, $NN^\dagger = N^\dagger N$. Self-adjoint and unitary operators are examples of normal operators. The spectral theorem states that for each normal operator N on H has a spectral resolution, that is, there exist distinct complex numbers $\lambda_1, \lambda_2, \ldots, \lambda_m$ and non-zero pairwise orthogonal projections $P_1, P_2, \ldots, P_m$ such that $\sum_{i=1}^m P_i = I$, and

$$N = \sum_{i=1}^m \lambda_i P_i.$$

Before coming on to the spectral theorem, we prove results for normal operators.

Lemma J.3.5 *If T ia normal, then ψ is an eigenvector of T with eigenvalue λ if and only if ψ is an eigenvector of $T^\dagger$ with eigenvalue λ^*.*

Proof:

Firsrt we show that an operator T is normal if and only if $\|T^\dagger\psi\| = \|T\psi\|$ for every ψ. Obviously $\|T^\dagger\psi\| = \|T\psi\|$ is equivalent to $\|T^\dagger\psi\|^2 = \|T\psi\|^2$, that is, $(T^\dagger, \psi T^\dagger\psi) =$

$(T\psi, T\psi)$. This in turn is equivalent to $(TT^\dagger\psi, \psi) = (T^\dagger T\psi, \psi)$. Finally this is eqivalent to $([TT^\dagger - T^\dagger T]\psi, \psi) = 0$.

Since T is normal, it is obvious that also is $T - \lambda I$ for any scalar λ,

$$(T - \lambda I)(T^\dagger - \lambda^* I) = TT^\dagger - \lambda^* T - \lambda T^\dagger + |\lambda|^2 I = (T^\dagger - \lambda^* I)(T - \lambda I).$$

Hence we have

$$\|T - \lambda I\| = \|T^\dagger - \lambda^* I\|$$

for all ψ, and the lemma follows at once.

□

Lemma J.3.6 *If T ia normal, then each M_i reduces T.*

Proof:

It is obvious that each M_i is mapped onto itself under T. By theorem J.3.2 it suffices to show that each M_i is invariant under $T^\dagger$. But this follows from lemma J.3.5, for if ψ_i is a vector in M_i, so $T\psi_i = \lambda_i \psi_i$, then $T^\dagger \psi_i = \lambda_i^* \psi_i$ is also in M_i.

□

We are now in a position to state the spectral theorem (the proof can be found in vol I, appendix F).

Theorem J.3.7 *Let T be an arbitrary operator on H. By ... we know that the distinct eigenvalues of T form a non-empty finite set of complex numbers. Let $\lambda_1, \lambda_2, \ldots, \lambda_m$ be these eigenvalues, let $M_1, M_2, \ldots, M_n$ be their corresponding eigenspaces, and let $P_1, P_2, \ldots, P_n$ be the projections onto these eigenspaces. The following statements are equivalent:*

i) The M_i's are pairwise orthogonal and span H.

ii) The P_i's orthogonal, $I = \sum_i^m P_i$ and $T = \sum_i^m \lambda_i P_i$.

iii) T is normal.

The expression

$$T = \lambda_1 P_1 + \lambda_2 P_2 + \cdots + \lambda_m P_m. \tag{J.12}$$

for T, when it exists, is called the **spectral resolution** of T. The spectral theorem in particular proves that if T is normal, then it has the spectral resolution (J.12).

J.3.4 Functional Calculus for Finite Dimension Vector Spaces

$$Tv_n = \lambda_n v_n \tag{J.13}$$

The collection of eigenvalues of T is denoted $\sigma(T)$,

$$\sigma(T) = \{\lambda_1, \lambda_2, \ldots, \lambda_m\} \tag{J.14}$$

an alternative (equivalent for finite dimensional vector spaces) definition is collection of λ's such that $T - \lambda\mathbf{1}$ is not invertible:

$$\sigma(T) = \{\lambda : T - \lambda\mathbf{1} \text{ is not inveretable}\} \tag{J.15}$$

$$I = P_1 + P_2 + \cdots + P_m. \tag{J.16}$$

$$T = \lambda_1 P_1 + \lambda_2 P_2 + \cdots + \lambda_m P_m \tag{J.17}$$

This is known as the spectral decomposition of the operator T. The space of functions on the spectrum $\sigma(T)$. Define $\Phi : C(\sigma(T)) \to \mathcal{B}(\mathcal{H})$

$$\Phi(f) = f(\lambda_1)P_1 + \cdots + f(\lambda_m)P_m. \tag{J.18}$$

Obviously we have

$$f(\sigma(T)) = \sigma(f(T)) \tag{J.19}$$

This relation does not hold in general for infinite vector spaces but only under certain conditions - the spectral mapping theorem??. The space of these functions can be equipped with a norm

$$\|f\|_\infty = \sup_{\lambda_i \in \sigma(T)} |f(\lambda_i)|. \tag{J.20}$$

$$\begin{array}{ll} \text{(a)} & \|\Phi(f)\| = \|f\|_\infty \\ \text{(b)} & \Phi(f+g) = \Phi(f) + \Phi(g) \\ \text{(c)} & \Phi(fg) = \Phi(f)\Phi(g) \\ \text{(d)} & \Phi(\overline{f}) = \Phi(f)^* \end{array} \tag{J.21}$$

(a)

$$\begin{aligned}\|\Phi(f)\| &= \sup_{\|v\|\le 1} \|f(T)v\| \\ &= \sup_{\lambda_i\in\sigma(T)} |f(\lambda_i)| \\ &= \|f\|_\infty \end{aligned} \tag{J.22}$$

$$\|f(T)\| = \sup\{|\lambda|\ ;\ \lambda\in\sigma(f(T))\} = \sup\{|\lambda|\ ;\ \lambda\in f(\sigma(T))\} = \|f\|. \tag{J.23}$$

(b)

$$\begin{aligned}\Phi(f+g) &= \big(f(\lambda_1)+g(\lambda_1)\big)P_1+\cdots+(f(\lambda_m)+g(\lambda_m))P_m \\ &= f(\lambda_1)P_1+\cdots+f(\lambda_m)P_m+g(\lambda_1)P_1+\cdots+g(\lambda_m)P_m \\ &= \Phi(f)+\Phi(g)\end{aligned} \tag{J.24}$$

(c)

$$\begin{aligned}\Phi(fg) &= f(\lambda_1)g(\lambda_1)P_1+\cdots+f(\lambda_m)g(\lambda_m)P_m \\ &= \big(f(\lambda_1)P_1+\cdots+f(\lambda_m)P_m\big)\big(g(\lambda_1)P_1+\cdots+g(\lambda_m)P_m\big) \quad \text{as } P_iP_j=\delta_{ij}P_i \\ &= \Phi(f)\Phi(g)\end{aligned} \tag{J.25}$$

(d)

$$\begin{aligned}\Phi(\overline{f}) &= \overline{f}(\lambda_1)P_1+\cdots+\overline{f}(\lambda_m)P_m \\ &= (f(\lambda_1)P_1+\cdots+f(\lambda_m)P_m)^* \qquad (P_i)^*=P_i\end{aligned} \tag{J.26}$$

we can go from $f\in C(\sigma(T))$ to $f(T)$. Functional calculus can be used to construct great many interesting operators operators for example, square roots, logs, and exponentials of operators, difference equations.

$$\sqrt{T}=\sqrt{\lambda_1}P_1+\cdots+\sqrt{\lambda_m}P_m \tag{J.27}$$

More standard notation

for each $A \subseteq \sigma(T)$, let

$$P(A) = \sum_{\lambda_i \in A} P_i. \tag{J.28}$$

for $v, w \in \mathcal{H}$

$$\mu_{v,w} = (P(A)v, w) \tag{J.29}$$

a measure on $\sigma(T)$

$$\int_{\sigma(T)} f(\lambda) dP(\lambda) \tag{J.30}$$

J.3.5 Decomposition of an N-Dimensional Space.

It follows from that the decomposition of an arbitrary vector $| >$ into vectors $\varphi(A)| >$ is unique. Therefore, an arbitrary vector either belongs entirely to one of the subspaces $\mathbf{S}^{(i)} (i = 1, 2, \ldots, L)$ or can be decomposed uniquely into vectors belonging to different subspaces $\mathbf{S}^{(i)}$. One says that $\mathbf{S}^{(i)}$ is a direct sum of the subspaces $\mathbf{S}^{(i)}$ and one writes

$$\mathbf{S}_N = \mathbf{S}^{(1)} \oplus \mathbf{S}^{(2)} \oplus \cdots \oplus \mathbf{S}^{(L)} \tag{J.31}$$

$$A = \sum_{i=1}^{m} \lambda_i P_i. \tag{J.32}$$

Continuity

Naive arguments can quickly lead to errors. It is complete - We need “enough” functions to do what we want.

J.3.6 Complications in Infinite Dimensional Spaces

The Spectral Theorem

$\psi_n^* \psi_n \not< \infty$ in general.

Even if $\sum_n \psi_n^* \psi_n < \infty$, there could be operators $\hat{A}$ such that

$$\phi_n := \hat{A}|\psi_n > \qquad \sum_n \phi_n^* \phi_n \to \infty \tag{J.33}$$

The maximal domain $\mathcal{D}_A$ is:

$$\mathcal{D}_A := \{|\psi > \in \mathcal{H} \; : \; \hat{A}|\psi > \in \mathcal{H}\} \tag{J.34}$$

scalars

$$\lim_{N\to\infty} \lim_{M\to\infty} \varphi_n^* A_{nm} \psi_m \neq \lim_{N\to\infty} \lim_{M\to\infty} \varphi_n^* A_{nm} \psi_m \tag{J.35}$$

making possible a mathematically rigorous treatment of the subject of Dirac's

finite dimensional spaces are automatically complete.

The norm enables us to define a distance, $d(\phi, \psi) := \|\phi - \psi\|$, between vectors. the state space $\mathcal{H}$ should be complete in the following sense: whenever $\{\psi_n \in \mathcal{H}\}$ is a sequence of vectors

$$\|\psi_m - \psi_n\| < \epsilon \tag{J.36}$$

for all $m, n > N_\epsilon$, then there exists a limit vector ψ such that

$$\|\psi_n - \psi\| \to 0 \tag{J.37}$$

Completedness is an important property to have around if you are counting on an equation having a solution.

$\psi_n := \sum_{r=0}^{n} (a_r \cos nx + b_r \sin nx)$

A metric space is **complete** if every Cauchy sequence in M converges to a point in M.

J.3.7 Cauchy Sequences

Proposition J.3.8 *A convergent sequence has a unique limit.*

Proof: Suppose $\{x_n\}$ converges to x and also to x', where $x \neq x'$. Put

$$\epsilon = \frac{1}{2}|x - x'|.$$

Since $\{x_n\}$ converges to x, there exists N_1 such that $|x_n - x| < \epsilon$ for all $n \geq N_1$. Similarly there exists N_2 such that $|x_n - x'| < \epsilon$ for all $n \geq N_2$. Put $N = \max\{N_1, N_2\}$. Then

$$|x - x'| = |x - x_N + x_N - x'| \leq |x - x_N| + |x_N - x'| < 2\epsilon = |x - x'|.$$

This contradiction shows that $x = x'$.

□

Theorem J.3.9 *Every bounded monotonic sequence converges.*

Proof: Suppose that $\{s_n\}$ is a monontonic increasing and bounded from above. By the completeness axiom, for the set S of its memebers $s = \sup S$ exists. We prove that $\{s_n\}$ converges to s. Let $\epsilon > 0$. Then $s - \epsilon$ is not an upper bound for S so

$$s_N > s - \epsilon$$

for some N. Since $\{s_n\}$ is a monontonic increasing,

$$s_n \geq s_N \text{ for all } n \geq N.$$

But

$$s_n \leq s \text{ for all } n,$$

since s is an upper bound for S. Hence for all $n \geq N$,

$$s \geq s_n \geq s_N > s - \epsilon,$$

and in particular

$$|s - s_n| < \epsilon.$$

The proof for $\{s_n\}$ is a monontonic decresing can be obtained from that proved to get the convergence of $-s_1, -s_2, -s_3, \ldots$, which implies the convergence of $s_1, s_2, s_3, \ldots$.

□

Definition A sequence $\{x_n\}$ is a Cauchy sequence if given $\epsilon > 0$, there exists N such that $m, n \geq N$ imply $|x_n - x_m| < 0$.

Lemma J.3.10 *Any Cauchy sequence is bounded.*

Proof: Take $\epsilon = 1$ in the Cauchy sequence. Thus there exists N such that $m, n \geq N$ imply

$$|x_n - x_m| < 1,$$

so for any $n \geq N$,

$$|x_n - x_N| < 1,$$

and hence

$$|x_n| < 1 + |x_N|.$$

Now

$$|x_n| \leq \max\{|x_1|, |x_2|, \ldots, |x_{N-1}|, 1 + |x_N|\}$$

for all n.

□

Theorem J.3.11 *A sequence $\{s_n\}$ of real numbers conveges if and only if it is a Cauchy sequence.*

Proof: Suppose that $\{s_n\}$ converges to s. Then given $\epsilon > 0$, there exists N such that $|s_n - s| < 0$ for $n \geq N$, so for $m, n \geq N$ we have

$$|s_n - s_m| = |s_n - s + s - s_m| \leq |s_n - s| + |s - s_m| < 2\epsilon.$$

Hence $\{s_n\}$ is a Cauchy sequence.

Suppose conversely that $\{s_n\}$ is a Cauchy sequence. We construct a monotonic sequence from the sequence $\{s_n\}$. For each m, let S_m denote a set of members of the sequence

$$S_m := \{s_n : n \geq m\}.$$

Obviously, each S_m is bounded as the whole set $S = S_1$ is. Hence supS_m exists. Let $u_m := \sup S_m$. Since $S_{m+1} \subset S_m$, it follows that $\sup S_{m+1} \leq \sup S_m$. Thus $\{u_m\}$ is monontonic decreasing.

By theorem J.3.9, converges, to s say.

Finally, we show that $\{s_m\}$ also converges to s. Given ϵ, there exists N_1 such that

$$|s_m - s_n| < \epsilon$$

for $m, n \geq N_1$ and N_2 such that

$$|s - u_m| < \epsilon$$

for $m \geq N_2$. Let $N = \max\{N_1, N_2\}$. By definition of u_N, $u_N - \epsilon$ is not an upper bound of S_N, so there exists $M \geq N$

Now for any $n \geq N$,

$$\begin{aligned} |s_n - s| &= |s_n - s_M + s_M - s_N + s_N - s| \\ &\leq |s_n - s_M| + |s_M - s_N| + |s_N - s| \\ &< 3\epsilon. \end{aligned}$$

Hence $\{s_n\}$ converges to s.

□

Definition A sequence of vectors $\{f_n\}$ in a normed vector space V is called a Cauchy sequence if given $\epsilon > 0$, there exists N such that $m, n \geq N$ imply $\|f_n - f_m\| < 0$.

J.3.8 Functions on Metric Spaces

First, $d(x, y)$ should be non-negative for any x, y in A, and it should only be zero if the x and y coincide. The distance should be the same whether we "measure" it from x to y or from y to x. The sum of the lengths of two sides of any triangle is not less than the length of the third. is the formal version of the triangle inequality.

(M1) $d(x, y) \geq 0; \quad d(x, y) = 0 \leftrightarrow x = y$

(M2) $d(x, y) = d(y, x)$ for all x, y in A,

(M3) $d(x, z) \leq d(x, y) + d(y, z)$ for all x, y, z in A.

Sequences of Functions

Definition The sequence (T_n) is said to *converge pointwise* to T provided

Obviously, uniform convergence is also pointwise convergence (to the same limit). The converse is generally false as can be seen from the example, so uniform convergence is **stronger** than pointwise convergence.

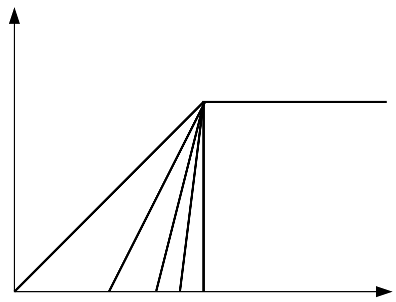

Figure J.1: The limit function of a sequence of continuous functions need not be continuous.

The primary use of uniform convergence is to ensure the continuity of the limit function.

Theorem Suppose that each T_n is continuous and (T_n) converges uniformally to the limit function T. Then, T will be continuous.

Proof: Take any point x in ??For all n and y

$$d(Tx, Ty) \leq d(Tx, Tx_n x) + d(T_n x, T_n y) + d(T_n y, Ty). \tag{J.38}$$

J.3.9 Topologies

We begin our exposition of topology by attaching names to properties that will be used very often and then introduce the shorthand that has been developed in the subject for talking about such things.

J.4 Function Spaces.

Length of a vector. The vector space consisting of position vectors $\mathbf{r}$ in 3-d Euclidian space with a distance

$$||\mathbf{r}|| = \sqrt{x^2 + y^2 + z^2} \tag{J.39}$$

Although each point of $\mathcal{C}$ is a function, in many ways $\mathcal{C}$ is like a Euclidean space. We may for instance, define a "distance" on on the space $\mathcal{C}$ as follows

$$\|f\| = \max_{x\in A} |f(x)| \tag{J.40}$$

The symbol $\psi \in V$ means that ψ is a member of the set of vectors in V.

$$\begin{aligned}
&\|\alpha f\| = \max_{x\in A} |\alpha f(x)| = |\alpha| \max_{x\in A} |f(x)| = |\alpha| \|\alpha f\|, \\
&\|f\| = \max_{x\in A} |f(x)| = 0 \;\Rightarrow\; f(x) = 0, \\
&\|f+g\| = \max_{x\in A} |f(x)+g(x)| \leq \max_{x\in A} |f(x)| + \leq \max_{x\in A} g(x)| = \|f\| + \|g\|.
\end{aligned} \tag{J.41}$$

J.4.1 Normed Spaces and Banach Spaces

Definition (Normed Space). A normed space X is a vector space with a norm defined on it. Here a norm is a (real or complex) vector is a real-valued function on X whose value at an $x \in X$ is denoted by

$$\|x\| \tag{J.42}$$

and which has the following properities

(i) $||x|| \geq 0$ and $||x|| = 0$ if and only if $x = 0$

(ii) For $c \in k$ and $\psi \in E$, $||c\psi|| = |c| \; ||\psi||$.

(iii) $||x + y|| \leq ||x|| + ||y||$. (triangle inequality)

If is endowed with a given norm we call a normed linear (or vector space) space.

A norm defines a metric induced by the norm which is given by

$$d(x, y) := \|x - y\| \tag{J.43}$$

and is called the *metric induced by the norm.* As a nice example of the space of continuous functions $\mathcal{C}$.

A special case of a normed space.

There is a general theorem in topology that tells us that any metric space can be completed by a procedure analogous to that for constructing the real numbers from the rationals.

Definition (Banach Space) A is a complete normed space (complete in the metric (J.43))

Examples of normed spaces.

The vector space consisting of position vectors $\mathbf{r}$ in 3-d Euclidian space with the norm

$$||\mathbf{r}|| = \sqrt{x^2 + y^2 + z^2} \tag{J.44}$$

forms a normed space. That $||\mathbf{r}||$ satisfies the conditions to be a norm is obvious. The triangle inequality is obvious from fig (J.4.3).

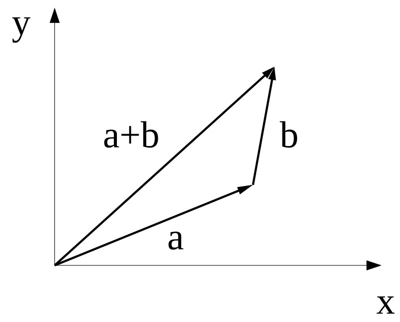

Figure J.2: Euclidean triangle inequaltiy.

This can be thought of as the prototype normed space.

Example 2 Space l^p. Is a Banach space with norm given by

$$\|x\| := \left(\sum_{j=1}^{\infty} |\xi_j|^p \right)^{1/p} \tag{J.45}$$

this norm induces the metric

$$d(x,y) \equiv \|x - y\| := \left(\sum_{j=1}^{\infty} |\xi_j - \eta_j|^p \right)^{1/p} \tag{J.46}$$

The set of sqaure integrable functions, that is functions for which $||\psi(x)|| < \infty$ where the norm is defined by $:= \int_{-\infty}^{\infty} dx|\psi(x)|^2$ forms a normed space.

The integrand is positive

$$||c\psi|| = \int_{-\infty}^{\infty} dx|c\psi(x)|^2 = |c|^2 \int_{-\infty}^{\infty} dx|\psi(x)|^2 = |c|^2 \; ||\psi|| \tag{J.47}$$

Consider the quantity

$$\begin{aligned}\int_{-\infty}^{\infty} dx|\psi(x) + \lambda\phi(x)|^2 &= \int_{-\infty}^{\infty} dx(\psi(x) + \lambda\phi(x))^*(\psi(x) + \lambda\phi(x)) \\ &= \int_{-\infty}^{\infty} dx \left[|\psi(x)|^2 + \lambda\psi(x)^*\phi(x) + \lambda^*\psi(x)\phi(x)^* + \lambda\lambda^*|\phi(x)|^2\right]\end{aligned} \tag{J.48}$$

$$\int_{-\infty}^{\infty} dx|\psi(x) + \phi(x)|^2 \leq \int_{-\infty}^{\infty} dx|\psi(x)|^2 + \int_{-\infty}^{\infty} dx|\phi(x)|^2 \tag{J.49}$$

There is a general theorem in topology that tells us that any metric space can be completed by a procedure analogous to that for constructing the real numbers from the rationals.

Strong convergence (or convergence in the mean)

$$||\psi - \psi_n|| \to 0 \tag{J.50}$$

So that, for any $\epsilon > 0$ there exists an integer n large enough that we can have

$$||\psi - \psi_n|| < \epsilon/2 \tag{J.51}$$

By the triangle inequality

$$||\psi_m - \psi_n|| \leq ||\psi - \psi_m|| + ||\psi - \psi_n|| < \epsilon \tag{J.52}$$

$$||\psi_m|| = ||(\psi_m - \psi_n) + \psi_n|| \leq ||\psi_m - \psi_n|| + ||\psi_n|| \tag{J.53}$$

$$\begin{aligned}||\psi_m - \psi_n|| &\geq ||\psi_m|| - ||\psi_n|| \\ ||\psi_m - \psi_n|| &\geq ||\psi_n|| - ||\psi_m||\end{aligned} \tag{J.54}$$

$$\boxed{||\psi_m - \psi_n|| \geq \big|\, ||\psi_m|| - ||\psi_n||\, \big|} \tag{J.55}$$

$$|\; ||\psi_m|| - ||\psi_n||\; | < \epsilon \tag{J.56}$$

$$|\; ||\psi|| - ||\psi_n||\; | < \epsilon \tag{J.57}$$

"if and only if" is a phrase used a lot in mathematics means you cannot have one without the other, in other words, they are equivalent.

Definition A metric space X is said to be **complete** if every Cauchy sequence in X converges to an element of X. Otherwise X is said to be **incomplete**.

Lemma J.4.1 *If a subsequence of a Cauchy sequence converges, then the whole sequence converges.*

Proof: Let $\{x_n\}$ be a Cauchy sequence in a normed vector space X, and let $\epsilon > 0$ be given. Then there is an N such that

$$\|x_n - x_m\| < \epsilon, \quad m, n > N.$$

Now if $\{x_n\}$ has a subsequence converging to $x \in X$, then there is an $m > N$ such that

$$\|x_m - x\| < \epsilon.$$

Thus,

$$\|x_n - x\| \leq \|x_n - x_m\| + \|x_m - x\| < 2\epsilon.$$

for all N.

$\square$

Definition A normed linear space X is said to be complete if it is complete as a metric space in the induced metric.

Definition A **Banach space** is a complete normed space.

Definition A **bounded linear transformation** from a normed linear space $(V_1, \|.\|_1)$ to a normed linear space $(V_2, \|.\|_2)$ is a function, T, from V_1 to v_2 satisfying

i) $T(\alpha v + \beta w) = \alpha T(v) + \beta T(w)$ for all $v, w \in V$ and for all $\alpha, \beta \in \mathbb{R}$ or $\mathbb{C}$.

ii) For some $C \geq 0$, $\|Tv\|_2 \leq C\|v\|_1$.

The smallest such C is called the **norm of** T, $\|T\|$. Thus

$$\|T\| := inf_{\|v\|_1=1} \|Tv\|_2.$$

Definition The **closure** $\overline{C}$ of a subset C in a topological space T is the union of C and all the limit points of C in T.

Lemma J.4.2 *A point $x \in T$ is in $\overline{C}$ if and only if every open set containing x contains a point of C.*

Proof:

□

Lemma J.4.3 *Let C be a subset of a metric space X and let $x \in X$. Then $x \in \overline{C}$ if and only if there exists a sequence $\{x_n\}$ in C converging to x.*

Proof: First suppose that $\{x_n\}$ is a sequence in C converging to x. Given $\epsilon > 0$, there exists N such that $x_n \in B_\epsilon(x)$ for al $n \geq N$. Since $x_n \in C$ for all n, this means that for every $\epsilon > 0$, $B_\epsilon(x) \cup C \neq \emptyset$. Hence $x \in \overline{C}$.

Now suppose that $x \in \overline{C}$. By definition of closure, there exists a point x_n in $C \cap B_{1/n}(x)$ for every $n \in \mathbb{N}$, and clearly $\{x_n\}$ converging to x.

□

J.4.2 Completion of a metric space

For example, the space of all real numbers is the completion of the space of all rational numbers.

Theorem J.4.4 *Every metric space of X has a completion. This completion is unique up to an isometric mapping carrying every point $x \in X$ into itself.*

Proof: Proof of the existence of the completion of X:

We call two Cauchy sequences $\{x_n\}$ and $\{y_n\}$ in X equivalent and write $\{x_n\} \sim \{y_n\}$ if

$$\lim_{n\to\infty} d(x_n, y_n) = 0.$$

Let the new space X^* be the family of equivalence classes of Cauchy sequences under this equivalence relation. Then define the distance between two arbitrary points $x^*, y^* \in X^*$ by the formular

$$d_1(x^*, y^*) = \lim_{n\to\infty} d(x_n, y_n), \tag{J.58}$$

where $\{x_n\}$ and $\{y_n\}$ are any representatives of x^* and y^* respectively. We must establish existence, independent of the choice of sequences $\{x_n\} \in x^*$, $\{y_n\} \in y^*$ and that it satisfies the three properties of a distance. Given any $\epsilon > 0$, it follows from the triangle inequality in X that

$$\begin{aligned} |d(x_n, y_n) - d(x_{n'}, y_{n'})| &\leq |d(x_n, y_n) - d(x_{n'}, y_n)| + |d(x_{n'}, y_n) - d(x_{n'}, y_{n'})| \\ &\leq d(x_n, x_{n'}) + d(y_n, y_{n'}) \\ &< \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon \end{aligned} \tag{J.59}$$

for sufficiently large n and n'. Therefore the sequence of real numbers $\{d(x_n, y_n)\}$ is Cauchy and hence has a limit. We show that it is independent of the representatives. Suppose

$$\{x_n\}, \{\tilde{x}_n\} \in x^*, \qquad \{y_n\}, \{\tilde{y}_n\} \in y^*.$$

Then

$$|d(x_n, y_n) - d(\tilde{x}_n, \tilde{y}_n)| \leq d(x_n, \tilde{x}_n) + d(y_n, \tilde{y}_n).$$

But

$$\lim_{n\to\infty} d(x_n, \tilde{x}_n) = \lim_{n\to\infty} d(y_n, \tilde{y}_n) = 0,$$

since $\{x_n\} \sim \{\tilde{x}_n\}$, $\{y_n\} \sim \{\tilde{y}_n\}$, and hence

$$\lim_{n\to\infty} d(x_n, y_n) = \lim_{n\to\infty} d(\tilde{x}_n, \tilde{y}_n).$$

Now we move onto proving the properties of a metric. It is obvious that $d_1(x^*, y^*) = d_1(y^*, x^*)$. That $d_1(x^*, y^*) = 0$ if and only if $x^* = y^*$ is an immediate consequence of the definition of equivalent Cauchy sequences. To verify the triangle inequality in X^* we start from the trianngle inequality

$$d(x_n, z_n) \le d(x_n, y_n) + d(y_n, z_n)$$

in the original space X and then take the limit as $n \to \infty$, obtaining

$$\lim_{n\to\infty} d(x_n, z_n) \le \lim_{n\to\infty} d(x_n, y_n) + \lim_{n\to\infty} d(y_n, z_n)$$

i.e.,

$$d_1(x^*, z^*) \le d_1(x^*, y^*) + d_1(y^*, z^*).$$

Now we prove that X^* is a completion of X. First we show that the mapping of X to X^* is isometric. Suppose that every point $x \in X$, we associate the class $x^* \in X^*$ of al Cauchy sequences converging to x. Let

$$x = \lim_{n\to\infty} x_n, \qquad y = \lim_{n\to\infty} y_n.$$

Then

$$d(x, y) = \lim_{n\to\infty} d(x_n, y_n),$$

while on the other hand

$$d_1(x^*, y^*) = \lim_{n\to\infty} d(x_n, y_n),$$

by definition. Therefore

$$d(x, y) = d_1(x^*, y^*),$$

and hence the mapping from X into X^* carrying x into x^* is isometric.

X can be regarded as a subset of X^*. It must be proved that

i) X is everywhere dense in X^*, i.e., $\overline{X} = X^*$; and

ii) X^* is complete.

To this end, given any point $x^* \in X^*$ and any $\epsilon > 0$, choose a Cauchy sequence $\{x_n\} \in x^*$. Let N be such that $d(x_n, x_{n'}) < \epsilon$ for all $n, n' > N$. Then

$$d(x_n, x^*) = \lim_{n' \to \infty} d(x_n, x_{n'}) \leq \epsilon$$

if $n > N$, i.e., every neighbourhood of the point x^* intersects X. If follows that $\overline{X} = X^*$.

Finally we show that X^* is complete, that is, we show that every Cauchy sequence $\{x_n^*\}$, consisting of points in X^*, converges to a point $x^* \in X^*$. Note by the very definition of X^*, any Cauchy sequence $\{x_n\}$ consisting of points in X converges to some point $x^* \in X^*$. As X is dense in X^*, for any $\{x_n^*\}$ we can find an equivalent sequence $\{x_n\}$ in X. Choose x_n to be any point of X such that $d(x_n, x_n^*) < 1/n$, the resulting sequence $\{x_n\}$ is equivalent to $\{x_n^*\}$. Now as

$$\begin{aligned} d(x_n, x_{n'}) &\leq d(x_n, x_n^*) + d(x_n^*, x_{n'}) \\ &\leq d(x_n, x_n^*) + d(x_n^*, x_{n'}^*) + d(x_{n'}^*, x_{n'}) \\ &< \frac{1}{n} + \frac{1}{n'} + d(x_n^*, x_{n'}^*) \end{aligned}$$

$\{x_n\}$ is Cauchy and so converges to a point $x^* \in X^*$. But then the sequence $\{x_n^*\}$ also converges to x^*.

Proof of uniqueness:

Say X^* and $X^{*\prime}$ are two completions of X. We wish to show that there is a one-to-one mapping $i : X^* \to X^{*\prime}$ such that $i(x) = x$ for all $x \in X$ and

$$d_1(x^*, y^*) = d_2(x^{*\prime}, y^{*\prime}) \tag{J.60}$$

where d_1 is the distance in X^* and d_2 is the distance in $X^{*\prime}$. The required mapping i is constructed as follows: Let x^* be an arbitrary point of X^*. Then, by definition of completion, there is a Cauchy sequence $\{x_n\}$ of points converging to x^*. The points of the sequence $\{x_n\}$ also belong to $X^{*\prime}$, where they form a Cauchy sequence. Therefore $\{x_n\}$ converges to a point $x^{*\prime} \in X^{*\prime}$, since $X^{*\prime}$ is complete. We see that $x^{*\prime}$ is independent of the choice of sequence.

If we set $i(x^*) = x^{*\prime}$, then i is the required mapping. In fact, $i(x) = x$ for all $x \in X$, since $x_n \to x \in X$, then obviously $x = x^* \in X^*$, $x^{*\prime} = x$. Moreover, suppose $x_n \to x^*$, $y_n \to y^*$ in X^*, while $x_n \to x^{*\prime}$, $y_n \to y^{*\prime}$ in $X^{*\prime}$. Then, if d is the distance in X,

$$d_1(x^*, y^*) = \lim_{n\to\infty} d_1(x_n, y_n) = \lim_{n\to\infty} d(x_n, y_n) \tag{J.61}$$

Now we prove that if $x_n \to x$, $y_n \to y$ as $n \to \infty$, then $d(x_n, y_n) \to d(x, y)$. So at the same time

$$d_2(x^{*\prime}, y^{*\prime}) = \lim_{n\to\infty} d_2(x_n, y_n) = \lim_{n\to\infty} d(x_n, y_n) \tag{J.62}$$

But (J.61) and (J.62) imply (J.60).

□

Definition Let $\mathcal{B}(X, Y)$ be the set of bounded linear operators from X to Y.

Theorem J.4.5 *If Y is a Banach space, so is $\mathcal{B}(X, Y)$.*

Proof: Suppose $\{A_n\}$ is a Cauchy sequence of operators in $\mathcal{B}(X, Y)$. Then for each $\epsilon > 0$ there is an N such that for all $m, n > N$

$$\|A_n - A_m\| < \epsilon.$$

Thus for each $x \neq 0$,

$$\|A_n x - A_m x\|_2 < \epsilon \|x\|_1, \quad m, n > N. \tag{J.63}$$

This shows that $\{A_n x\}$ is a Cauchy sequence in Y. Since Y is complete, there is a $y_x \in Y$ such that $A_n x \to y_x$ in Y. Define the operator A from X to Y by $Ax = y_x$.

Then A is linear. Note that we have

$$\begin{aligned}
\|A(\alpha x + \beta x') - \alpha A(x) - \beta A(x)'\|_2 &\leq \|A(\alpha x + \beta x') - A_n(\alpha x_n + \beta x'_n)\|_2 \\
&\quad + \|\alpha A_n(x_n) + \beta A_n(x'_n) - \alpha A(x) - \beta \tilde{T}(x')\|_2 \\
&\leq \|A(\alpha x + \beta x') - A_n(\alpha x + \beta x')\|_2 \\
&\quad + |\alpha| \, \|A_n(x) - A(x)\|_2 \\
&\quad + |\beta| \, \|A_n(x') - A(x')\|_2.
\end{aligned}$$

Since the right-hand-side tends to zero as $n \to \infty$, we conclude that $\tilde{T}$ is linear.

Let $m \to \infty$ in (J.63). Then

$$\|A_n x - Ax\| < \epsilon \|x\|, \quad n > N.$$

Hence,

$$\|Ax\| \leq \epsilon \|x\| + \|A_n x\| \leq (\epsilon + \|A_n\|)\|x\|, \quad n > N.$$

This shows that A is bounded. Moreover,

$$\|A_n - A\| \leq \epsilon, \quad n > N.$$

Hence,

$$\|A_n - A\| \to 0 \text{ as } n \to \infty.$$

□

Theorem J.4.6 *(The* **bounded linear transformation theorem***). Suppose T is a bounded linear transformation from a normed linear space $(V_1, \|.\|_1)$ to a complete normed linear space $(V_2, \|.\|_2)$. Then T can be uniquely extended to a bounded linear transformation (with the same bound), $\tilde{T}$, from the completion of V_1 to $(V_2, \|.\|_2)$.*

Proof: Proof of existence: Let $\overline{V}_1$ be the completion of V_1. For each $x \in \overline{V}_1$, there is a sequence of elements $\{x_n\}$ in V_1 with $x_n \to x$ as $n \to \infty$. Since x_n converges, it is Cauchy so given $\epsilon > 0$, we can find N so that for $n, m > N$ such that

$$\|x_n - x_m\|_1 \leq \epsilon / \|T\|$$

.

Then

$$\|Tx_n - Tx_m\|_2 = \|T(x_n - x_m)\|_2 \leq \|T\| \, \|x_n - x_m\|_1 \leq \epsilon$$

which proves that Tx_n is a Cauchy sequence in V_2. Since V_2 is complete, $Tx_n \to y$ for some $y \in V_2$. Set $\tilde{T}x = y$.

We must show that this defintion is independent of the sequence $x_n \to x$ chosen. We know that $x \in \overline{V}_1$ is independent of the Cauchy sequence $\{x_n\}$. $\{Tx_n\}$ is a Cauchy sequence in V_2. We must show that $\{Tx_n\}$ and $\{Tx'_n\}$ have the same convergence point in V_2. As $\{x_n\}$ and $\{x'_n\}$ are Cauchy, there is N such that for all $n > N$, $\|x_n - x\|_1 \leq \epsilon/2\|T\|$ and $\|x'_n - x\|_1 \leq \epsilon/2\|T\|$. Therefore we have

$$\begin{aligned}
\|Tx_n - Tx'_n\|_2 &= \|T(x_n - x'_n)\|_2 \\
&\leq \|T\| \, \|x_n - x'_n\|_1 \\
&\leq \|T\| \, (\|x_n - x\|_1 + \|x - x'_n\|_1) \\
&\leq \epsilon.
\end{aligned}$$

We conclude that $\lim_{n\to\infty} Tx_n = \lim_{n\to\infty} Tx'_n$.

The operator $\tilde{T}$ is bounded with same bound as

$$\begin{aligned}
\|\tilde{T}x\|_2 &= \lim_{n\to\infty} \|\tilde{T}x_n\|_2 \\
&\leq \lim_{n\to\infty} C\|x_n\|_1 \\
&= C\|x\|_1.
\end{aligned} \tag{J.64}$$

Proof of linearity: Note that we have

$$\begin{aligned}
\|\tilde{T}(\alpha x + \beta x') - \alpha\tilde{T}(x) - \beta\tilde{T}(x')\|_2 &\leq \|\tilde{T}(\alpha x + \beta x') - T(\alpha x_n + \beta x'_n)\|_2 \\
&\quad + \|\alpha T(x_n) + \beta T(x'_n) - \alpha\tilde{T}(x) - \beta\tilde{T}(x')\|_2 \\
&\leq \|\tilde{T}(\alpha x + \beta x') - T(\alpha x_n + \beta x'_n)\|_2 \\
&\quad + |\alpha| \, \|T(x_n) - \tilde{T}(x)\|_2 \\
&\quad + |\beta| \, \|T(x'_n) - \tilde{T}(x')\|_2.
\end{aligned}$$

Since the right-hand-side tends to zero as $n \to \infty$, we conclude that $\tilde{T}$ is linear.

Proof of uniqueness:

$$\begin{aligned}
\|\tilde{T} - \tilde{T}'\| &= \sup_{x\in\overline{V}_1 - \{0\}} \frac{\|(\tilde{T} - \tilde{T}')x\|_2}{\|x\|_1} \\
&= \sup_{x\in\overline{V}_1 - \{0\}} 0 \\
&= 0
\end{aligned}$$

We conclude that $\tilde{T}$ is unique.

□

J.4.3 Inner Product Spaces and Hilbert spaces

Definition: A linear vector space upon which has an inner product

and which has the folowing properities

(i) $< x, x > \geq 0$ and $||x|| = 0$ if and only if $x = 0$

(ii) For $c \in k$ and $\psi \in E$, $||c\psi|| = |c|\ ||\psi||$.

(iii) $||x + y|| \leq ||x|| + ||y||$. (triangle inequality)

If is endowed with a given norm we call a normed linear (or vector space) space.

Definition: A *Hilbert space* is a complex inner product space which is complete. The state space of QM is aussumed to be a Hilbert space.

Definition: A *separable Hilbert space* is a Hilbert space that contains a countable dense subset.

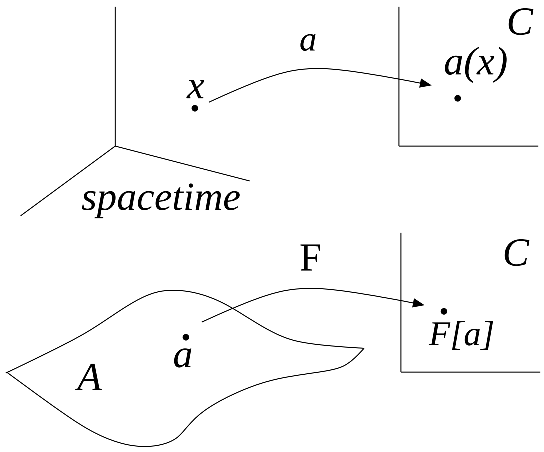

Figure J.3: Normed, Banach, inner product space, Hilbert space.

Orthogonal complements

Lemma J.4.7 *Let M be a linear subspace of an inner product space $\mathcal{H}$ and let $x \in \mathcal{H}$. Then $x \in M^{\perp}$ if and only if*

$$\|x - y\| \leq \|x\| \tag{J.65}$$

for all $y \in M$.

Suppose that (J.65) holds. Pick any $y \in M$ and $\alpha \in \mathbb{C}$. Then $\alpha y \in M$

$\square$

Lemma J.4.8 *Let M be a subspace of a separable Hilbert space. The set of all vectors which are orthogonal to every vector in M is called the* **orthogonal complement** *$M^\perp$ of M.*

Proof:

Evidently $M^\perp$ is a linear manifold; if ψ and ϕ are vectors in $M^\perp$ and α is a scalar, then for any vector χ in M

$$(\chi, \psi + \phi) = (\chi, \psi) + (\chi, \phi) = 0$$

and

$$(\chi, \alpha\phi) = \alpha(\chi, \psi) = 0$$

so $\psi + \phi$ and $\alpha\psi$ are in $M^\perp$. In fact $M^\perp$ is a closed subspace, for if a sequence of vectors ψ_n in $M^\perp$ converges to a limit vector ψ, then for any χ in M as the inner product is continuous

$$(\chi, \psi) = \lim_{n\to\infty} (\chi, \psi_n) = 0,$$

so ψ is in $M^\perp$.

$\square$

This implies $\overline{M} = M^{\perp\perp}$

Proof: If $x \in M^\perp$ then, for any $y \in M$, x and y are orthogonal so that

$$\|x - y\|^2 = \|x\|^2 + \|y\|^2 \geq \|x\|^2.$$

$\square$

Corollary J.4.9 *Let M be a linear subspace of a Hilbert space $\mathcal{H}$, then M is dense in $\mathcal{H}$ (that is $\overline{M} = \mathcal{H}$) if and only if $M^{\perp} = \{0\}$.*

J.4.4 Linear Functionals

A linear functional F assigns a scalar $F(\phi)$ to each vector ϕ such that for any to vectors ϕ and ψ and a scalar A

$$F(\phi + \psi) = F(\phi) + F(\psi) \tag{J.66}$$

$$F(a\phi) = aF(\phi). \tag{J.67}$$

An example of a functional would be

$$F(\phi) = (\varphi, \phi), \tag{J.68}$$

(but not all functionals will be of this form!)

Linear functionals form a linear space with the sum $F_1 + F_2$

To get a similar result for infine-dimensional spaces we need the concept of continuity. Recall that a function f of a complex variable is continuous if $f(z_n) \to f(z)$ whenever $z_n \to z$. The same property defines continuity for linear functionals; a linear functional F is *continuous* if $F(\chi_n)$ converges to $F(\chi)$ for any sequence of vectors χ_n which converges to a limit vector χ.

J.4.5 Complete Orthonormal Systems of Vectors in $\mathcal{H}$

Definition (Gram-Schmidt orthogonalization process). Let $X = \{x_1, \ldots, x_n\}$ be any basis in V. We shall construct a complete orthornormal set $Y = \{y_1, \ldots, y_n\}$ such that each y_j is a linear combination of the x_j's.

As X is linearly independent, $x_1 \neq 0$ and we may take $y_1 = x_1/\|x_1\|$. Now put

$$y_2 = \frac{(x_2 - \alpha_1 y_1)}{\|x_2 - \alpha_1 y_1\|}.$$

Thus setting

$$(y_1, y_2) = \frac{(y_1, x_2) - \alpha_1}{\|x_2 - \alpha_1 y_1\|} = 0$$

implies $\alpha_1 = (y_1, x_2)$. Therefore

$$y_2 = \frac{x_2 - (y_1, x_2)y_1}{\|x_2 - \alpha_1 y_1\|}.$$

Note that

$$x_2 - (y_1, x_2)y_1 = x_2 - \frac{(y_1, x_2)}{\|x_1\|} x_1 \neq 0,$$

since X is linearily independent. Thus we have found two vectors, y_1 and y_2, such that $(y_i, y_j) = \delta_{ij}$, $i, j = 1, 2$. We continue inductively. When this process is repeated n times, we shall have generated an orthonormal set of n basis vectors.

An orthonormal sequence $\{e_n\}$ in a Hilbert space H is complete if the only member of H which is orthogonal to every e_n is the zero vector.

Lemma J.4.10 *Every separable Hilbert space has a complete orthonomal sequence.*

Proof: Let $\mathcal{H}$ be a separable Hilbert space, and let $\{g_n\}$ be a dense sequence in H. Remove from this sequence any element which is a linear combination of the proceding g_j's via a Gram-Schmidt process. Let S_n be a subspace of spanned by $g_1, \ldots, g_n$, and let φ_n be an element of S_n having norm one and orthogonal to S_{n-1}. Such an element exists by a corollary of the Riesz lemma, corollary J.7.4. Clearly, the sequence $\{\varphi_n\}$ is orthonormal. It is complete since the original sequence $\{g_n\}$ is contained in

$$W = \bigcup_{n=1}^{\infty} S_n,$$

and each element in W is a linear combination of a finite number of the φ_n's.

□

Theorem J.4.11 *For each continuous linear functional F on a separable Hilbert space there is a unique vector ψ_F in the space of such states that $F(\phi) = (\psi_f, \phi)$ for every vector ϕ.*

Proof: Let the set of vectors $\phi_1, \phi_2, \ldots, \phi_k, \ldots$ be an orthonormal basis. Consider the infinite linear combination

$$\psi_F = \sum_{n=1}^{\infty} F(\phi_n)^* \phi_n. \tag{J.69}$$

If this converges to a vector ψ_F, then for any vector $\phi = \sum_{k=1}^{\infty}(\phi, \phi_k)\phi_k$ we have

$$\begin{aligned} F(\phi) &= \sum_{k=1}^{\infty}(\phi_k, \phi)F(\phi_k) \\ &= (\sum_{k=1}^{\infty} F(\phi_k)^* \phi_k, \phi) \\ &= (\psi_F, \phi) \end{aligned} \tag{J.70}$$

We must prove that ψ_F defines a vector, that is, we must prove $\|\psi_F\|$ is finite. We will then know what the vector is and it will be unique. As

$$\begin{aligned} \|\psi_F\|^2 &= \|\sum_{k=1}^{\infty} F(\phi_k)^* \phi_k\|^2 \\ &= (\sum_{j=1}^{\infty} F(\phi_j)^* \phi_j, \sum_{k=1}^{\infty} F(\phi_k)^* \phi_k) \\ &= \sum_{k=1}^{\infty} |F(\phi_k)|^2 \end{aligned} \tag{J.71}$$

we must show that $\sum_{k=1}^{\infty} |F(\phi_k)|^2$ converges. Let

$$\psi_n := \sum_{k=1}^{n} F(\phi_k)^* \phi_k.$$

Then

$$F(\psi_n) = \sum_{j=1}^{\infty} \big(\phi_j, \sum_{k=1}^{n} F(\phi_k)^* \phi_k\big) F(\phi_j) = \sum_{k=1}^{n} |F(\phi_k)|^2 = \|\psi_n\|^2.$$

We show that this remains finite as $n \to \infty$. We do this by proving if it doesn't then the continuity of F is violated. To this end set $\chi_n := (1/\|\psi_n\|^2)\psi_n$. Then $\|\chi_n\| = 1/\|\psi_n\|$ and

$$F(\chi_n) = \frac{1}{\|\psi_n\|^2} F(\psi_n) = 1.$$

But $F(0) = 0$ because F is linear. Therefore continuity would be violated if $\chi_n \to 0$ as $n \to \infty$. Since F is continuous we cannot have that $\|\chi_n\|$ converge to zero and hence $\|\psi_n\|$ does not diverge as $\to \infty$. Therefore ψ_F exists.

$\square$

Theorem J.4.12 *Any two complete, orthonormal systems in a Hilbert space have the same cardinality.*

Proof: We need only consider the case where the space $\mathcal{H}$ is a non-separable.

Let $\mathfrak{m}$ and $\mathfrak{n}$ be any two cardinal numbers (finite or infinite). The statement that $\mathfrak{m} \leq \mathfrak{n}$ is defined to mean the following: if X and Y are sets with $\mathfrak{m}$ and $\mathfrak{n}$ elements, then there exits a one-to-one mapping of X into Y.

Suppose that two orthonormal systems M and N of cardinality $\mathfrak{m}$ and $\mathfrak{n}$ respectively are complete.

For every element $f \in N$ an element $e \in M$ can be found such that $(e, f) \neq 0$. Having found such an element e for the element f, we pick out from N all the elements which are not orthogonal to e. There can be no more than a countable set of these elements by the inequality

$$|(e, f')|^2 + |(e, f'')|^2 + |(e, f''')|^2 + \cdots \leq \|e\|^2,$$

which holds for any set of elements in N, for if there were more than a countable number of terms the summation would not converge. We enumerate the elements which are not orthogonal to e by

$$f_1, f_2, \ldots, f_n \qquad (1 \leq n \leq \infty).$$

We denote the set of these elements by S. Since the set N is not countable, an eleement f' can be found in it which is different from all the elements of the set S and which is therefore orthogonal to e. We choose an element e' of M for which $(e', f') \neq 0$, and then pick out from N all the elements which are not orthogonal to e' and which do not appear in S. Let us denote them as

$$f_1', f_2', \ldots, f_n' \qquad (1 \leq n' \leq \infty)$$

and the set of these elements by S'.

This process may be continued indefinitely - transfinitely - and brings us a collection of sets $S_\alpha \subset N$ which have the following properties:

a) each element of N appears in just one of the sets S_α;

b) each set S_α corresponds to some element $e_\alpha \in M$, and different sets S_α, S_β correspond to different elements e_α, e_β;

c) each set S_α contains at least one element and does not contain more than a countable set of elements.

Let $\mathfrak{p}$ be the cardinality of the collection of sets S_α (by property a), $\mathfrak{n}$ is the same as the cardinality of the index set α). There is a theroem of set theory which tells us that the cardinality of the collection of a disjoint class of sets, each of which is countably infinite, is equal to the cardinaity of the index set of the class. If each set S_α is infinite then,

$$\mathfrak{n} = \mathfrak{p}.$$

Since the equation $\mathfrak{n} = \mathfrak{p}$ clearly holds in the other extreme case where each set S_α contains just one element, it follows that it holds in general.

On the other hand, it is clear that

$$\mathfrak{p} \leq \mathfrak{m}.$$

Therefore

$$\mathfrak{n} \leq \mathfrak{m}.$$

Interchanging the roles of M and N, we find in the same way that

$$\mathfrak{m} \leq \mathfrak{n}.$$

Consequently

$$\mathfrak{m} = \mathfrak{n}.$$

□

J.4.6 Projection Operators

From the defintion of a projection operator it is easy to see that

i) $P^2 = P$,

ii) $P^\dagger = P$.

Proof:

$\square$

Theorem J.4.13 *If P is an operator defined everywhere in $\mathcal{H}$, and such that, for all $h_1, h_2 \in \mathcal{H}$,*

i) $(P^2 h_1, h_2) = (Ph_1, h_2)$,

ii) $(Ph_1, h_2) = (h_1, Ph_2)$,.

then there is a subspace $G \subseteq \mathcal{H}$ such that P is the operator of projection on to G.

Proof: The operator P is bounded. This follows from the Hellinger-Toeplitz theorem, but is proved more easily by a direct argument,

$$\|Ph\|^2 = (Ph, Ph) = (P^2 h, h) = (Ph, h),$$

and therefore

$$\|Ph\|^2 \leq \|Ph\| \, \|h\|$$

and so $\|Ph\| \leq \|h\|$, i.e., the operator P is bounded and its norm is not greater than 1.

Let G denote the set of all vectors $g \in \mathcal{H}$ for which

$$Pg = g$$

It is clear that G is a linear manifold, and we now show that it is closed, so that it is also a susbspace. Let $g_n \in G$ $(n = 1, 2, 3, \ldots)$, so that

$$g_n = Pg_n,$$

and therefore, for any $g \in \mathcal{H}$,

$$Pg - g_n = Pg - Pg_n = P(g - g_n),$$

from which

$$\|Pg - g_n\| \leq \|g - g_n\|.$$

We thus see that $g_n \to g$ implies that $Pg = g$, and this shows that G is closed.

For any $h \in \mathcal{H}$ the vector $Ph = g$ belongs to G because $P(Ph) = Ph$. The element $P_G h$ also belongs to G. So it suffices to prove that

$$(Ph - P_G h, g') = 0$$

or

$$(Ph, g') = (P_G h, g')$$

for all $g' \in G$. But this follows from the fact that

$$\begin{aligned} (Ph, g') &= (h, Pg') = (h, g') \\ (P_G h, g') &= (h, P_G g') = (h, g'), \end{aligned}$$

therefore $P = P_G$.

$\square$

Theorem J.4.14 *If $\{P_k\}_{k=1}^{\infty}$ is an infinite monotonic sequence of projection operators, then as $k \to \infty$, P_k converges strongly to some projection operator P*

Proof:

$$\lim_{k \to \infty} P_k = P$$

the limit exists in the sense of strong convergence. Moreover, for any k and for any $f, g \in \mathcal{H}$,

$$(P_k f, P_k g) = (P_k f, g) = (f, P_k g).$$

Therefore in the limit $k \to \infty$

$$(Pf, Pg) = (Pf, g) = (f, Pg).$$

This implies $P = P^{\dagger} = P^2$, and so P is a projection operator.

□

Theorem J.4.15 *If a sequence $\{P_k\}_{k=1}^{\infty}$ of projection operators converges weakly to some projection operator P, then it also converges strongly to P*

Proof: By assumption, for any $h \in \mathcal{H}$,

$$(P_k h, h) \to (Ph, h).$$

Therefore

$$\|P_k h\| \to \|Ph\|,$$

and since the sequence of vectors $\{P_k h\}_{k=1}^{\infty}$ converges weakly to Ph, it must also converge strongly to Ph, by theorem ??.

□

Definition The **aperture** of two linear subspaces M_1, M_2 in $\mathcal{H}$ is the norm of the difference of the operators P_1, P_2 which project $\mathcal{H}$ on to the closure of these subspaces, denoted by $\theta(M_1, M_2)$. Thus

$$\theta(M_1, M_2) := \|P_1 - P_2\| \equiv \|P_2 - P_1\|.$$

□

Since for any $h \in \mathcal{H}$ the vectors $(\mathbb{I} - P_1)h$ and $P_1 h$ are orthogonal it is easy to see that the vectors $P_2(\mathbb{I} - P_1)h$ and $(\mathbb{I} - P_2)P_1 h$ are also orthogonal. So by the identity

$$P_2 - P_1 = P_2(\mathbb{I} - P_1) - (\mathbb{I} - P_2)P_1$$

it follows that

$$\begin{aligned}\|(P_2 - P_1)h\|^2 &= \|P_2(\mathbb{I} - P_1)h\|^2 + \|(\mathbb{I} - P_2)P_1 h\|^2 \\ &\leq \|(\mathbb{I} - P_1)h\|^2 + \|P_1 h\|^2 \\ &= \|h\|^2.\end{aligned}$$

This inequality shows that the aperature of two subspaces does not exceed 1:

$$\theta(M_1, M_2) \leq 1.$$

Moreover, we see that $\|(P_2 - P_1)h\|$ attains its maximum value of 1 if and only if one subspace contains a non-zero vector orthogonal to the other subspace. This fact is employed in the following theorem.

Theorem J.4.16 *If the aperture of two subspaces M_1, M_2 is less than 1, then the two subspaces have the same dimension:*

$$dim\, M_1 = dim\, M_2.$$

Proof: It suffices to show the inequality

$$\dim M_2 > \dim M_1.$$

implies that there is a non-zero vector in $\overline{M}_2$ which is orthogonal to $\overline{M}_1$. Consider the linear subspace

$$G = P_2 \overline{M}_1$$

Obviously

$$\dim G \leq \dim \overline{M}_1 < \dim \overline{M}_2$$

Therefore there is a non-zero vector in $\overline{M}_2$ orthogonal to G. This vector will be orthogonal to the whole subspace $\overline{M}_1$, since the subspace $\overline{M}_2$ is orthogonal to $\overline{M}_1 \ominus G$.

□

An alternative defintion of the aperture of two linear subspaces M_1, M_2 is:

$$\theta(M_1, M_2) = \max\left\{\sup_{f\in\overline{M}_2\,,\,\|f\|=1} \|(\mathbb{I} - P_1)f\|, \sup_{g\in\overline{M}_1\,,\,\|g\|=1} \|(\mathbb{I} - P_2)g\|\right\}. \tag{J.72}$$

Proof:

Proof of (J.72). By the original defintion of the aperture,

$$\begin{aligned}
\theta(M_1, M_2) &= \sup_{h\in\mathcal{H}} \frac{\|(P_2 - P_1)h\|}{\|h\|} \\
&= \sup_{h\in\mathcal{H}} \frac{\sqrt{\|P_2(\mathbb{I} - P_1)h\|^2 + \|(\mathbb{I} - P_2)P_1h\|^2}}{\|h\|}.
\end{aligned} \tag{J.73}$$

Say we allow h to only transverse $\overline{M}_1$ instead of the whole of $\mathcal{H}$, then the right-hand side of (J.73) is either unchanged or is decreased, and so

$$\begin{aligned}
\theta(M_1, M_2) &\geq \sup_{h\in\overline{M}_1} \frac{\sqrt{\|P_2(\mathbb{I} - P_1)h\|^2 + \|(\mathbb{I} - P_2)P_1h\|^2}}{\|h\|}. \\
&= \sup_{h\in\overline{M}_1} \frac{\|(\mathbb{I} - P_2)h\|}{\|h\|} =: \rho_2.
\end{aligned}$$

Similarly it may be shown that

$$\theta(M_1, M_2) \geq \sup_{h\in\overline{M}_2} \frac{\|(\mathbb{I} - P_1)h\|}{\|h\|} =: \rho_1.$$

Thus

$$\theta(M_1, M_2) \geq \max\{\rho_1, \rho_2\}.$$

It remains to prove

$$\theta(M_1, M_2) \leq \max\{\rho_1, \rho_2\}.$$

Notice that, by definition of ρ_2,

$$\|(\mathbb{I} - P_2)P_1h\|^2 \leq \rho_2^2\|P_1h\|^2. \tag{J.74}$$

On the other hand,

$$
\begin{array}{rcl}
\|P_2(\mathbb{I} - P_1)h\|^2 & = & (P_2(\mathbb{I} - P_1)h, P_2(\mathbb{I} - P_1)h) \\
 & = & (P_2(\mathbb{I} - P_1)h, (\mathbb{I} - P_1)h) \\
 & = & ((\mathbb{I} - P_1)P_2(\mathbb{I} - P_1)h, (\mathbb{I} - P_1)h) \\
 & \leq & \|(\mathbb{I} - P_1)P_2(\mathbb{I} - P_1)h\| \cdot \|(\mathbb{I} - P_1)h\|
\end{array}
$$

and therefore by definition of ρ_1,

$$\|P_2(\mathbb{I} - P_1)h\|^2 \leq \rho_1 \|P_2(\mathbb{I} - P_1)h\| \cdot \|(\mathbb{I} - P_1)h\|$$

i.e.,

$$\|P_2(\mathbb{I} - P_1)h\| \leq \rho_1 \|(\mathbb{I} - P_1)h\|. \tag{J.75}$$

From (J.74) and (J.75) we conclude that

$$
\begin{array}{rcl}
\|(\mathbb{I} - P_2)P_1 h\|^2 + \|P_2(\mathbb{I} - P_1)h\|^2 & \leq & \rho_2^2 \|P_1 h\|^2 + \rho_1^2 \|(\mathbb{I} - P_1)h\|^2 \\
 & \leq & \max\{\rho_1, \rho_2\} \left[\|P_1 h\|^2 + \|(\mathbb{I} - P_1)h\|^2\right] \\
 & = & \|h\|^2 \cdot \max\{\rho_1, \rho_2\},
\end{array}
$$

and (J.73) gives the required result

$$\theta(M_1, M_2) \leq \max\{\rho_1, \rho_2\}.$$

□

J.4.7 Spectral Decomposition

We now reexpress these results in a way more to make convenient the transition to infinite-dimension case. For a self-adjoint and unitary operators the eigenvalues are ordered in a natural way. For a self-adjoint operator this is so because the eigenvalues are real numbers and for a unitary operators the eigenvalues are represented by points on the unit circle of the complex plane.

Let us first consider self-adjoint operators. We assume that $\lambda_1 < \lambda_2 < \cdots < \lambda_m$, and we use the P_i's to define new projections:

$$
\begin{aligned}
E_{\lambda_0} &= 0 \\
E_{\lambda_1} &= P_1 \\
E_{\lambda_2} &= P_1 + P_2; \\
&\vdots \\
E_{\lambda_m} &= P_1 + P_2 + \cdots + P_m.
\end{aligned} \tag{J.76}
$$

$$
\begin{aligned}
A &= \lambda_1 P_1 + \lambda_2 P_2 + \cdots \lambda_m P_m \\
&= \lambda_1 (E_{\lambda_1} - E_{\lambda_0}) + \lambda_2 (E_{\lambda_2} - E_{\lambda_1}) + \cdots \lambda_m (E_{\lambda_m} - E_{\lambda_{m-1}}) \\
&= \sum_{i=1}^{m} \lambda_i (E_{\lambda_i} - E_{\lambda_{i-1}}).
\end{aligned} \tag{J.77}
$$

$$
A = \sum_{i=1}^{m} \lambda_i \Delta E_{\lambda_i} \tag{J.78}
$$

Proof:

□

Theorem J.4.17 *For each self-adjoint operator there is a unique spectral family of projection operators P_x such that*

Proof:

□

Theorem

$$
(\phi, A\varphi) = \int_{\infty}^{\infty} x d(\phi, P_x \varphi) \tag{J.79}
$$

for all vectors φ and ϕ.

The spectral decomposition of A

$$
A = \int_{-\infty}^{\infty} dE_x \tag{J.80}
$$

$$\begin{aligned}(\phi,\psi) &= (\phi,P_1\psi)+(\phi,P_2\psi)+\cdots(\phi,P_m\psi)\\ &= (\phi,(E_{\lambda_1}-E_{\lambda_0})\psi)+(\phi,(E_{\lambda_2}-E_{\lambda_1})\psi)+\cdots(\phi,(E_{\lambda_m}-E_{\lambda_{m-1}})\psi)\\ &= \sum_{i=1}^{m}(\phi,\Delta E_{\lambda_i}\psi).\end{aligned}\tag{J.81}$$

$$\begin{aligned}(\phi,A\psi) &= (\phi,\lambda_1P_1\psi)+(\phi,\lambda_2P_2\psi)+\cdots(\phi,\lambda_mP_m\psi)\\ &= (\phi,(E_{\lambda_1}-E_{\lambda_0})\psi)+(\phi,(E_{\lambda_2}-E_{\lambda_1})\psi)+\cdots(\phi,(E_{\lambda_m}-E_{\lambda_{m-1}})\psi)\\ &= \sum_{i=1}^{m}\lambda_i(\phi,\Delta E_{\lambda_i}\psi).\end{aligned}\tag{J.82}$$

for any vectors φ and ϕ we have

$$(\phi,\psi)=\int_{-\infty}^{\infty}d(\phi,E_x\psi)\tag{J.83}$$

and

$$(\phi,A\psi)=\int_{-\infty}^{\infty}xd(\phi,E_x\psi).\tag{J.84}$$

Unitary operators can be treated in a similarly. For a unitary operator U let its eigenvalues be $u_i=e^{i\theta_i}$ labelled in the order

$$0<\theta_1<\theta_2<\cdots<\theta_{m-1}<\theta_m\leq 2\pi\tag{J.85}$$

For each real number x let

$$E_x=\sum_{\theta_i\leq x}P_i\tag{J.86}$$

This is the progection operator onto the space spanned by all eigenvectors for eigenvalues $e^{i\theta_i}$ with $\theta_i\leq x$. If $x\leq 0$, then $E_x=0$. If $x\geq 2\pi$, then $E_x=1$. Evidently E_x increses by incriments P_i the same as for the Hermitian operator with eigenvalues θ_i. For

$$U=\sum_{i=1}^{m}u_iP_i=\sum_{i=1}^{m}e^{i\theta_i}P_i\tag{J.87}$$

we can write

$$U = \int_0^{2\pi} e^{ix} dE_x. \tag{J.88}$$

For any vectors ψ and ϕ we have

$$(\phi, U\psi) = \int_0^{2\pi} e^{ix} d(\phi, E_x\psi). \tag{J.89}$$

Theorem J.4.18 *Let A be a self-adjoint operator with spectral decomposition*

$$A = \int_{-\infty}^{\infty} x dE_x.$$

Then E_x jumps in value at $x = a$ if and only if a is an eigenvalue of A.

Proof: First we prove that if there is a discontinuity from the right at $x = a$ that a is an eigenvalue. Take $\epsilon < \epsilon' < 0$ and any vector ψ,

$$\begin{aligned}
\|(E_a - E_{a-\epsilon})\psi - (E_a - E_{a-\epsilon'})\psi\|^2 &= \|(E_a - E_{a-\epsilon})\psi\|^2 + \|(E_a - E_{a-\epsilon'})\psi\|^2 \\
&\quad -\big(\psi, (E_a - E_{a-\epsilon})(E_a - E_{a-\epsilon'})\psi\big) \\
&\quad -\big(\psi, (E_a - E_{a-\epsilon'})(E_a - E_{a-\epsilon})\psi\psi\big) \\
&= \|(E_a - E_{a-\epsilon})\psi\|^2 + \|(E_a - E_{a-\epsilon'})\psi\|^2 \\
&\quad -2\big(\psi, (E_a - E_{a-\epsilon'})\psi\big) \\
&= \|(E_a - E_{a-\epsilon})\psi\|^2 - \|(E_a - E_{a-\epsilon'})\psi\|^2
\end{aligned}$$

because $E_a - E_{a-\epsilon}$ and $E_a - E_{a-\epsilon'}$ are projection operators and $E_a - E_{a-\epsilon} \geq E_a - E_{a-\epsilon'}$. This shows that $\|(E_a - E_{a-\epsilon})\psi\|^2$ is monotonically decreasing as $\epsilon \to 0$ and by ... shows that the vectors $(E_a - E_{a-\epsilon})\psi$ have the Cauchy property. Because a Hilbert space is complete by definition, $(E_a - E_{a-\epsilon})\psi$ must converge to a limit vector as $\epsilon \to 0$. We denote the limit vector by ψ_a. We are supposing that there is a discontinuity and hence take ψ_a to be non-zero. If $x \geq a$, then

$$E_x(E_a - E_{a-\epsilon}) = E_xE_a - E_xE_{a-\epsilon} = E_a - E_{a-\epsilon}.$$

If $x < a$, then for ϵ small enough that $x < a - \epsilon$

$$E_x(E_a - E_{a-\epsilon}) = E_xE_a - E_xE_{a-\epsilon} = E_x - E_x = 0.$$

Therefore $E_x\psi_a = 0$ for $x < a$ and $E_x\psi_a = \psi_a$ for $x \geq a$, so

$$\begin{aligned}
(\phi, A\psi_a) &= \int_{-\infty}^{\infty} x d(\phi, E_x\psi_a) \\
&= \int_{-\infty}^{\infty} x \frac{d}{dx}(\phi, E_x\psi_a) dx \\
&= \int_{-\infty}^{\infty} x(\phi, E_x\psi_a)\delta(x-a) dx \\
&= (\phi, a\psi_a)
\end{aligned}$$

for any vector ϕ. Putting $\phi = A\psi_a - a\psi_a$ we have

$$(A\psi_a - a\psi_a, A\psi_a - a\psi_a) = \|A\psi_a - a\psi_a\|^2 = 0$$

so

$$A\psi_a = a\psi_a.$$

Thus a is an eigenvalue and ψ_a is an eigenvector.

Now we prove that if a is an eigenvalue that E_x jumps in value at $x = a$. For any vectors ψ and ϕ we have

$$
\begin{aligned}
(\phi, A^2\psi) &= (\phi, A(A\psi)) \\
&= \int_{-\infty}^{\infty} x d(\phi, E_x A\psi) \\
&= \int_{-\infty}^{\infty} x d(E_x\phi, A\psi) \\
&= \int_{-\infty}^{\infty} x d_x \int_{-\infty}^{\infty} y d_y (E_x\phi, E_y\psi) \\
&= \int_{-\infty}^{\infty} x d_x \int_{-\infty}^{\infty} y d_y (\phi, E_x E_y\psi) \\
&= \int_{-\infty}^{\infty} x d_x \int_{-\infty}^{x]} y d_y (\phi, E_x E_y\psi) + \int_{-\infty}^{\infty} x d_x \int_{(x}^{\infty} y d_y (\phi, E_x E_y\psi) \\
&= \int_{-\infty}^{\infty} x d_x \int_{-\infty}^{x]} y d_y (\phi, E_y\psi) + \int_{-\infty}^{\infty} x d_x \int_{(x}^{\infty} y d_y (\phi, E_x\psi) \\
&= \int_{-\infty}^{\infty} x^2 d(\phi, E_x\psi)
\end{aligned}
$$

because $E_x E_y = E_y$ for $y \leq x$, and for $y > x$

$$(\phi, E_x E_y) = (\phi, E_x\psi)$$

which is independent of y.

By assumption a is an eigenvalue of A and let I_a be the projection onto the subspace of eigenvectors. Then

$$\int_{-\infty}^{\infty} (x^2 - 2ax - a^2) d(\phi, E_x I_a\psi) = (\phi, [A^2 - 2aA + a^2] I_a\psi) = 0$$

for any vectors ψ and ϕ. For $\phi = I_a\psi$ we have

$$(\phi, E_x I_a\psi) = (I_a\psi, E_x^2 I_a\psi) = \|E_x I_a\psi\|^2$$

and

$$\int_{-\infty}^{\infty} (x-a)^2 d\|E_x I_a\psi\|^2 = 0.$$

Because $\|E_x I_a\psi\|^2$ is a monotonically increasing function of x, we see that it can only change value at $x = a$. But $\|E_x I_a\psi\|^2 \to 0$ as $x \to -\infty$ and $\|E_x I_a\psi\|^2 \to \|I_a\psi\|^2$ as

$x \to +\infty$. Therefore $\|E_x I_a \psi\|^2 = 0$ if $x < a$, and $\|E_x I_a \psi\|^2 = \|I_a \psi\|^2$ if $x \geq a$ for any vector ψ. This implies that $E_x I_a = 0$ for $x < a$ and $E_x I_a = I_a$ for $x \geq a$. For any vector ψ we have

$$(E_a - E_{a-\epsilon}) I_a \psi = (E_a I_a - E_{a-\epsilon} I_a) = I_a \psi.$$

Thus E_x jumps in value at $x = a$.

□

Dirac Notation

A vector $|\Psi> \in \mathcal{H}$ is called a **ket** and to this vector we can associate a linear functional $\omega_{|\Psi>} \equiv <\Psi|$ called a **bra** and defined by means of a scalar product. This acts on elements of the Hilbert space to give a complex number.

$$\omega_{|\Psi>} : \mathcal{H} \to C \tag{J.90}$$

and for typical ket Φ we write

$$\omega_{|\Psi>}(|\Phi>) < |\Psi>, |\Phi>>_{\mathcal{H}} \equiv <\Psi|\Phi> . \tag{J.91}$$

Conversely, to each bra $<\Psi| \in \mathcal{H}^*$ we can associate a ket $|\Psi> \in \mathcal{H}$. Moreover, by virtue of the Theorem ?? just proven, every element $\omega \in \mathcal{H}^*$ uniquely determines a vector $\Psi_\omega \in \mathcal{H}$ such that

$$\omega(|\Phi>) = <\Psi_\omega|\Phi> \quad \text{for all } \; |\Phi> \in \mathcal{H}. \tag{J.92}$$

The vector

we thus have a one-to-one correspondence between $|\Psi> \in \mathcal{H}$ and $<\Psi| \in \mathcal{H}^*$. We can identify (isometry)

• **Every ket has a corresponding bra**

If one defines the norm of $\omega \in \mathcal{H}^*$ by $||\omega|| := \sup|\omega(f)|$, where the supremum is taken over all unit vectors $f \in \mathcal{H}$, then .... one-to-one and *onto* i.e. a **bijection** $\mathcal{H} \to \mathcal{H}^*$ is antilinear and is norm-preserving, that is it represents an **isometry**.

The idea of an operator $\hat{A}$ is something tht acts on a ket $|\psi>$ and transforms it into another ket $|\psi'>$ of the same space and when it acts on a bra it transforms it into another bra:

$$\hat{A}|\psi>=|\psi'>,\quad <\psi|\hat{A}=<\psi'|. \tag{J.93}$$

The introduction of a dual vector space is only necessary for generalized vectors.

J.5 Distributions (or Generalized Functions)

J.5.1 Introduction

The Dirac delta-function can not be defined as a function but as a functional on a suitable space of functions.

We note that distributions play an essential role in in quantum field theory, where quantum fields are defined as operator-valued distributions. According to the LQG approach, the surface states of a black hole, which are the statistical mechanical origin of black hole entropy, must be the pull-back of distributional fields of the bulk spacetime.

J.5.2 Mathematical Formulism of Quantum Mechanics

Schwartz space

$$h_\alpha = \sqrt{\frac{\pi^{|\alpha|}}{\alpha!}} Z^{*\alpha} h_0$$

$$\begin{aligned} Z_j^* f(x) &= x_j f(x) - \frac{1}{2\pi}\frac{\partial f}{\partial x_j} \\ &= -\frac{1}{2\pi} e^{\pi x^2} \frac{\partial}{\partial x_j}(e^{-\pi x^2} f(x)) \end{aligned}$$

so that

$$\begin{aligned} Z_j^{*2} f(x) &= \left(x_j - \frac{1}{2\pi}\frac{\partial}{\partial x_j}\right)\frac{-1}{2\pi} e^{\pi x^2}\frac{\partial}{\partial x_j}(e^{-\pi x^2} f(x)) \\ &= \left(-\frac{1}{2\pi}\right)^2 e^{\pi x^2}\left(\frac{\partial}{\partial x_j}\right)^2 (e^{-\pi x^2} f(x)) \end{aligned}$$

and

$$Z_j^{*\alpha} f(x) = \left(-\frac{1}{2\pi}\right)^{|\alpha|} e^{\pi x^2} \left(\frac{\partial}{\partial x}\right)^{\alpha} (e^{-\pi x^2} f(x))$$

explicityly

$$h_\alpha(x) = \frac{2^{n/4}}{\sqrt{\alpha!}} \left(\frac{-1}{2\sqrt{\pi}}\right)^{|\alpha|} e^{\pi x^2} \left(\frac{\partial}{\partial x}\right)^{\alpha} e^{-\pi x^2} \tag{J.94}$$

The eigenfunctions of $\mathcal{F}$ are obtained by taking tensor products of the one dimensional Hermite functions.

For each $\alpha \in \mathbb{N}^n$ we define

$$\Phi_\alpha(x) = \varphi_{\alpha_1}(x_1)\varphi_{\alpha_2}(x_2)\ldots\varphi_{\alpha_n}(x_n).$$

Since

$$\begin{aligned} [Z_j, Z_k^*]f(x) &= \Big[x_j + \frac{1}{2\pi}\frac{\partial}{\partial x_j}, x_k - \frac{1}{2\pi}\frac{\partial}{\partial x_k}\Big] f(x) \\ &= \pi^{-1}\delta_{jk} \end{aligned}$$

we obviouslsy have

$$[Z_j, Z^{*\alpha}] = \pi^{-1}\alpha_i Z^{*(\alpha - 1_j)}$$

$$Z_j h_0 = 2^{n/4} Z_j e^{\pi x^2} = 2^{n/4}(x_j + \frac{1}{2\pi}\frac{\partial}{\partial x_j})e^{-\pi x^2} = 0$$

Lemma J.5.1 *$\{h_\alpha\}$ is an orthonormal basis for $L^2(\mathbb{R}^n)$.*

Proof: We have

$$(h_\alpha, h_\beta) = \sqrt{\frac{\pi^{|\alpha|+|\beta|}}{\alpha!\beta!}}(h_0, Z^\alpha Z^{*\beta} h_0).$$

Consider $\alpha_j > \beta_j$

$$\begin{aligned} Z_j^{\alpha_j} Z_j^{*\beta_j} h_0 &= Z_j^{\alpha_j-1} Z_j^{*\beta_j} Z_j h_0 + \pi^{-1}\alpha_j Z_j^{\alpha_j-1} Z_j^{*\beta_j-1} h_0 \\ &= \pi^{-2}\alpha_j(\alpha_j-1) Z_j^{\alpha_j-2} Z_j^{*\beta_j-2} h_0 \\ &= \pi^{-\beta_j}\alpha_j(\alpha_j-1)\dots(\alpha_j-\beta_j+1) Z_j^{\alpha_j-\beta_j} h_0 = 0 \end{aligned}$$

If $\alpha_j = \beta_j$

$$Z_j^{\alpha_j} Z_j^{*\alpha_j} h_0 = \pi^{-\alpha_j}\alpha_j! h_0$$

So we have

$$(h_\alpha, h_\beta) = \delta_{\alpha\beta}\|h_0\|_2^2 = \delta_{\alpha\beta}.$$

Now we prove completeness. If $g \in L^2(\mathbb{R}^n)$ and

$$(g, h_\alpha) = 0 \quad \text{for all} \quad \alpha,$$

Now the function $H_n(x) = e^{\pi x^2} h_\alpha(x)$ is a polynomial of degree $|\alpha|$ called the αth Hermite polynomial. Obviously, every polynomial of degree less than equal to k on $\mathbb{R}^n$ is some linear combination of Hermite polynomials of degree less than equal to k, hence the condition $(g, h_\alpha) = 0$ for all α is equivalent to $(g, P(x)e^{-\pi x^2}) = 0$ for all polynomials P. But it then follows that

$$\int g(x) e^{-\pi x^2} e^{2\pi i x\zeta} dx = \sum_{j=1}^{\infty} \int g(x) e^{-\pi x^2} \frac{(2\pi i x\zeta)^j}{j!} dx = 0.$$

By the Fourier uniqueness, $g(x)e^{-\pi x^2} = 0$ a.e., and hence $g = 0$.

□

$$\begin{aligned} \mathcal{F} Z_j^* f &= \int dx e^{2\pi i k x}(X_j - iD_j) f(x) \\ &= \int dx e^{2\pi i k x}(x_j - \frac{1}{2\pi}\frac{d}{dx_j}) f(x) \\ &= (i\frac{1}{2\pi}\frac{d}{dk_j} - ik_j)\int dx e^{2\pi i k x} f(x) \\ &= -iZ_j^* \mathcal{F} f \end{aligned}$$

Recalling

$$\mathcal{F}h_0 = h_0$$

$$\begin{aligned}\mathcal{F}h_\alpha &= \sqrt{\frac{\pi^{|\alpha|}}{\alpha!}}\mathcal{F}Z^{*\alpha}h_0 \\ &= (-i)^{|\alpha|}\sqrt{\frac{\pi^{|\alpha|}}{\alpha!}}Z^{*\alpha}h_0 \\ &= (-i)^{|\alpha|}h_\alpha\end{aligned}$$

Then it follows that

$$\mathcal{F}\Phi_\alpha = (-i)^{|\alpha|}\Phi_\alpha$$

where $|\alpha| = \sum_{j=1}^n \alpha_j$. Moreover, $\{\Phi_\alpha : \alpha \in \mathbb{N}^n\}$ is an orthonormal basis for $L^2(\mathbb{R}^n)$. Thus every $f \in L^2(\mathbb{R}^n)$ has the expansion

$$f = \sum_{\alpha\in\mathbb{N}^n}(f, \Phi_\alpha)\Phi_\alpha$$

the series being convergent in the L^2 norm.

Definition A subspace W of $L^2(\mathbb{R}^n)$ is said to be invariant under $\mathcal{F}$ if $\mathcal{F}f \in W$ whenever $f \in L^2(\mathbb{R}^n)$.

$\square$

As $\{\Phi_\alpha : \alpha \in \mathbb{N}^n\}$ is an orthonormal basis for $L^2(\mathbb{R}^n)$ it follows that $f \in L^2(\mathbb{R}^n)$ if and only if

$$\sum_{\alpha\in\mathbb{N}^n}|(f, \Phi_\alpha)|^2 < \infty.$$

Since

$$(\mathcal{F}f, \Phi_\alpha) = (-i)^{|\alpha|}(f, \Phi_\alpha)$$

any subspace of $L^2(\mathbb{R}^n)$ defined in terms of its behaviour of $|(f, \Phi_\alpha)|$ will be invariant under the Fourier transform. We can define a whole family of invariant subspaces. Indeed, for each $s > 0$ define $W^s_H(\mathbb{R}^n)$ to be a subspace of $L^2(\mathbb{R}^n)$ consisting of functions f for which

$$\|f\|^2_{2,s} = \sum_{\alpha \in \mathbb{N}^n} (2|\alpha| + n)^2 |(f, \Phi_\alpha)|^2 < \infty.$$

Definition Let $\mathcal{S}(\mathbb{R}^n) = \cap_{s>0} W^s_H(\mathbb{R}^n)$. This is called the Schwartz space and its members are called Schwartz functions.

□

Theorem J.5.2 *The Schwartz space*

(i) is a dense subspace of $L^2(\mathbb{R}^n)$*;*

(ii) $\mathcal{S}(\mathbb{R}^n)$ *is invariant under* $\mathcal{F}$*;*

(iii) $\mathcal{F} : \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n)$ *is one to one and onto.*

Proof: The density follows from the fact that finite linear combinations of Hermite functions form a subspace of $\mathcal{S}(\mathbb{R}^n)$ which is dense in $L^2(\mathbb{R}^n)$.

The inverse follows from that of each $W^s_H(\mathbb{R}^n)$.

As $\mathcal{F}(\mathcal{F}^* f) = f$ surjectivity follows.

□

The original motivation for RHS was to provide a rigorous mathematical context for Dirac's bra-and-ket formulation of quantum mechanics.

$$\sqrt{\frac{n}{\pi}} e^{-nx^2} \quad \text{for } n = 1, 2, \ldots \tag{J.95}$$

$$\lim_{n \to \infty} D_n(x) = \begin{cases} \infty & x = 0 \\ 0 & x \neq 0 \end{cases} \tag{J.96}$$

$$\int_{-\infty}^{\infty} D_n(x) dx = 1 \tag{J.97}$$

for all n

Dirac formulism on concrete mathematical

A distribution $\chi(x)$ is the limit of a sequence of good functions $h(x)$ such that

$$\lim_{n\to\infty}\int_{-\infty}^{\infty} h_n(x)g(x)dx := \int_{-\infty}^{\infty}\chi(x)g(x)dx \tag{J.98}$$

exists i.e. $< \infty$. Lebesgue integral?????

The product of two generalized functions is not defined. the convergence of

$$\int_{-\infty}^{\infty} a_n(x)g(x)dx \quad \text{and} \quad \int_{-\infty}^{\infty} a_n(x)g(x)dx \tag{J.99}$$

as $n \to \infty$, does not imply the convergence of

$$\int_{-\infty}^{\infty}[a_n(x)b_n(x)]g(x)dx \tag{J.100}$$

products of distributional valued operators in Field theory need to be regularized to make mathematical sence.

Theorem Let the sequnce $f_n(x)$ $(n = 1, 2, \ldots$ define a generalized function $\varphi(x)$. Then the sequence of Fourier transforms of the sequence $f_n(x)$

$$F_n = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} f_n(x)dx \tag{J.101}$$

defines a generalized function $\Phi(t)$, which is called the Fourier transform of $\varphi(x)$

$$\int_{-\infty}^{\infty} F_n(t)g(t)dt \tag{J.102}$$

Let

$$g(t) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} G(x)e^{itx}dt \tag{J.103}$$

$$\int_{-\infty}^{\infty} F_n(t)g(t)dt = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_n(x)g(t)e^{itx}dtdx \tag{J.104}$$

and also

$$\int_{-\infty}^{\infty} f_n(x)G(x)dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_n(x)g(t)e^{itx}dtdx \tag{J.105}$$

Since the RHS of exist, we have

$$\int_{-\infty}^{\infty} F_n(t)g(t)dt = \int_{-\infty}^{\infty} f_n(x)G(x)dx \tag{J.106}$$

Formally

$$\Phi(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \varphi(x)e^{-itx}dt \tag{J.107}$$

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \Phi(t)e^{itx}dt \tag{J.108}$$

$$G(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} g(x)e^{-itx}dx \tag{J.109}$$

a derivative of the mth order of $G(t)$ is

$$\frac{d^m G(t)}{dt^m} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (-ix)^m g(x)e^{-itx} \tag{J.110}$$

$$\begin{aligned} \left|\frac{d^m G(t)}{dt^m}\right| &= \left|\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (-ix)^m g(x)e^{-itx}dx\right| \\ &\leq \frac{1}{\sqrt{2\pi}} \frac{1}{|t|^n} \int_{-\infty}^{\infty} \left|\frac{d^n}{dx^n}[x^m g(x)]\right| dx \end{aligned} \tag{J.111}$$

Dirac δ function

$$D_n(x) = \sqrt{\frac{n}{\pi}} e^{-nx^2} \quad (n = 1, 2, \dots) \tag{J.112}$$

for a good function $g(x)$,

$$\int_{-\infty}^{\infty} \delta(x)g(x)dx = g(0) \tag{J.113}$$

$$\delta(y(x)) = \sum_{i=1} \frac{\delta(x - x_i)}{\left|\frac{dy}{dx}\right|_{x=x_i}} \tag{J.114}$$

We define distributions to be abstract linear functionals defined on Schwartz space, i.e. linear maps from $\mathcal{S}(\mathbb{R}^n)$ to the complex numbers. The delta function is then the linear functional

$$\phi \mapsto \phi(0).$$

The Fourier transform has a natural extension to $\mathcal{S}'(\mathbb{R}^n)$. If f is a Schwartz function, then

$$(\mathcal{F}u, f) = \int\int f(\zeta)e^{ix\cdot\zeta}\phi(x)dxd\zeta = (u, \mathcal{F}f)$$

Thus we define, for any distribution, the Fourier transform $\mathcal{F}u$ by

$$(\mathcal{F}u, f) = (u, \mathcal{F}f)$$

where the brackets here stands for the action of a tempered distribution on a Schwartz function.

$\delta-$sequences

convolutions of $\delta-$sequences.

$$\sigma_{m,n}(x) := \int_{-\infty}^{\infty} \alpha_n^*(t)\beta_m(x - t)dt \tag{J.115}$$

$$\lim_{m,n\to\infty} \int_{-\infty}^{\infty} \sigma_{m,n}(x)g(x)dx = g(0)? \tag{J.116}$$

$$\begin{aligned}
\lim_{m,n\to\infty} \int_{-\infty}^{\infty} \sigma_{m,n}(x)g(x)dx &= \lim_{m,n\to\infty} \int_{-\infty}^{\infty} (\int_{-\infty}^{\infty} \alpha_n^*(t)\beta_m(x - t)dt)\ g(x)dx \\
&= \lim_{m\to\infty} \int_{-\infty}^{\infty} \left(\lim_{n\to\infty} \int_{-\infty}^{\infty} \alpha_n^*(t)\beta_m(x - t)dt\right)\ g(x)dx \\
&= \lim_{m\to\infty} \int_{-\infty}^{\infty} \beta_m(x)g(x)dx \\
&= g(0).
\end{aligned} \tag{J.117}$$

$$\lim_{n\to\infty}\int_{-\infty}^{\infty} f_n(x)dx = \int_{-\infty}^{\infty}\lim_{n\to\infty} f_n(x)dx \tag{J.118}$$

where

$$f_n(x) := g(x)\int_{-\infty}^{\infty} \alpha_n^*(t)\beta_m(x-t)dt \tag{J.119}$$

Is $f_n(x)$ uniformally continuous.

J.5.3 Gel'fand triple

Most of the operators one wants to use are in practice are defined on a common dense domain, S, consisting of the space of infinetly differentiable wavefunctions, ψ, for which $\hat{x}^j\hat{p}^k\psi$ is normalizable for all positive integers j and k. This actually too small to be of much use, but it has a very large dual space, S', called the space of *tempered distributions*, to which most interesting operators can be transposed.

J.6 Linear Operators

We have seen that in a particular basis $\mathbf{e}_i$ linear operators can be described in terms of their components with respect ot the basis. These components can be displayed as an array of numbers - a matrix.

A *linear operator* A on a vector space assigns to each vector ψ a vector $A\psi$ such that

$$A(\psi+\phi) = A\psi + A\phi \tag{J.120}$$

and

$$A(a\psi) = aA\psi \tag{J.121}$$

for any vectors ψ and ϕ and a scalar a.

Linear operators form a with sum $A+B$ of two operators A and B and a scalar multiple aA of an operator A and a scalar A defined by

$$(A+B)\psi = A\psi + B\psi \tag{J.122}$$

and

$$(aA)\psi = aA\psi \tag{J.123}$$

for every vector ψ.

J.6.1 Mappings

Let X and Y be sets and $A \subset X$ any subset. A mapping T from A into Y is defined by an association with each $x \in A$ a single $y \in Y$, written and called the image of x with respect to T.

Definition The set A is called the domain of defintion of T or just the **domain** of T and is denoted $D(T)$, and we write

□

$$\begin{array}{rcl} T : D(T) & \rightarrow & Y \\ x & \mapsto & Tx \end{array}$$

Definition The **range** $\mathrm{Ran}(T)$ of T is the set of all images, that is,

$$\mathrm{Ran}(T) = \{y \in Y : y = Tx \text{ for some } x \in D(T)\}.$$

□

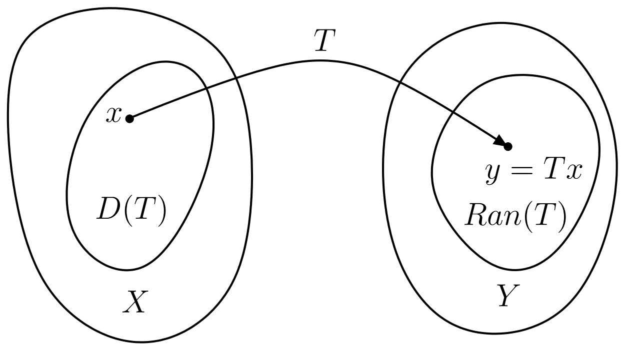

Figure J.4: Visualization of a mapping.

Definition The **inverse image** of a $y_0 \in Y$ is the set of all $x \in D(T)$ such that $Tx = y_0$.

□

Definition The mapping T^{-1} is **injective**, an **injection**, or **one-to-one** if for every $x_1, x_2 \in D(T)$,

$$x_1 \neq x_2 \qquad \text{implies} \qquad Tx_1 \neq Tx_2;$$

that is, different points in $D(T)$ have different images, so that the inverse image of any point in $\mathrm{Ran}(T)$ is a single point.

□

If a mapping is injective different points in the domain have different images, that is, if for any $x_1, x_2 \in D(T)$,

$$x_1 \neq x_2 \qquad \Rightarrow \qquad Tx_1 \neq Tx_2; \tag{J.124}$$

equivalently,

$$Tx_1 = Tx_2 \qquad \Rightarrow \qquad x_1 = x_2. \tag{J.125}$$

In this case there exists a maping $T^{-1} : \mathrm{Ran}(T) \to D(T)$ $y_0 \mapsto x_0$, which maps every $y_0 \in \mathrm{Ran}(T)$ onto that $Tx_0 = y_0$.

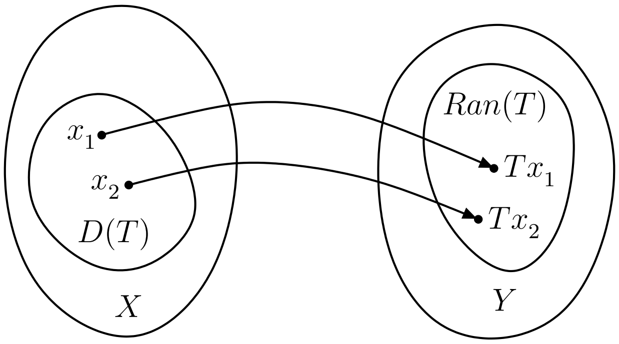

Figure J.5: Injective mapping.

Definition The mapping T^{-1} is **surjective**, a **surjection**, or a mapping of $D(T)$ **onto** Y if $\mathrm{Ran}(T) = Y$. Clearly,

$$\begin{array}{rcl} D(T) & \to & \mathrm{Ran}(T) \\ x & \mapsto & Tx \end{array}$$

$\square$

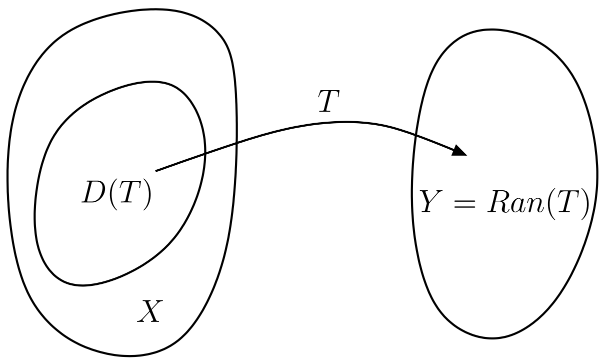

Figure J.6: Surjective mapping.

Definition The mapping T^{-1} is called the **inverse** of T.

$\square$

Theorem J.6.1 *Let X and Y be vector spaces. Let $T : D(T) \to Y$ be a linear operator with domain $D(T) \subset X$ and range $Ran(T) \subset Y$. Then*

The inverse $T^{-1} : Ran(T) \to D(T)$ exists if and only if

$$Tx = 0 \qquad \Rightarrow \qquad x = 0.$$

Proof: Suppose that $Tx = 0$ implies $x = 0$. Let $Tx_1 = Tx_2$. Since T is linear,

$$T(x_1 - x_2) = Tx_1 - Tx_2 = 0,$$

so that $x_1 - x_2 = 0$ by assumption. Hence $Tx_1 = Tx_2$ implies $x_1 = x_2$, and T^{-1} exists by (J.125).

Conversely, if T^{-1} exists, then by (J.125) holds. Putting $x_2 = 0$ in (J.125) we obtain

$$Tx_1 = T0 = 0 \qquad \Rightarrow \qquad x_1 = 0.$$

□

J.6.2 Bounded Operators

Definitions. A linear operator is *continuous* if $A\psi_n \to A\psi$ for any sequence of vectors ψ_n which converge to a limit vector ψ. A linear operator is *bounded* if there is a positive number b such that $||A|| \leq b||\psi||$ for every vector ψ ; the smallest number with this property is called the *norm* of A and is denoted by $||A||$.

Theorem A linear operator is continuous if and only if it is bounded.

Proof. Let A be a bounded linear operator. If the sequqnce of vectors ψ_n converges to a limit vector ψ, then

$$||A\psi - A\psi_n|| = ||A(\psi - \psi_n)|| \leq ||A|| \, ||\psi - \psi_n|| \to 0 \tag{J.126}$$

so $A\psi_n \to A\psi$ as $n \to \infty$. Thus A is continuous.

To complete the proof we suppose a linear operator A is not bounded and show that it is not continuous. As A is not bounded there does not exist a number b such that $||A|| \leq b||\phi||$ for every vector ϕ. Therefore for each positive integer n there must be a vector ψ_n such that $||A\psi_n|| > n||\psi_n||$. Let $\chi_n = (1/n||\psi_n||)\psi_n$. Then $||\chi_n|| = 1/n$ so $\chi_n \to 0$ as $n \to \infty$. But $||A\chi_n|| > 1$ so $A\chi_n \not\to o$ as $n \to \infty$. Thus A is not continuous.

Adjoints

The *adjoint* of a bounded linear operator A is defined by letting

$$(\phi, A^\dagger\psi) = (A\phi, \psi) \tag{J.127}$$

for all vectors ψ and ϕ. For each vector ψ a linear functional F is defined by $F(\phi) = (\psi, A\phi)$. This linear functional is continuous because A is bounded; if a sequence of vectors χ_n converges to a limit vector χ, then

$$|F(\chi) - F(\chi_n)| = |(\psi, A[\chi - \chi_n])| \leq ||\psi|| \, ||A|| \, ||\chi - \chi_n|| \to 0 \tag{J.128}$$

as $n \to \infty$. This implies (by theorem) that there exists a unique vector, which we call $A^\dagger\psi$, such that $F(\phi) = (A^\dagger\psi, \phi)$ or $(A^\dagger\psi, \phi) = (\psi, A\phi)$.

That $A^\dagger$ is a linear operator is evident form the definition. It is also bounded.

$$\hat{A}(a_1|\psi_1>+a_2|\psi_2>) = a_1\hat{A}|\psi_1>+a_2\hat{A}|\psi_2>, \tag{J.129}$$

J.7 Three Main Theorems in Functional Analysis

Hahn-Banach theorem guarantees there exists a linear functional over the whole vector space.

J.7.1 Uniform Boundedness Theorem

Theorem J.7.1 (Uniform boundedness Theorem) *Let X be a Banach space and Y be a normed space. Let the family $\{A_i\}_{i\in I} \subset L(X,Y)$ be pointwise bounded, i.e. for every $x \in X$ there exists a constant $C(x) > 0$ such that $\|A_i x\|_Y \leq C(x)$ for all $i \in I$. The family of norms $\{\|\{A_i\|\}_{i\in I}$ is bounded.*

J.7.2 The Open Mapping Theorem

Theorem J.7.2 *Let F and G be two Banach spaces. A continuous linear map $L :\to F$ which takes an open subset to an open subset. If it is continuous and one to one, it is a homeomorphism*

J.7.3 The Hahn-Banach Theorem

dual spaces. guarantees rich supply of functionals for Banach (and normed) spaces

The closest point property

$$\|x+y\|^2 + \|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2. \tag{J.130}$$

Definition: A subset $A$of a real or complex vector space is *convex* for if, for all$a, b \in A$ and all λ such that $0 < \lambda < 1$, the point $\lambda a + (1-\lambda)b$ belongs to A.

There is a unique point in $y \in A$ such that

$$\|x-y\| = \inf_{a\in A} \|x-a\| \tag{J.131}$$

Theorem: (The closest point property)

Let

$$M = \inf_{a \in A} \|x - a\| \tag{J.132}$$

$$\|x - y_n\|^2 < M^2 + \frac{1}{n}. \tag{J.133}$$

We are interested in establishing an inequality for the quantity $\|y_n - y_m\|$ when we know that $\|x - y_n\|$

$$\begin{aligned} \|y_n - y_m\|^2 + \|2x - y_n - y_m\|^2 &\equiv \|x - y_n - (x - y_m)\| + \|x - y_n + (x - y_m)\| \\ &= 2\|x - y_n\|^2 + 2\|x - y_m\|^2 \\ &< 4M^2 + 2\left(\frac{1}{n} + \frac{1}{m}\right) \end{aligned} \tag{J.134}$$

$$\|y_n - y_m\|^2 < 4M^2 + 2\left(\frac{1}{n} + \frac{1}{m}\right) - 4\left\|x - \frac{y_n + y_m}{2}\right\|^2. \tag{J.135}$$

Since A is convex and $y_n, y_m \in A$, $(y_n + y_m)/2 \in A$, therefore

$$4\left\|x - \frac{y_n + y_m}{2}\right\|^2 \geq M^2 \tag{J.136}$$

$$\|y_n - y_m\|^2 < 2\left(\frac{1}{n} + \frac{1}{m}\right) \tag{J.137}$$

it is a Cauchy sequence.

Suppose that $w \in A$ and $\|x - w\| = M$. Then $(y+)/2 \in A$, so that

$$\left\|x - \frac{y + w}{2}\right\| \geq M. \tag{J.138}$$

On applying the parallelogram law to $x - w$, $x - y$ we obtain

$$\begin{aligned} \|y - w\|^2 &= 2\|x - w\| + 2\|x - y\| - \left\|x - \frac{y + w}{2}\right\|^2 \\ &\leq 2M^2 + 2M^2 - 4M^2 = 0. \end{aligned} \tag{J.139}$$

Thus $y = w$, proving uniqueness.

Definition The space $\mathcal{L}(\mathcal{H}, \mathbf{C})$ is called the **algebraic dual space** of $\mathcal{H}$ and is denoted $\mathcal{H}^*$. The elements of $\mathcal{H}^*$ are **linear functionals**.

the Riesz lemma

Lemma J.7.3 *For each $T \in \mathcal{H}^*$, there is a unique $y_T \in \mathcal{H}$ such that $T(x) = (y_t, x)$ for all $x \in \mathcal{H}$. In addition $\|y_T\| = \|T\|_{\mathcal{H}*}$.*

Proof:

Define $y_T := \overline{T(x_0)}\|T\|$

$$\|T\| = \sup_{\|x\|=1} \|T(\mathbf{x})\| \tag{J.140}$$

Note that any element $y \in \mathcal{H}$ can be written

$$y = \underbrace{\left(y - \frac{T(y)}{T(x_0)}x_0\right)}_{\in \mathcal{N}} + \underbrace{\frac{T(y)}{T(x_0)}x_0}_{\sim x_0} \tag{J.141}$$

Since the functions $(y_T, .)$ and $T(.)$ are linear and agree on $\mathcal{N}$ and x_0, they must agree for any $y \in \mathcal{H}$. Thus $T(x) = (y_T, x)$ for all $x \in \mathcal{H}$.

To prove $\|T\|_{\mathcal{H}^*} = \|y_T\|_{\mathcal{H}}$

$$\begin{aligned} \|\mathbf{T}\| = \sup_{\|x\|=1} |T(x)| &\equiv \sup_{\|x\|=1} |(y_T, x)| \\ &\leq \sup_{\|x\|=1} \|y_T\|\|x\| = \|y_T\| \end{aligned} \tag{J.142}$$

□

Example

$$\|x\| := \sqrt{(x^2 + y^2 + z^2)} \tag{J.143}$$

$$T(\mathbf{x}) \Longleftrightarrow \begin{pmatrix} a & b & 0 \\ c & d & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \tag{J.144}$$

$$\|T\| = \sup_{\sqrt{(x^2+y^2+z^2)}=1} \sqrt{[(ax+by)^2 + (cx+dy)^2]} \tag{J.145}$$

Corollary J.7.4 *If N is a closed subspace of a Hilbert space H but is not the whole of H, then there is an element $y \neq 0$ in H which is orthogonal to N.*

Proof:

$\square$

Hahn-Banach Theorem

Theorem J.7.5 (The real Hahn-Banach Theorem)

$$F(x_1) \geq \frac{1}{\alpha}[p(\alpha x_1 + x) - f(x)] = p(x_1 + \frac{x}{\alpha}) - f(\frac{x}{\alpha}) = p(x_1 + z) - f(z) \tag{J.146}$$

we need Zorn's lamma. Let E be the collection of all extensions $\tilde{\ell}(x) < \varphi(x)$. Then E can be partially ordered by inclusion with respect to the domain and every linear chain has an upper bound (defined on the union of all domains). Hence there is a maximal element $\tilde{\ell}$ by Zorn's lemma.

Theorem J.7.6 (The complex Hahn-Banach Theorem)

functional f on M of V such that

$$Re\ f(u) \leq p(u), \quad u \in M. \tag{J.147}$$

There is a linear functional F on the whole of V such that

$$\begin{aligned} F(u) &= f(u), \quad u \in M. && \text{(J.148)} \\ |F(u)| &\leq p(u), \quad u \in V. && \text{(J.149)} \end{aligned}$$

Proof.

define the real valued functional

$$f_1(u) = \text{Re } f(u) \tag{J.150}$$

Then by (J.147),

$$f_1(u) \leq p(u), \quad u \in M. \tag{J.151}$$

By the real Hahn-Banach theorem there is a real functional $F_1(u)$ on V such

$$\begin{aligned} F_1(u) &= f_1(u), \quad u \in M. && \text{(J.152)} \\ F_1(u) &\leq p(u), \quad u \in V. && \text{(J.153)} \end{aligned}$$

$$f_1(iu) = \text{Re } f(iu) = \text{Re } if(u) = -\text{Im } f(u) \tag{J.154}$$

$$f(u) = f_1(u) - if_1(iu). \tag{J.155}$$

Let

$$F(u) = F_1(u) - iF_1(iu), \quad u \in V. \tag{J.156}$$

For it to be linear with respect to complex scalars, i.e. $F((\alpha + i\beta)u) = (\alpha + i\beta)F(u)$, we need only check that $F(iu) = iF(u)$,

$$F(iu) = F_1(iu) - iF_1(-u) = i[F_1(u) - iF_1(iu)] = iF(u) \tag{J.157}$$

Note $F(u) = f(u)$ for $u \in M$. We must now show (J.149)

$$|F(u)| = e^{-i\theta}F(u) = F(e^{-i\theta}u) = F_1(e^{-i\theta}u) \leq p(e^{-i\theta}u) = p(u). \tag{J.158}$$

as the imaginary part of $F(e^{-i\theta}u)$ is zero and $|e^{-i\theta}| = 1$.

Theorem J.7.7 (Hahn-Banach) *Let Y be a subspace of a normed linear space X and f a bounded linear functional on Y. There exists a bounded linear extension F of f on X, such that $\|F\| = \|f\|$.*

J.7.4 Adjoint of Bounded Operators

Theorem J.7.8 *Let A be a bounded operator in $\mathcal{B}(X,Y)$ where X,Y are Hilbert spaces. There exists a unique operator $A^\dagger \in \mathcal{B}(Y,X)$ such that*

$$(x, A^\dagger y)_X = (Ax, y)_Y \tag{J.159}$$

for all $x \in X$ $y \in Y$.

Proof: Consider any $y \in Y$. The linear functional on X

$$L(x) = (Ax, y)_Y$$

is continuous by the Schwartz inequality: If a sequence of vectors x_n converges to a limit vector x, then

$$\begin{aligned} |L(x) - L(x_n)| &= |L(x - x_n)| \\ &= |(A(x - x_n), y)_Y| \\ &\leq \|A\| \, \|x - x_n\| \, \|y\| \end{aligned}$$

so $L(x_n) \to L(x)$ as $n \to \infty$. Thus L is continuous. So by the Riesz lemma, there is a unique element $z \in X$ such that

$$(Ax, y)_Y = (x, z)_X$$

for all $x \in X$. We define $A^\dagger y$ to be z.

Let us show that $A^\dagger$ is linear. For any $y, z \in Y$ and $\alpha, \beta \in \mathbb{C}$ we have, for any $x \in X$,

$$\begin{aligned} (x, A^\dagger(\alpha y + \beta z))_X &= (Ax, \alpha y + \beta z)_Y \\ &= \overline{\alpha}(Ax, y)_Y + \overline{\beta}(Ax, z)_Y \\ &= \overline{\alpha}(x, A^\dagger y)_X + \overline{\beta}(x, A^\dagger z)_X \\ &= (x, \alpha A^\dagger y + \beta A^\dagger z)_X \end{aligned}$$

implying

$$A^\dagger(\alpha y + \beta z) = \alpha A^\dagger y + \beta A^\dagger z$$

To see that $A^\dagger$ is bounded note that for any $y \in Y$,

$$\begin{aligned} \|A^\dagger y\|^2 &= (A^\dagger y, A^\dagger y) \\ &= (AA^\dagger y, y) \\ &\le \|AA^\dagger y\| \, \|y\|, \end{aligned}$$

by the Cauchy-Schwartz inequality. Since $\|AA^\dagger y\| \le \|A\| \, \|A^\dagger y\|$, this implies

$$\|A^\dagger y\|^2 \le \|A\| \, \|A^\dagger y\| \, \|y\|.$$

If $\|A^\dagger y\| > 0$ we may divide through to obtain

$$\|A^\dagger y\| \le \|A\| \, \|y\|.$$

and since the inequality holds trivially if $\|A^\dagger y\| = 0$, it is always true. It follows that $A^\dagger$ is bounded, and

$$\|A^\dagger\| \le \|A\|.$$

Finally we show that $A^\dagger$ is determined uniquely by (J.159). Say we have $B, B' \in \mathcal{B}(X, Y)$ such that

$$(x, By)_X = (x, B'y)_X$$

for all $x \in X$, $y \in Y$, then

$$\begin{aligned} 0 &= (x, By)_X + (-1)(x, B'y)_X \\ &= (x, By)_X + (x, -B'y)_X \\ &= (x, By - B'y)_X \end{aligned}$$

holds for all $x \in X$ then in particular it holds when $x = By - B'y$. Thus $(By - B'y, By - B'y) = 0$ and we conclude $By = B'y$ for all $y \in Y$, so $B = B'$.

$\square$

Theorem J.7.9 *For any $A \in \mathcal{B}(X, Y)$, $A^{\dagger\dagger} = A$ and $\|A^\dagger\| = \|A\|$ for any pair X, Y of Hilbert spaces.*

Proof: Set $B = A^\dagger$ then $A^{\dagger\dagger} \equiv B^\dagger$. We have $B \in \mathcal{B}(Y,X)$ and $B^\dagger \in \mathcal{B}(X,Y)$. On applying the general defintion of the adjoint, we have for $x \in Y$, $y \in X$,

$$\begin{aligned}
(x, B^\dagger y)_Y &= (Bx, y)_X \\
&= (A^\dagger x, y)_X \\
&= (y, A^\dagger x)^*_X \\
&= (Ay, x)^*_Y \\
&= (x, Ay)_Y
\end{aligned}$$

and hence

$$A^{\dagger\dagger} y = B^\dagger y = Ay,$$

so that $A^{\dagger\dagger} = A$.

We know we have

$$\|A^\dagger\| \leq \|A\|.$$

Since this applies to any bounded linear operator we may replace A by $A^\dagger$ to obtain

$$\|A^{\dagger\dagger}\| \leq \|A^\dagger\|.$$

and, as we have established, $A^{\dagger\dagger} = A$,

$$\|A\| \leq \|A^\dagger\|.$$

Hence

$$\|A\| = \|A^\dagger\|.$$

$\square$

J.7.5 Unbounded Operators

In quantum mechanics we often work with operators which are not bounded. For example of such an operator consider the space of $L^2(-\infty, \infty)$ of sqaure-integrable functions $\psi(x)$ for $-\infty < x < \infty$, and let $\mathcal{Q}$ be the linear operator defined by $(\mathcal{Q}\psi)(x) = x\psi(x)$. This operator has the Hermitian property that

$$(\phi, \mathcal{Q}\psi) = \int_{-\infty}^{\infty} \phi(x)^* x\psi(X)dx = \int_{-\infty}^{\infty} [x\phi(x)]^* \psi(X)dx = (\mathcal{Q}\phi, \psi) \tag{J.160}$$

when the integrals converge. It as not bounded, because

$$||\mathcal{Q}\psi||^2 = \int_{-\infty}^{\infty} |x\psi(x)|^2 dx \tag{J.161}$$

can be any number of times larger than

$$||\psi||^2 = \int_{-\infty}^{\infty} |\psi(x)|^2 dx. \tag{J.162}$$

In fact $\mathcal{Q}$ is not even defined for all vectors; there are vectors ψ for which $x\psi(x)$ does not define a vector because $\int_{-\infty}^{\infty} |x\psi(x)|^2 dx$ is not finite.

It is impossible to define an unbounded Hermitian operator for all vectors. That this is true is a general fact shown by the following.

Theorem J.7.10 **(Hellinger-Toeplitz theorem).** *If a linear operator T is defined on all of a complex Hilbert space $\mathcal{H}$ satisfying*

$$(Tx, y) = (x, Ty)$$

for all $x, y \in \mathcal{H}$, then T is bounded.

Proof: Suppose T is not bounded so that we can find a sequence $\{y_n\}$ such that

$$\|y_n\| = 1 \qquad \text{and} \qquad \|Ty_n\| \to \infty.$$

Consider the functional f_n defined by

$$f_n := (Tx, y_n) = (x, Ty_n)$$

where $n = 1, 2, \ldots$. Each f_n is defined on all of $\mathcal{H}$ and is linear Note each f_n is a linear operator $f_n : \mathcal{H} \to \mathbb{C}$. For each n the functional f_n is bounded since

$$|f_n(x)| = |(x, Ty_n)| \leq \|Ty_n\| \, \|x\|.$$

Therefore $\{f_n\}$ is a sequence of bounded linear operators $f_n : \mathcal{H} \to \mathbb{C}$. Moreover, since

$$|f_n(x)| = |(Tx, y_n)| \leq \|Tx\|.$$

for every fixed $x \in \mathcal{H}$, the sequence $\{f_n(x)\}$ is bounded, i.e.,

$$\|f_n x\| := \sup_{\|x\|=1} |f_n(x)| \leq k_x$$

By the uniform boundedness theorem J.7.1 we can conclude that $\{\|f_n\|\}$ is bounded, say,

$$\|f_n\| \leq k \qquad \text{for all } \ n.$$

This implies that for every $x \in \mathcal{H}$ we have

$$|f_n(x)| \leq \|f_n\| \, \|x\| \leq k\|x\|$$

and, taking $x = Ty_n$, we arrive at

$$\|Ty_n\|^2 = (Ty_n, Ty_n) = |f_n(Ty_n)| \leq k\|Ty_n\|.$$

Hence $\|Ty_n\| \leq k$, which contradicts the original assumption $\|Ty_n\| \to \infty$.

□

Definition of the adjoint of an unbounded operator

Theorem J.7.11 *Every self-adjoint operator is a maximal symmetric operator.*

The domain of an unbounded operator

On the other hand an unbounded operator can not be deifned as self-adjoint for all vectors.

The set of all vectors ψ for which $A\psi$ is defined is called the *domain* of the operator A. We can assume that this is a linear manifold; if $A\psi$ $A\phi$ are defined, then

$$A(\psi + \phi) = A\psi + A\phi \tag{J.163}$$

and

$$A(b\psi) = bA\psi \tag{J.164}$$

define $A(\psi + \phi)$ and $A(b\psi)$.

Until now we have assumed that the domain of every operator being considered is the whole space. In the absence of this assumption some of our definitions have to be refined.

Thus $A = B$ implies that A and B have the same domain. The domain of $A + B$ is the set of all vectors which are both in the domain of A and the domain of B.

The sum of two operators A and B is given by $(A + B)u = Au + Bu$,

$$D(A + B) = D(A) \cap D(B).$$

The domain of AB is the set of all vectors ψ in the domain of B such that $B\psi$ is in the domain of A.

The composition of an operator A with an operator B is given by $(AB)u = A(Bu)$, where $D(AB) = \{u \in D(B) | Bu \in D(A)\}$

Two operators A and B are said to be equal if $D(A) = D(B)$ and for $u \in D(A)$, $Au = Bu$. A is said to be a restriction of B if $D(A) \subseteq D(B)$ and for $u \in D(A)$, $Au = Bu$.

If $D(B) \subseteq D(A)$ and for $u \in D(B)$ $Au = Bu$, then we say that A is an extension of B.

Definition A set of vectors, for example the domain of an operator, is called *dense* if for every vector ψ there is a sequence of vectors ψ_n in the set such that $\psi_n \to \psi$.

Definition If an operator has a dense domain, we can define $A^\dagger$ as follows. The domain of $A^\dagger$ is the set of all vectors ψ for which there is a vector $A^\dagger\psi$ such that

$$(\phi, A^\dagger\psi) = (A\phi, \psi) \tag{J.165}$$

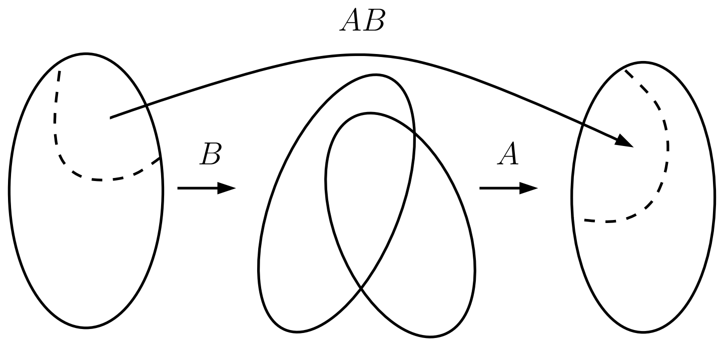

Figure J.7: $\mathrm{Ran}(B) \cap D(A)$

for every vector ϕ in the domain of A.

This defines $A^\dagger\psi$. For suppose χ and χ' are vectors that

$$(\phi, \chi) = (\phi, \chi') \tag{J.166}$$

for ever vector ϕ in the domain of A. Then, because the domain of A is dense, there is a sequence of vectors ϕ_n such that $\phi_n \to \chi - \chi'$ and

$$(\phi_n, \chi - \chi') = 0 \tag{J.167}$$

Therefore $\chi - \chi' = 0$, because

$$(\phi_n, \chi - \chi') \to (\chi - \chi', \chi - \chi'). \tag{J.168}$$

Example: A dense operator.

We consider the operator $\mathcal{Q}$ of Example. Let the domain of $\mathcal{Q}$ be the set of all vectors $\psi(x)$ such that $x\psi(x)$ is square-integrable. We prove this to be dense.

For any vector $\psi(x)$ let the sequance of vectors $\psi_n(x)$ be defined by

$$\begin{aligned} \psi_n(x) = \psi(x) \ \text{ for } \ -n \le x \le n \\ \psi_n(x) = 0 \ \text{ for } \ |x| > n. \end{aligned} \tag{J.169}$$

Then $x\psi_n(x)$ is square integrable, so $\psi_n(x)$ is in the domain of $\mathcal{Q}$ for each $n = 0, 1, 2, \dots$, and

$$||\psi - \psi_n||^2 = \int_{-\infty}^{-n} |\psi(x)|^2 dx + \int_{n}^{-\infty} |\psi(x)|^2 dx \to 0 \tag{J.170}$$

as $n \to \infty$, so $\psi - \psi_n$. Equation() in example ?? shows that $\mathcal{Q}$ symmetric. Therefore $\mathcal{Q}^\dagger$ is an extension of $\mathcal{Q}^\dagger$.

The example above demonstrates that a symmetric operator A can fail to be self-adjoint only if $A^\dagger$ is an extension of A to a larger domain.

Theorem J.7.12 *Let A be a Hermitian operator on a Hibert space $\mathcal{H}$. All eigenvalues of A are real, and eigenvectors of A corresponding to distinct eigenvalues are orthogonal.*

Proof: Suppose that λ is an einvalue of A and φ a corresponding eigenvector: then $A\varphi = \lambda\varphi$ and $\varphi \neq 0$. By self-adjointness of A,

$$\begin{aligned} 0 &= (A\varphi, \varphi) - (\varphi, A\varphi) \\ &= (\lambda\varphi, \varphi) - (\varphi, \lambda\varphi) \\ &= (\lambda - \overline{\lambda})\|\varphi\|^2 \end{aligned}$$

Since $\varphi \neq 0$, $\lambda = \overline{\lambda}$, and so λ is real. Now let λ, μ be distinct eigenvalues of A and let φ, ψ be corresponding eigenvectors. Then

$$\begin{aligned} 0 &= (A\varphi, \psi) - (\varphi, A\psi) \\ &= (\lambda\varphi, \psi) - (\varphi, \mu\psi) \\ &= (\lambda - \mu)(\varphi, \psi) \end{aligned}$$

Since μ is real and $\lambda \neq \mu$,

$$\lambda - \overline{\mu} = \lambda - \mu \neq 0.$$

Hence $(\varphi, \psi) \neq 0$, as required.

□

However note that A need not have any eigenvalues!

J.7.6 Operators of Finite Rank

Definition The **rank** of an operator is defined as the dimension of its range.

As an example, if $A \in \mathcal{L}(X, Y)$, ($\mathcal{L}(X, Y)$ being the space of all linear operators from X to Y), it is of finite rank if it maps X into some finite dimensional subspace Z of Y. Z is a normed space with respect to the restriction of the norm of Y.

□

Finite rank operators have the characteristic property that they can be expressed in the form

$$Tf = \sum_{k=1}^{n} (f, g_k) h_k, \tag{J.171}$$

where n is the dimension of Ran$T$, $\{h_k\}_{k=1}^n$ is some basis in Ran$T$, and $\{g_k\}_{k=1}^n$ is some finite system of vectors which does not depend on f.

We prove the part that every finite rank operator T can be expressed in the form (J.171). Choose any orthonormal basis $\{h_k\}_{k=1}^n$ in Ran$T$. For any $f \in \mathcal{H}$ we shall have

$$Tf = \sum_{k=1}^{n} \alpha_k h_k,$$

where the numbers α_k can be found from

$$\alpha_k = (Tf, h_k).$$

Thus the α_k's are linear functionals of f and by Riesz's lemma, elements g_k ($k = 1, 2, \ldots, n$) can be found such that the α_k's can be expressed as

$$\alpha_k = (f, g_k).$$

□

J.7.7 Square Roots of a Positive operator

Define

$$T_1 \leq T_2 \quad \text{if and only if} \quad (T_1 x, x) \leq (T_2 x, x)$$

for all $x \in \mathcal{H}$.

Definition A bounded self-adjoint operator $T : \mathcal{H} \to \mathcal{H}$ is said to be **positive**, written $T \geq 0$ if and only if

$$(Tx, x) \geq 0.$$

□

Lemma J.7.13 *If two bounded self-adjoint linear operators A and B on a Hilbert space $\mathcal{H}$ are positive and commute ($AB = BA$), then their product AB is positive.*

Proof: We prove the lemma by introducing a sequence of opeators defined by

$$A_1 = \frac{A}{\|A\|}, \qquad A_{n+1} = A_n - A_n^2 \qquad (n = 1, 2, \ldots). \tag{J.172}$$

First we porve by induction that

$$0 \leq A_n \leq I. \tag{J.173}$$

We see that the equality holds for $n = 1$,

$$(A_1 x, x) = \frac{1}{\|A\|}(Ax, x) \leq \frac{1}{\|A\|}\|Ax\| \, \|x\| \leq \|x\|^2 = (Ix, x).$$

Suppose (J.173) holds for fixed n,

$$0 \leq A_n \leq I, \quad \text{or} \quad 0 \leq I - A_n \leq I.$$

Since A_n is self-adjoint, for every $x \in \mathcal{H}$ and $y \in A_n x$ we obtain

$$
\begin{aligned}
(A_n^2(I-A_n)x,x) &= ((I-A_n)A_nx,A_nx) \\
&= ((I-A_n)y,y) \geq 0.
\end{aligned}
$$

This proves

$$A_n^2(I-A_n) \geq 0.$$

Similarly,

$$A_n(I-A_n)^2 \geq 0.$$

Addition results in

$$0 \leq A_n^2(I-A_n) + A_n(I-A_n)^2 = A_n - A_n^2 = A_{n+1}$$

Hence $0 \leq A_{n+1}$. That $A_{n+1} \leq I$ follows from $A_n^2 \geq 0$ and $I - A_n \geq 0$,

$$0 \leq I - A_n + A_n^2 = I - A_{n+1}.$$

Thus we have proved $0 \leq A_{n+1} \leq I$.

Now we move on to show that $(ABx,x) \geq 0$ for all $x \in \mathcal{H}$.

$$
\begin{aligned}
A_1 &= A_1^2 + A_2 \\
&= A_1^2 + A_2^2 + A_3 \\
&\cdots \\
&= A_1^2 + A_2^2 + \cdots + A_n^2 + A_{n+1}
\end{aligned}
$$

Since $A_{n+1} \geq 0$, this implies

$$A_1^2 + \cdots + A_n^2 = A_1 - A_{n+1} \leq A_1. \tag{J.174}$$

By the definition of $\leq$ and the self-adjointsness of the A_j's implies

$$\sum_{j=1}^{n} \|A_jx\|^2 = \sum_{j=1}^{n}(A_jx,A_jx) = \sum_{j=1}^{n}(A_j^2x,x) \leq (A_1x,x).$$

Since n is arbitrary, the infinite series $\|A_1x\|^2+\|A_2x\|^2+\cdots$ converges. Hence $\|A_nx\| \to 0$ and so $A_nx \to 0$.

$$\left(\sum_{j=1}^{n} A_j^2\right) x = (A_1 - A_{n+1})x \to A_1x \quad (n \to \infty). \tag{J.175}$$

Set $z_n = \sum_{j=1}^{n} A_j^2 x$ and $z = \lim_{n\to\infty} z_n$, by the continuity of the inner product we have

$$\lim_{n\to\infty} \sum_{j=1}^{n} (BA_j^2x, x) = \lim_{n\to\infty} (Bz_n, x) = (Bz, x) = (BA_1x, x).$$

All the A_j's commute with B since they are sums and products of $A_1 = \|A\|^{-1}A$, and A and B commute .For every $x \in \mathcal{H}$ and $y_j = A_jx$,

$$\begin{aligned}
(ABx, x) &= \|A\|(BA_1x, x) \\
&= \|A\| \lim_{n\to\infty} \sum_{j=1}^{n} (BA_j^2x, x) \\
&= \|A\| \lim_{n\to\infty} \sum_{j=1}^{n} (A_jBA_jx, x) \\
&= \|A\| \lim_{n\to\infty} \sum_{j=1}^{n} (By_j, y_j) \geq 0,
\end{aligned}$$

that is, $(ABx, x) \geq 0$.

□

Lemma J.7.14 *Let $\{A_n\}$ be a sequence of bounded self-adjoint operators on a Hilbert space $\mathcal{H}$ such that*

$$A_1 \leq A_2 \leq \cdots \leq A_n \leq \cdots \leq K \tag{J.176}$$

where K is a bounded self-adjoint operator on $\mathcal{H}$. Suppose that any A_j commutes with K and every A_k. Then $\{A_n\}$ strongly convergent ($A_nx \to Ax$ for all $x \in \mathcal{H}$) with the limit operator linear, bounded and self-adjoint and satisfying $A \leq K$

Proof: Consider the operators

$$B_n = K - A_n.$$

Clearly B_n is self-adjoint. We have

$$B_m^2 - B_n B_m = (B_m - B_n)B_m = (A_n - A_m)(K - A_m).$$

Let $m < n$. Then $A_n - A_m$ and $K - A_m$ are positive by (J.176). Sine these operators commute, their product is also positive by lemma J.7.13. Hence $B_m^2 \geq B_n B_m$ for $m < n$. In a similary manner we can show that $B_n B_m \geq B_n$ for $m < n$.

Together

$$B_m^2 \geq B_n B_m \geq B_n^2 \qquad (m < n).$$

By the self-adjointnes of the B_n's we thus have

$$(B_m^2 x, x) \geq (B_n B_m x, x) \geq (B_n^2 x, x) = (B_n x, B_n x) = \|B_n\|^2 \geq 0. \tag{J.177}$$

This shows that $(B_m^2 x, x)$ with fixed x is a monotone decreasing sequence of nonnegative numbers. Hence it converges.

We now show that $\{A_n\}$ converges. By assumption, every A_n commutes with every A_m and with K. Hence the B_j's all commute. These operators are self-adjoint. Since $-2(B_m B_n x, x) \leq -2(B_n^2 x, x)$, where $m < n$, we obtain

$$\begin{aligned}
\|B_m x - B_n x\|^2 &= ((B_m x - B_n)x, (B_m x - B_n)x) \\
&= ((B_m x - B_n)^2 x, x) \\
&= (B_m^2 x, x) - 2(B_n B_m x, x) + (B_n^2 x, x) \\
&\leq (B_m^2 x, x) - (B_n^2 x, x)
\end{aligned}$$

From thus we see that the sequence $\{B_n\}$ is Cauchy. It converges since $\mathcal{H}$ is complete. Now $A_n = K - B_n$. Hence $\{A_n\}$ also converges.

Since $\{A_n x\}$ converges, it is bounded for every $x \in \mathcal{H}$. The uniform boundednes theorem implies that A is bounded.

□

Definition Let $T : \mathcal{H} \to \mathcal{H}$ be a positive bounded self-adjoint operator on a Hilbert space $\mathcal{H}$. Then a bounded self-adjoint operator A satisfying

$$A^2 = T$$

is called the **square root** of T. If in addition $A \geq 0$, then A is called a **positive square root** of T and is denoted by

$$A = T^{1/2}.$$

$\square$

In the following theorem it is shown that $T^{1/2}$ exists and is unique.

Theorem J.7.15 *Every positive bounded self-adjoint operator on a Hilbert space $\mathcal{H}$ has a unique posititve square root.*

Proof: First we will prove that if the theorem holds for $T \leq I$, it holds in general. If $T = 0$, we can take $A = T^{1/2} = 0$. Let $T \neq 0$. By the Schwarz inequality

$$(Tx, x) \leq \|Tx\| \, \|x\| \leq \|T\| \, \|x\|^2.$$

Dividing by $\|T\| \neq 0$ and setting

$$S = \frac{1}{\|T\|} T,$$

we obtain

$$(Sx, x) \leq \|x\|^2 = (Ix, x),$$

that is, $S \leq I$. Asumming that S has a unique positive square root $B = S^{1/2}$, we have $B^2 = S$ and we see that the unique square root of $T = \|T\| S$ is $\|T\|^{1/2} B$ because

$$(\|T\|^{1/2} B)^2 = \|T\| B^2 = \|T\| S = T.$$

We now establish the theorem under the condition $T \leq I$. To prove the existence of the operator $A = T^{1/2}$ we introduce a sequence of operators defined by

$$A_{n+1} = A_n + \frac{1}{2}(T - A_n^2), \qquad n = 0, 1, \ldots \tag{J.178}$$

where $A_0 = 0$. We will prove $A_n x \to T^{1/2} x$.

Since $A_0 = 0$ $A_1 = \frac{1}{2}T$, $A_2 = T - \frac{1}{8}T^2$, etc. Each A_n is a polynomial in T. Hence the A_n's are self-adjoint and all commute, and they also commute with every operator that commutes with T. We will prove

$$\begin{aligned} A_n &\leq I, \qquad n = 0, 1, \ldots &\text{(J.179)}\\ A_n &\leq A_{n+1}, \qquad n = 0, 1, \ldots &\text{(J.180)}\\ A_n x \to Ax, &\qquad A = T^{1/2} &\text{(J.181)}\end{aligned}$$

We have that $A_0 \leq I$. Let $n > 0$. Since $I - A_{n-1}$ is self-addjoint,

$$(I - A_{n-1})^2 \geq 0.$$

But $I - T \geq 0$ so we have

$$\begin{aligned} 0 &\leq \frac{1}{2}(I - A_{n-1})^2 + \frac{1}{2}(I - T) \\ &= I - A_{n-1} - \frac{1}{2}(T - A_n^2) \\ &= I - A_n. \end{aligned}$$

We now prove by induction that $A_n \leq A_{n+1}$ for all $n \geq 0$. From (J.178) we have $0 = A_0 \leq A_1 = \frac{1}{2}T$. We assum that $A_{n-1} \leq A_n$ for fixed n and show it implies $A_n \leq A_{n+1}$. By (J.178)

$$\begin{aligned} A_{n+1} - A_n &= A_n + \frac{1}{2}(T - A_n^2) - A_{n-1} - \frac{1}{2}(T - A_{n-1}^2) \\ &= (A_n - A_{n-1})[I - \frac{1}{2}(A_n + A_{n-1})]. \end{aligned}$$

It follows from (J.179) that $A_n \leq A_{n+1}$.

As $\{A_n\}$ satisfies $A_n \leq A_{n+1}$ and $A_n \leq I$, lemma J.7.14 implies the existence of a bounded self-adjoint linear operator A such that $A_n x \to Ax$ for all $x \in \mathcal{H}$. Since $\{A_n x\}$ converges, (J.178) gives

$$A_{n+1}x - A_n x = \frac{1}{2}(Tx - A_n^2 x) \to 0$$

as $n \to \infty$. Hence

$$Tx - A^2 x = 0 \qquad \text{for all} \qquad x \in \mathcal{H}$$

$\square$

J.7.8 Polar decomposition

Recall, a linear operator $A : X \to Y$ is called an isometry if and only if for all $x, y \in X$ the inner product is invariant: $< x, y >=< Ax, Ay >$. If A is a surjective isometry, then for every $y \in Y$ the equation $Ax = y$ has a solution. Then

$$AA^* y = A(A^* A)x = Ax = y,$$

hence $AA^* = \mathbb{I}$.

We are interested in the case where $X = Y = \mathcal{H}$.

Definition An operator $W \in \mathcal{B}(\mathcal{H})$ is said to be a partial isometry if there are closed subspaces $\mathcal{K}, \mathcal{L} \subset \mathcal{H}$ such that $W : \mathcal{K} \to \mathcal{L}$ is an isometry onto $\mathcal{L}$ and $W : \mathcal{K}^\perp \to \{0\}$. $\mathcal{K}$ is called the initial subspace and $\mathcal{L}$ the final subspace of W.

Theorem J.7.16 (Polar decomposition) *Let $A \in \mathcal{B}(\mathcal{H})$. Then there exists a partial isometry*

$$W : \overline{Ran|A|} \to \overline{RanA}$$

(where Ran stands for the range), such that A can be written uniquely as $A = W|A|$ where $|A|$ is the positive square root of $A^* A$.

Proof:

$$\|Ax\|^2 = < Ax, Ax > = < x, A^* Ax > = < x, |A|^2 x >= \| \, |A|x\|^2.$$

$$|A|(x - y) = 0 \quad \text{if and only if} \quad A(x - y) = 0.$$

Thus there is a unique unitary operator W_0 from the range of $|A|$ to the range of A and

$$Ax = W_0|A|x$$

an isometric operator $W_0 : \mathrm{Ran}|A| \to imA$ which can be extended uniquely to an isometric operator $W : \overline{\mathrm{Ran}|A|} \to \overline{\mathrm{Ran}A}$ by

$$Wx = \lim W_0 x_k$$

for $x = \lim x_k$. Hence

$$A = W|A|,$$

We extend W to a partial isometry on $\mathcal{H}$ by defining $Wy = 0$ if y is orthogonal to the range of $|A|$.

□

The uniqueness of the polar decomposition follows from $A = VB$, hence, $A^* = B^*V^*$ and $|A|^2 = A^*A = VB^*V^* = |B|^2 = B^2$, hance $B = |A|$, since the root is unique.

Suppose that U is unitary and commutes with A. Then U also commutes with $|A|$ and we have

$$A = UAU^* = UWU^*U|A|U^* = UWU^*|A|.$$

This is another polar decomposition of A and so, by uniqueness, we see that $W = UWU^*$, that is, U also commutes with W.

J.7.9 Compact Operators

Definition A linear operator K from X to Y is said to be **compact** if it transforms every bounded sequence of vectors $\{x_n\} \in X$ (i.e., vectors whose norms are all less than some fixed constant) into a sequence $Kx_n \in Y$ containing a convergent subsequence. The set of compact operators from X to Y is denoted $K(X, Y)$.

Notice that compact operators are always bounded. Otherwise, there would be a sequence $\{x_n\}$ such that $\|x_n\| \leq C$, while $\|Kx_n\| \to \infty$. Then $\{Kx_n\}$ could not have a convergent subsequence. The sum of two compact operators is always compact, as is the product of a scalar and a compact operator. Hence, $K(X, Y) \subset \mathcal{B}(X, Y)$.

All finite-rank operators are compact, for they take a bounded sequence into a bounded sequence in a finite-dimensional Banach space.

The product of a compact operator with a bounded operator is a compact operator. Suppose that K is compact and B is a bounded operator. Consider KB acting on $\{x_n\}$, where $\|x_n\| < C$ for all n. Then K acts on the sequence $\{Bx_n\}$, which is bounded because B is a bounded operator. Since K is compact, there must be a subsequence of $\{KBx_n\}$ which converges to an element of the space, so KB is compact. Now consider BK acting on $\{x_n\}$. Since K is compact, $\{Kx_n\}$ posses a subsequence which converges. Since B is bounded, it is continuous, so $\{BKx_n\}$ has a related convergent subsequence.

Lemma J.7.17 *If $\{g_k\}_{k=1}^{\infty}$ is an infinite orthonormal sequences of vectors in X, and if*

$$Kg_k = \beta_{k0}g_0 + \beta_{k1}g_1 + \cdots + \beta_{kk}g_k \tag{J.182}$$

where K is a compact operator in X, then

$$\lim_{k\to\infty} \beta_{kk} = 0.$$

Proof: Let $n > m$; then

$$\|Kg_n - Kg_m\|^2 \geq |\beta_{kk}|^2.$$

If there were only a finite number of β_{kk}'s such that $|\beta_{kk}| \geq \epsilon$ for all $\epsilon > 0$ then β_{kk} would tend to 0 as $k \to \infty$. Say β_{kk} does not tend to 0 as $k \to \infty$, then there must be an infinite sequence of subscripts

$$n_1 < n_2 < n_3 < \ldots$$

such that

$$|\beta_{n_jn_j}| \geq \delta > 0, \qquad j = 1, 2, 3, \ldots$$

and therefore

$$\|Kg_{n_k} - Kg_{n_j}\|^2 \geq \delta^2 > 0.$$

This implies that the infinite sequence of vectors $\{Kg_{n_j}\}_{j=1}^{\infty}$ does not contain any convergent subsequences, and as the sequence $\{g_{n_j}\}_{j=1}^{\infty}$ is bounded, (of course, $\|g_{n_j}\| = 1$ for all j), this contradicts the defintion of a compact operator.

$\square$

Lemma J.7.18 *If $\lambda \neq 0$, if K is a compact operator and if $\{f_n\}_{n=0}^{\infty}$ is an infinite sequence of vectors which satisfy the relations*

$$Kf_0 = \lambda f_0 = 0, \quad Kf_n = \lambda f_n = f_{n-1} \qquad (n = 1, 2, 3, \dots),$$

then $f_0 = 0$ (and hence also $f_n = 0$ for $n = 1, 2, 3, \dots$).

Proof: First we show that if $f_0 \neq 0$ implies there are no linearly dependent vectors among the f_k's. For suppose that among the vectors

$$f_0, f_1, f_2, \dots, f_k, \dots$$

there are some elements which are linear combinations of the earlier ones. Let f_n be the first such element, i.e. the first of the f_k's such that

$$f_n = \alpha_0 f_0 + \alpha_1 f_1 + \cdots + \alpha_{n-1} f_{n-1}, \tag{J.183}$$

then, applying the operator K to both sides, we obtain

$$\lambda f_n + f_{n-1} = \alpha_0 \lambda f_0 + \alpha_1(\lambda f_1 + f_0) + \cdots + \alpha_{n-1}(\lambda f_{n-1} + f_{n-2}).$$

But by (J.183), this implies

$$f_{n-1} = \alpha_1 f_0 + \alpha_2 f_1 + \cdots + \alpha_{n-1} f_{n-2}$$

which contradicts that f_n is te first elements that is a linear combination of its predecessors.

Now, we assume the lemma is false by taking the vectors $f_0, f_1, f_2, \dots$ to be linearly independent, and so we can orthonormalize this sequence. Let the orthonormalized vectors be

$$
\begin{array}{rcl}
g_0 & = & \alpha_{00} f_0, \\
g_1 & = & \alpha_{10} f_0 + \alpha_{11} f_1, \\
\vdots & & \vdots \qquad \vdots \\
g_k & = & \alpha_{k0} f_0 + \alpha_{k1} f_1 + \cdots + \alpha_{kk} f_k \\
\vdots & & \vdots \qquad \vdots \qquad\qquad \vdots
\end{array}
\tag{J.184}
$$

We observe that

$$
\begin{aligned}
K g_k &= \alpha_{k0} \lambda f_0 + \alpha_{k1}(\lambda f_1 + f_0) + \cdots + \alpha_{kk}(\lambda f_k + f_{k-1}) \\
&= \alpha_{k1} \lambda f_0 + \alpha_{k2} \lambda f_1 + \cdots + \alpha_{kk} f_{k-1} + \lambda(\alpha_{k0} f_0 + \alpha_{k1} f_1 + \cdots + \alpha_{kk} f_k) \\
&= \alpha_{k1} \lambda f_0 + \alpha_{k2} \lambda f_1 + \cdots + \alpha_{kk} f_{k-1} + \lambda g_k \\
&= \beta_{k0} g_0 + \beta_{k1} g_1 + \cdots + \beta_{k,k-1} g_{k-1} + \lambda g_k.
\end{aligned}
$$

contradicting lemma J.7.17, showing that the hypothesis that $f_0 \neq 0$ is impossible.

□

Theorem J.7.19 *Every compact operator K in X can have, for any $\rho > 0$, only a finite nunmber of linearly independent eigenvectors corresponding to eigenvalues greater than $|\rho|$.*

Proof: Assuming the ocntrary, we suppose that there is an infinite set of linearly independent vectors f_n $(n = 1, 2, 3, \dots)$ for which

$$
K f_n = \lambda_n f_n, \qquad |\lambda_n| > \rho > 0 \qquad n = 1, 2, 3, \dots
$$

Orthonormalizing the sequence $\{f_n\}_{n=1}^{\infty}$, we obtain an orthonormal sequence of vectors

$$
\begin{array}{rcl}
g_1 & = & \alpha_{11} f_1, \\
g_2 & = & \alpha_{21} f_0 + \alpha_{22} f_2, \\
\vdots & & \vdots \qquad \vdots \\
g_k & = & \alpha_{k1} f_1 + \alpha_{k1} f_2 + \cdots + \alpha_{kk} f_k \\
\vdots & & \vdots \qquad \vdots \qquad\qquad \vdots
\end{array}
$$

Then

$$\begin{aligned} Kg_k - \lambda_k g_k &= \alpha_{k1}(\lambda_1 - \lambda_k) + \cdots + \alpha_{k1}(\lambda_{k-1} - \lambda_k)\alpha_{k,k-1} f_{k-1} \\ &= \beta_{k1} g_1 + \beta_{k2} g_2 + \cdots + \beta_{k,k-1} g_{k-1}, \end{aligned} \tag{J.185}$$

and thus

$$Kg_k = \beta_{k1} g_1 + \beta_{k2} g_2 + \cdots + \beta_{k,k-1} g_{k-1} + \lambda_k g_k.$$

Thus the assumption that $|\lambda_k| > \rho > 0$ $(k = 1, 2, 3, \dots)$ contradicts lemma J.7.17.

□

Corollary J.7.20 *The eigenvalues of a compact operator K in X can have a limit point only at the point 0.*

Say we have an infinite sequence of eiegenvalues $\{\lambda_n\}$ of a compact operator K tending to λ where $|\lambda| > 0$. Then for any ϵ satifying $|\lambda| > \epsilon > 0$ there is some n_0 such that $|\lambda - \lambda_n| < \epsilon$ for all $n \geq n_0$. However this contradicts theorem J.7.19, for pick any ρ such that $0 < \rho < |\lambda| - \epsilon$, by the theorem there can only be a finite number of eigenvalues with modulus greater than $|\rho|$. Thus, if the eigenvalues of K are infinite in number, they must comprise a sequence tending to zero.

□

Corollary J.7.21 *The number of independent eigenvectors corresponding to any non-zero eigenvalue of a compact operator in X is finite. In other words, the multiplicity of each non-zero eigenvalue of a compact operator in X is finite.*

Given any non-zero eigenvalue λ, by theorem J.7.19, for any ρ such that $|\lambda| < \rho < 0$ there are only a finite number of linearly independent eigenvectors with eigenvalues whose modulus is greater than ρ. Hence the mutilplicity of λ must be finite.

□

Corollary J.7.22 *Every compact operator in X has not more than a countable set of linearly independent eigenvectors corresponding to non-zero eigenvalues.*

Choose $\rho_n = 1/n$ for $n \in \mathbb{N}$. Let N_n be the *finite* (possibly zero) number of linearly independent eigenvectors with eigenvalue of modulus greater than ρ_n take away the number with modulus greater than ρ_{n-1}. Then the number of linearly independent eigenvectors corresponding to non-zero eigenvalues is $\sum_{n\in\mathbb{N}} N_n$ which is at most a countable set.

□

Theorem J.7.23 *If $\|K_j - K\| \to 0$ as $j \to \infty$, and if K_j is a compact transformation for all n, then K is a compact operator.*

Proof:

$$\begin{aligned}\|K\phi_n - K\phi_m\| &\le \|K\phi_n - K_j\phi_n\| + \|K_j\phi_n - K_j\phi_n\| + \|K_j\phi_m - K\phi_m\| \\ &\le \|K - K_j\| \cdot \|\phi_n\| + \|K_j\phi_n - K_j\phi_n\| + \|K_j - K\| \cdot \|\phi_m\|.\end{aligned}$$

Since the ϕ_n are bounded,

$$\|K\phi_n - K\phi_m\| \le C\|K - K_j\| + \|K_j\phi_n - K_j\phi_n\|.$$

Because $\|K_j - K\| \to 0$, we can pick some j sufficiently larlge so that

$$\|K_j - K\| < \epsilon/(2C).$$

Then when $m, n > N$ the second term on the right above can be made smaller than $\epsilon/2$. Thus for $m, n > N$, we have

$$\|K\phi_n - K\phi_m\| < \epsilon.$$

□

Any transformation with square-integrable kernel can be approximated arbitrarily closely in norm by a kernel of finite rank.

Things are especially nice if $X = Y$ is a separable Hilbert space.

Theorem J.7.24 *For each compact operator K of a Hilbert space, on can find a sequence of operators of finite rank converging to K in norm.*

Proof:

Define the operator

$$P_n u = \sum_{i=1}^{n} (u, \varphi_i)\varphi_i \tag{J.186}$$

If the assertion were false, there would be an operator $K \in K(H)$ and number $\delta > 0$ such that

$$\|K - F\| \geq \delta. \tag{J.187}$$

for all operators F of finite rank. Now

$$F_n u := P_n K u = \sum_{i=1}^{n} (Ku, \varphi_i)\varphi_i$$

is an operator of finite rank. By (J.187) for each n there is a $u_n \in H$ satifying

$$\|u_n\| = 1, \quad \|(K - F_n)u_n\| \geq \delta/2. \tag{J.188}$$

Since the sequence $\{u_n\}$ is bounded, Ku_n has a subsequence converging to an element $w \in H$. But

$$\begin{aligned} \|(K - F_n)u_n\| &\leq \|(1 - P_n)(Ku_n - w)\| + \|(1 - P_n)w\| \\ &\leq \|(Ku_n - w)\| + \|(1 - P_n)w\|. \end{aligned} \tag{J.189}$$

□

We see that compact operators are in a sence the closest analogue of finite dimensional matrices and in fact the finite rank operators form a dense subset with respect to the operator norm $\|T\|$.

Compact Self-Adjoint Operators

Lemma J.7.25 *If A is a self-adjoint operator on a Hilbert spacce $\mathcal{H}$ then*

$$\|A\| = \sup_{\|x\|=1} |(Ax, x)|.$$

Proof: for any $x \in \mathcal{H}$ such that $\|x\| = 1$,

$$\begin{aligned} |(Ax, x)| &\le \|Ax\| \, \|x\| \\ &\le \|A\| \, \|x\|^2 \\ &\le \|A\|. \end{aligned}$$

Hence

$$\|A\| \ge \sup_{\|x\|=1} |(Ax, x)|.$$

Let $m \equiv \sup_{\|x\|=1} |(Ax, x)|$, so that

$$|(Au, u)| \le m\|u\|^2$$

for any $u \in \mathcal{H}$. Then for any x, y,

$$(A(x \pm y), (x \pm y)) = (Ax, x) \pm 2Re(Ax, y) + (Ay, y).$$

Thus

$$\begin{aligned} 4Re(Ax, y) &\le (A(x + y), (x + y)) - (A(x - y), (x - y)) \\ &\le m(\|x + y\|^2 + \|x - y\|^2) \\ &= 2m(\|x\|^2 + \|y\|^2) \end{aligned} \tag{J.190}$$

by the parallelogram law. Replace x by λx where $|\lambda| = 1$ and λ is chosen so that $\lambda(Ax, y) \ge 0$. We obtain

$$|(Ax, y)| \le \frac{1}{2} m(\|x\|^2 + \|y\|^2).$$

Suppose that $Ax \ne 0$. Since the above holds for all $x, y \in \mathcal{H}$ we may choose y to be given by

$$y = \frac{\|x\|}{\|Ax\|} Ax.$$

Since $\|y\| = \|x\|$ the inequality yields

$$\|x\| \, \|Ax\| \leq m\|x\|^2.$$

This holds trivially when $Ax = 0$, hence is true for all x. From the definition of the operator norm

$$\begin{aligned} \|A\| &\leq m \\ &= \sup_{\|x\|=1} |(Ax, x)|. \end{aligned}$$

□

Theorem J.7.26 *Let K be a compact self-adjoint operator on a Hilbert space $\mathcal{H}$. Either $\|K\|$ or $-\|K\|$ is an eigenvalue of K.*

Proof: We have by the above

$$\|K\| = \sup_{\|x\|=1} |(Kx, x)|.$$

Since the inner product (Kx, x) is real, by definition of the least upper bound a sequence of (normalized) vectors $\{g_n\}_{n=1}^{\infty}$ can be found such that the limit

$$\lim_{n\to\infty} (Kg_n, g_n)$$

exists and is equal to $\|K\|$ or -$\|K\|$. We shall call this limit λ.

Since K is a compact operator there is a subsequence $\{g_{n_i}\}_{i=1}^{\infty}$ of $\{g_n\}_{n=1}^{\infty}$ such that $\{Kg_{n_i}\}_{i=1}^{\infty}$ is a convergent sequence. Denote its limit by y,

$$Kg_{n_i} \to y.$$

As $\{(Kg_n, g_n)\}$ is Cauchy

$$|(Kg_n, g_n) - (Kg_m, g_m)| < \epsilon, \quad m, n > N.$$

The sequence $\{(Kg_n, g_n)\}$ has a subsequence converging to $\lim_{i\to\infty}(y, g_{n_i}) =: \mu$, then there is an $n_i > N$ such that

$$|(Kg_{n_i}, g_{n_i}) - (Kg_n, g_n)| < \epsilon.$$

Thus,

$$|(Kg_n, g_n) - \mu| \leq |(Kg_n, g_n) - (Kg_{n_i}, g_{n_i})| + |(Kg_{n_i}, g_{n_i}) - \mu| < 2\epsilon.$$

for all $n > N$. It follows $\mu = \lambda$.

Now since

$$\|Kg_{n_i} - \lambda g_{n_i}\|^2 = \|Kg_{n_i}\|^2 - 2\lambda(Kg_{n_i}, g_{n_i}) + \lambda^2, \tag{J.191}$$

we have

$$\lim_{i\to\infty} \|Kg_{n_i} - \lambda g_{n_i}\|^2 = \|y\|^2 - 2\lambda^2 + \lambda^2 = \|y\|^2 - \lambda^2 \tag{J.192}$$

But

$$\|Kg_{n_i}\| \leq \|K\| \, \|g_{n_i}\| = \|K\| = |\lambda|,$$

and consequently

$$\|y\| \leq |\lambda|.$$

But since the left-hand side of (J.192) is non-negative $\|y\| \geq |\lambda|$, it follows that $\|y\| = |\lambda|$, and therefore

$$\lim_{i\to\infty} \|Kg_{n_i} - \lambda g_{n_i}\| = 0. \tag{J.193}$$

But

$$\|y - \lambda g_{n_i}\| \leq \|Kg_{n_i} - y\| + \|Kg_{n_i} - \lambda g_{n_i}\|$$

Thus by (J.191) and (J.193) it follows that $\lim_{i\to\infty} g_{n_i}$ exists and is equal to y/λ. Then (J.193) can be rewritten as

$$Ky - \lambda y = 0.$$

□

Lemma J.7.27 *Let M be a closed linear subspace of a Hilbert space $\mathcal{H}$ and let M be invariant under a bounded linear operator T on $\mathcal{H}$ (that is, $TM \subseteq M$). Then $M^\perp$ is invariant under $T^\dagger$.*

Proof: Consider $x \in M^\perp$ and $m \in M$. Then $Tm \in M$ and hence $(Tm, x) = 0$. Thus $(m, T^\dagger x) = 0$ for all $m \in M$, so that $T^\dagger x \in M^\perp$.

$\square$

Theorem J.7.28 *Let K be a compact self-adjoint operator on a Hilbert space $\mathcal{H}$. There exists a finite or infinite orthonormal sequence $\{e_n\}$ of eigenvectors of K, corresponding to non-zero eigenvalues*

$$\lambda_1, \lambda_2, \lambda_3, \ldots \qquad (|\lambda_1| \geq |\lambda_2| \geq |\lambda_3| \geq \ldots$$

We obtain the eigenvectors inductively. By theorem J.7.26 K has an eigenvalue $\lambda_1 = \pm\|K\|$, and we may pick a corresponding eigenvector e_1 of unit norm.

$$Ke_1 = \lambda_1 e_1$$

Since the linear span of e_1 is invariant under K, by lemma J.7.27, its orthogonal complement $\mathcal{H}_2$ is invariant under $K^\dagger(= K)$ (let us redenote K as K_1). Thus $K_1\mathcal{H}_2 \subseteq \mathcal{H}_2$ and so we may define K_2 to be the restriction of K to $\mathcal{H}_2$. Clearly K_2 is compact, and if $x, y \in \mathcal{H}_2$ then

$$\begin{aligned} (K_2^\dagger x, y) &= (x, K_2 y) = (x, Ky) \\ &= (K^\dagger x, y) = (x, Ky). \end{aligned}$$

Since $Kx \in \mathcal{H}_2$, it follows that $Kx = K_2^\dagger x$ for all $x \in \mathcal{H}_2$. Thus $K_2^\dagger$ is the restriction of K to $\mathcal{H}_2$, that is, $K_2^\dagger = K_2$. Thus K_2 is a self-adjoint operator on $\mathcal{H}_2$. So if K_2 is not identically zero, we may apply theroem J.7.26 again to obtain an eigenvalue $\lambda_2 = \pm\|K_2\|$ of K_2, and hence of K, corresponding to a unit eigenvector e_2 of K_2,

$$K_2 e_2 = \lambda_2 e_2$$

Since $e_2 \in \mathcal{H}_2$, we have $(e_2, e_1) = 0$ and

$$|\lambda_2| = \sup_{\|f\|=1, f\in\mathcal{H}_2} |(K_1 f, f)| \le \sup_{\|g\|=1} |(K_1 g, g)| = |\lambda_1|$$

We continue this inductive construction as long as $K_n \neq 0$. If we encounter m such that $K_m = 0$ the construction stops. Then for an arbitrary $f \in \mathcal{H}$,

$$f - \sum_{j=1}^{m-1} (f, e_j) e_j \in \mathcal{H}_m.$$

Denoting the left-hand side by g, we have

$$\begin{aligned} 0 &= K_m g = K g \\ &= K g - \sum_{j=1}^{m-1} (f, e_j) K e_j \end{aligned} \tag{J.194}$$

That is, for any $f \in \mathcal{H}$

$$\begin{aligned} K g &= \sum_{j=1}^{m-1} (f, e_j) K e_j \\ &= \sum_{j=1}^{m-1} \lambda_j (f, e_j) e_j, \end{aligned} \tag{J.195}$$

and the assertion is proved with finite sum.

If the process does not terminate, we obtain an infinite, orthonormal sequence $\{e_k\}_{k=1}^{\infty}$. Consier an arbitray $f \in \mathcal{H}$ and let

$$g_n = f - \sum_{j=1}^{n-1} (f, e_j) e_j.$$

Then $g_n \in \mathcal{H}_n$. Applying Pythagoras' theorem to the relation

$$f = g_n + \sum_{j=1}^{n-1} (f, e_j) e_j$$

we see that

$$\|f\|^2 = \|g_n\|^2 + \sum_{j=1}^{n-1} |(f, e_j)|^2$$

so that

$$\|g_n\| \le \|f\|.$$

Now

$$\begin{aligned} \|Kg_n\| &= \|K_n g_n\| \le \|K_n\| \, \|g_n\| \\ &\le |\lambda_n| \, \|f\|. \end{aligned}$$

That is,

$$\left\| Kf - \sum_{j=1}^{n-1} \lambda_j (f, e_j) e_j \right\| \le |\lambda_n| \, \|f\|$$

By corollary J.7.20, $\lambda_n \to 0$ as $n \to \infty$. Hence

$$Kf = \sum_{j=1}^{\infty} \lambda_j (f, e_j) e_j$$

for all $f \in \mathcal{H}$.

□

Corollary J.7.29 *(Hilbert-Schmidt theorem) Let K be a compact operator on a separable infinite-dimensional Hilbert space $\mathcal{H}$. There exists a complete orthonormal sequence $\{e_n\}_{n \in \mathbb{N}}$ of eigenvectors of K with real eigenvalues λ_n which converge to zero. For any $f \in \mathcal{H}$,*

$$Kf = \sum_{n=1}^{\infty} \lambda_n (f, e_n) e_n.$$

Proof: By theorem J.7.28 there is a finite or infinte orthonormal sequence $\{x_n\}$ such that, for all $f \in \mathcal{H}$,

$$Kf = \sum_n \lambda_n (f, x_n) x_n. \tag{J.196}$$

We may suppose that each $\lambda_n \neq 0$. Let $\{\varphi_n\}$ be a complete orthonormal system in the Hilbert space Ker K, the kernel of K. Then each φ_m is an eigenvector of K corresponding to the eigenvector 0. As x_n is an eigenvector corresponding to a different eigenvalue, x_n is orthogonal to φ_n by theorem ??. Thus $\{x_n\} \cup \{\varphi_n\}$ is a countable orthonormal set in $\mathcal{H}$. For any $f \in \mathcal{H}$, (J.196) implies that

$$f - \sum_n (f, x_n) x_n \in \mathrm{Ker} K,$$

and since $\{\varphi_n\}$ is an orthonrmal basis of Ker K,

$$f - \sum_n (f, x_n) x_n = \sum_m (f, \varphi_m) \varphi_m.$$

Thus $\{x_n\} \cup \{\varphi_n\}$ is a complete orthonormal set in $\mathcal{H}$. Since dim $\mathcal{H} = \infty$ so this set is infinite, so it may be indexed by $\mathbb{N}$, as $\{e_n\}_{n \in \mathbb{N}}$.

□

Theorem J.7.30 *Let K be a compact operator of a Hilbert space $\mathcal{H}$. Then there exist orthonormal but not necessarily complete systems $\{x_n\}_{n=1}^N$ and $\{y_n\}_{n=1}^N$ as well as a sequence of positive real numbers $\{\lambda_n\}_{n=1}^N$ converging to zero, called the singular values of K such that*

$$K = \sum_{n=1}^N \lambda_n (x_n, .) y_n. \tag{J.197}$$

Proof: Since K is compact so is $K^\dagger K$. To see this, let $\{f_n\}_{n=1}^\infty$ be any sequence of bounded vectors ($\|f_n\| < C$ for all n). Since

$$\begin{aligned} \|K^\dagger f_n - K^\dagger f_m\|^2 &= (K^\dagger(f_n - f_m), K^\dagger(f_n - f_m)) \\ &= (KK^\dagger(f_n - f_m), (f_n - f_m)) \\ &\leq \|K^\dagger\| \, \|Kf_n - Kf_m\| \, \|f_n - f_m\|, \end{aligned}$$

$K^\dagger$ is bounded and $\|f_n - f_m\| < 2C$, it follows

$$\lim_{n,m\to\infty} \|K^\dagger f_n - K^\dagger f_m\| = 0,$$

i.e., the sequence $K^\dagger\{f_n\}_{n=1}^\infty$ converges. Since the product of two compact operators is compact $K^\dagger K$ is compact.

As $K^\dagger K$ is compact and self-adjoint, by the Hilbert-Schmidt theorem, there is an orthonormal set $\{x_n\}_{n=1}^N$ such that

$$K^\dagger K x_n = \mu_n x_n \qquad \text{with } \mu_n \neq 0$$

and so that $K^\dagger K$ is the zero operator on the subspace orthogonal to $\{x_n\}_{n=1}^N$.

Since $K^\dagger K$ is positive, each $\mu_n > 0$. Let λ_n be the positive square root of μ_n and set $y_n = Kx_n/\lambda_n$. As

$$\begin{aligned}
(y_n, y_m) &= \frac{1}{\lambda_n\lambda_m}(K^\dagger K x_n, x_m) \\
&= \frac{\mu_n}{\lambda_n\lambda_m}(x_n, x_m) = \frac{\mu_n}{\lambda_n\lambda_m}\delta_{mn} = \delta_{mn}
\end{aligned}$$

the vectors y_n are orthonormal. Now we prove the claimed expansion for Kf. Consider

$$\begin{aligned}
K^\dagger K f &= \sum_n \mu_n (f, x_n) x_n \\
&= \sum_n (f, x_n)(K^\dagger K x_n).
\end{aligned}$$

Set

$$g = f - \sum_n (f, x_n) x_n,$$

then

$$\begin{aligned}
\|Kg\|^2 &= (Kg, Kg) \\
&= (K^\dagger K g, g) \\
&= 0. \qquad \text{(J.198)}
\end{aligned}$$

Therefore

$$Kf = \sum_n (f, x_n) K x_n = \sum_n \lambda_n (f, x_n) y_n.$$

□

The proof shows that the singular values of K are precisely the eigenvalues of $|K| = \sqrt{K^\dagger K}$.

Theorem J.7.31 *Let X be a separable Hilbert space and T a bounded positive operator. Let $\{b_n\}$ be any orthonormal basis. Then*

$$Tr(T) := \sum_n^{\infty} < b_n, T b_n >$$

is independent of the orthonormal basis and is called the trace of T. It satisfies the following properties:

(1) $Tr(T_1 + T_2) = Tr(T_1) + Tr(T_2)$,

(2) $Tr(\lambda T) = \lambda Tr\ (T)$,

(3) $Tr(UTU^{-1}) = Tr(T)$ for all unitary operators U and

(4) $T \geq 0$ implies $Tr(T) \geq 0$ for all positive T_I and $\lambda \in \mathbb{C}$.

Proof: Given an orthonormal basis $\{\varphi_n\}_{n=1}^{\infty}$, define $\mathrm{Tr}_\varphi(A) = \sum_{n=1}^{\infty} (\varphi_n, A\varphi_n)$. If $\{\psi_n\}_{n=1}^{\infty}$ is another orthonormal basis then

$$\begin{aligned}
(\psi_l, A^{1/2}\varphi_n) &= (\psi_l, \sum_{m=1}^{\infty} c_m \psi_m) \\
&= \sum_{m=1}^{\infty} c_m (\psi_l, \psi_m) = c_l.
\end{aligned}$$

Using this in the trace formula,

$$\begin{aligned}
\mathrm{Tr}_{\varphi}(A) &= \sum_{n=1}^{\infty}(\varphi_n, A\varphi_n) = \sum_{n=1}^{\infty}\|A^{1/2}\varphi_n\|^2 \\
&= \sum_{n=1}^{\infty}(A^{1/2}\varphi_n, A^{1/2}\varphi_n) \\
&= \sum_{n=1}^{\infty}(\sum_{m=1}^{\infty} c_m\psi_m, A^{1/2}\varphi_n) \\
&= \sum_{n=1}^{\infty}\left(\sum_{m=1}^{\infty}|(\psi_m, A^{1/2}\varphi_n)|^2\right) \\
&= \sum_{m=1}^{\infty}\left(\sum_{n=1}^{\infty}|(A^{1/2}\psi_m, \varphi_n)|^2\right) \\
&= \sum_{m=1}^{\infty}\|A^{1/2}\psi_m\|^2 \\
&= \sum_{m=1}^{\infty}(\psi_m, A\psi_m) \\
&= \mathrm{Tr}_{\psi}(A)
\end{aligned}$$

Properties 1), 2),tracetheorem and 4) are obvious. To prove 3) we note that $\{\varphi_n\}$ is an orthonormal basis, then so is $\{U\varphi_n\}$. Thus,

$$\mathrm{Tr}(UAU^{-1}) = \mathrm{Tr}_{(U\varphi)}(UAU^{-1}) = \mathrm{Tr}_{\varphi}(A) = Tr(A)$$

□

Two important subsets of compact operators based on the trace are the following.

Definition A bounded operator T on a separable Hilbert space X is called Hilbert-Schmidt if and only if $\mathrm{Tr}(T^{\dagger}T) < \infty$. We denote the family of Hilbert-Schmidt operators by $B_2(X)$.

□

Definition A bounded operator T on a separable Hilbert space X is called **trace class** if and only if $\mathrm{Tr}(|T|) < \infty$. We denote the family of trace class operators by $B_1(X)$.

For any $T \in B_1(X)$ we extend the trace by $\mathrm{Tr}(T) := \sum_n < b_n, Tb_n >$ which is independent of the orthogonal basis (b_n). The trace satisfies $\mathrm{Tr}(T^{\dagger}) = \overline{\mathrm{Tr}(T)}$ and $\mathrm{Tr}(AT) = \mathrm{Tr}(TA)$ for $T \in B_1(X)$ and $A \in B_2$.

$\square$

A bounded operator A on a separable Hilbert space $\mathcal{H}$ is said to be trce class if $|A| = \sqrt{A^\dagger A}$ has finite trace, that is $Tr(|A|) := \sum_I < b_I, |A| b_I >< \infty$ where e_I is any orthonormal basis for $\mathcal{H}$. Then $\mathrm{Tr}(A)$ is independent of the basis chosen.

An operator A is called Hilbert-Schmidt if $A^\dagger$ is trace class. Trace class operators are dense in a Hilbert space with inner product $< A, B >= \mathrm{Tr} A^\dagger B$ and its completion are the Hilbert-Schmidt operators.

Mixed states or density matrices are positive trace class operators of unit trace.

Hilbert-Schmidt operators naturally appear in Bogol'ubov transformations

J.7.10 Hilbert-Schmidt operators

We define the Hilbert-Schmidt norm $\|A\|_2$ as

$$\|A\|_2 := \left[\sum_{n=1}^{\infty} \|Af_n\|^2 \right]^{1/2}$$

which, by theorem J.7.31, when finite, is independent of the orthonormal basis in $\mathcal{H}$.

Theorem J.7.32 *Hilbert-Schmidt operators are compact.*

Proof: Let A be a bounded linear operetor $A : X \to Y$ and $\{e_n\}$ be a complete orthnormal sequence in X. We shall show that A is compact by expressing it as a norm limit of a sequence of finite rank operators. Define A_k for $k \in \mathbb{N}$ by

$$A_k \left(\sum_{n=1}^{\infty} x_n e_n \right) = \sum_{n=1}^{k} x_n A e_n$$

where

$$x = \sum_{n=1}^{\infty} x_n e_n \tag{J.199}$$

is an arbitrary element of X. Thus A_k agrees with A in the span of $e_1, \ldots, e_k$ and is zero on the span of the remaining e_n's. the rank of A_k is at most k, and so A_k is compact. For x as in (J.199),

$$
\begin{aligned}
(A - A_k)x &= \sum_{n=1}^{\infty} x_n e_n - \sum_{n=1}^{k} x_n A e_n \\
&= \sum_{n=k+1}^{\infty} x_n e_n
\end{aligned}
$$

Hence

$$
\begin{aligned}
\|(A - A_k)x\| &= \sum_{n=k+1}^{\infty} |x_n| \, \|Ae_n\| \\
&\leq \left\{ \sum_{n=k+1}^{\infty} |x_n|^2 \right\}^{1/2} \left\{ \sum_{n=k+1}^{\infty} \|Ae_n\|^2 \right\}^{1/2} \\
&\leq \|x\| \left\{ \sum_{n=k+1}^{\infty} \|Ae_n\|^2 \right\}^{1/2}
\end{aligned}
$$

It follows from the definition of the operator norm that

$$
\|(A - A_k)x\| \leq \left\{ \sum_{n=k+1}^{\infty} \|Ae_n\|^2 \right\}^{1/2} .
$$

The sum is the tail of a convergent series, and hence tends to zero as $k \to \infty$.

$\square$

Lemma J.7.33 *Let A be an operator such that the series $\sum_{n=1}^{\infty} \|Af_n\|^2$ converges for some orthonormal basis $f_1, f_2, \ldots$ in $\mathcal{H}$. Then*

$$
\|A\| \leq \|A\|_2
$$

Proof: Since we may take any unit vector as f_1,

$$
\|Af_1\| \leq \left[\sum_{n=1}^{\infty} \|Af_n\|^2 \right]^{1/2}
$$

from which it follows

$$\|A\| = \sup_{\|f\|=1} \|Af\| \le \|A\|_2.$$

□

Theorem J.7.34 *In order that a the operator A be of Hilbert-Schmidt type, it is necessary and sufficient that $\sum_{n=1}^{\infty} \lambda_n^2$ converge.*

Proof: First we prove necessity. Let $\{e_n\}$ be an orthonormal basis in $\mathcal{H}$ consisting of eigenvectors of $A^\dagger A$. Then $\lambda_n = (A^\dagger A e_n, e_n) = \|Ae_n\|$. As the trace is independent of the basis in $\mathcal{H}$, $\sum_{n=1}^{\infty} \|Af_n\|^2 = \sum_{n=1}^{\infty} \|Ae_n\|^2 = \sum_{n=1}^{\infty} \lambda_n^2$. As the trace is finite by the definition of a Hilbert-Schmidt operator, the series $\sum_{n=1}^{\infty} \lambda_n^2$ is finite.

Now we prove that convergence of the series $\sum_{n=1}^{\infty} \lambda_n^2$ is sufficient for A to be a Hilbert-Schmidt operator.

First we prove it is compact.

Let us denote by A_k the finite rank operator

$$A_k f = \sum_{n=1}^{k} (Af, e_n) e_n.$$

Then

$$\|A - A_k\|^2 \le \|A - A_k\|_2^2 = \sum_{n=1}^{\infty} \|(A - A_k)e_n\|^2 = \sum_{n=k+1}^{\infty} \lambda_n^2.$$

From the convergence of the series $\sum_{n=1}^{\infty} \lambda_n^2$ it follows that $\lim_{k\to\infty} \|A - A_k\| = 0$. Therefore A is the limit in the operator norm of a sequence of finite rank operators. Since finite rank operators are compact, the operator A is also compact.

Since $\lambda_n = \|Ae_n\|$ where $e_1, e_2, \ldots$ is an orthonormal basis consisting of eigenvectors of $A^\dagger A$, and as the series $\sum_{n=1}^{\infty} \lambda_n^2$ converges, the series $\sum_{n=1}^{\infty} \|Af_n\|^2$, where $f_1, f_2, \ldots$ is any orthonormal basis of $\mathcal{H}$, also converges and so A is a Hibert-Schmidt operator.

□

Definition We call the number $\|A\|_2$ the **Hilbert-Schmidt** norm of A.

$\square$

Obviously the Hilbert-Schmidt norm is finite for Hilbert-Schmidt operators and only for such operators. It satifies the basic properties of norms:

$$\|A+B\|_2 \leq \|A\|_2 + \|B\|_2$$

for

$$\begin{aligned}
\|A+B\|_2 &= \left[\sum_{n=1}^{\infty} \|(A+B)f_n\|^2\right]^{1/2} \\
&\leq \left[\sum_{n=1}^{\infty} (\|Af_n\| + \|Bf_n\|)^2\right]^{1/2} \\
&\leq \left[\sum_{n=1}^{\infty} \|Af_n\|^2\right]^{1/2} + \left[\sum_{n=1}^{\infty} \|Bf_n\|^2\right]^{1/2} = \|A\|_2 + \|B\|_2
\end{aligned}$$

and

$$\|\lambda A\|_2 = |\lambda| \|A\|_2.$$

The space of Hilbert-Schmidt operators is closed in the Hilbert-Schmidt norm.

$$\|A\|_2 = \left[\sum_{n=1}^{\infty} \|Af_n\|^2\right]^{1/2} = \left[\sum_{n=1}^{\infty}\sum_{k=1}^{\infty} |(Af_n, h_k)|^2\right]^{1/2}.$$

From this it follows that the space of Hilbert-Schmidt operators is isomophic to the space of infinite matrices $\|a_{nk}\|$ for which the series $\sum_{n=1}^{\infty}\sum_{k=1}^{\infty} |a_{nk}|^2$ converges. As is well know, this form a Hilbert space.

J.7.11 Trace Class Operators

Lemma J.7.35 *A positive operator is self-adjoint.*

Proof: Let T be positive, that is

$$(Tx, x) \geq 0$$

indicating it is real number, therefore

$$(Tx, x) = (Tx, x)^* = (x, Tx) = (T^\dagger x, x).$$

This implies

$$(T(x + \lambda y), x + \lambda y) = (T^\dagger(x + \lambda y), x + \lambda y)$$

for any complex number λ. The expansion of both sides is displayed below

$$\begin{aligned}
(T(x + \lambda y), x + \lambda y) &= (Tx, x) + |\lambda|^2 (Ty, y) + \lambda(Tx, y) + \lambda^*(Ty, x) \\
(T^\dagger(x + \lambda y), x + \lambda y) &= (T^\dagger x, x) + |\lambda|^2 (T^\dagger y, y) + \lambda(T^\dagger x, y) + \lambda^*(T^\dagger y, x).
\end{aligned}$$

From which we have

$$\lambda(Tx, y) + \lambda^*(Ty, x) = \lambda(T^\dagger x, y) + \lambda^*(T^\dagger y, x).$$

Taking λ to be 1 then i, and then subtracting the results gives

$$(Tx, y) = (T^\dagger x, y).$$

□

Lemma J.7.36 *In order that a compact positive operator T be trace class, it is necessary and sufficient that $\sum_{n=1}^{\infty} \lambda_n$ converge.*

Proof: Let T be a positive-definte trace class opertor. We introduce the operator $T^{1/2}$, setting

$$T^{1/2} e_n = \lambda_n^{1/2} e_n,$$

where $\{e_n\}$ is an orthonormal basis consisting of eigenvectors of T, and λ_n are corresponding eigenvalues. In view of the self-adjointness of $T^{1/2}$

$$\sum_{n=1}^{\infty} \|T^{1/2} f_n\|^2 = \sum_{n=1}^{\infty} (T^{1/2} f_n, T^{1/2} f_n) = \sum_{n=1}^{\infty} (T f_n, f_n).$$

Since

$$\sum_{n=1}^{\infty} \|T^{1/2}e_n\|^2 = \sum_{n=1}^{\infty}(Te_n, e_n) < \infty,$$

$T^{1/2}$ is a Hilbert-Schmidt operator. Therefore, for any orthonormal basis $f_1, f_2, \dots$ we have

$$\sum_{n=1}^{\infty} \lambda_n = \sum_{n=1}^{\infty} \|T^{1/2}e_n\|^2 = \sum_{n=1}^{\infty} \|T^{1/2}f_n\|^2.$$

Therefore for any orthonormal basis $f_1, f_2, \dots$ in $\mathcal{H}$

$$\sum_{n=1}^{\infty} \lambda_n = \sum_{n=1}^{\infty}(Tf_n, f_n) < \infty$$

from which it follows that $\sum_{n=1}^{\infty} \lambda_n$ is finite for the operator T.

Conversely, let T be a compact positive-definite operator for which $\sum_{n=1}^{\infty} \lambda_n$ is finite. We choose a basis $e_1, e_2, \dots$ in $\mathcal{H}$ consisting of eigenvectors of T, with corresponding eigenvalues $\lambda_1, \lambda_2, \dots$. Then

$$\sum_{n=1}^{\infty}(Te_n, e_n) = \sum_{n=1}^{\infty} \lambda_n < \infty$$

from which it follows that T is a trace class operator.

□

Theorem J.7.37 *The product AB of two Hilbert-Schmidt operators is a trace class operator. Conversely, every trace class operator is the product of two Hilbert-Schmidt operators.*

Proof: Suppose that B maps $\mathcal{H}_1$ into $\mathcal{H}_2$ and A maps $\mathcal{H}_2$ into $\mathcal{H}_3$, and let $AB = UT$ be the decomposition of the operator AB into the product of a positive-definite operator T, acting on the space $\mathcal{H}_1$ and an isometric operator U which maps the range of T into the sapce $\mathcal{H}_3$. We denote by $e_1, e_2, \dots$ an orthonormal basis in $\mathcal{H}_1$ consisting of eigenvectors of T, and by $h_1, h_2, \dots$ the orthonormal system in $\mathcal{H}_3$ consisting of the vectors $h_n = Ue_n$, $\lambda_n \neq 0$. Then for $\lambda_n \neq 0$ we have

$$\begin{aligned}
\lambda_n &= (Te_n, e_n) = (UTe_n, Ue_n) \\
&= (Be_n, A^\dagger h_n) \\
&\le \|Be_n\| \, \|A^\dagger h_n\| \\
&\le \frac{1}{2}[\|Be_n\|^2 + \|A^\dagger h_n\|^2]
\end{aligned} \tag{J.200}$$

where the last inequality follows from $(\|Be_n\| - \|A^\dagger h_n\|)^2 \ge 0$. If A and B are Hilbert-Schmidt operators, then the series $\sum_{n=1}^{\infty} \|Be_n\|^2$ and $\sum_{n=1}^{\infty} \|A^\dagger h_n\|^2$ converge and so the series $\sum_{n=1}^{\infty} \lambda_n$ converges. Thus we have proved that the product AB of two Hilbert-Schmidt operators is a trace class operator.

Now, let A be a trace class operator, and UT be the decomposition into the product of a positive-definite operator T and an isometric operator U. Then as was shown above, the operator $T^{1/2}$ is a Hilbert-Schmidt operator. Since

$$\sum_{n=1}^{\infty} \|UT^{1/2}e_n\|^2 = \sum_{n=1}^{\infty} \|T^{1/2}e_n\|^2$$

$UT^{1/2}$ is also a Hilbert-Schmidt operator. As $A = (UT^{1/2})T^{1/2}$, A is the product of two Hilbert-Schmidt operators.

□

Corollary J.7.38 *The space of trace class operators forms a two-sided * ideal in the space of bounded operators, i.e.*

1) The adjoint $A^\dagger$ of a trace class operator A is a trace class operator.

2) The product AB of any bounded linear operator A with a trace class opertor B is a trace class operator. The analogous assertion holds for the product BA.

Proof of 1): With $A = UT^{1/2}T^{1/2}$, then $A^\dagger = T^{1/2}(UT^{1/2})^\dagger$. The operator $(UT^{1/2})^\dagger$, as the adjoint of a Hilbert-Schmidt operator is also Hilbert-Schmidt. Consequently A is a trace class operator.

Proof of 2): If B is a trace class operator then $B = UT^{1/2}T^{1/2}$, where $UT^{1/2}$ and $T^{1/2}$ are Hilbert-Schmidt operators. Consequently,

$$AB = (AUT^{1/2})T^{1/2}$$

is the product of two Hilbert-Schmidt operators, and therefore is a trace class operator. The analogous assertion holds for the product BA, since $(A^\dagger B^\dagger)^\dagger$

□

Lemma J.7.39 *Every trace class operator*

$$\sum_{n=1}^{\infty}(ABe_n, e_n) = \sum_{n=1}^{\infty}(BAe_n, e_n) \tag{J.201}$$

Proof:

□

Corollary J.7.40 *In order that a compact A be trace class, it is necessary and sufficient that $\sum_{n=1}^{\infty}\lambda_n$ converge.*

Proof: We prove that for any unitary opertor V and any trace class operator $A = UT$ we have

□

Corollary J.7.41 *Every trace class operator is a Hilbert-Schimdt operator.*

Proof: Since the converge of the series $\sum_{n=1}^{\infty}\lambda_n^2$ follows from the convergence of $\sum_{n=1}^{\infty}\lambda_n$, every trace class operator is a Hilbert-Schimdt operator.

□

J.8 Convergence of Sequences of Operators and Functionals

Definition

weak convergence in a Hilbert space.

We say that a sequence of vectors $\{f_n\}$ conveges weakly to a vector f, if for every $h \in \mathcal{H}$,

$$\lim_{n\to 1}(f_n, h) = (f, h).$$

weak convergence does not imply strong convergence

For let $\{e_n\}_{n=1}^{\infty}$ be any infinite orthonormal sequence of vectors in $\mathcal{H}$. Since

$$\sum_{k=1}^{\infty} |(h, e_k)|^2 \leq (h, h),$$

it follows that, for any $h \in \mathcal{H}$,

$$\lim_{k \to 1} (e_k, h) = 0.$$

Thus the sequence $\{e_n\}_{n=1}^{\infty}$ converges weakly to the vector 0, but this sequence does not converge strongly, because

$$\|e_m - e_n\|^2 = 2 \qquad (m \neq n)$$

and therefore $\|e_m - e_n\|$ does not tend to 0 as $m, n \to \infty$.

Convergence of sequence of operators. Let X and Y be normed spaces. A sequence (T_n) of operators $T_n \in \mathcal{B}(X, Y)$ is said to be

(*i*) if (T_n) converges in the norm on $\mathcal{B}(X, Y)$

$$\|T_n - T\| \to 0 \tag{J.202}$$
$$\|T_n x - Tx\| \to 0 \quad \text{for all } x \in X \tag{J.203}$$
$$\|f(T_n x) - f(Tx)\| \to 0 \quad \text{for all } x \in X \text{ and for all } f \in Y'. \tag{J.204}$$

T is called the uniform, strong and weak operator limit of T_n, respectively.

for the other four topologies sequencies are not sufficient. the resulting topologies are not first countable, and so the closure of a subset N of $\mathcal{B}(\mathcal{H})$ is generally larger than the set of all limit points of sequencies in N. Rather, the closure of N is the set of all limit points of generalised sequencies (nets) in N.

Definition In terms of nets.

A net $\{T_\alpha\}$ converges to T in the strong operator topology if T_α converges to $T_\alpha \xi$ for every $\xi \in \mathcal{H}$.

A net $\{T_\alpha\}$ converges to T in the weak operator topology if $(T_\alpha \xi, \eta) \to (T\xi, \eta)$ converges to $T_\alpha \xi$ for all $\xi, \eta \in \mathcal{H}$.

□

Definition Strong and weak* convergence of a sequence of functionals. Let (f_n) be a sequence of bounded linear functionals on a normed space X. Then:

(a) Strong convergence of (f_n) means that there is an $f \in X'$ such that $\|f_n - f\| \to 0$.

(b) weak* convergence of (f_n) means that there is an $f \in X'$ such that $f_n(x) \to f(x)$ for all $x \in X$.

□

the norm topology is characterized by convergence of sequences. The other topologies requires the use of a generalization of sequences called nets. This we will come to after we have more topology.

for example when wish to investigate semi-classical issues in LQG when we do not yet have a physical inner product.

Definition The ultrastrong topology on $\mathcal{B}(\mathcal{H})$ is that given by the open neighbourhood base

$$\mathcal{N}(A;(x)_1^\infty,\epsilon) = \{B \in \mathcal{B}(\mathcal{H}) : |\sum_{i=1}^{\infty} \|(A-B)x_i\|^2| < \epsilon\} \tag{J.205}$$

for $A \in \mathcal{B}(\mathcal{H})$, $\epsilon > 0$ and any sequences in $\mathcal{H}$ satisfying $\sum_{i=1}^{\infty} \|x_i\|^2 < \infty$.

□

J.9 More Topology

we consider projective limits of infinite families of finite dimensional topological and measurable spaces.

J.9.1 Basis and Subbasis for a Topology

The notion of a basis for a finite dimensional vector space: It is a minimal collection of vectors that spans the vector space. If you know a basis, you can always recover the vector space. A similar notion of a basis exists for a topology as a family of open subsets that 'span' the topology.

Definition A basis for a given topology $\mathcal{T}$ on X is a subcollection $\mathcal{B}$ of $\mathcal{T}$ such that any open set in X is a union of members of $\mathcal{B}$.

$\square$

Definition A subbasis for a given topology is $\mathcal{T}$ on X is a subcollection $\mathcal{S}$ of $\mathcal{T}$ such that the collection of all finite intersections of members of $\mathcal{S}$ is a basis for $\mathcal{T}$. Hence, any open subset of X is a union of finite intersections of members in $\mathcal{T}$.

$\square$

Definition If $Y \subset X$, and $X, \mathcal{T}$ is a topological space, then relative topology on Y consists of the sets

$$\{A \cap Y | A \in \mathcal{T}\}. \tag{J.206}$$

$\square$

J.9.2 Hausdorff Spaces

Definition A *Hausdorff space* if the following is true: If $p \in X$, $q \in X$, and $p \neq q$, then p has a neighborhood U and q has a neighborhood V such that $U \cap V$.

$\square$

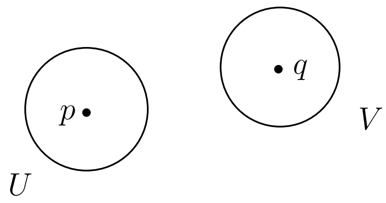

Figure J.8: hausdorff.

Theorem J.9.1 *Metric spaces are Hausdorff.*

Proof: If p and q are distinct points in a metric space, then $r = d(p,q) \neq 0$. Thus the open balls $B_{r/3}(p)$ and $B_{r/3}(q)$ are disjoint neighborhoods of p and q.

$\square$

Theorem J.9.2 *In a Hausdorff space, a convergent sequence has at most one limit point.*

Proof: Suppose we have a convergent sequence $\{x_n\}$ with two different limit points y and z. As $y \neq z$, there are disjoint neighbourhoods U and V of y and z, respectively. Since $x_n \to y$, there is $m \in \mathbb{N}$ such that $x_n \in U$ whenever $n \geq m$. Similarly, there is $m' \in \mathbb{N}$ such that $x_n \in V$ whenever $n \geq m'$. Let $l \in \mathbb{N}$ such that $m \leq l$ and $m' \leq l$. Then $x_l \in U, V$ i.e. $x_l \in U \cap V$, which is impossible since $U \cap V = \emptyset$. It follows that $y = z$.

□

Lemma J.9.3 *In Hausdorff spaces one point sets are closed.*

Proof: One point of a topological space is cloesd if and only if its complement is open which is true if and only if each point x' different from x has an open neighbourhood which does not contain x. Clearly, for any Hausdorff space one point sets are closed.

□

Theorem J.9.4 *The finite product of Hausdorff spaces is Hausdorff.*

Proof: Let $X := \prod_i X_i$ be the product of Hausdorff spaces X_i. If $x \neq x'$ then we must have $x_{i_0} \neq x'_{i_0}$ for at least one index, i_0. Since X_{i_0} is a Hausdorff space, x_{i_0} and x'_{i_0} can be separated by open sets in X_{i_0}. These two disjoint open subsets of X_{i_0} give rise to two disjoint sets in the defining open subbase for X, each of which contains one of the points x and x'.

□

Lemma J.9.5 *In a Hausdorff space, any point and disjoint compact subspace can be separated by open sets, that is, they have disjoint open neighbourhoods.*

Proof: Let X be a Hausdorff space, x a point in X, and C a compact subspace of X which does not contain x. We construct a disjoint pair of open sets U and V such that $x \in U$ and $C \subseteq V$. Let y be a point in C. Since X is Hausdorff, x and y have disjoint open neighbourhoods U_x and V_y. If we allow y to vary over C, we obtain a class of V_y's whose union contains C. Since C is compact, there is some finite subcover, i.e, a subclass $\{V_1, V_2, \ldots V_n\}$ such that $C \subseteq \cup_{i=1}^n V_i$. If $U_1, U_2, \ldots, U_n$ are the open sets of x which correspond to the V_i's, then any V_i is disjoint from the intersection of the U_i's so

$$U := \bigcap_{i=1}^{n} U_i$$

and

$$V := \bigcup_{i=1}^{n} V_i$$

are such that $x \in U$, $C \subseteq V$, and $U \cap V = \emptyset$.

$\square$

Lemma J.9.6 *Every compact subspace of a Hausdorff space is closed.*

Proof: Let C be a compact subspace of a Hausdorff space X. We prove C is closed by showing that X/C is open. Let x be any point in X/C. By lemma J.9.5, x has an open neighbourhood G such that $x \in G \subseteq X/C$. This shows that X/C is a union of open sets and is therefore open itself.

$\square$

Definition X is *locally compact* if every point of X has a neighbourhood whose closure is compact.

$\square$

The compact subsets of a Euclidean space $\mathbb{R}^n$ are precisely those that are closed and bounded. From this it follows easily that $\mathbb{R}^n$ is a localy compact Hausdorff space. Also, every metric space is locally compact.

Definition A collection of functions $\mathcal{C}$ on a (topological) space X is said to separate its points if and only if for any $x_1 \neq x_2$ we can find $f \in \mathcal{C}$ such that $f(x_1) \neq f(x_2)$.

$\square$

The only if part of the definition says given the values assumed by each and every function in the collection $\mathcal{C}$ exists a unique point $p \in X$ for which the functions take their given values.

Lemma J.9.7 *Let X be a topological space and $\mathcal{C} \subset C(X)$ a collection of continuous functions X which separate the points of X. Then the topology of X is Hausdorff.*

Proof of Lemma J.9.7:

Let $x_1, x_2 \in X$ with $x_1 \neq x_2$ be any two distinct points. Since $\mathcal{C}$ separates points we find $f \in \mathcal{C}$ with $f(x_1) \neq f(x_2)$. Let $d := |f(x_2) - f(x_1)| > 0$. Since f is continuous at points x_1 and x_2, for any $\epsilon > 0$ we find a neighbourhood $U_1(\epsilon)$ of x_1 such that

$$|f(x) - f(x_1)| < \epsilon$$

for any $x \in U_1(\epsilon)$ and a neighbourhood $U_2(\epsilon)$ of x_2 such that

$$|f(x) - f(x_2)| < \epsilon$$

for any $x \in U_2(\epsilon)$. Now we consider any one point $x \in X$ and show it cannot be in both $U_1(\epsilon)$ and $U_2(\epsilon)$ at the same time if we take ϵ to be small enough. We have

$$d = |f(x_2) - f(x_1)| \leq |f(x_2) - f(x)| + |f(x_1) - f(x)|$$

for any $x \in X$. Thus

$$d - \epsilon < |f(x_2) - f(x)| \text{ for any } x \in U_1(\epsilon)$$

and

$$d - \epsilon < |f(x_1) - f(x)| \text{ for any } x \in U_2(\epsilon).$$

Take x to be a point in $U_1(\epsilon) \cap U_2(\epsilon)$. If we choose $\epsilon < d/2$ the previous two inequlities give

$$d/2 < |f(x_2) - f(x)| \text{ and } d/2 < |f(x_1) - f(x)|.$$

Now, the point x cannot satisfy both these conditions at the same time, hence

$$U_1(\epsilon) \cap U_2(\epsilon) = \emptyset.$$

□

J.9.3 Urysohns's Lemma

Definition A **normal space** is a space such that one point sets are closed and any two disjoint closed sets A, B are contained in disjoint open sets.

□

Theorem J.9.8 *Every compact Hausdorff space is normal.*

Proof: By lemma J.9.3 in Hausdorff spaces one point sets are closed.

Let X be a compact Hausdorff space, and A and B disjoint closed subsets of X. We must construct a pair of disjoint open sets U and V such that $A \subset U$ and $B \subset V$. Since X is compact, A and B are disjoint compact subspaces of X. Let x be a point of A. By lemma J.9.5, x and B have disjoint neighbourhoods U_x and V_B. If we allow x to vary over A, we obtain a class of U_x's whose union contains A. Since A is compact, there is a finite subcover, i.e., a subclass, $\{U_1, U_2, \ldots, U_n\}$ such that $A \subseteq \cup_{i=1}^n U_i$. Let $V_1, V_2, \ldots, V_n$ be the open neighbourhoods of B which are correspond to the U_i's, it is clear that $U := \cup_{i=1}^n U_i$ and $V := \cap_{i=1}^n V_i$, are disjoint open neighbourhoods of A and B.

□

Lemma J.9.9 *A topological space X is normal if and only if for every closed subset $A \subset X$ and every open subset $B \subset X$ containing A, there exists an open set U such that*

$$A \subset U \subset \overline{U} \subset B.$$

Proof: Suppose that X is normal and A and B are as given in the lemma. Then the sets A and X/B are closed and disjoint. By the normailty of X, there exist open disjoint sets U and V such that $A \subset U$ and $X/B \subset V$. Then

$$\overline{U} \subset X/V \subset B,$$

so that U has the properties of the lemma.

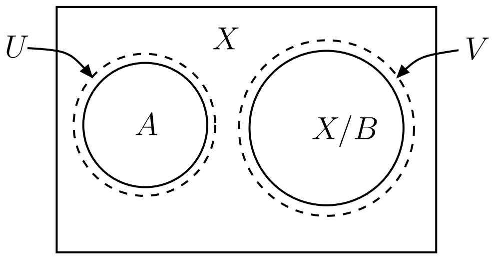

Figure J.9:

Conversely, let A and B be closed disjoint subsets of X. Then $V = X/B$ is open and $A \subset V$. By hypothesis there exists an open set U such that $A \subset U \subset \overline{U} \subset V$. Then U and $X/\overline{U}$ are disjoint open sets satisfying $A \subset U$ and $B \subset X/\overline{U}$ ($\overline{U} \subset X/B$ implies $B \subset X/\overline{U}$). So X is normal.

□

Lemma J.9.10 (Urysohn's Lemma). *In a normal topological space $(X, \mathcal{T})$ there is for each pair A, B of disjoint closed sets a continuous function $f : X \to [0,1]$ that is 0 on A and 1 on B.*

Proof: For the proof recall that a dyadic rational number is a number which can be written in the form $p = \frac{m}{2^n}$ with n, m being integers. Set $V = X/B$, an open set which contains the closed set A. By the previous lemma, there exists an open set $U_{1/2}$ such that

$$A \subset U_{1/2} \subset \overline{U_{1/2}} \subset V. \tag{J.207}$$

Applying lemma J.9.9 again to the open set $U_{1/2}$ containing A, we obtain an open set $U_{1/4}$ such that

$$A \subset U_{1/4} \subset \overline{U_{1/4}} \subset U_{1/2}$$

and to the open set V containing $\overline{U_{1/2}}$, we obtain the open set $U_{3/4}$ such that

$$\overline{U_{1/2}} \subset U_{3/4} \subset \overline{U_{3/4}} \subset V,$$

that is,

$$A \subset U_{1/4} \subset \overline{U_{1/4}} \subset U_{1/2} \subset \overline{U_{1/2}} \subset U_{3/4} \subset \overline{U_{3/4}} \subset V.$$

Continuing in this way, we associate to every such dyadic rational number $p \in (0,1)$ of the form $p = m/2^n$ (where $n = 1, 2, \ldots$ and $m = 1, 2, \ldots 2^n - 1$) an open subset $U_p \subset U$, such that

$$A \subset U_p \subset \overline{U_p} \subset U_q \subset \overline{U_q} \subset V$$

for $0 < p < q < 1$. Now we construct a function f which us continuous and such that the sets ∂U_p are level sets of f on which f assumes the value p. Define $f(x) = 0$ if x is in every U_p and

$$f(x) = \sup\{p : x \notin U_p\}$$

otherwise. It is clear that the values of f lie in $[0,1]$, and that $f(A) = 0$ and $f(B) = 1$. It remains to show that f is continuous. All intervals of the form $[0, a)$ and $(a, 1]$, where $0 < a < 1$, constitute an open subbase for $[0,1]$. It therefore suffices to show that

$f^{-1}([0,a))$ and $f^{-1}((a,1])$ are open. It is easy to see that $f(x) < a$ means that x is in some U_p for $p < a$. From this it follows that

$$f^{-1}([0,a)) = \{x : f(x) < a\} = \bigcup_{p<a} U_p,$$

which is an open set. Similarly, $f(x) > a$ means x is outside of $\overline{U_p}$ for some $p > a$.

$$f^{-1}((a,1]) = \{x : f(x) > a\} = \bigcup_{p>a} X/\overline{U_p},$$

which is an open set.

□

J.9.4 Products of Topological Spaces

Topological Space in Terms of Closed Sets

From De Morgan's laws and the definition of a topology in terms of opens sets

i) $X \in \mathcal{F}$ and $\emptyset \in \mathcal{F}$

ii) $\mathcal{F}$ is closed under finite union

iii) $\mathcal{F}$ is closed under arbitrary intersection.

DeMorgan'a Laws: Let A and B be subsets of X then

1) $X\backslash(A \cup B) = (X\backslash A) \cap (X\backslash B)$ and

2) $X\backslash(A \cap B) = (X\backslash A) \cup (X\backslash B)$.

open subsets $A_1, A_2, \ldots, A_n$ - finite intersctions of open sets, in terms of closed sets

$$X\backslash(\underbrace{A_1 \cap A_2 \cap \cdots \cap A_n}_{\text{finite intersection of open sets}}) = \underbrace{(X\backslash A_1) \cup (X\backslash A_2) \cup \cdots \cup (X\backslash A_n)}_{\text{finite union of closed sets}} \qquad \text{(J.208)}$$

$$\bigcap_{\alpha} (X\backslash A_\alpha) = X\backslash(\bigcup_{\alpha} A_\alpha) \qquad \text{(J.209)}$$

Since $\bigcup_\alpha A_\alpha$ is an open set $\bigcap_\alpha (X\backslash A_\alpha)$ is a closed set.

Compactness in Terms of Closed Sets

Recall that a topological space X is said to be compact if every open covering of X has a finite subcovering. This can be expressed in terms of closed sets instead, based on the notion of the *finite intersection property*:

Definition (The finite intersection property (FIP)) A nonempty family of sets $\mathcal{E}$ of a set X is said to have the finite intersection property (FIP for short) if the intersection of each nonempty finite subfamily $\mathcal{F}$ is nonempty, that is,

$$\bigcap_{E\in\mathcal{F}} E \neq \emptyset \quad \text{for every finite family } \mathcal{F} \subset \mathcal{E}. \tag{J.210}$$

Proposition J.9.11 *The space X is compact if every family of closed sets with the finite intersection property has nonempty intersection.*

Proposition J.9.12 *A topological space is compact if and only if every family of closed sets with the FIP has nonempty intersection.*

This means that for a family of sets $\mathcal{E}$

$$\begin{aligned} \bigcap_{E\in\mathcal{F}} E \;&\neq\; \emptyset \quad \text{for every finite family of closed sets } \mathcal{F} \subset \mathcal{E}, \\ &\text{implies } \bigcap_{E\in\mathcal{E}} E \neq \emptyset. \end{aligned} \tag{J.211}$$

X is compact

Proof. We show that the following statements are all equivalent

(i) every open cover of X has a finite subcover

(ii) if $\mathcal{U}$ is a family of open sets such that each finite subfamily fails to cover X, then $\mathcal{U}$ fails to cover X

(iii) if $\mathcal{U}$ is a family of open sets such that $\{X - U : U \in \mathcal{U}\}$ has the FIP, then $\cap\mathcal{U} \neq \emptyset$

(iv) every family of closed sets with the FIP has nonempty intersection.

It is obvious that (i) implies (ii). Note that any collection of opens sets which contains a finite cover of X is again a cover of X. With this trivial observation, it is easy to see (ii) implies the only open covers of X are those which contain finite subcovers. Hence (i) $\Leftrightarrow$ (ii). If a finite subfamily fails to cover X then

$$(X - U_{i_1}) \cap (X - U_{i_2}) \cap \cdots \cap (X - U_{i_n}) \neq \emptyset$$

If a collection of open sets $\mathcal{U}$ fail to cover X then $\cap\mathcal{U} \neq \emptyset$. So (iii) is just a restatement of (ii). Hence (ii) $\Leftrightarrow$ (iii) . Any closed set, by definition, is the complement, $X - U$, of an open set U. Any family of closed sets is given by the collection $\{X - U_i\}$ where the U_i's belongs to certain family of open sets $\mathcal{U}$, and so it follows that (iii) $\Leftrightarrow$ (iv).

$\square$

Proposition J.9.13 *For a family of sets, the finite intersection property implies the complete intersection property for closures.*

This means that for a family of sets $\mathcal{E}$

$$\begin{aligned} \bigcap_{E \in \mathcal{F}} E &\neq \emptyset \quad \text{for every finite family } \mathcal{F} \subset \mathcal{E}, \\ &\text{implies } \bigcap_{E \in \mathcal{E}} \overline{E} \neq \emptyset. \end{aligned} \tag{J.212}$$

Proof:

Assume (J.211). If the finite family of sets $\mathcal{F}$ in (J.212) has nonempty intersection then so does the closure of each of its members. Then by condition (J.211) we have $\bigcap_{E \in \mathcal{E}} \overline{E} \neq \emptyset$. Hence (J.211) implies (J.212). Conversly, (J.211) is just a special case of (J.212) where all the sets in the family $\mathcal{F}$ are closed. Hence (J.211) $\Leftrightarrow$ (J.212).

Lemma J.9.14 *FIP is closed under union.*

In Tynchonov's theorem we will...

...and we need only show the intersection over closures of elements of D is nonempty

J.9.5 Tychonov's Theorem

Tychonov's Theorem is possibly the deepest theorem in point set topology.

Tychonov Topology

We know the topological space when we know a subbase of open sets by arbitray unions and finite intersections. Since the preimages of open sets under contiuous functions are open by definition, we obtain a topology once we know which functions are continuous.

Lemma J.9.15 *the spaces are topological spaces $(X_\alpha, \mathcal{T}_\alpha)$ are topological spaces and we have maps $f_\alpha : X \to X_\alpha$. Then there is a smallest topology $\mathcal{T}$ on X for which all the maps are contiuous.*

A topology $\mathcal{T}$ on X makes all the functions f_α continuous if and only if it contains

$$\mathcal{D} = \{f_\alpha^{-1}(U) \ : \ U \in \mathcal{T}_\alpha, \ \alpha \in A\}. \tag{J.213}$$

The weakest topology for which all f_α are continuous is generated by including the finite intersection and arbitray union of the open sets in $\mathcal{D}$, that is, the collection $\mathcal{D}$ forms a subbasis for the topology $\mathcal{T}$.

Definition The product topoloy is the smallest topology on the product for which all projections π_α are continuous.

□

The sets formed by finite intersection of the $\pi_i^{-1}(U_i)$'s are bases for the product topology If we have a family of topological spaces $(X_i, \mathcal{T}_i)_{i\in\mathcal{I}}$, then we define the product topology on $\times_{i\in\mathcal{I}} X_i$ as the topology with a basis

$$\bigcap_{i=1}^{n} \pi_i^{-1}(U_i) \tag{J.214}$$

where i runs over a finite number of elements of indexing set $\mathcal{I}$.

Note that for $\cap_{i_j}^{n} \pi^{-1}(U_{i_j})$, without loss of generality we can assume $i_1, i_2, \ldots, i_n$ all different because if there were repetitions like

$$\cdots \cap \pi_{i_k}^{-1}(U) \cap \pi_{i_k}^{-1}(V) \ldots$$

we can replce each by

$$\cdots \cap \pi_{i_k}^{-1}(U \cap V) \cap \ldots$$

and so eliminate repetitions.

Definition A subfamily $\mathcal{C}$ of a family of $\mathcal{E}$ of sets is called a chain if it is totally ordered by inclusion, that is, for any two members C_i, C_j of $\mathcal{C}$ one is contained in the other or they are equal,

$$C_i \subseteq C_j \quad \Longleftrightarrow \quad C_i \succeq C_j \tag{J.215}$$

□

Definition A family of $\mathcal{E}$ of sets is said to be closed under union if, whenever $\mathcal{C}$ is a chain in $\mathcal{E}$, we have $\cup\mathcal{C} \in \mathcal{E}$.

□

Definition A member M of $\mathcal{E}$ is said to be maximal if for any $X \in \mathcal{E}$ we have $M \subseteq X$ implies $M = X$. That is, a member is maximal if it is not a proper subset of any other memeber in the family.

□

The following Lemma, which can be derived from the Axion of choice,

Lemma J.9.16 (Zorn's Lemma). *Any family of sets closed under unions of chains has a maximal member.*

Zorn's Lemma is used in the proof of Tychonov's Theorem.

Theorem J.9.17 (Tychonov's Theorem). *If all X_α is compact, then $X = \prod_{\alpha \in A} X_\alpha$ is compact in the product topology.*

Proof of Theorem J.9.17:

Let $\{X_i : i \in I\}$ be a family of compact spaces. To show their product X is compact it suffices to show that, if $\mathcal{E}$ is any family of subsets of X with the FIP, then

$$\bigcap \overline{\mathcal{E}} \neq \emptyset, \quad \text{that is,} \quad \bigcap_{E \in \mathcal{E}} \overline{E} \neq \emptyset. \tag{J.216}$$

Consider the collection of all such families of subsets of X, we denote it $\mathbf{\Omega}$. Take some subcollection $\mathcal{E}_i$ of $\mathbf{\Omega}$, this defines a *subset* of $\mathbf{\Omega}$. We can define a partial order on such subsets. Take any two $\mathcal{E}_1, \mathcal{E}_2 \in \mathbf{\Omega}$, we write $\mathcal{E}_1 \subseteq \mathcal{E}_2$ if for every $E_i \in \mathcal{E}_1$ there exists

$E_j \in \mathcal{E}_2$ such that $E_i \subseteq E_j$. A subfamily $\mathcal{C}$ of $\mathbf{\Omega}$ is a chain if it is totally ordered by this inclusion. A member $\mathcal{M}$ of $\mathbf{\Omega}$ is maximal if for any $\mathcal{E} \in \mathbf{\Omega}$ we have $\mathcal{M} \subseteq \mathcal{E}$ implies $\mathcal{M} = \mathcal{E}$.

It is easy to verify that $\mathbf{\Omega}$ is closed under unions of chains, i.e., $\cup\mathcal{C} \in \mathbf{\Omega}$.

By Zorn's Lemma, there exists a maximal family satisfying FIP and containing $\mathcal{E}$. Let $\mathcal{M}$ be the collection of all families of subsets of X which contain $\mathcal{E}$ and have the FIP (- no collection satisfying the FIP properly contains $\mathcal{M}$). We need only show the intersection over closures of elements of $\mathcal{M}$ is nonempty, i.e.

$$\bigcap \overline{\mathcal{M}} \neq \emptyset, \quad \text{that is,} \quad \bigcap_{E \in \mathcal{M}} \overline{E} \neq \emptyset, \tag{J.217}$$

as this will imply (J.216). First of all we want to show that $\mathcal{M}$, being a maximal FIP, satisfies the two following conditions:

1) any finite intersection of elements of $\mathcal{M}$ is contained in $\mathcal{M}$, i.e., $E_1, E_2 \in \mathcal{M}$ implies $E_1 \cap E_2 \in \mathcal{M}$ and

2) any subset A of X intersecting every element of $\mathcal{M}$ is also contained in $\mathcal{M}$, i.e., for $A \subset X$ such that $A \cap E \neq \emptyset$ for all $E \in \mathcal{M}$ implies $A \in \mathcal{M}$.

To verify 1): If $E_1, E_2 \in \mathcal{M}$, then clearly the $\mathcal{M} \cup (E_1 \cap E_2)$ is a member of $\mathbf{\Omega}$; since it includes $\mathcal{M}$, and $\mathcal{M}$ is maximal in $\mathbf{\Omega}$, it must coincide with $\mathcal{M}$, so that $E_1 \cap E_2$ must be a member of $\mathcal{M}$.

To verify 2): Suppose that A is a subset of X which meets every member of $\mathcal{M}$. Then, for each finite subset $\{E_1, \dots, E_n\}$ of $\mathcal{M}$ we have by 1) $E_1 \cap \cdots \cap E_n \in \mathcal{M}$, so that $A \cap (E_1 \cap \cdots \cap E_n) \neq \emptyset$. Therefore $\mathcal{M} \cup \{E\}$ has the FIP, and so is a member of $\mathbf{\Omega}$ including $\mathcal{M}$. The latter's maximality implies then $\overline{U}\hat{A} \in \mathcal{M}$.

Now for each $\alpha \in I$, FIP for $\mathcal{M}$ implies FIP for $\pi_\alpha(\mathcal{M}) = \{\pi_\alpha(E) \mid E \in \mathcal{M}\}$. Since each X_α is compact we have

$$\bigcap_{E \in \mathcal{M}} \overline{\pi_\alpha(E)} \neq \emptyset.$$

This means for all α there exists $x_\alpha \in X_\alpha$, for all U_α and for all $E \in \mathcal{M}$

$$U_\alpha \cap \pi_\alpha(E) \neq \emptyset. \tag{J.218}$$

But this is equivalent to $\pi_\alpha^{-1}(U_\alpha) \cap E \neq \emptyset$, therefore from 2) we get $\pi_\alpha^{-1}(U_\alpha) \in \mathcal{M}$. Thus from 1) we get

$$\bigcap_{k=1}^{n} \pi_{\alpha_k}^{-1}(U_{\alpha_k}) \in \mathcal{M} \tag{J.219}$$

so that for all $E \in \mathcal{M}$

$$\bigcap_{k=1}^{n} \pi_{\alpha_k}^{-1}(U_{\alpha_k}) \bigcap E \neq \emptyset. \tag{J.220}$$

Comparing with (J.214), this means that $x = (x_\alpha)$ belongs to $\cap_{E\in\mathcal{M}}\overline{E}$. Thus

$$x \in \bigcap_{E\in\mathcal{M}} \overline{E} \neq \emptyset. \tag{J.221}$$

Finally, since $\mathcal{E} \subseteq \mathcal{M}$, it follows that $\bigcap_{E\in\mathcal{M}} \overline{E} \subseteq \bigcap \overline{\mathcal{E}}$, so that $\bigcap \overline{\mathcal{E}} \neq \emptyset$.

□

The proof follows trivially form the properties of nets on product spaces that is technically clearer.

J.10 Nets

A net does for general topological spaces that a sequence does for metric (and metrizable) spaces. The notions of closedness, continuity and compactness can be formulated in terms of nets - nets fully describe the structure of topological spaces.

At first we will see that nets are a straightforward generalsation of sequencies and things proved earlier with sequencies are easily reproved with nets - we will try to avoid much repetion. However, generalisation is not so easy when we need the notion of a *subnet*, for example when we come to compactness.

One easily shows that if a net converges (a fucntion is continuous) in a certain toplogy, then it does so in any weaker (stronger) toplogy. In our applications direct products of topological spaces are of fundamental imprtance.

"eventually greater than" - the concept of a directed set: First we need the notion of partial order.

Definition A binary relation $\preceq$ on a set X is said to be a **pre-order** iff

(*i*) $p \preceq p$ for all $p \in X$

(*ii*) $p \preceq q$ and $q \preceq r$ imply $p \leq r$ for all $p, q, r \in X$.

If it is also true that for $p, q \in X$,

$p \preceq q$ and $q \preceq p$ then $p = q$

then X is said to be a **partially ordered set** (or poset).

$\square$

So a poset consists of a set with a binary relation that indicates that, for certain pairs of elements in the set, one of the elements preorders the other. In a poset some pairs of elements bear the preorder relationship but other pairs don't.

Not all pairs of elements of L need to be in relation but if they, L is said to be totally ordered or linearly ordered.

Definition A pre-ordered set X is said to be **directed** if and only if each pair of members of X has an upperbound. That is, if $p, q \in X$, then there exists $s \in X$ such that $p \preceq s, q \preceq s$.

$\square$

Examples of directed sets

Example (Reinmann integrals)

(i) Partial order relation. Two partitions are ordered $P_1 \prec P_2$ if one or more of subintervals in the coarser partition P_1 is a finite union of subintervals of the finer partition P_2. Put another way, a sub-interval of P_2 breaks up some of the subintervals of P_1.

(ii) Common refinement. Given any two partitions P_1, P_2 there is a common refinement; P_3 such that $P_1 \preceq P_3$ and $P_2 \preceq P_3$. This property makes the family of partitions a partially ordered and directed set. (Directed towards refinement of partition.)

Example Consider the set Γ of all graphs with oriented edges. Whenever each edge of γ can be .... of edges of γ'. The set Γ is a directed set with respect to the relation

$$\gamma' \preceq \gamma \tag{J.222}$$

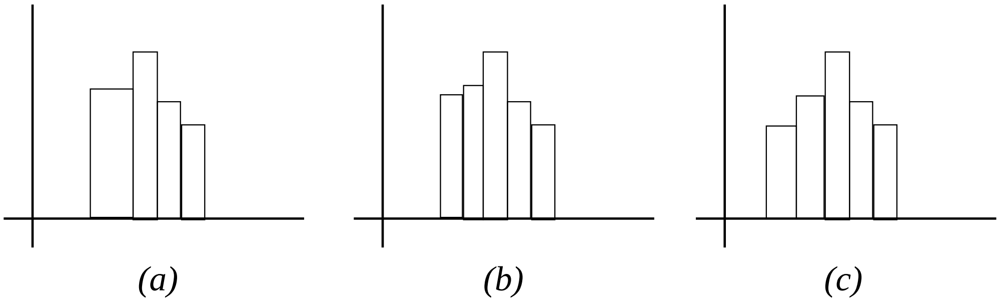

Figure J.10: netsReiman: Partions of Reinmann integrals. The partition in (b), P_2, is a finer partition than that in (a), P_1.

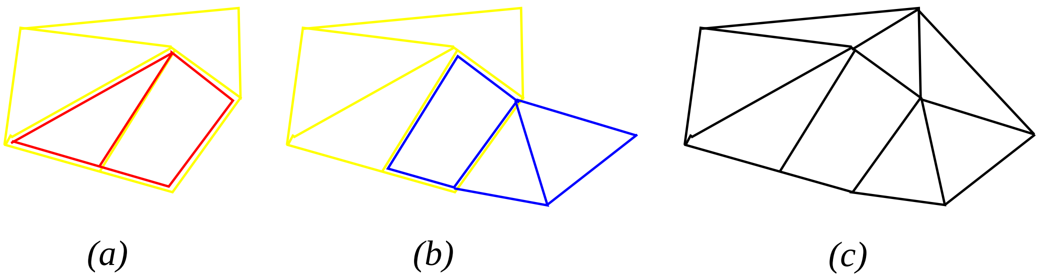

Figure J.11: (a) graph set Γ. (a) The red graph γ_R is contained in the yellow graph γ_Y so that $\gamma_R \prec \gamma_Y$ (b) Neither graph is contained in the other so γ_Y and the blue graph γ_B are incomparible. (c) Both the graphs γ_R and γ_B are contained this larger graph γ_L, i.e., $\gamma_Y \prec \gamma_L$ and $\gamma_B \prec \gamma_L$.

Example (Cellular decomposition)

(i) Partial order relation. Two cellular decompositions are related $C_1 \preceq C_2$ if any cell in the coarser decomposition C_1 is a finite union of cells of the finer decomposition C_2.

(ii) Common refinement. Given any two cellular decompositions C_1, C_2 there is a common refinement; C_3 such that $C_1 \preceq C_3$ and $C_2 \preceq C_3$. This property makes the family of cellular decompositions a partially ordered and directed set. (Directed towards refinement.)

Definition Here we define a **neighbourhood**, N, of a point x in a topological space if and only if there is an ope set in the topology such that $x \in U$ and $U \subseteq N$.

$\square$

A sequence in a set X is a mapping from $\mathbb{N}$ into X, we generalise this to directed sets

Definition A net in a topological space $(X, \mathcal{T})$ is a mapping from a directed set I into X, denoted $(x_\alpha)_{\alpha \in I}$.

□

Examples of a net

A sequence is a net whose domain is the set of positive integers with the usual ordering.

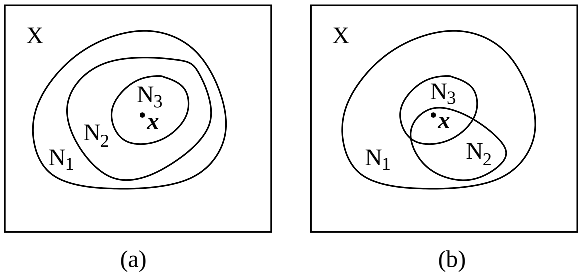

Figure J.12: (a) Inverse set inclusion: $N_3 \subset N_2 \subset N_1$. (b) $N_3 \not\subset N_2$ still so that $N_2 \subset N_1$ and $N_3 \subset N_1$.

Given $x \in (X, \mathcal{T})$ select in any way an element from each neigbourhood N of x; then (x_N) is a net in X.

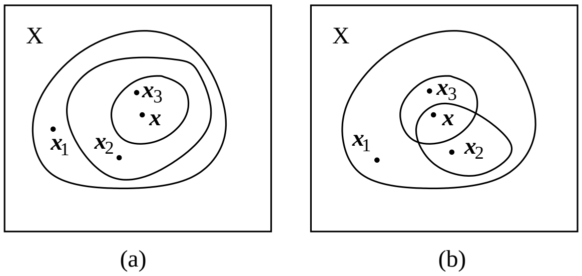

Figure J.13: Example of a net.(a) $x_1 \preceq x_2 \preceq x_3$. (b) We have that x_2 and x_3 are incomparable, $x_3 \| x_2$, but $x_1 \preceq x_2$ and $x_1 \preceq x_3$.

J.10.1 Net Convergence and Closure

We say that $(x_\alpha)_{\alpha \in I}$ converges to the point $x \in X$ if for any neighbourhood U of x $(x_\alpha)_{\alpha \in I}$ is eventually in U. x is called the limit of $(x_\alpha)_{\alpha \in I}$, and we write $x_\alpha \to x$.

Definition A **net** (x_α) in $(X, \mathcal{T})$ *converges to a limit* l if for each neighbourhood N of l, there exists some $\alpha_N \in A$ such that $x_\alpha \in N$ for all $\alpha \succeq \alpha_N$.

□

Proposition J.10.1 *Let $\mathcal{N}_x$ be a neighbourhood base at a point x in a topological space $(X, \mathcal{T})$ and suppose that, for each $U \in \mathcal{N}_x$, x_U is a given point in U. Then the net $(x_U)_{\mathcal{N}_x}$ converges to x, where $\mathcal{N}_x$ is partially ordered by reverse inclusion.*

Proof:

□

Theorem J.10.2 *Given $(X, \mathcal{T})$, $A \subseteq X$, $p \in X : p \in \overline{A}$ if and only if there exists a net in A converging to p.*

Proof: We know that a point $x \in X$ belongs to $\overline{A}$ if and only if every neighbourhood of x meets A. Suppose that $(a_\alpha)_{\alpha \in I}$ is a net in A such that $a_\alpha \to x$. By defintion of convergence, $(a_\alpha)_{\alpha \in I}$ is eventually in every neighbourhood of x, so $x \in \overline{A}$.

Now suppose that $x \in \overline{A}$. Let $\mathcal{U}_x$ be the collection of all neighbourhoods of x ordered by reverse inclusion (i.e. for $U, V \in \mathcal{U}_x$, $U \preceq V$ if and only if $V \subseteq U$). Then $\mathcal{U}_x$ is a directed set. We know that for each $U \in \mathcal{U}_x$ the set $U \cap A$ is non-empty. Let a_U be any element of $U \cap A$. So $a_U \to x$.

□

Theorem J.10.3 *A set E in a topological space $(X, \mathcal{T})$ is closed if and only if no net in E can converge to a point in X/E.*

Proof: Suppose E is closed and let (x_α) be any net in E such that (x_α) is eventually in every neighbourhood of some point p, say. Then for each open set U containing p, there exists α such that $x_\alpha \in U$. Thus $U \cap E \neq \emptyset$ for each open set U containing p. It follows that $p \in \overline{E}$, and since E is closed, $p \in E$.

Now suppose that any net in E can not be eventually in every neighbourhood of a point not belonging to E. Let p be any point in $\overline{E}$. Then for each open set U containing p, $U \cap E \neq \emptyset$. Let $x_U \in U \cap E$ for each open set U containing p, then $(x_U)_{U \in I}$ is a net in E, where I is the family of neighbourhoods of p ordered by reverse inclusion, such that x_U converges to p. But by hypothesis the convergence point must lie in E and we conclude that E is closed.

□

Theorem J.10.4 *In a Hausdorff space, a convergent net has at most one limit point.*

Proof: Similar as the proof of Theorem J.9.2. If $y \neq z$, there are disjoint open neighbourhoods U and V of y and z, respectively. Since $x_\alpha \to y$, there is $\beta \in I$ such that $x_\alpha \in U$ whenever $\alpha \succeq \beta$. Similarly, there is $\beta' \in I$ such that $x_\alpha \in V$ whenever $\alpha \succeq \beta'$. Let $\gamma \in I$ be such that $\beta \preceq \gamma$ and $\beta' \preceq \gamma$. Then $x_\gamma \in U \cap V$, which is impossible since $U \cap V = \emptyset$.

□

Theorem J.10.5 *Let X and Y be topological spaces. A mapping $f : X \to Y$ is continuous if and only if whenever $(x_\alpha)_I$ is a net in X convergent to x then the net $(f(x_\alpha))_I$ converges to $f(x)$.*

Proof: Suppose that $f : X \to Y$ is continuous and suppose that $(x_\alpha)_I$ is a net in X such that $x_\alpha \to x$. We wish to show that $(f(x_\alpha))_I$ converges to $f(x)$. To see this, let V be any open set in Y with $f(x) \in V$. Since f is continuous, $f^{-1}(V)$ is an open set in X containing x. But since $(x_\alpha)_I$ convergess to x, it is eventually in $f^{-1}(V)$, and so $(f(x_\alpha))_I$ is eventually in V. We conclude $f(x_\alpha) \to f(x)$.

Conversely, suppose that $f(x_\alpha) \to f(x)$ whenever $x_\alpha \to x$. Let V be open in Y. We must show that $f^{-1}(V)$ is open in X. If this is not true then there is a point $x \in f^{-1}(V)$ which is not an interior point. So that every open set containing x meets $X/f^{-1}(V)$. This means there is a net $(x_\alpha)_I$ in $X/f^{-1}(V)$ converging to x. By hypothesis, it follows that $f(x_\alpha) \to f(x)$. In particular, $(f(x_\alpha))_I$ is eventually in V, that is, $(x_\alpha)_I$ is eventually in $f^{-1}(V)$. But then $(x_\alpha)_I$ cannot be eventually in $X/f^{-1}(V)$. From this contracdition we conclude that $f^{-1}(V)$ is, indeed, open and therefore f is continuous.

□

J.10.2 Nets and Compactness

A subset of a metric space is compact if and only if any sequence in K has a subsequence which converges to an element of K.

What is the generalization, a subnet, of the notion of the subsequence of a sequence for a net?

Definition A map $F : J \to I$ between directed sets J and I is said to be cofinal if for any $\alpha \in I$ there is some $\beta' \in J$ such that $F(\beta) \succeq \alpha$ whenever $\beta \succeq \beta'$. In other words, F is eventually greater than any given $\alpha \in I$.

□

Definition Suppose that (x^α) is a net indexed by I and that $F : J \to I$ is a cofinal map from the directed set J into I. The net $(y_\beta)_{\beta \in J} = (x_{F(\beta)})_{\beta \in J}$ is said to be a subnet of (x^α)

□

Definition Let $(x_\alpha)_{\alpha \in A}$ be any net and let $\alpha_0 \in A$. The α_0^{th} **tail** of the net is the set $\{x_\alpha : \alpha \succeq \alpha_0\} = x([\alpha_0, \ ))$. We denote it by $x(\alpha_0 \to)$.

□

Definition The net $(x_\alpha)_{\alpha \in I}$ is said to be frequently in the set A if, for any given $\gamma \in I$, $x_\alpha \in A$ for some $\alpha \in I$ with $\alpha \succeq \gamma$.

□

Definition A point x is a cluster point of the net $(x_\alpha)_{\alpha \in I}$ if $(x_\alpha)_{\alpha \in I}$ is frequently in any neighbourhood of x.

□

Lemma J.10.6 *A function $f : X \to Y$ between topological spaces is continuous if for every convergent net (x_α) in X, the net $(f(x_\alpha))$ is convergent in Y.*

Proposition J.10.7 *Let $(x_\alpha)_I$ be a net in the space X and let $\mathcal{A}$ be a family of subsets of X such that*

(i) $(x_\alpha)_I$ is frequently in each member of $\mathcal{A}$;

(ii) for any $A, B \in \mathcal{A}$ there is $C \in \mathcal{A}$ such that $C \subseteq A \cap B$.

Then there is a subnet $(x_{F(\beta)})_J$ of the net $(x_\alpha)_I$ such that $(x_{F(\beta)})_J$ is eventually in each member of $\mathcal{A}$.

Proof: Let us equip $\mathcal{A}$ with ordering by reverse inclusion, i.e for any $A, B \in \mathcal{A}$

$$A \preceq B \quad \text{implies} \quad B \subseteq A.$$

Condition (ii) makes $\mathcal{A}$ equipped with this partial ordering a directed set, as for given any $A, B \in \mathcal{A}$, there is $C \in \mathcal{A}$ with $C \subseteq A \cap B$ meaning $C \succeq A$ and $C \succeq B$.

From this directed set we now construct a second directed set that will be the index set for the subnet. Let $\mathcal{E}$ denote the collection of pairs $(\alpha, A) \in I \times \mathcal{A}$ such that $x_\alpha \in A$, i.e

$$\mathcal{E} = \{(\alpha, A) : \alpha \in I,\ A \in \mathcal{A},\ x_\alpha \in A\}.$$

Define

$$(\alpha', A') \preceq (\alpha'', A'') \quad \text{to mean} \quad \alpha' \preceq \alpha'' \text{ in } I \text{ and } A' \preceq A'' \text{ in } \mathcal{E}.$$

This is a partial ordering for $\mathcal{E}$. We shall now see that condition (i) makes this into a directed set. Given $(\alpha', A'), (\alpha'', A'') \in \mathcal{E}$, there is $\alpha \in I$ with $\alpha \succeq \alpha', \alpha''$, and there is $A \in \mathcal{A}$ with $A \succeq A', A''$ ($\mathcal{A}$ being a directed set). Now, by (i) (x_α) is frequently in A and therefore there is $\gamma \succeq \alpha \in I$ such that $x_\gamma \in A$. Thus $(\gamma, A) \in \mathcal{E}$ and $(\gamma, A) \succeq (\alpha', A')$, $(\gamma, A) \succeq (\alpha'', A'')$ and hence $\mathcal{E}$ is a directed set.

Now we construct a cofinal map from $\mathcal{E}$ to I. Define $F : \mathcal{E} \to I$ by

$$F((\alpha, A)) = \alpha.$$

Let $\alpha \in I$ be given. For any $A_1 \in \mathcal{A}$ there is $\alpha_1 \succeq \alpha$ such that $x_{\alpha_1} \in A_1$ (by (i)). Set $\beta' = (\alpha_1, A_1)$, $\beta' \in \mathcal{E}$ and $F(\beta') = \alpha_1 \succeq \alpha$. So whenever we have $\beta \equiv (\alpha_2, A_2)$ satisfying

$$\beta \succeq \beta'$$

we have by $F(\beta) = \alpha_2 \succeq \alpha_1 \succeq \alpha$,

$$F(\beta) \succeq \alpha.$$

This shows that F is cofinal and therefore $(x_{F((\alpha,A))})_{\mathcal{E}}$ is a subnet of $(x_\alpha)_I$.

It remains to show that this subnet is eventually in every member of $\mathcal{A}$. Let $A \in \mathcal{A}$ be given. Then, by (i), there is $\alpha \in I$ such that $x_\alpha \in A$ and so $(\alpha, A) \in \mathcal{E}$. For any $(\alpha', A') \in \mathcal{E}$ with $(\alpha', A') \succeq (\alpha, A)$, we have

$$x_{F((\alpha',A'))} = x_{\alpha'} \in A' \subseteq A.$$

Thus $(x_{F((\alpha,A))})_{\mathcal{E}}$ is eventually in A.

□

Theorem J.10.8 *A point x in a topological space X is a cluster point of the net $(x_\alpha)_I$ if and only if some subnet converges to x.*

Proof: Suppose that x is a cluster point of the net $(x_\alpha)_I$ and let $\mathcal{N}_x$ denote the collection of all neighbourhoods of x. If $A, B \in \mathcal{N}_x$, so is their intersection $A \cap B$. As the net $(x_\alpha)_I$ is frequently in each member of $\mathcal{N}_x$, by the preceding proposition, there is a subnet $(y_\beta)_J$ eventually in each member of $\mathcal{N}_x$, that is, the subnet $(y_\beta)_J$ converges to x.

Conversely, suppose that $(y_\beta)_{\beta\in J} = (x_{F(x)})_{\beta\in J}$ is a subnet of $(x_\alpha)_I$ converging to x. Let N be any neighbourhood of x. Then there is $\beta_0 \in J$ such that $x_{F(\beta)} \in N$ whenever $\beta \succeq \beta_0$. Since F is cofinal, for any given $\gamma \in I$ there is $\beta' \in J$ such that $F(\beta) \succeq \gamma$ whenever $\beta \succeq \beta'$. Let $\beta \succeq \beta_0$ and $\beta \succeq \beta'$. Then

$$F(\beta) \succeq \gamma \ \text{ and } \ y_\beta = x_{F(\beta)} \in N.$$

Hence $(x_\alpha)_I$ is frequently in N and we conclude that x is a cluster point of the net $(x_\alpha)_I$.

□

Theorem J.10.9 $(X, \mathcal{T})$ *is compact if and only if in X, every net has (at least one) convergent subnet.*

Proof: Suppose every net has a convergent subnet. Let $\{V_\alpha\}_{\alpha\in I}$ be an open cover of X with no finte subcover. Let $\mathcal{F}$ be the collection of all finite subsets of $\{V_\alpha\}_{\alpha\in I}$ ordered by set-theoretical inclusion:

$$\mathcal{U}_1 \preceq \mathcal{U}_2 \ \text{ if and only if } \ \mathcal{U}_1 \subseteq \mathcal{U}_2$$

for $\mathcal{U}_1, \mathcal{U}_2 \in \mathcal{F}$. Given any two $\mathcal{U}_1, \mathcal{U}_2 \in \mathcal{F}$ and an arbitrary third $\mathcal{U}_3 \in \mathcal{F}$ we have

$$\mathcal{U}_1, \mathcal{U}_2 \subseteq \mathcal{U}_1 \cup \mathcal{U}_2 \cup \mathcal{U}_3 \in \mathcal{F} \quad \text{that is} \quad \mathcal{U}_1, \mathcal{U}_2 \preceq \mathcal{U}_1 \cup \mathcal{U}_2 \cup \mathcal{U}_3 \in \mathcal{F}$$

therefore it forms a directed set. For each $\mathcal{U} = \{V_{\alpha_1}, \ldots, V_{\alpha_m}\} \in \mathcal{F}$, there exists an

$$x_{\mathcal{U}} \notin \bigcup_{j=1}^{m} V_{\alpha_j}$$

otherwise, $\{V_\alpha\}_{\alpha\in I}$ would have a finite subcover. By hypothesis, the net $(x_{\mathcal{U}})_{x_{\mathcal{U}}\in\mathcal{F}}$ has a convergent subnet, equivalently a cluster point, x, by theorem J.10.8. Now, since $\{V_\alpha\}_{\alpha\in I}$ is a cover of X, there is some α' such that $x \in V_{\alpha'}$, and which $(x_{\mathcal{U}})_{x_{\mathcal{U}}}$ is frequently in.

Thus, by definition of frequently in, for any F' in the directed set $\mathcal{F}$, there is $F \succeq F' \in \mathcal{F}$ such that $x_F \in V_{\alpha'}$. In particular, if we take $F' = \{V_{\alpha'}\}$, we deduce that there is $F = \{V_{\alpha_1}, \ldots, V_{\alpha_k}\}$ such that $F \succeq \{V_{\alpha'}\}$, that is, $\{V_{\alpha'}\} \subseteq F$, and such that $x_F \in V_{\alpha'}$. But $x_F \notin F$, by construction. This contradiction implies that $(X, \mathcal{T})$ is compact.

Now, suppose that $(X, \mathcal{T})$ is compact and let $(x_\alpha)_{\alpha \in I}$ be a net in X. We suppose that $(x_\alpha)_{\alpha \in I}$ has no cluster points. Then, for any $x \in X$, there is an open neighbourhood U_x of x and $\alpha_x \in I$ such that $x_\alpha \notin U_x$ whenever $\alpha \succeq \alpha_x$ As the topological space is compact, the family $\{U_x : x \in X\}$ has an open cover of X for a finite number of points $x_1, \ldots, x_n \in X$, i.e. $X = \cup_{i=1}^n U_{x_i}$. Since I is a directed set there is $\alpha \succeq \alpha_i$ for each $i = 1, \ldots, n$. But then $x_\alpha \notin U_{x_i}$ for $i = 1, \ldots, n$, which is impossible since the U_{x_i}'s are an open cover for X.

□

Corollary J.10.10 *Compactness is preserved by continuous maps.*

J.10.3 Universal Nets

Universal nets will be used in the proof of Tychonov's Theorem.

Definition A net (x^α) on a set X is called *universal* if for every subset Y of X, either (x^α) is eventually either only in Y or only in X/Y.

□

Note that if (x_α) is a universal net, $A \subseteq X$, and (x_α) is frequently in A, then (x_α) is eventually in A.

Proposition J.10.11 *If a universal net has a cluster point, then it converges to the cluster point. In particular, a universal net in a Hausdorff space can have at most one cluster point.*

Proof: Suppose that x is a cluster point of the univrsal net $(x_\alpha)_{\alpha \in I}$, that is, for each neighbourhood N of x, $(x_\alpha)_{\alpha \in I}$ is frequently in N. So $(x_\alpha)_{\alpha \in I}$ is eventually in every neighbourhood of x and so converges to x. The last part of the proposition follows because in a Hausdorff space a net can converge to at most one point.

□

The following lemma is where Zorn's lemma comes into the proof of Tychonoff's theorem via nets.

Lemma J.10.12 *Let $(x_\alpha)_{\alpha \in I}$ be a net in a toplogical space X. Then there is a family $\mathcal{C}$ of subsets of X such that*

(i) $(x_\alpha)_{\alpha \in I}$ is frequently in each member of $\mathcal{C}$;

(ii) if $A, B \in \mathcal{C}$ then $A \cap B \in \mathcal{C}$;

(iii) for any $A \subset X$, either $A \in \mathcal{C}$ or $X/A \in \mathcal{C}$

Proof: Let Φ denote the collection of families of subsets of X satisfying the conditions (i) and (ii):

$$\Phi = \{\mathcal{F} : \mathcal{F} \text{ satifies (i) and (ii)}\}.$$

Φ is non-empty since $\{X\} \in \Phi$ ($\emptyset \notin \Phi$). The collection of Φ is partially ordered by set inclusion:

$$\mathcal{F}_1 \preceq \mathcal{F}_2 \text{ if and only if } \mathcal{F}_1 \subseteq \mathcal{F}_2$$

for $\mathcal{F}_1, \mathcal{F}_2 \in \Phi$.

Let $\{\mathcal{F}_\gamma\}$ be a totally ordered family in Φ, (recall a totally ordered family $\mathcal{F}_\gamma$ is a family in which any two subsets $A_1, A_2 \in \mathcal{F}_\gamma$ satisfy either $A_1 \preceq A_2$ or $A_2 \preceq A_1$). Put $\hat{\mathcal{F}} = \bigcup_\gamma \mathcal{F}_\gamma$. We shall show that $\hat{\mathcal{F}}$ itself is in Φ and so Zorn's lemma apllies.

If $A \in \hat{\mathcal{F}}$, then there is some γ such that $A \in \mathcal{F}_\gamma$, and so (x_α) is frequently in A (by (i) for $\mathcal{F}_\gamma$) and so condition (i) holds for $\hat{\mathcal{F}}$. Now, for any $A, B \in \hat{\mathcal{F}}$, there is γ_1 and γ_2 such that $A \in \mathcal{F}_{\gamma_1}$, and $B \in \mathcal{F}_{\gamma_2}$. Suppose, without loss of generality, that $\mathcal{F}_{\gamma_1} \preceq \mathcal{F}_{\gamma_2}$. Then $A, B \in \mathcal{F}_{\gamma_2}$, and therefore $A \cap B \in \mathcal{F}_{\gamma_2} \subseteq \hat{\mathcal{F}}$ (by (ii) for $\mathcal{F}_{\gamma_2}$), and we see that condition (ii) is satisfied for $\hat{\mathcal{F}}$. Thus $\hat{\mathcal{F}} \in \Phi$ as claimed.

By Zorn's lemma J.9.16, we conclude that Φ has a maximal element, $\mathcal{C}$, say. We shall show that $\mathcal{C}$ also satisfies condition (iii).

To see this, first suppose that it is true that (x_α) is frequently in $A \cap B$ for all $B \in \mathcal{C}$. Define $\mathcal{F}'$ by

$$\mathcal{F}' = \{C \subseteq X : A \cap B \subseteq C, \text{ for some } B \in \mathcal{C}\}.$$

We shall show that under the supposition that $\mathcal{F}' \in \Phi$. If $C \in \mathcal{F}'$ implies that $A \cap B \subseteq C$ for some $B \in \mathcal{C}$ and so (x_α) is frequently in C. Also, if $C_1, C_2 \in \mathcal{F}'$, then there are B_1 and B_2 in $\mathcal{C}$ such that $A \cap B_1 \subseteq C_1$ and $A \cap B_2 \subseteq C_2$. Obviously $A \cap (B_1 \cap B_2) \subseteq C_1 \cap C_2$.

Since $B_1 \cap B_2 \in \mathcal{C}$, (as $\mathcal{C}$ satisfies (ii) being in Φ), we deduce that $C_1 \cap C_2 \in \mathcal{F}'$. Thus $\mathcal{F}' \in \Phi$.

However, it is clear that $A \in \mathcal{F}'$ and also that if $B \in \mathcal{C}$ then $B \in \mathcal{F}'$. But $\mathcal{C}$ is maximal in Φ, and so $\mathcal{F}' = \mathcal{C}$ and we conclude that $A \in \mathcal{C}$, and (iii) holds.

Now suppose that it is false that (x_α) is frequently in every $A \cap B$, for $B \in \mathcal{C}$. Then there is some $B_0 \in \mathcal{C}$ such that (x_α) is not frequently in $A \cap B_0$, and so must be frequently in $X/(A \cap B_0)$. Thus for any α_0 there is $\alpha \succeq \alpha_0$ such that $x_\alpha \in X/(A \cap B_0)$ for all $\alpha \succeq \alpha_0$. That is, (x_α) is eventually in $X/A \cap B_0$. Let us set $\tilde{A} \equiv X/A \cap B_0$. As (x_α) is eventually in $\tilde{A}$, (x_α) must be frequently in $\tilde{A} \cap B$ for every $B \in \mathcal{C}$. Define $\mathcal{F}''$ by

$$\mathcal{F}'' = \{C \subseteq X : \tilde{A} \cap B \subseteq C, \text{ for some } B \in \mathcal{C}\}.$$

Thus, as above, we deduce that $\tilde{A} \in \mathcal{C}$. Furthermore, for any $B \in \mathcal{C}$, $B \cap B_0 \in \mathcal{C}$ and so $\tilde{A} \cap B \cap B_0 \in \mathcal{C}$. But

$$\begin{aligned}
\tilde{A} \cap B \cap B_0 &= \big(X/(A \cap B_0)\big) \cap (B \cap B_0) \\
&= \big((X/A) \cup (X/B_0)\big) \cap B \cap B_0 \\
&= \{(X/A) \cap B \cap B_0\} \cup \underbrace{\{(X/B_0) \cap B \cap B_0\}}_{=\emptyset} \\
&= (X/A) \cap B \cap B_0
\end{aligned}$$

and so we see that (x_α) is frequenlty in $(X/A) \cap B \cap B_0$ and hence is frequently in $(X/A) \cap B$ for any $B \in \mathcal{C}$. Define $\mathcal{F}'''$ by

$$\mathcal{F}''' = \{C \subseteq X : (X/A) \cap B \subseteq C, \text{ for some } B \in \mathcal{C}\}.$$

Again, by the above argument, we deduce that $(X/A) \in \mathcal{C}$. However, it is clear that $(X/A) \in \mathcal{F}'''$ and also that if $B \in \mathcal{C}$ then $B \in \mathcal{F}'''$. But $\mathcal{C}$ is maximal in Φ, and so $\mathcal{F}''' = \mathcal{C}$ and we conclude that $(X/A) \in \mathcal{C}$, and (iii) holds. This proves the claim and completes the proof of the lemma.

□

Theorem J.10.13 *Every net has a universal subnet.*

Proof: Let $(x_\alpha)_I$ be any net in X, and let $\mathcal{C}$ be a family of subsets as given by the previous lemma. Then by proposition J.10.7, $(x_\alpha)_I$ has a subnet $(y_\beta)_J$ which is eventually in each member of $\mathcal{C}$. But, by (iii) of the previous lemma, for any $A \subseteq X$, either $(y_\beta)_J$ is in $A \in \mathcal{C}$ or is in $X/A \in \mathcal{C}$, hence the subnet $(y_\beta)_J$ is either eventually in A or eventually in X/A, hence $(y_\beta)_J$ is universal.

□

Lemma J.10.14 *If (x_α) is a universal net in a space X, and $f : X \to Y$ is a function, then $f(x_\alpha)$ is a universal net in Y.*

Take any $A \subseteq Y$ and suppose that the net $f(x_\alpha)$ is frequently in A. Then (x_α) is frequently in $f^{-1}(A)$. It follows that (x_α) is eventually in $f^{-1}(A)$ since (x_α) is universal. Thus $f(x_\alpha)$ is eventually in A and so $f(x_\alpha)$ is universal.

□

Theorem J.10.15 *A topological space is compact if and only if every universal net converges.*

Proof: Suppose $(X, \mathcal{T})$ is a compact toplogical space and that (x_α) is a universal net in X. Since X is compact (x_α) has a convergent subnet, with limit point $x \in X$, say (by theorem J.10.9). The universal net (x_α) is frequently in any neighbourhood of x, for if it wasn't it wouldn't have a subnet converging to x. So x is a cluster point of the univrsal net (x_α) and therefore the net itself (x_α) converges to x (by proposition J.10.11).

Now, suppose that every universal net in X converges. Let (x_α) be any net in X. Then (x_α) has a subnet which is a universal subnet and must therefore converge.

□

Corollary J.10.16 *A subset C of a topological space is compact if and only if every universal net in C converges in C.*

The subset C of the topological space $(X, \mathcal{T})$ is compact if and only if it is compact with respect to the induced topology $\mathcal{T}_C = \{U \cap C : U \in \mathcal{T}\}$ on C. The result now follows by applying the previous theorem to $(C, \mathcal{T}_C)$.

□

J.10.4 Proof of Tychonov's Theorem Using Nets

The following, proven earlier in this section, facts about nets. For a map $f : X \to Y$ between topological spaces the net $f(x_\alpha)$ in Y is universal whenever (x_α) is universal in

X with no restriction on f and a topological space X is compact if and only if every universal net converges.

Proof of Theorem J.9.17

Let $(x_\alpha) = (x_{\alpha,l})_{l\in I}$ be any universal net in $X_\infty = \prod_{l\in I} X_l$. By Lemma J.10.14 the net $(\pi_l(x_\alpha)) = (x_\alpha)$ is universal in X_l. Since X_l is compact, it converges to some x_l. Define $x := (x_l)_{l\in I}$. By definition of the Tychonov topology, $x_\alpha \to x$ if and only if $x_{\alpha,l} \to x_l$ for any $l \in I$ whence (x_α) converges.

J.10.5 Quotient Topologies

"In our discussion of the gauge orbit of connections we will deal with the quotient of connections by the set of gauge tansformations whihc is a topological which is a too-logical spce again. The resulting quotient space carries a natural topology, the quotient topology."

Definition Let X, Y be a toplogical spaces and $p : X \to Y$ a surjection (onto). The map p is said to be a **quotient map** provided that $V \subset Y$ is open in Y if and only if $p^{-1}(V)$ is open in X.

□

As p is surjective the inverse map p^{-1} is well defined.

Notice the requirement for p to be a quotient map is stronger that that it be continuous which would only require that $p^{-1}(V)$ is open in X whenever V is open in Y (but not vice versa).

Lemma J.10.17 *Let X and Y be two sets, $p : X \to Y$ a surjection, and V_1, V_2 any two subsets of Y then*

$$p^{-1}(V_1) \cap p^{-1}(V_2) = p^{-1}(V_1 \cap V_2) \;\; \textit{and} \;\; p^{-1}(V_1) \cup p^{-1}(V_2) = p^{-1}(V_1 \cup V_2)$$

Proof: Consider any two subsets $V_1, V_2 \subset Y$. Suppose $x \in p^{-1}(V_1) \cap p^{-1}(V_2) \subset X$. Then there exist $y_1 \in V_1$, $y_2 \in V_2$ such that $y_1 = p(x) = y_2$, that is, $y_1 = y_2 \in V_1 \cap V_2$ so that actually $x \in p^{-1}(V_1 \cap V_2)$. We conclude

$$p^{-1}(V_1) \cap p^{-1}(V_2) \subset p^{-1}(V_1 \cap V_2)$$

On the other hand, let $x \in p^{-1}(V_1 \cap V_2) \subset X$ such that $x \in p^{-1}(y)$. Since $y \in V_1 \cap V_2$ we have $p^{-1}(y) \in p^{-1}(V_1)$ and $p^{-1}(y) \in p^{-1}(V_2)$, thus $x \in p^{-1}(V_1) \cap p^{-1}(V_2)$. We conclude

$$p^{-1}(V_1 \cap V_2) \subset p^{-1}(V_1) \cap p^{-1}(V_2),$$

so that

$$p^{-1}(V_1) \cap p^{-1}(V_2) = p^{-1}(V_1 \cap V_2)$$

and by taking complements we have

$$p^{-1}(V_1) \cup p^{-1}(V_2) = p^{-1}(V_1 \cup V_2).$$

$\square$

Theorem J.10.18 *If X is a toplogical space, Y a set and $p : X \to Y$ a surjection then there exists a unique topology on Y with respect to which p is a quotient map.*

Proof: Let

$$\mathcal{T}' = \{V : V \subseteq Y \text{ and } p^{-1}(V) \text{ is open in } X\}. \tag{J.223}$$

It is immediate that $\emptyset, Y \in \mathcal{T}'$ as $p^{-1}(\emptyset) = \emptyset$ and $p^{-1}(Y) = X$.

Let V and W be in $\mathcal{T}'$. Then $p^{-1}(V \cap W) = p^{-1}(V) \cap p^{-1}(W)$ is open in X and hence $V \cap W \in \mathcal{T}'$.

Let $\{V_\alpha : \alpha \in A\}$ be a family of sets in $\mathcal{T}'$. Then $p^{-1}(\bigcup_{\alpha \in A} V_\alpha) = \bigcup_{\alpha \in A} p^{-1}(V_\alpha)$ is open in X and hence $\bigcup_{\alpha \in A} V_\alpha \in \mathcal{T}'$. It follows that $(Y, \mathcal{T}')$ is a toplogical space.

That p is quotient follows directly from the defintion of the topology $\mathcal{T}'$.

$\square$

The topology $\mathcal{T}'$ is called the quotient topology on Y induced by p.

Definition Let X be a topological space and let $[X]$ be a **partition** of X (i.e. a collection of mutually disjoint subsets of X whose union is X). Denote by $[x]$, $x \in X$ the subset of X in that partition of X that contains x. Equip $[X]$ with the quotient toplogy induced by the map $[\,] : X \to [X];\ x \mapsto [x]$. Then $[X]$ is called the **quotient space** of X.

$\square$

Quotient spaces naturally arise if we have a group action $\lambda : G \times X \to X;\ (g, x) \to \lambda_g(x) := \lambda(g, x)$ on the topological space X and define $[x] := \{\lambda_g(x) : g \in G\}$ to be the orbit of x. The orbits clearly define a partition of X.

Lemma J.10.19 *Let X be a compact topological space, Y a set and $p : X \to Y$ a surjection. Then Y is compact in the quotient topology.*

Proof: Let $\mathcal{V}$ be an open cover for Y. By defintion of the quotient topology, $p^{-1}(V)$ is open in X and $\mathcal{U} := \{p^{-1}(V) : V \in \mathcal{V}\}$ covers X because

$$\bigcup_{U \in \mathcal{U}} U = \bigcup_{V \in \mathcal{V}} p^{-1}(V) = p^{-1}(\bigcup_{V \in \mathcal{V}} V) = p^{-1}(Y) = X$$

since p is a surjection and $\mathcal{V}$ covers Y. Therefore $\mathcal{U}$ is an open cover of X. Since X is compact there is a finite subcover $\{p^{-1}(V_k)\}_{k=1}^{N}$ of X so that

$$X = \bigcup_{k=1}^{N} p^{-1}(V_k) = p^{-1}(\bigcup_{k=1}^{N} V_k) = p^{-1}(Y)$$

Thus $Y = \bigcup_{k=1}^{N} V_k$, that is, $\{p^{-1}(V_k)\}_{k=1}^{N}$ is a finite subcover of $\mathcal{V}$ and so Y is compact.

$\square$

Lemma J.10.20 *Let X be a Hausdorff space and $\lambda : G \times X \to X$ a continupos group action on X (i.e., λ_g defined by $\lambda_g(x) := \lambda(g, x)$ is continuous for any $g \in G$). Then the quotient space $X/G := \{[x] : x \in X\}$ defined by the orbits $[x] = \{\lambda_g(x); g \in G\}$ is Hausdorff in the quotient topology.*

Proof:

$\square$

Theorem J.10.21 *Let X, Y be topological spaces and let G be a group acting (not necessarily continuously) on them via λ, λ' respectively. If $f : X \to Y$ is a homeomorphism with respect to which the actions λ, λ' are equivariant then f extends as a homeomorphism to the quotient spaces $X/G, Y/G$ in their respective quotient topologies.*

Proof:

$\square$

J.10.6 Limit Spaces

Definition A **projective system** $(X_\alpha, \varphi_{\beta\alpha}, D)$ is a directed set D with a collection $\{X_\alpha : \alpha \in D\}$ of Hausdorff spaces, and continuous functions

$$\varphi_{\beta\alpha} : X_\beta \to X_\alpha$$

such that if $\alpha \preceq \beta \preceq \gamma$ in D, then

$$\varphi_{\gamma\alpha} = \varphi_{\beta\alpha} \circ \varphi_{\gamma\beta}, \quad \text{and} \quad \varphi_{\alpha\alpha} = \mathbf{1}_{X_\alpha}$$

for each $\alpha \in D$. Each X_α is called a **factor space**.

□

Definition Let $(X_\alpha, \varphi_{\beta\alpha}, D)$ be a projective system of spaces and let

$$X_\infty = \prod_{\alpha \in D} X_\alpha.$$

For $\alpha \preceq \beta$ in D, let

$$S_\alpha^\beta = \{x \in X_\infty : \varphi_{\beta\alpha} \circ \pi_\beta(x) = \pi_\alpha(x)\}$$

where $\pi_\beta : X_\infty \to X_\beta$ is a projection. Let

$$S_\beta = \bigcap \{S_\alpha^\beta : \alpha \preceq \beta\}$$

for each $\beta \in D$, and let X be the subset of X_∞ defined by

$$X = \bigcap \{S_\beta : \beta \in D\}.$$

The space X with the subspace topology of X_∞ is called the **projective limit** of the system $(X_\alpha, \varphi_{\beta\alpha}, D)$ and is denoted $\overline{X}$.

□

Theorem J.10.22 *If $(X_\alpha, \varphi_{\beta\alpha}, D)$ is a projective system of spaces, then the projective limit of $\overline{X}$ of X_α is a closed subspace of X_∞.*

Proof: Let $\alpha \preceq \beta$ in D. We shall show that S_α^β is closed. Let $x \in X/S_\alpha^\beta$. Then $\varphi_{\beta\alpha}(x_\beta) \neq x_\alpha$. Let U and V be disjoint open subsets of X_α such that $\varphi_{\beta\alpha}(x_\beta) \neq x_\alpha$ and $x_\alpha \in V$. Since $\varphi_{\beta\alpha}$ is a continuous function from X_β to X_α there exists an open subset W in X_β such that $\varphi_{\beta\alpha}(W) \subseteq U$. Let $H \subseteq X$ such that

$$\pi_\gamma(H) = \begin{cases} V & \text{if } \gamma = \alpha : \\ W & \text{if } \gamma = \beta; \text{ and} \\ X_\gamma & \text{otherwise .} \end{cases}$$

□

Theorem J.10.23 *The projective limit of nonempty compact spaces is nonempty and compact.*

Proof: Since closed subspaces of compact spaces are compact in the subspace topology we conclude that $\overline{X}$ is compact in the subspace topology induced by X_∞.

□

Theorem J.10.24 *Both $X_\infty, \overline{X}$ are Haudorff spaces.*

Proof: Let $x \neq x'$ be points in X_∞. There is at least one $\alpha_0 \in D$ such that

$$\pi_{\alpha_0}(x) \neq \pi_{\alpha_0}(x').$$

Since X_{α_0} is Hausdorff there are disjoint open neighbourhoods $U_{\alpha_0}, U'_{\alpha_0} \subset X_{\alpha_0}$ of $\pi_{\alpha_0}(x), \pi_{\alpha_0}(x')$ repectively. Let

$$U := \pi_{\alpha_0}^{-1}(U_{\alpha_0}) \text{ and } U' := \pi_{\alpha_0}^{-1}(U'_{\alpha_0}).$$

Since the topology of X_∞ is generated by the continuous functions $\pi_\alpha : X_\infty \to X_\alpha$ from the topology of X_α, it follows that U and U' are open in X_∞. Obviously U and U' are neighbourhoods of x and x' respectively. Finally, $U \cap U' = \emptyset$ since $\pi_{\alpha_0}(U \cap U') = U_{\alpha_0} \cap U'_{\alpha_0} = \emptyset$ so that U and U' are disjoint open neighbourhoods of $x \neq x'$ and thus X_∞ is Haudorff.

To see that $\overline{X}$ is Hausdorff, let $x \neq x'$ be points in $\overline{X}$, then there are respective disjoint open neighbourhoods $U, U' \in X_\infty$, then $U \cap \overline{X}$ and $U' \cap \overline{X}$ are disjoint open neighbourhoods in $\overline{X}$ by definition of the subspace topology.

□

The compactness of each Hausdorff space will imply that every X_α is compact and Haudorff. Now on a direct product space (independent of the cardinality of the index set) in which each factor space is compact and Haudorff one can naturally define a topology, the Tychonov topology, such that X_∞ is itself compact. If $\overline{X}$ is closed in X_∞ then $\overline{X}$ will be compact and Hausdorff as well in the subspace toplogy. However, for compact Hausdorff spaces powerful measure theoretic theorems hold which enable one to equip to relavent infinite dimensional spaces associated to background independent gauge theories with the structure of the so-called $\sigma-$algebra and to develop measure theory thereon.

J.10.7 Locally Convex Spaces

Locally convex spaces play an important role in the theory of distributions which typically arise as solutions of constraints.

Definition A **seminorm** on a vector space X is a map $\rho : X \to [0, \infty)$ such that

1) $\rho(x + y) \leq \rho(x) + \rho(y)$ sublinerity

and

2) $\rho(\lambda x) = |\lambda| \rho(x)$ homogeneity

for all $x, y \in X$, $\lambda \in (C)$.

□

Unlike a norm, a seminorm is not positive definite, i.e. it does not have the property that $\rho(x) = 0$ implies $x = 0$.

It is possible to characterize seminorms geometrically. Let B_1 be the unit ball,

$$B_1 = \{x \in V : p(x) \leq 1\}$$

Proposition J.10.25 *Suppose that p is homogeneous. Then, it is subliear if and only if its unit ball, B_1, is a convex subset of V.*

Proof:

First, let us suppose that p is sublinear, and prove B_1 must be convex. Let $x, y \in B_1$, and let k be real number between 0 and 1. We show that $kx + (1-k)y$ is also in B_1. By assumption,

$$p(kx + (1-k)y) \leq kp(x) + (1-k)p(y).$$

As this is the weighted average of the two numbers $p(x) \leq 1$, $p(y) \leq 1$, we have

$$kx + (1-k)y \in B_1.$$

Conversely, suppose that the unit ball is convex. Given, $x, y \in V$, we must show that

$$p(x + y) \leq p(x) + p(y).$$

There are three cases we need to consider. First $p(x) = p(y) = 0$ where x and y need not be the zero elements. By homogeneity, for every $k > 0$ we have

$$kx, ky \in B_1,$$

and hence as B_1 is assumed to be convex

$$\frac{k}{2}x, \frac{k}{2}y \in B_1,$$

as well. By homogeneity, again,

$$p(x + y) \leq \frac{2}{k}.$$

Since the above is true for all positive k, we conclude that

$$p(x + y) = 0.$$

Next suppose $p(x) = 0$, but that $p(y) \neq 0$, we show that it turns out that we have

$$p(x + y) = p(y).$$

Due to homogeneity, we may without loss of generality assume that

$$p(y) = 1.$$

For k such that $0 \leq k < 1$ we have

$$kx + ky = (1-k)\frac{kx}{1-k} + ky.$$

The right hand side is an element of B_1 as

$$\frac{kx}{1-k}, y \in B_1$$

Hence

$$kp(x+y) \leq 1,$$

and since this holds for k arbitrarily close to 1 we conclude

$$p(x+y) \leq p(y).$$

Now apply the same argument to $x' = -x$ and $y' = x + y$ to obtain

$$p(x' + y') \leq p(y')$$

or

$$p(y) \leq p(x+y).$$

Hence

$$p(x+y) = p(y).$$

Finally, suppose that $p(x) \neq 0, p(y) \neq 0$. Then

$$\frac{x}{p(x)}, \frac{y}{p(y)} \in B_1,$$

and hence so is

$$\frac{p(x)}{p(x)+p(y)}\frac{x}{p(x)} + \frac{p(y)}{p(x)+p(y)}\frac{y}{p(y)} = \frac{x+y}{p(x)+p(y)}.$$

By homogeneity, we conclude that

$$p(x+y) \leq p(x) + p(y).$$

□

Definition A family of seminorms $(\rho_I)_{I\in\mathcal{I}}$ are said to **separate points** if $\rho_I = 0$ for all $I \in \mathcal{I}$ implies $x = 0$.

□

Definition A **locally convex space** is a vextor space X with a family of seminorms $(\rho_I)_{I\in\mathcal{I}}$ separating points.

□

Its **natural topology** is the weakest topology in which all the ρ_I and the operation of addition is continuous, i.e. the weakest topology such that

1) if $V \subset [0, \infty)$ is open then $\rho_I^{-1}(V)$ is open in X for all $I \in \mathcal{I}$ or, equivalently in terms of nets, if $x_\alpha \to x$ then $\rho(x_\alpha) \to \rho_I(x)$ for all $I \in \mathcal{I}$;

2) if $x_\alpha \to x$ and $y_\alpha \to y$, then $x_\alpha + y_\alpha \to x + y$

The natural topology is the intersection of all topologies on X with respect to which all the ρ_I's are continuous (equivalently, it is made up of the subsets $\rho_I^{-1}(V)$ where V is any open set in $[0, \infty)$, along with their arbitrary unions and finite intersections). This clearly defines a topology on X, and is weaker than any topology which has this property.

Definition Let $(\rho_I)_{I\in\mathcal{I}}$ be a family of seminorms on a vector space X. An $\mathcal{I}-$**open ball** of center x_0 is the set of points $x \in X$ which satisfy a finite number of inequalities

$$\rho_I(x - x_0) < \epsilon_I, \qquad \epsilon_I > 0, \quad I \in \mathcal{I}.$$

□

Recall that an open base for X is a class of open sets with the property that every open set is a union of sets in this class. The condition can also be expressed in the following equivalent form: if U is an arbitrary non-empty set and x is a point in U, then there is a set B in the open base such that $x \in B \subseteq U$.

Theorem J.10.26 *Let X be a locally convex vector space. Consider subsets U of X satisfying the condition, for each $x_0 \in U$, there are ϵ and $\rho_1, \dots, \rho_n \in (\rho_I)_{I \in \mathcal{I}}$ such that*

$$\left\{ x \in X : \max_{j=1,\dots,n} \rho_j(x - x_0) < \epsilon \right\} \subset U.$$

The collection of such subsets of X is a topology, i.e.

(i) $\emptyset$ and X are included;

(ii) if $(U_\alpha)_\alpha$ are in this family of sets, then so is $\cup_\alpha U_\alpha$;

(iii) if $U_1, \dots, U_n$ are in this family of sets, then so is $U_1 \cap \dots \cap U_n$.

Proof: We only prove (iii).

Let $x_0 \in U_1 \cap \dots \cap U_n$. For each $k = 1, \dots, n$, there are $\epsilon_k > 0$ and $\rho_1^{(k)}, \dots, \rho_{n_k}^{(k)} \in (\rho_I)_{I \in \mathcal{I}}$ such that

$$V_k := \left\{ x \in X : \max_{j=1,\dots,n_k} \rho_j^{(k)}(x - x_0) < \epsilon_k \right\} \subset U_k.$$

Let $\epsilon := \min\{\epsilon_1, \dots, \epsilon_k\}$, and note that

$$\begin{aligned} \left\{ x \in X : \max_{k=1,\dots,n} \max_{j=1,\dots,n_k} \rho_j^{(k)}(x - x_0) < \epsilon_k \right\} &\subset V_1 \cap \dots \cap V_n \\ &\subset U_1 \cap \dots \cap U_n. \end{aligned}$$

Similarly, conditions (i) and (ii) are also satisfied by the subsets U's defined at the beginning of the theorem. Hence, they form a topology and are therefore themselves open sets of the topology.

□

Theorem J.10.27 *Let X be a locally convex vector space. Then a net $(x_\alpha)_\alpha$ in X converges to $x_0 \in X$ in the topology defined in theorem J.10.26 if and only if $\rho_I(x_\alpha - x) \to 0$ for each $I \in (\rho_I)_{I \in \mathcal{I}}$.*

Proof: Suppose that $x_\alpha \to x$ in the topology. Fix $\epsilon > 0$ and $I \in (\rho_I)_{I \in \mathcal{I}}$. Then $U := \{x \in X : \rho(x_\alpha - x_0) < \epsilon\}$ is an open neighbourhood of x_0. Hence, there is an index α_0 such that $x_\alpha \in U$, i.e. $\rho(x_\alpha - x_0) < \epsilon$ for all $\alpha \succ \alpha_0$. Hence, $\rho(x_\alpha - x_0) \to 0$.

Conversely, suppose that $\rho(x_\alpha - x) \to 0$ for all $I \in \mathcal{I}$. Let U be a neighbourhood of x_0, i.e. there is an open set $V \subset U$ with $x_0 \in V$. By the definition, there are $\epsilon > 0$ and $\rho_1, \ldots, \rho_n \in (\rho_I)_{I \in \mathcal{I}}$ such that

$$\left\{ x \in X : \max_{j=1,\ldots,n} \rho_j(x - x_0) < \epsilon \right\} \subset V.$$

Since $\rho(x_\alpha - x) \to 0$ for $j = 1, \ldots, n$, there is an index α_0 such that

$$\rho_j(x_\alpha - x_0) < \epsilon \qquad (j = 1, \ldots, n,\ \alpha \succ \alpha_0).$$

Thus $x_\alpha \in V \subset U$ for all $\alpha \succ \alpha_0$.

□

By theorem J.10.27, when $x_\alpha \to x$ and $y_\alpha \to y$, then $x_\alpha + y_\alpha \to x + y$ if and only if when $\rho_I(x_\alpha) \to \rho_I(x)$ and $\rho_I(y_\alpha) \to \rho_I(y)$, then $\rho_I(x_\alpha + y_\alpha) \to \rho_I(x + y)$ for each $I \in (\rho_I)_{I \in \mathcal{I}}$.

Theorem J.10.28 *The topology defined in theorem J.10.26 is the natural topology.*

Proof: This is a direct consequence of the previous theorem. It implies the topology defined in theorem J.10.26 is necessary and sufficient for all the ρ_I's to be continuous. That addition is continuous in this topology follows from the first property of seminorms $(\rho_I(x + y) \leq \rho_I(x) + \rho_I(y))$. By

$$\rho_I(x_\alpha + y_\alpha - x - y) \leq \rho_I(x_\alpha - x) + \rho_I(y_\alpha - y),$$

we see that $x_\alpha \to x$ and $y_\alpha \to y$ implies $x_\alpha + y_\alpha \to x + y$.

□

Theorem J.10.29 *A family of seminorms which separates points of X defines a Hausdorff topology.*

Proof: Let $x, y \in X$, $x \neq y$. Then there exists a seminorm ρ such that $\rho(x - y) \neq 0$, i.e. such that $\rho(x - y) > \epsilon$. The nieghbourhoods of x and y defined respectively by

$$\rho(z - x) < \epsilon/2 \ \text{ and } \ \rho(z' - y) < \epsilon/2$$

are disjoint for if we had $z = z'$ at all the three inequalities would violate the triangle inequality

$$\rho(x-z+z-y) \leq \rho(z-x)+\rho(z-y).$$

Therefore if the family of seminorms separates points the corresponding topology must be Hausdorff.

Conversely, for the topology of a family of seminorms to be Hausdorff this family must separate points of X. For if there were an $x \in X$, $x \neq 0$, such that $\rho(x) = 0$ for every $\rho \in (\rho_I)_{I \in \mathcal{I}}$, then every ball of centre 0, and hence every neighbourhood of 0, would contain x; thus 0 and x would not be separated.

□

We show that a locally convex space X whose underlying family of seminorms is countable can be equipped with the following metric

$$d(x,y) = \sum_{n=1}^{\infty} 2^{-n} \frac{\rho_n(x-y)}{1+\rho_n(x-y)} \tag{J.224}$$

Note that for real number x such that $0 \leq x \leq \infty$ by

$$\frac{x}{1+x} = 1 - \frac{1}{1+x} \tag{J.225}$$

we have

$$0 \leq \frac{x}{1+x} \leq 1$$

and hence

$$0 \leq \frac{\rho_n(x-y)}{1+\rho_n(x-y)} \leq 1.$$

This is the reason for the factor 2^{-n} in the summation; it guarantees convergence. We must demonstrate that (J.224) has the properties of a metric. As $p_n(x) \geq 0$ we obviously have $d(x,y) \geq 0$. If $d(x,y) = 0$ then we must have that $p_n(x-y) = 0$ for all n, but as $\{p_n\}$ seperates points this implies $x = y$. As $p_n(x) = p_n(-x)$ we obviusly have $d(x,y) = d(y,x)$. By (J.225) the function $x/(1+x)$ is monotonically increasing and so we can write for real numbers $a, b, c \geq 0$ with $a \leq b + c$ that

$$\begin{aligned}\frac{a}{1+a} &\leq \frac{b+c}{1+b+c}\\ &= \frac{b}{1+b+c}+\frac{c}{1+b+c}\\ &\leq \frac{b}{1+b}+\frac{c}{1+c}. \end{aligned} \tag{J.226}$$

This inequality underlies the definition of the metric (J.224), it implies, via $p_n(x-z) = p_n(x-y+y-z) \leq p_n(x-y)+p_n(y-z)$, that

$$\frac{p_n(x-z)}{1+p_n(x-z)} \leq \frac{p_n(x-y)}{1+p_n(x-y)} + \frac{p_n(y-z)}{1+p_n(y-z)}.$$

and hence

$$d(x,z) \leq d(x,y)+d(y,z),$$

completing the properties defining a metric. Note it is obvious that this metric is translatinally invariant. This metric generates the same toplogy as the family of seminorms.

Proof:

On the one hand, one has for $\epsilon \in [0,1]$ and $N \in \mathbb{N}$:

$$d(x) \leq \epsilon 2^{-N}$$

implies

$$\frac{1}{2^n}\frac{p_n(x)}{1+p_n(x)} \leq \epsilon 2^{-N} \quad \text{for all } k$$

this in turn implies

$$\frac{p_n(x)}{1+p_n(x)} \leq \epsilon 2^{-N} \quad \text{for } k \leq N$$

and finally

$$p_n(x) \leq \frac{\epsilon}{1-\epsilon} \quad \text{for } k \leq N$$

On the other hand if $2^{-N} < \epsilon/2$, $p_n(x) \leq \epsilon/2$ for $k \leq N$ implies

$$\begin{aligned}
d(x) &= \sum_{n=1}^{N} \frac{1}{2^n} \frac{p_n(x)}{1+p_n(x)} + \sum_{n=N+1}^{\infty} \frac{1}{2^n} \frac{p_n(x)}{1+p_n(x)} \\
&\leq \sum_{n=1}^{N} \frac{1}{2^n} p_n(x) + \sum_{n=N+1}^{\infty} \frac{1}{2^n} \\
&\leq \frac{\epsilon}{2} \sum_{n=1}^{N} \frac{1}{2^n} + \frac{1}{2^N} \\
&\leq \frac{\epsilon}{2}(1 - \frac{1}{2^N}) + \frac{\epsilon}{2} \\
&\leq \epsilon
\end{aligned} \tag{J.227}$$

□

Theorem J.10.30 *Let X be a locally convex space. The following are equivalent:*

1) X is metrizable.

2) 0 has a countable neighbourhood base.

3) The topology on X is generated by some countable family of seminorms.

Proof: That 1) implies 2) is a property of any metric space.

□

Definition If it completed in this metric, it is called a **Frechet space**.

□

An important application is the following. Consider the space $\mathbb{R}^n$ with coordinates x_k and let $\alpha = (\alpha_1, \ldots, \alpha_n)$, $\alpha_k = 0, 1, 2, \ldots$ and $|\alpha| = \sum_{k=1}^{n} \alpha_k$. Set $\partial_\alpha; = \partial^{|\alpha|}/(\partial x_1^{\alpha_1} \ldots \partial x_n^{\alpha_n})$ and $x^\alpha = x_1^{\alpha_1} \ldots x_n^{\alpha_n}$. The space of smooth functions on $\mathbb{R}^n$ of rapid decrease $\mathcal{S}(\mathbb{R}^n)$ consists of those smooth functions f for which

$$\rho_{\alpha,\beta}(f) := \sup_x |x^\alpha \partial_\beta f(x)| < \infty$$

for all α, β. They fall off together with their derivatives faster than any polynomial at infinity. One can show that this space with the countable family of seminorms $\rho_{\alpha,\beta}$ is a Frechet space.

Definition The topological dual of $\mathcal{S}(\mathbb{R}^n)$, denoted $\mathcal{S}'(\mathbb{R}^n)$, is called the space of **tempered distributions**.

$\square$

J.11 Stone-Weierstrass Theorem

Important results are the real and complex Stone-Weierstrass theorems. We start with the Weierstrass approximation theorem as it is used in the proof of the Stone-Weierstrass Theorem.

J.11.1 Weierstrass Approximation Theorem

Theorem J.11.1 (Weierstrass Approximation Theorem) *Let f be a continuous real function defined on a closed interval $[a, b]$ and let $\epsilon > 0$. Then there is a polynomial p such that $\|f(x) - p(x)\|_\infty = sup_{x\in[a,b]}|f(x) - p(x)| < \epsilon$*

Without loss of generality, we let $a = 0$ and $b = 1$. The n-th Berstein polynomials associated with $f \in \mathcal{C}([0, 1])$ are defined as

$$B_n(x) = \sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k} f\left(\frac{k}{n}\right). \tag{J.228}$$

First we will find a couple of identities, these will be the main tools for proving the theorem.

$$\sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k} = [x + (1-x)]^n = 1 \tag{J.229}$$

Differentiating (J.229) with respect to x gives

$$\sum_{k=0}^{n} \binom{n}{k} x^{k-1} (1-x)^{n-k-1} [k(1-x) - (n-k)x] = 0.$$

and then multiplying by $x(1-x)$ gives us

$$\sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k} (k - nx) = 0.$$

We now differentiate this with respect to x, considering $x^k(1-x)^{n-k}$ as one of the two factors in applying the product rule,

$$\sum_{k=0}^{n}\left(\begin{array}{c}n\\k\end{array}\right)x^{k-1}(1-x)^{n-k-1}(k-nx)^2-\sum_{k=0}^{n}\left(\begin{array}{c}n\\k\end{array}\right)x^k(1-x)^{n-k}n=0.$$

We then multiply by $x(1-x)/n^2$ to give us upon simply rearangement

$$\sum_{k=0}^{n}\left(\begin{array}{c}n\\k\end{array}\right)x^k(1-x)^{n-k}\left(x-\frac{k}{n}\right)^2=\frac{x(1-x)}{n}. \tag{J.230}$$

By using (J.229), we see that

$$f(x)-B_n(x)=\sum_{k=0}^{n}\left(\begin{array}{c}n\\k\end{array}\right)x^k(1-x)^{n-k}\left[f(x)-f\left(\frac{k}{n}\right)\right]$$

and hence

$$|f(x)-B_n(x)|\leq\sum_{k=0}^{n}\left(\begin{array}{c}n\\k\end{array}\right)x^k(1-x)^{n-k}\left|f(x)-f\left(\frac{k}{n}\right)\right| \tag{J.231}$$

Since f is uniformly continuous, for any $\epsilon>0$ there is $\delta>0$ such that $|f(x)-f(y)|<\epsilon/2$ when $|x-y|<\delta$ for $0\leq x,y\leq 1$. In particular,

$$|f(x)-f(k/n)|<\epsilon/2 \;\text{ when }\; |x-k/n|<\delta. \tag{J.232}$$

We now spilt the sum on the right of (J.231) into two parts, a sum over those terms for which $|x-k/n|<\delta$ and a sum over the remaining terms:

$$\begin{aligned}
|f(x)-B_n(x)| &\leq \Big(\sum_{|x-k/n|<\delta}+\sum_{|x-k/n|\geq\delta}\Big)\left(\begin{array}{c}n\\k\end{array}\right)x^k(1-x)^{n-k}\left|f(x)-f\left(\frac{k}{n}\right)\right| \\
&\leq \Big(\sum_{|x-k/n|<\delta}\frac{\epsilon}{2}+\sum_{|x-k/n|\geq\delta}\left|f(x)-f\left(\frac{k}{n}\right)\right|\Big)\left(\begin{array}{c}n\\k\end{array}\right)x^k(1-x)^{n-k} \\
&\leq \frac{\epsilon}{2}+\sum_{|x-k/n|\geq\delta}\left(\begin{array}{c}n\\k\end{array}\right)x^k(1-x)^{n-k}\left|f(x)-f\left(\frac{k}{n}\right)\right|
\end{aligned} \tag{J.233}$$

Where we have used (J.232) and that $\sum_{|x-k/n|<\delta} \begin{pmatrix} n \\ k \end{pmatrix} x^k(1-x)^{n-k} \leq 1$ by (J.229) (remember $0 \leq x \leq 1$). We complete the proof by showing the second term can be made less than $\epsilon/2$ independent of x. As f is a continuous function of a closed interval on the real line, it is bounded. Since f is bounded there exists a positive number K such that $|f(x)| \leq K$ for all $x \in [0,1]$ so that $|f(x) - f(k/n)| \leq 2K$. It follows that

$$|f(x) - B_n(x)| \leq \frac{\epsilon}{2} + 2K \sum_{|x-k/n|\geq\delta} \begin{pmatrix} n \\ k \end{pmatrix} x^k(1-x)^{n-k} \tag{J.234}$$

if n is sufficiently large, the second term can be made less than $\epsilon/2$. Because we are summing over k such that $|x - k/n| \geq \delta$, we have

$$\begin{aligned} 2K\delta^2 \sum_{|x-k/n|\geq\delta} \begin{pmatrix} n \\ k \end{pmatrix} x^k(1-x)^{n-k} &\leq 2K \sum_{|x-k/n|\geq\delta} \begin{pmatrix} n \\ k \end{pmatrix} x^k(1-x)^{n-k} \left(x - \frac{k}{n}\right)^2 \\ &\leq 2K\frac{x(1-x)}{n} \end{aligned}$$

where we have used (J.230). The maximum value of $x(1-x)$ on $[0,1]$ is $1/4$, so the previous inequality leads to

$$2K \sum_{|x-k/n|\geq\delta} \begin{pmatrix} n \\ k \end{pmatrix} x^k(1-x)^{n-k} \leq \frac{K}{2\delta^2 n}. \tag{J.235}$$

If we take n to be any integer greater than $K/\delta^2\epsilon$, then from (J.234) we have

$$|f(x) - B_n(x)| < \epsilon$$

for all x in $[0,1]$, which proves the theorem.

□

Second Proof the Weierstrass Approximation Theorem

The set P of all polynomials on $[a,b]$ is identical with the set of functions which can be built 1 and x by applying multiplication, multiplication by a number, and addition. That is, P is a subalgebra of $\mathcal{C}[a,b]$ generated by $\{1, x\}$.

Instructive in the generalisation to the Stone-Weierstrass theorem.

Given an arbitrary function in $C[a,b]$,

$$g(z) = a\frac{z-y}{x-y} + b\frac{z-x}{y-x}. \tag{J.236}$$

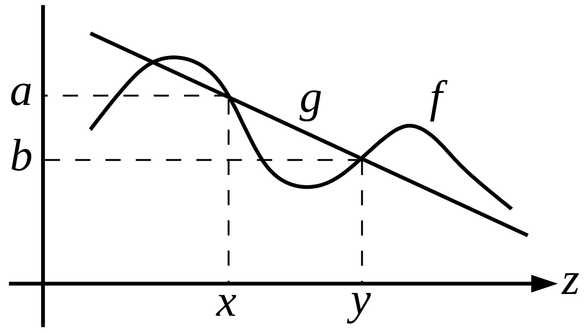

Figure J.14: polyExists.

As it is uniformally continuous, there is $\|f-p\| < \epsilon$

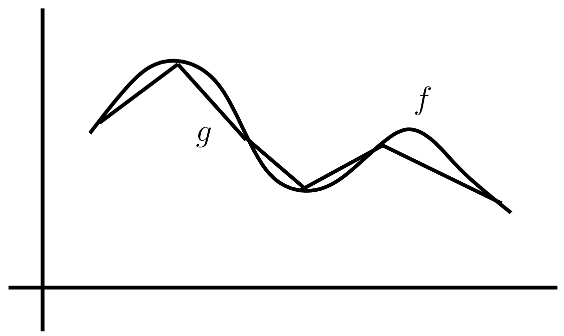

Figure J.15: polygonApp.

□

J.11.2 Stone-Weierstrass Theorem

a generalisation of Theorem J.11.1 to the space of continuous functions on an arbitrary compact Hausdorff space X

Suppose $\mathcal{B}$ is a set of functions on X.

i. $\mathcal{B}$ **separates points** in that given any two different points x and y in X there exists a function g in $\mathcal{B}$ with $g(x) \neq g(y)$.

ii. $\mathcal{B}$ **vanishes at no point** in that for each point $x \in X$ there is a $g \in \mathcal{B}$ such that $g(x) \neq 0$.

Clearly x separates points of $[a,b]$ and the unit function nowhere vanishes on $[a,b]$.

Theorem J.11.2 (Stone-Weierstrass Theorem (real case)) *Suppose $\mathcal{B}$ is an algebra of real functions on a compact Hausdorff space X. If $\mathcal{B}$ separates points on X and if $\mathcal{B}$ vanishes at no point of X, then $\mathcal{B} = C(X,\mathbb{R})$. In other words, if $f \in C(X,\mathbb{R})$ then given $\epsilon > 0$ there is a function $p \in \mathcal{B}$ such that $\|f - p\| < \epsilon$.*

The proof is split up into three lemmas.

$$\begin{aligned} (f \vee g)(x) &= \max\{f(x), g(x)\} \\ (f \wedge g)(x) &= \min\{f(x), g(x)\} \end{aligned} \tag{J.237}$$

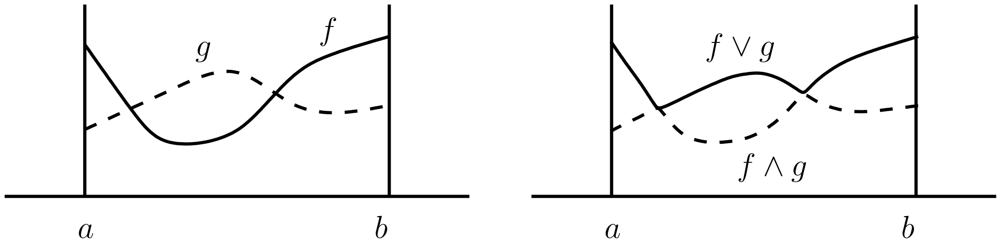

Figure J.16: fgmaxmin. The meaning of $f \vee g$ and $f \wedge g$ for two continuous functions on a closed interval $[a,b]$ of the real line.

$$\begin{aligned} (f \vee g)(x) &= \frac{f + g + |f - g|}{2} \\ (f \wedge g)(x) &= \frac{f + g - |f - g|}{2} \end{aligned} \tag{J.238}$$

Definition (A lattice of functions) If for any $f, g \in \mathcal{B}$, $\max\{f,g\} \in \mathcal{B}$ and $\min\{f,g\} \in \mathcal{B}$ then $\mathcal{B}$ is called a lattice.

Lemma J.11.3 *Let X be an arbitrary topological space. Then every closed subalgebra of $\mathcal{C}(X,\mathbb{R})$ is also a closed sublattice of $\mathcal{C}(X,\mathbb{R})$.*

By (J.238) it suffices to prove that if $f \in \mathcal{B}$ then so is $|f|$. The maximum value of $f(x)$ in X is $\|f\|$. By the Weierstrass approximation theorem there exists a polynomial

$p'(t) = a_o + a_1 t + \cdots + p_n t^n$ with the property that $\big||t| - p'(t)\big| < \epsilon/2$ for every t on the closed interval $[-\|f\|, \|f\|]$. In particular $|a_0| = p'(0) < \epsilon/2$. We define the polynomial $p(t) = p'(t) - a_0$, (so that $p(t) = 0$). This satisfies,

$$\begin{aligned} \big|\, |t| - p(t)\big| &< \frac{\epsilon}{2} + |a_0| \\ &< \epsilon. \end{aligned} \tag{J.239}$$

Now define the function

$$p(f(x)) = a_1 f(x) + \cdots + a_n f^n(x). \tag{J.240}$$

Since $|f(x)| \leq \|f\|$ or all $x \in X$, we can substitute t for $f(x)$ in (J.239) and obtain

$$\big|\, |f(x)| - p(f(x))\big| < \epsilon \quad \text{for all } x \in X. \tag{J.241}$$

The function $p(f) = \sum_{n=1}^{k} a_n f^n$ is in $\mathcal{B}$ since $\mathcal{B}$ is an algebra. Since $\mathcal{B}$ is closed, the fact that $|f|$ can be approximated by the function $p(f)$ shows that $|f| \in \mathcal{B}$.

□

Note, inductively, if $f_1, \ldots, f_n \in \mathcal{B}$ then

$$\begin{aligned} f_1 \vee f_2 \vee \cdots \vee f_n = \max\{f_1, \ldots, f_n\} &\in \mathcal{B} \\ f_1 \wedge f_2 \wedge \cdots \wedge f_n = \max\{f_1, \ldots, f_n\} &\in \mathcal{B}. \end{aligned} \tag{J.242}$$

Proof of Theorem J.11.2

Let f be an arbitrary function in $C(X, \mathbb{R})$. $\epsilon > 0$. We construct a function g such that $f(z) - \epsilon g(z) < f(z) + \epsilon$ for all points z in X $\|f - g\| < \epsilon$. We split this up into three steps.

Step 1. Suppose x and y are distinct points in X and a and b are constants. Since $\mathcal{B}$ separates points there exists a function g in $\mathcal{B}$ such that $g(x) \neq g(y)$. Let f

$$f(z) = a\frac{g(z) - g(y)}{g(x) - g(y)} + b\frac{g(z) - g(x)}{g(y) - g(x)}. \tag{J.243}$$

Then $f \in \mathcal{A}$ and $f(x) = a$ and $f(y) = b$. For example in Weierstrass theorem f would be (J.236).

In the second part we show that given any function $f \in C(X, \mathbb{R})$, a point $x \in X$, and $\epsilon > 0$, there is a function $g_x \in \mathcal{B}$ such that

$$
\begin{aligned}
g_x(x) &= f(x) \\
g_x(t) &> f(t) - \epsilon, \quad \text{for } t \in X.
\end{aligned} \tag{J.244}
$$

Let $y \in X$ be distinct from x. Given $c_1 = f(x)$ and $c_2 = f(y)$ we know there is a function $h_y \in \mathcal{A}$ with the property that

$$
h_y(x) = f(x) \quad \text{and} \quad h_y(y) = f(y). \tag{J.245}
$$

The function $h'_y(t) := h_y(t) - f(t)$ is continuous and $h_y(y) - f(y) = 0$. There is an open interval G_y about y such that $|h_y(t) - f(t)| < \epsilon$ for all $t \in G_y$. This implies

$$
h_y(t) > f(t) - \epsilon, \quad t \in G_y. \tag{J.246}
$$

Since X is compact there are $y_1, \ldots, y_n$ such that $G_{h_1}, \ldots, G_{h_n}$ covers X. Let

$$
g_x(t) = h_{h_1}(t) \vee \cdots \vee h_{h_n}(t). \tag{J.247}
$$

$$
g_z = f_1 \wedge f_2 \wedge \cdots \wedge f_n \tag{J.248}
$$

is a function in L such that $g_x(x) = f(x)$ and $g_x(z) < f(z) + \epsilon$ for all points z in X.

$$
g = g_1 \vee g_2 \vee \cdots \vee g_m \tag{J.249}
$$

g is a function in $\mathcal{A}$ with the property that

$$
f(z) - \epsilon < g(z) < f(z) + \epsilon \tag{J.250}
$$

□

We next turn to the complex case, which is the most important for our purposes. The conditions under which a subalgebra of $C(X, \mathbb{C})$ equals $C(X, \mathbb{C})$.

A complex algebra involves multiplication and addition of complex functions and multiplication by complex numbers.

the conjugate function $\overline{f}(x) = \overline{f(x)}$. if the algebra contains its conjugate

Theorem J.11.4 (Stone-Weierstrass Theorem (complex case)) *Let X be a compact Hausforff space, and let $\mathcal{A}$ be a closed subalgebra of $C(X, \mathbb{C})$ such that:*

(i) $\mathbf{1} \in \mathcal{A}$

(ii) if $f \in \mathcal{A}$, then $\overline{f} \in \mathcal{A}$, that is, closed under complex conjugation;

(iii) if $x, y \in X$ with $x \neq y$, there is $f \in \mathcal{A}$ such that $f(x) \neq f(y)$.

Then $\mathcal{A} = C(X, \mathbb{C})$.

Proof of Theorem J.11.4

Take any complex function $f = u+iv \in C(X, \mathbb{C})$, where u and v are real valued continuous functions. We wish to demonstrate that there exists a function $f_1 = u_1 + iv_1 \in \mathcal{A}$ such that given any $\epsilon > 0$

$$\|f - f_1\| \leq \epsilon.$$

Observe

$$\|f - f_1\| = \|u - u_1 + i(v - v_1)\| \leq \|u - u_1\| + \|v - v_1\|.$$

If we show that there exists $u_1, v_1 \in \mathcal{A}$ such that $\|u - u_1\| \leq \epsilon/2$ and $\|v - v_1\| \leq \epsilon/2$, then we will have proved the theorem. The real functions in $\mathcal{A}$ form a closed subalgebra $\mathcal{A}_{\mathbb{R}}$ of $C(X, \mathbb{C})$. Since $\mathcal{A}$ is an algebra which contains the conjugate of each of its functions, $u = (f + \overline{f})/2$ and $v = (f - \overline{f})/2i$ are in $\mathcal{A}_{\mathbb{R}}$. With these remarks, it should be easy to see that to prove the complex Stone-Weierstrass theorem it is sufficient to show that $\mathcal{A}_{\mathbb{R}}$ satisfies the conditions of the real Stone-Weierstrass theorem. First we show that $\mathcal{A}_{\mathbb{R}}$ separates points. Take $x \neq y \in X$. Since $\mathcal{A}$ separates points, there exists a function $f \in \mathcal{A}$ which has different values at the two points,

$$u(x) + iv(x) = f(x) \neq f(y) = u(y) + iv(y)$$

implying $u(x) \neq u(y)$ or $v(x) \neq v(y)$ or both, so $\mathcal{A}_{\mathbb{R}}$ separates points. Next we prove that $\mathcal{A}_{\mathbb{R}}$ contains a non-zero constant function. Since $\mathcal{A}$ is an algebra which contains the conjugate of each of its functions, $g\overline{g} = |g|^2$ is a non-zero constant in $\mathcal{A}_{\mathbb{R}}$.

□

J.12 Introduction to Measure Theory and Lebesgue Integration

The Riemann integral we split the x coordinate into small intervals and approximating $f(x)$ in every interval by its maximum and minimum. The problem with this approach is that the difference between the maximum and minimum will only tend to zero, as the interval gets smaller, if $f(x)$ is sufficiently well behaved.

J.13 Measure

J.13.1 Null Sets

J.13.2 Outer Measure

J.13.3 Lebesgue-Measure

Definition A set $E \subseteq \mathbb{R}$ is (Lebesgue-) measurable if for every set $A \subseteq \mathbb{R}$ we have

$$m^*(A) = m^*(A \cap E) + m^*(A \cap E^c) \tag{J.251}$$

where $E^c := \mathbb{R}/E$.

$\square$

Suppose

$$E_1 \cap E_2 = \emptyset, \quad E_1, E_2 \in \mathcal{E}.$$

we show that

$$E_1 \cup E_2 \in \mathcal{E}$$

and

$$m^*(E_1 \cup E_2) = m^*(E_1) + m^*(E_2).$$

Proof:

It is countably subadditive

$$m^*(\bigcup_{k=1}^{\infty} E_k) \leq \sum_{k=1}^{\infty} m^*(E_k)$$

From

$$A = (A \cap E) \cup (A \cap E^c)$$

we have

$$m^*(A) \leq m^*(A \cap (E_1 \cup E_2)) + m^*(A \cap (E_1 \cup E_2)^c)$$

we will prove the inequality the other way round.

As $E_1 \in \mathcal{E}$ we have

$$m^*(A) = m^*(A \cap E_1) + m^*(A \cap E_1^c) \tag{J.252}$$

Apply the same condition for E_2 but with A replaced by $A \cap E_1^c$:

$$\begin{aligned} m^*(A \cap E_1^c) &= m^*((A \cap E_1^c) \cap E_2) + m^*((A \cap E_1^c) \cap E_2^c) \\ &= m^*(A \cap (E_1^c \cap E_2)) + m^*(A \cap (E_1^c \cap E_2^c)) \end{aligned}$$

Since E_1 and E_2 are disjoint,

$$E_1^c \cap E_2 = E_2.$$

By de Morgan'law

$$E_1^c \cap E_2^c = (E_1 \cup E_2)^c.$$

Therefore we have

$$m^*(A \cap E_1^c) = m^*(A \cap E_2) + m^*(A \cap (E_1 \cup E_2)^c)$$

Inserting this into ()

$$m^*(A) = m^*(A \cap E_1) + m^*(A \cap E_2) + m^*(A \cap (E_1 \cup E_2)^c) \qquad \text{(J.253)}$$

Now by the subaddativity property of m^*

$$\begin{aligned} m^*(A \cap E_1) + m^*(A \cap E_2) &\geq m^*((A \cap E_1) \cup (A \cap E_2)) \\ &= m^*(A \cap (E_1 \cup E_2)) \end{aligned}$$

$$m^*(A) \geq m^*(A \cap (E_1 \cup E_2)) + m^*(A \cap (E_1 \cup E_2)^c)$$

Hence $E_1 \cup E_2 \in \mathcal{E}$.

Finally put $A = E_1 \cup E_2$ in () to get

$$m^*(E_1 \cup E_2) = m^*(E_1) + m^*(E_2).$$

□

Theorem J.13.1 *(i)* $\mathbb{R} \in \mathcal{E}$,

(ii) if $E \in \mathcal{E}$ *then* $E^c \in \mathcal{E}$,

(iii) if $E_n \in \mathcal{E}$ *for all* $n = 1, 2, \dots$ *then* $\cup_{n=1}^{\infty} E_n \in \mathcal{E}$.

Moreover, if $E_n \in \mathcal{E}$ *for all* $n = 1, 2, \dots$ *and the* E_n*'s are pairwise disjoint then*

$$m^*(\bigcup_{n=1}^{\infty} E_n) = \sum_{n=1}^{\infty} m^*(E_n) \qquad \text{(J.254)}$$

Proof:

(i) Let $A \subseteq \mathbb{R}$. Note that $A \cap \mathbb{R} = A$, $\mathbb{R}^c = \emptyset$, so that $A \cap \mathbb{R}^c = \emptyset$.

$$\begin{aligned} m^*(A) &= m^*(A) + 0 \\ &= m^*(A) + m^*(\emptyset) \\ &= m^*(A \cap \mathbb{R}) + m^*(A \cap \mathbb{R}^c) \end{aligned}$$

(ii) Suppose $E \in \mathcal{E}$

$$
\begin{aligned}
m^*(A) &= m^*(A \cap E) + m^*(A \cap E^c) \\
&= m^*(A \cap E^c) + m^*(A \cap E) \\
&= m^*(A \cap E^c) + m^*(A \cap (E^c)^c)
\end{aligned}
$$

where we have used $(E^c)^c = E$.

□

J.14 Intregration

To circumvent these difficulties, the idea is to portion the range instead of the domain. Fig(J.14).

We considering instead of an interval, the set of x for which $f(x)$ lies between two numbers $a < b$. Now we need the size of the set of these x, that is, the size of the preimage $f^{-1}((a,b))$.

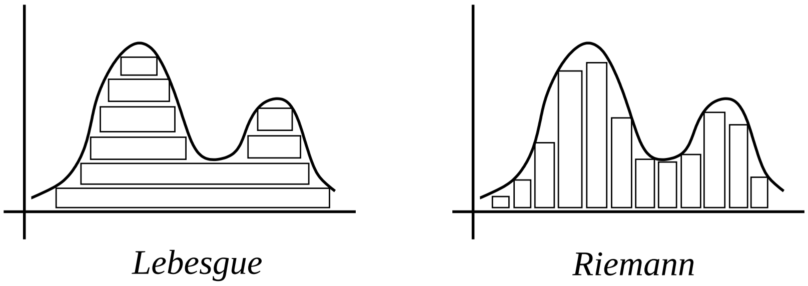

Figure J.17: LebRien.

What are reasonable to take for subsets? There is no reasponable way of adding up an uncountable set of numbers each of which is zero. subsets that satisfy the union of any two subsets is also in the family, the intersection of any two subsets is also in the family, the completment of any subset is also in the family.

Borel Sets

$$\begin{aligned}
[a,b) &= \bigcap_{n=1}^{\infty}(a-\frac{1}{n},b), \\
[a,b] &= \bigcap_{n=1}^{\infty}(a-\frac{1}{n},b+\frac{1}{n}), \\
(a,b] &= \bigcap_{n=1}^{\infty}(a,b+\frac{1}{n}).
\end{aligned} \tag{J.255}$$

Characteristic function of a set $A \subset X$ is a mapping defined by

$$\chi_A(x) = \begin{cases} 1, & \text{if } x \in A \\ 0, & \text{if } x \notin A \end{cases} \tag{J.256}$$

A finite linear combination of characteristic functions of semi-open intervals

$$f = \sum_{k=1}^{N} \alpha_k \mu(I_k) \tag{J.257}$$

is called a *step function*.

The Lebesgue integral of a step function is defined

$$\int \sum_{k=1}^{N} \alpha_k \chi_{I_k} dx = \sum_{k=1}^{N} \alpha_k \mu(I_k) \tag{J.258}$$

Lebesgue integral

$$(1) \quad \sum_{k=1}^{\infty} \int |f_k| dx < \infty, \tag{J.259}$$

$$(2) \quad f(x) = \sum_{k=1}^{\infty} f_k(x) \;\; \text{for all } x \in \mathbb{R} \text{ such that } \;\; \sum_{k=1}^{\infty} |f_k(x)| < \infty. \tag{J.260}$$

The *Lebesgue integral* of f is then defined by

$$\int f dx := \sum_{k=1}^{\infty} \int f_k dx \tag{J.261}$$

In R we need a notion of length. We need to define a function $\mu : \mathcal{P}(R) \to [0, \infty]$ to measure length of as many sets as possible, such that

$$\mu((a,b]) = b - a \text{ and } \mu(A \sqcup B) = \mu(A) + \mu(B), \ \ \mu\left(\sqcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n) \qquad \text{(J.262)}$$

However this cannot be done for all subsets of $\mathcal{P}$.

$$(a,b] = \bigcup_{n=1}^{\infty} \left[a + \frac{(b-a)}{n}, b\right]. \qquad \text{(J.263)}$$

$$(a,b) = \bigcup_{n=1}^{\infty} \left(a, b - \frac{(b-a)}{n}\right]. \qquad \text{(J.264)}$$

To say a property holds almost everywhere with respect to a measure μ means that there is a set $E \in$ with $\mu(E) = 0$ such that the property holds for all $x \in X \backslash E$.

A measure on a set A is an association of nonnegative numbers to subsets of A

We will denote the union of pairwise disjoint subsets $A_1, \ldots, A_n$

$$\bigsqcup_{i=1}^{n} A_i. \qquad \text{(J.265)}$$

Definition Let X be a set. Then a collection of sets $\mathcal{L} \subseteq \mathcal{P}(X)\mathcal{P}$ is a semi-ring if

(i) $\emptyset \in \mathcal{L}$,

(ii) if $A, B \in \mathcal{L}$ then $A \cap B \in \mathcal{L}$,

(iii) if $A, B \in \mathcal{L}$ then there is an $n \in N$ and there are sets $A_1, A_2, \ldots, A_n \in \mathcal{L}$ such that A_i are pairwise disjoint and $A \backslash B = \bigcup_{i=1}^{n} A_i$,

Definition Let X be a set, let $R \subseteq \mathcal{P}(\mathcal{X})$. Then R is a ring of subsets of X if

(i) $\emptyset \in R$;

(ii) if $A, B \in R$ then $A \cap B$, $A \cup B$ and $A \backslash B$ are all in R.

□

Definition Let X be a set and let $\mathcal{F} \subseteq \mathcal{P}(X)$. Then $\mathcal{F}$ is a $\sigma-$**field** of subsets of X if $\mathcal{F}$ satisfies

(i) $\emptyset, X \in \mathcal{F}$,

(ii) for all $A, B \in \mathcal{F}$

(iii) whenever $A_1, A_2, A_3, \cdots \in \mathcal{F}$, then $\cup_{n=1}^{\infty} A_n \in \mathcal{F}$.

Proof:

(i) Each $\sigma-$field contains $\emptyset$ and X. Thus $\emptyset \in$ the intersection and $X \in \cap_{\gamma\in\Gamma}\mathcal{F}_\gamma$.

(ii) for all $A, B \in \mathcal{F}$

(iii) whenever $A_1, A_2, A_3, \cdots \in \cap_{\gamma\in\Gamma}\mathcal{F}$, then $\cup_{n=1}^{\infty} A_n \in \mathcal{F}$.

$$A_1 \cap A_2 = X \backslash (\ (X\backslash A_1) \cup (X\backslash A_2)\ )\,, \tag{J.266}$$

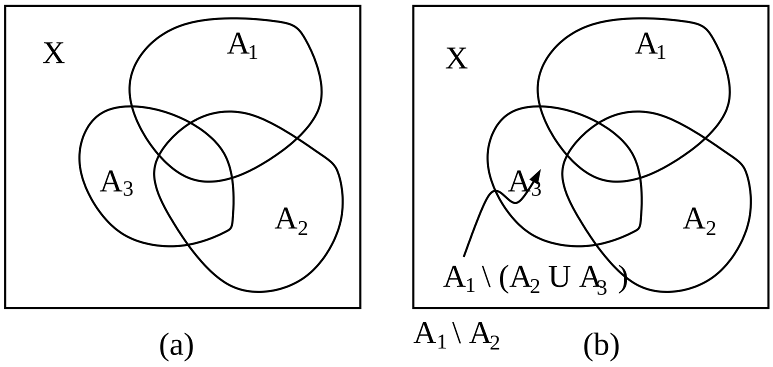

Figure J.18: (a) . (b) $A_1 \cap A_2 = X \backslash (\ (X\backslash A_1) \cup (X\backslash A_2)\ )$

Iterating this we have

$$\bigcap_{k=1}^{n} A_k = X \backslash \left(\bigcup_{k=1}^{n} (X\backslash A_k) \right), \tag{J.267}$$

Taking the limit $k \to \infty$

$$\bigcap_{n=1}^{\infty} A_n = X \backslash \left(\bigcup_{n=1}^{\infty} (X\backslash A_n) \right), \tag{J.268}$$

J.14.1 Measures and Measure Spaces

Definition Let X be a set, let $\mathcal{C} \subseteq \mathcal{P}$ which contains the empty set, $\emptyset \in \mathcal{C}$, and let $\mu : \mathcal{C} \to [0, \infty]$. Then μ is a measure on $\mathcal{C}$ if

(i) $\quad \mu(\emptyset) = 0$,

(ii) if $A_1, A_2, \ldots,$ is a sequence of pairwise disjoint sets in $\mathcal{C}$ such that $\sqcup_{n=1}^{\infty} A_n$ is in $\mathcal{C}$, then

$$\mu\left(\bigsqcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n). \tag{J.269}$$

Definition: A *function* $m : \mathcal{U} \to \mathbb{R}_+ \cup \{\infty\}$ *is called a measure if*
1. $m(A) \geq 0$ *for any* $A \in \mathcal{U}$ *and* $\mu(\phi) = 0$*,*
2. if $(A_i)_{i \geq 1}$ *is a disjoint family of sets in* $\mathcal{U}$ $(A_i \cap A_j = 0$ *for any* $i \neq j)$ *such that* $\bigsqcup_{i=1}^{\infty} \in \mathcal{U}$*, then*

$$\mu\left(\bigsqcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu(A_i). \tag{J.270}$$

This important property is called *countability additivity* or σ*-additivity* of the measure μ.

Definition

A measurable space is a pair $(X, \mathcal{F})$ where X is a set and $\mathcal{F}$ is a $\sigma-$field of subsets of X.

A measure space is a triple $(X, \mathcal{F}, \mu)$ where $\mathcal{F}$ is a $\sigma-$field on X, and

$$\mu : \mathcal{F} \to [0, \infty] \text{ is a measure.} \tag{J.271}$$

Measures of Rings

(i) Monotonicty of μ: If $A, B \in \mathcal{U}$ and $B \subset A$ then $\mu(A) \leq \mu(A)$.

Proof: $A = (A/B) \sqcup B$ implies that

$$\mu(A) = \mu(A \backslash B) + \mu(B). \tag{J.272}$$

Since $\mu(A \backslash B) \geq 0$ it follows that $\mu(A) \geq \mu(B)$.

(ii) Subtractivity of μ. If $A, B \in \mathcal{U}$ and $B \subset A$ and $\mu < \infty$ then $\mu(A \backslash B) = \mu(A) - \mu(B)$.

Proof: In (i) we proved that

$$\mu(A) = \mu(A \backslash B) + \mu(B). \tag{J.273}$$

If $\mu(B) < \infty$ then

$$\mu(A) - \mu(B) = \mu(A \backslash B). \tag{J.274}$$

with counting measure on N.

Set

$$\begin{aligned} A &= \{2, 4, 6, \dots\} \\ B &= \{\text{primes}\} \end{aligned} \tag{J.275}$$

$A \cap B = \{2\},\ \mu(A \cap B) = 1,\ \mu(A) = \mu(B) = \mu(A \backslash B) = \infty$ so that $\mu(A) - \mu(A \backslash B)$ is not defined.

(iii) If $A, B \in \mathcal{U}$ and $\mu(A \cap B) < \infty$ then $\mu(A \cup B) = \mu(A) + \mu(B) - \mu(A \cap B)$.

Proof

$A \cap B \subset A$, $\mu(A \cap B) \subset B$, therefore

$$A \cup B = (A \backslash (A \cap B)) \sqcup B. \tag{J.276}$$

Since $\mu(A \cap B) < \infty$ one has

$$\mu(A \cup B) = (\mu(A) - \mu(A \cap B)) + \mu(B). \tag{J.277}$$

Properties of measures

(i) If $\bigcup_{n=1}^{\infty} A_n \in R$, then

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \lim_{n \to \infty} \mu\left(\bigcup_{k=1}^{n} A_k\right). \tag{J.278}$$

(ii) If $\mu(A_1) < \infty$ $\bigcap_{n=1}^{\infty} A_n \in R$, then

$$\mu\left(\bigcap_{n=1}^{\infty} A_n\right) = \lim_{n \to \infty} \mu\left(\bigcap_{k=1}^{n} A_k\right). \tag{J.279}$$

Proof

(i) Set $A := \cup_{n=1}^{\infty} A_n$. As the sets $A_1, A_2, \dots$ are not pairwise disjoint.

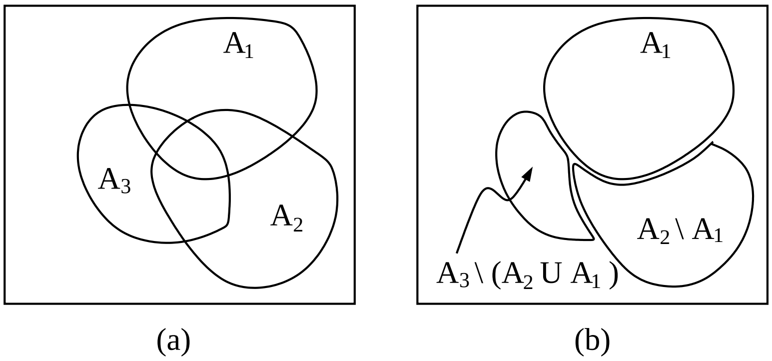

Figure J.19: (a) . (b) $B_1 = A_1$, $B_2 = A_2 \backslash A_1$, $B_3 = A_3 \backslash (A_2 \cup A_1)$. $\cup_{k=1}^3 A_k = \sqcup_{k=1}^3 B_k$

Set $B_1 := A_1$ and $B_n := A_n \backslash \cap_{k=1}^{n-1} A_k$ (for $n > 1$). Then each $B_n \in R$, the sets B_n are pointwise disjoint,

$$\bigcup_{k=1}^{n} A_k = \bigsqcup_{k=1}^{n} B_k \text{ for all } n, \tag{J.280}$$

and $A = \bigsqcup_{k=1}^{\infty} B_k$. Thus

$$\begin{aligned} \mu(A) &= \sum_{n=1}^{\infty} \mu(B_n) \\ &= \lim_{n\to\infty} \left(\sum_{k=1}^{n} \mu(B_k) \right) \\ &= \lim_{n\to\infty} \left(\mu \left(\bigsqcup_{k=1}^{n} B_k \right) \right) \\ &= \lim_{n\to\infty} \left(\mu \left(\bigcup_{k=1}^{n} A_k \right) \right). \end{aligned} \tag{J.281}$$

(ii) Suppose that $\mu(A_1) < \infty$.

set

$$C_n = A_1 \backslash A_n \text{ for all } n. \tag{J.282}$$

Then $C_n \in R$ and

$$\begin{aligned}\bigcap_{n=1}^{\infty} A_n &= A_1 \backslash \bigcup_{n=1}^{\infty} (A_1 \backslash A_n) \\ &= A_1 \backslash \bigcup_{n=1}^{\infty} C_n. \end{aligned} \tag{J.283}$$

If we do not have $A_1 \supseteq A_2 \supseteq A_3 \supseteq \ldots$ we can work with $A_1, A_1 \cap A_2, A_1 \cap A_2 \cap A_3, \ldots$

Nice Phrases

Suppose X is a set equipped with a collection $\mathcal{M}$ of subsets $E \subset X$ such that $\mathcal{M}$ is closed under countable unions and complements; that is, whenever $E \subset \mathcal{M}$, so is $X \backslash E$, and whenever we have a sequence E_j of sets in $\mathcal{M}$, then $\cup_{j=1}^{\infty}$ is in $\mathcal{M}$. Then $\mathcal{M}$ is said to be a sigma-algebra on X.

(The word "sigma" refers to sum, meaning union, while the word "algebra" indicates that $\mathcal{M}$ is defined in terms of certain operations, in this case unions and complements of sets).

(AN ENTROPY PRIMER CHRIS HILLMAN)

σ-algebra

Theorem J.14.1 *If $\mathcal{F}$ is any collection of subsets of X, there exists a smallest σ-algebra $\mathcal{M}^*$ in X such that $\mathcal{F} \in \mathcal{M}^*$.*

2^X is an algebra so that there always is an algebra. We need to prove that the intersection of any two algebras also forms and algebra.

Definition A measure μ defined on the σ-algebra of all Borel sets in a locally compact Hausdorff space X is called a *Borel measure* on X.

□

J.14.2 Measurable functions

Let X be a set, $\mathcal{U}$ a σ-algebra on X.

Definition A pair $(X,\mathcal{U})$ is called measurable space.

Let f be a function defined on a measurable space $(X,\mathcal{U})$, with values i the extended real number system. The function f is called measurable if the set

$$\{x : f(x) > a\} \tag{J.284}$$

is a measurable for every real a.

Step functions (simple functions)

Definition A real valued function $f : X \to R$ is called simple function if it takes only a finite number of distinct values.

$$\chi_E(x) = \begin{cases} 1 & \text{if } x \in E \\ 0 & \text{otherwise} \end{cases} \tag{J.285}$$

Theorem J.14.2 *A simple function $f = \sum_{j=1}^{n} c_j \chi_{E_j}$ is measurable if and only if all the sets E_j are measurable.*

J.15 Main Results of Integration Theory

J.15.1 Dominating Convergence

Theorem J.15.1 *(Monotone Convergence Theorem)*

Let (X, f, μ) be a measure space, let

$$f_n : X \to [0, \infty] \tag{J.286}$$

be a sequence of measurable functions with

$$0 \leq f_1(x) \leq f_2(x) \leq \ldots \text{ for all } x \in X. \tag{J.287}$$

Suppose

$$f(x) = \lim_{n\to\infty} f_n(x) \text{ for all } x \in X. \tag{J.288}$$

Then f is measurable and

$$\int_X f d\mu = \lim_{n\to\infty} \int_X f_n d\mu. \tag{J.289}$$

Lemma J.15.2 (Fatou's Lemma).

Theorem J.15.3 (Dominating Convergence Theorem)

The Lebesgue dominating convergence theorem we can see when the integral of the limit of functions is equal to the limit of integrals,

$$\lim_{n\to\infty} \int f_n = \int \lim_{n\to\infty} f_n \tag{J.290}$$

If

$$f(x) \to \lim_{n\to\infty} \sum_{k=1}^{n} f_k(x), \tag{J.291}$$

then it says we can intergrate term by term, i.e.,

$$\int f(x) d\alpha = \lim_{n\to\infty} \sum_{k=1}^{n} \int f_k(x) d\alpha \tag{J.292}$$

$$|f(x) - f_n(x)| \leq \epsilon_n \tag{J.293}$$

proof:

$$\int (f(x) - \epsilon_n) d\alpha \leq \underline{\int} f_n d\alpha \leq \overline{\int} f d\alpha \leq \int (f(x) + \epsilon_n) d\alpha \tag{J.294}$$

$$0 \leq \overline{\int} f_n d\alpha - \underline{\int} f_n d\alpha \leq 2\epsilon_n [\alpha(b) - \alpha(a)] \tag{J.295}$$

$$|f(x) - f_n(x)| \leq \epsilon_n [\alpha(b) - \alpha(a)] \tag{J.296}$$

J.15.2 Fubini

Theorem J.15.4 *(Fubini)*

Let (X_1, τ_1, μ_1) and (X_2, τ_2, μ_2) be two measure spaces in which μ_1 and μ_2 are $\sigma-$finite.

There is a unique measure μ on $X_1 \times X_2$ with the property that

$$\mu(A \times B) = \mu_1(A) \times \mu_2(B) \tag{J.297}$$

for all measurable rectangles $A \times B$. This measure is called the product measure *and we write $\mu = \mu_1 \times \mu_2$. Fubini's theorem is the most important result for product measures.*

Theorem J.15.5 *If f is an integrable function on $X_1 \times X_2$, then $x \mapsto f(x, y)$ is an integrable function of X for a.e. y, $y \mapsto f(x, y)$ is an integrable function of y for a.e. x, and*

$$\int f \, d(\mu_1 \times \mu_2) = \int \left[\int f \, d\mu_1 \right] d\mu_2 = \int \left[\int f \, d\mu_2 \right] d\mu_1 \tag{J.298}$$

$$\int_X f d\mu = \int_{X_1} \left[\int_{X_2} f(\omega_1, \omega_2) \mu_2(d\omega_2) \right] \mu_1(d\omega_1). \tag{J.299}$$

J.15.3 Radon-Nikodym

a real Borel measure ν and the Lebesgue measure m.

$$A \mapsto \nu(A) = \int_A f \, dm \tag{J.300}$$

defines a Borel measure ν on $(\mathbb{R}, \mathcal{M})$.

Definition Let X be a set and $\mathcal{F}$ be a $\sigma-$field of its subsets. Suppose that ν and μ are measures on $(X, \mathcal{F})$. We say that ν is absolutely continuous with respect to μ if $\mu(A) = 0$ implies $\nu(A) = 0$ for $A \in \mathcal{F}$

how it connects with the notion of continuity of real functions.

Definition If there is a set $E \in \mathcal{F}$ such that $\lambda(F) = \lambda(E \cap F)$ for every $F \in \mathcal{F}$ then λ is concentrated on E.

If two measures μ, ν are concentrated on disjoint subsets of X, we say that they are mutually singular and write $\mu \perp \nu$.

The two main results will follow as simple consequence of the following result:

Theorem J.15.6 *Let μ, ν be $\sigma-$finite measures. Then there exists a unique (a.e.) non-negative function f and a set N of measure zero, such that*

$$\nu(A) = \nu(A \cap N) + \int_A f d\mu. \tag{J.301}$$

Proof. We first assume μ, to be finite measures. Let $\alpha = \mu + \nu$ and consider the Hilbert space $L^2(X, d\alpha)$. Then

$$\ell(h) = \int_X h d\nu \tag{J.302}$$

is a bounded linear functional by Cauchy-Schwarz:

$$\begin{aligned} |\ell(h)|^2 &= \Big| \int_X h d\nu \Big|^2 \leq \Big(\int |1|^2 d\nu \Big) \Big(\int |h|^2 d\nu \Big) \\ &= \nu(X) \Big(\int |h|^2 d\alpha \Big) = \nu(X) \|h\|^2. \end{aligned} \tag{J.303}$$

Hence by the Riesz lemma (Theorem N.-19) there exists an $g \in L^2(X, d\alpha)$ such that

$$\ell(h) = \int_X hg \; d\alpha \tag{J.304}$$

By construction

$$\nu(A) = \int \chi_A \; d\nu = \int \chi_A g \; d\alpha = \int_A g \; d\alpha \tag{J.305}$$

In particular, g must be positive a.e. (to see this take A to be the set where g is negative). Furthermore, let $N = x | g(x) \geq 1$, then

$$\nu(N) = \int_N g \; d\alpha \geq \mu(N) + \nu(N), \tag{J.306}$$

which shows $\mu(N) = 0$. Now set

$$f = \frac{g}{1-g}\chi_{N'}, \tag{J.307}$$

$N' = X \backslash N$. Then, since (N.-19) implies $d\nu = g d\alpha$ respectively $d\mu = (1-g)d\alpha$, we have

$$\begin{aligned} \int_A f d\mu &= \int \chi \frac{g}{1-g}\chi_{N'} d\mu \\ &= \int \chi_{A\cap N'} g \, d\alpha \\ &= \nu(A \cap N') \end{aligned} \tag{J.308}$$

as desired. Clearly f is unique, since if there is a second function $\tilde{f}$, then $\int_A (f - \tilde{f}) d\mu = 0$ for every A shows $f - \tilde{f} = 0$ a.e..

To see the $\sigma-$finite case, observe that $X_n \subseteq X$, $\mu(X_n) < \infty$ and $Y_n \subseteq X, (Y_n) < \infty$ implies $X_n Y_n \subseteq X$ and $\alpha(X_n \, Y_n) < \infty$. Hence when restricted to $X_n \cap Y_n$ we have sets N_n and functions f_n. Now take $N = \cup N_n$ and choose f such that $f|_{X_n} = f_n$ (this is possible since$f_{n+1}|_{X_n} = f_n$ a.e.). Then $(N) = 0$ and

$$\nu(A \cap N') = \lim_{n\to\infty} \nu(A \cap (X_n \backslash N)) \lim_{n\to\infty} \int_{A\cap X_n} f \, d\mu = \int f \, d\mu, \tag{J.309}$$

which finishes the proof.

□

Theorem J.15.7 (Lebesgue decomposition) *Let λ, μ be $\sigma-$finite measures on $(X, \mathcal{F})$. Then λ can be expressed uniquely as a sum of two measures, $\lambda = \lambda_a + \lambda_s$ where $\lambda_a \gg \mu$ and $\lambda_s \perp \mu$.*

□

Theorem J.15.8 (Radon-Nikodym)

$$\nu(A) = \int_A f \, d\mu \tag{J.310}$$

for every $A \in \Sigma$.

□

Definition (Radon-Nikodym derivative) The function f of theorem J.15.8 is called the Radon-Nikodym derivative

$$\frac{d\nu}{d\mu} \tag{J.311}$$

of ν with respect to μ.

Some Definitions and Theorems

Theorem J.15.9 *Let $(X,\mathcal{U},\mu)$ be a measure space. Let $\mathcal{U}'$ be a collection of all $S \subset X$ such that there exists $U, V \in \mathcal{U}$ with $U \subset S \subset V$ and $\mu(V-U) = 0$ (in particular $\mathcal{U} \subset \mathcal{U}'$). Define $\mu'(S) = \mu(U)$ in that case. Then $(X,\mathcal{U}',\mu')$ is a measure space again, called the completion of $(X,\mathcal{U},\mu)$.*

Definition A set Y in X a measure space $(X,\mathcal{U},\mu)$ is called thick or a support for μ provided that for any measurable set $U \in \mathcal{U}$ the condition $U \cap Y = \emptyset$ implies $\mu(U) = 0$, A support for μ will be denoted supp(μ).

□

quantization of background independent theories of connecctions with local degrees of freedom for example of above defition!!!

J.15.4 Riesz Representation

Theorem J.15.10 *(Riesz Representation Theorem)*

i)

Let X be a locally compact Hausdorff space and let $\Lambda : C_0(X) \to \mathbb{C}$ be a positive linear functional one the space of continuous, complex-valued functions of compact support in X. Then there exists a $\sigma-$algebra $\mathcal{U}$ on X which contains the Borel $\sigma-$algebra and a unique positive measure μ on $\mathcal{U}$ such that Λ is represented by μ, that is,

$$\Lambda(f) = \int_X d\mu(x) f(x) \text{ for all } f \in C_0(X). \tag{J.312}$$

Moreover, μ has the following properties:

1) $\mu(K) < \infty$ if $K \subset X$ is compact.

2) For every $S \in \mathcal{U}$ property N.-19 holds.

3) For every open $S \in \mathcal{U}$ with $\mu(S) < \infty$ property N.-19.

4) If $S' \subset S \in \mathcal{U}$ and $\mu(S) = 0$ then $S' \in \mathcal{U}$

ii)

If, in addition to i), X is $\sigma-$compact then μ has the following properties

5) μ is regular

6) For any $S \in \mathcal{U}$ and every $\epsilon > 0$ there exists a closed set C and an open set O such that $C \subset S \subset O$ and $\mu(O - C) < \epsilon$.

Proof:

□

Non-Measurable Sets

J.16 Spectral Decomposition of Infinite-Dimensional Spaces

J.16.1 Spectral Integrals

should we put this in LoopRep???

Recall . In the finite dimensional case the spectrum is $\sigma(T) = \{\lambda_1, \ldots, \lambda_n\}$.

$$\begin{aligned} A &= \lambda_1 P_1 + \lambda_2 P_2 + \cdots \lambda_m P_m \\ &= \lambda_1 (E_{\lambda_1} - E_{\lambda_0}) + \lambda_2 (E_{\lambda_2} - E_{\lambda_1}) + \cdots \lambda_m (E_{\lambda_m} - E_{\lambda_{m-1}}) \\ &= \sum_{i=1}^{m} \lambda_i (E_{\lambda_i} - E_{\lambda_{i-1}}). \end{aligned} \tag{J.313}$$

$$A = \sum_{i=1}^{m} \lambda_i \Delta E_{\lambda_i} \tag{J.314}$$

for each $B \subseteq \sigma(T)$, let

$$P(B) = \sum_{\lambda_i \in B} P_i. \tag{J.315}$$

$$P : \mathcal{P}(\sigma(T)) \to \mathcal{B}(\mathcal{H}) \tag{J.316}$$

a "$\mathcal{H}$-projection valued measure" on $\sigma(T)$.

Definition Let $\mathcal{H}$ be a complex Hilbert space, and suppose $(X, \mathcal{F})$ is a measurable space. function P from $\mathcal{F}$ into orthogonal projections in $B(\mathcal{H})$ so that

(1) $P(X) = I$, and
(2) if $\{E_n\}_{n=1}^{\infty} \subseteq \mathcal{F}$ are pairwise disjoint, then

$$P\left(\bigcup_{n=1}^{\infty} E_n\right) = \sum_{n=1}^{\infty} P(E_n) \tag{J.317}$$

in the strong operator topology.

$$\mu_{\xi,\eta}(E) = (P(E)\xi, \eta) \tag{J.318}$$

Properties:

i) $P(A)P(B) = P(A \cap B)$;

ii) $P(\emptyset) = 0$;

iii) $P(A) \leq P(B)$ for $A \subset B$ where for two projections $P_1 \leq P_2$ means that $P_1\mathcal{H} \subset P_2\mathcal{H}$.

$$\begin{aligned} P(A \cup B) + P(A \cap B) &= P(A) + P(B) \\ P(A \cup B)P(A) + P(A \cap B)P(A) &= P(A) + P(B)P(A) \\ P(A) + P(A \cap B) &= P(A) + P(B)P(A) \end{aligned} \tag{J.319}$$

Just for ordinary measures, the set of $B \in \mathcal{F}$ for which $P(B) = 0$ is a $\sigma-$algebra - we may enlarge $\mathcal{F}$ if necessary

J.16.2 Spectral Measure

we derived from the projection property $E(B)^2 = E(B)$ that $E(B_1)E(B_2) = E(B_1 \bigcap B_2)$, $E(\emptyset) = 0$ and $E(B_1) \leq E(B_2)$ for $B_1 \subset B_2$ where for two projections $P_1 \leq P_2$ means that $P_1\mathcal{H} \subset P_2\mathcal{H}$.

Given a p.v.m P and a unit vector $\Omega \in \mathcal{H}$ we define the spectral measure

$$\mu_\Omega(B) := <\Omega, P(B)\Omega>_{\mathcal{H}} := \int_B d\mu_\Omega(x) \tag{J.320}$$

Theorem J.16.1 *Let a be a self-adjoint operator on a Hilbert space. Then there exists a* p.v.m. *E on the measure space $(\mathbb{R}, \mathcal{B}_{Borel})$ such that*

$$a = \int_{\mathbb{R}} x dE(x) \tag{J.321}$$

In order to construct E from a we notice that for each measurable, bounded set B the function χ_B has support in a closed, finite interval containing B, therefore it can be approximated pointwise by polynomials due to the Weierstrass theorem

$$\chi_B(x) \approx \sum x^n$$

On the other hand, the function x can be approximated arbitrarily well by simple functions of the form $f(x) = \sum_k x_k \chi_{B_k}(x)$ where $\cup_n B_k = \mathbb{R}$ is a collection of disjoint intervals and $x_k \in B_k$.

$$\int x d\mu_{\Omega_1, a^{n-1}\Omega_2} = x = \lim \sum_k x_k \chi_{B_k}(x)$$

In order to construct E from a we notice that for each measurable, bounded set B the function χ_B has support in a closed, finite interval containing B, therefore it can be approximated pointwise by polynomial due to Weierstrass theorem. On the other hand, the function x can be approximated arbitrarily well by simple functions of the form $f(x) = \sum_k \chi_{B_k}(x)$ where $\cup_n B_k = \mathbb{R}$ is a collection of disjoint intervals and $x_k \in B_k$. Therefore for Ω_2 in the domain of a^n

$$
\begin{aligned}
<\Omega_1, a^n\Omega_2> &= \int x d\mu_{\Omega_1, a^{n-1}\Omega_2}(x) \\
&= \lim \sum_{k_1} x_{k_1} \;\; <\Omega_1, E(B_{k_1}) a^{n-1}\Omega_2> \\
&= \lim \sum_{k_1} x_{k_1} \;\; <E(B_{k_1})\Omega_1, a^{n-1}\Omega_2> \\
&= \lim \sum_{k_1} x_{k_1} \int x \; d\mu_{E(B_{k_1})\Omega_1, a^{n-2}\Omega_2}(x) \\
&= \lim \sum_{k_1,k_1} x_{k_1} x_{k_2} \;\; <E(B_{k_2})E(B_{k_1})\Omega_1, a^{n-2}\Omega_2> \\
&= \cdots \\
&= \lim \sum_{k_1 \ldots k_n} x_{k_1} \ldots x_{k_n} \;\; <E(B_{k_1} \cap \cdots \cap B_{k_n})\Omega_1, \Omega_2> \\
&= \lim \sum_{k} x_k^n <\Omega_1, E(B_{k_1})\Omega_2> \\
&= \int x^n d\mu_{\Omega_1,\Omega_2}(x)
\end{aligned} \tag{J.322}
$$

we conclude that for every measurable set B

$$
<\Omega_1, \chi_B(a)\Omega_2> = \int \chi_B(x) \; d\mu_{\Omega_1,\Omega_2}(x) = <\Omega_1, E(B)\Omega_2> \tag{J.323}
$$

for all Ω_1, Ω_2 since $E(B)$ is a bounded operator. Thus

$$
E(B) = \chi_B(a) \tag{J.324}
$$

are the spectral projections associated with a self-adjoint operator. If we know the representation of a on H then we have to aproximate $\chi_B(a)$ by polynomials and then can construct the $E(B)$. In particular we conclude that for every measurable function and Ω_2 in the domain of $f(a)$

$$
<\Omega_1, f(a)\Omega_2> = \int f(x) \; d\mu_{\Omega_1,\Omega_2}(x) \tag{J.325}
$$

since $f(E) := f(a)$ if $f(E) = \sum_n z_n E(B_n)$. Formula (J.325) is sometimes referred to as the functional calculus. Combining (??) ans (??) we have

$$
\chi_{(-\infty,\lambda]}(a) = \theta(\lambda - a) = E((-\infty, \lambda]) = \int_{-\infty}^{\lambda} dE(x) = E(\lambda) - E(\infty) = E(\lambda) \tag{J.326}
$$

where the inegration constant $E(-\infty) = 0$ follows from the fact that $E((-\infty, -\infty)) = E(\emptyset) = 0$ by defintion.

We remark that the spcetral theorem holds without making any separability assumptions, that is, it holds also when $\mathcal{H}$ does not have a countable basis.

J.17 Measure Theory on Compact Spaces

There is a one-to-one correspondence between Baire measures and regular Borel measures.

Theorem J.17.1 *(the Riesz-Markov theorem) Let X be a compact Hausdorff space. For any positive linear functional ℓ on $C(X)$ there is a unique Baire measure μ on X with*

$$\ell(f) = \int f d\mu.$$

Proof:

Since μ is inner regular, that is,

$$\mu(Y) = \sup\{\mu(C) : C \subset Y, C \text{ compact}\},$$

we need only find $\mu(C)$ for compact C to recover μ.

It is claimed that

$$\mu(C) = \inf\{\ell_\mu(f) : f \in C(X), f \geq \chi_C\}.$$

Since μ is positive, it is clear that

$$\mu(C) \leq \ell_\mu(f) \qquad \text{with} \qquad f \geq \chi_C;$$

thus we need only show that, given ϵ, we can find

$$f \in C(X) \quad \text{with} \quad \chi_C \leq f \quad \text{and} \quad \ell_\mu(f) \leq \mu(C) + \epsilon.$$

Since μ is outer regular, given ϵ we can find O open with $\mu(O/C) < \epsilon$ and $C \subset O$. By Urysohn's lemma, we can find $f \in C(X)$ with $0 \leq f \leq 1$, $f(x) = 1$ if $x \in C$ and $f(x) = 0$ if $x \in X/O$. Thus

$$\ell(f) \leq \mu(O) < \mu(C) + \epsilon.$$

$\square$

J.17.1 Spectral Theorem for Bounded Self-Adjoint Operators

Definition A vector $\psi \in \mathcal{L}(\mathcal{H})$ is called a **cyclic vector** for A if finite linear combinations of the elemnets $\{A^n\psi\}_{n=1}^{\infty}$ are dense in $\mathcal{H}$.

Theorem J.17.2 *(spectrum theorem - multiplication operator form) Let A be a bounded self-adjoint operator on $\mathcal{H}$, a separable Hilbert space. Then, there exist measures $\{\mu_n\}_{n=1}^{N}$ $(N = 1, 2, \dots)$ or $\infty)$ on $\sigma(A)$ and a unitary operator*

$$U : \mathcal{H} \to \oplus_{n=1}^{N} L_2(\mathbb{R}, d\mu_n) \tag{J.327}$$

so that

$$(UAU^{-1}\psi)_n(\lambda) = \lambda\psi_n(\lambda) \tag{J.328}$$

Proof:

$\square$

J.18 Extensions of Unbounded Operartors

Can they be defined for unbounded operators?

it would be useful if some of the results that have been proved for bounded operators would hold for unbounded ones. many do apparently

J.18.1 Semi-bounded Operators

A symmetric transformation S is said to be *semi-bounded* if there exists a real quantity c such that

$$(S\psi, \psi) \geq c(\psi, \psi) \tag{J.329}$$

for all ψ in $\mathcal{D}[S]$; it is said to be *upper semi-bounded* if the opposite inequality is valid. If in particular,

$$(S\psi, \psi) \geq 0, \tag{J.330}$$

we shall say, following the definition set down for bounded transformations, that S is *positive.*

For a positive self-adjoint transformation A, the spectral decomposition can be deduced from that for a bounded self-adjoint transformation. This is done with the aid of a linear transformation of the semi-axis $\lambda \geq 0$ into a finite segment of the μ-axis, for example, the transformation

$$\mu = \frac{\lambda - 1}{\lambda + 1}, \tag{J.331}$$

which carries the semi-axis $\lambda \geq 0$ into the segment $-1 \leq \lambda < 1$.

$$B = (A - I)(A + I)^{-1} \tag{J.332}$$

Since C is the inverse of a self-adjoint transformation, it is also self-adjoint, and since it is bounded in its domain

$$I - 2C = (A + I)C2C = (A - I)C = B \tag{J.333}$$

is also self-adjoint and bounded, and since $0 \leq C \leq I$, we have $||B| \leq 1$.

Let

$$B \int_{-1-0}^{1} \mu dF_{\mu} \tag{J.334}$$

be the spectral decomposition of B. Since the transformation $I - B = 2C$ possesses an inverse (namely $(A + I)/2$, the value 1 is not a characteristic value of B, hence F_{μ} is a continuous function of $\mu = 1$, that is, $F_{1-0} = F_1 = I$. Consequently we have

$$A = (I + B)(I - B)^{-1} = \int_{-1-0}^{1} \frac{1 + \mu}{1 - \mu} dF_{\mu} = \in_{-0}^{\infty} \lambda dE_{\lambda}, \tag{J.335}$$

where

$$E_\lambda = F_\mu \text{ for } \frac{1-\lambda}{1+\lambda}; \tag{J.336}$$

J.18.2 Closure

Definition An operator is **closed** whenever a sequence of vectors φ_n in the domain of A converges to a limit vector and a sequence of vectors $A\varphi_n$ converges to a limit vector ϕ, then φ is in the domain of A and $\lim_{n\to\infty} A\varphi_n = \phi$.

□

Note that this property is different from the following property of bounded linear operator. If a linear operator A is bounded and thus continuous, and if $\{x_n\}$ is a sequence in $D(A)$ which converges in $D(A)$, then $\{Ax_n\}$ also converges. This need not hold for a closed operator.

Consider an infinite linear combination $\varphi = \sum_{k=1}^{\infty} x_k\phi_k$ where all the vectors are in the domain of A. Then each partial sum $\varphi_n = \sum_{k=1}^{n} x_k\phi_k$ is in the domain of A and

$$A\varphi_n = \sum_{k=1}^{n} x_k A\phi_k \tag{J.337}$$

The sequence of vectors φ_n converges to a limit vector φ. Suppose the sequence of vectors $A\varphi_n$ converges to a limit vector

$$\phi = \sum_{k=1}^{n} x_k A\phi_k. \tag{J.338}$$

If A is closed, then $\sum_{k=1}^{n} x_k\phi_k$ is in the domain of A and

$$A\sum_{k=1}^{\infty} x_k\phi_k = \sum_{k=1}^{\infty} x_k A\phi_k \tag{J.339}$$

Theorem J.18.1 *Let A be an operator with dense domain. It's adjoint $A^\dagger$ is closed. (Which has the immediate corrollary that every self-adjoint operator is closed).*

Proof: Suppose a sequence of vectors ψ_n in the domain of $A^\dagger$ converges to a limit vector ψ, (i.e. $lim_{n\to\infty}\psi_n \to \psi$), and the sequence of vectors $A^\dagger\psi_n$ converges to a limit vector χ, (i.e. $lim_{n\to\infty}A^\dagger\psi_n \to \chi$). Because of the continuity of the inner product we obviously have

$$(\phi, A^\dagger \psi_n) \to (\phi, \chi) \tag{J.340}$$

and

$$(A\phi, \psi_n) \to (A\phi, \psi). \tag{J.341}$$

where ϕ is taken to be in the domain of A so that $A\phi$ is well defined.

Now, because

$$(\phi, A^\dagger \varphi_n) = (A\phi, \psi_n), \tag{J.342}$$

equations (J.340) and (J.18.2) are equal, so that,

$$(\phi, \chi) = (A\phi, \psi). \tag{J.343}$$

It follows that ψ is in the domain of $A^\dagger$ and $A^\dagger\psi = \chi$. This completes the proof.

□

Definition We say the operator A is an **extension** of the operator B if $D(B) \subset D(A)$ and $A\psi = B\psi$ for $\psi \in D(B)$.

□

Definition Let A and B be operators on $\mathcal{H}$. If $\Gamma(A) \supset \Gamma(B)$, then A is said to be an **extension** of B and we write $A \supset B$. This is equivalent to the earlier definition, $A \supset B$ if and only if $D(A) \supset D(B)$ and $A\varphi = B\varphi$ for all $\varphi \in D(B)$.

□

Definition T is said to be **closable** if it has a closed extension to $(\overline{T})$, that is, $D(T) \subset D(\overline{T})$ and $\overline{T} = T$ in $D(T)$. The smallest closed extension is called the **closure** $\overline{T}$.

If A has a closed extension (i.e., an extension of its domain such that the associated graph is closed and such that the extended operator coincides with the original one on the original domain) it is called closable and the smallest such extension of A is called the closure $\overline{A}$ of A.

□

Definition Given a not neccessarily densely defined operator A with domain $D(A)$ consider the set

$$\Gamma(A) := \{[\psi, A\psi] : \psi \in D(A)\} \subset \mathcal{H} \oplus \mathcal{H},$$

called the **graph** of A.

$$< [\psi_1, \psi_2], [\psi_1', \psi_2'] > := < \psi_1, \psi_1' > + < \psi_2, \psi_2' > \tag{J.344}$$

The operator is closed if its graph is closed in the metric induced by the inner product. The relation to the first definition of a closed operator comes from the norm induced from the inner product (J.344),

$$\|\varphi - \varphi_n, \phi - A\varphi_n\|^2 \;\;=\;\; \|\varphi - \varphi_n\|^2 + \|\phi - A\varphi_n\|^2$$

We find $\delta > 0$ such that $\|\varphi - \varphi_n, \phi - A\varphi_n\|^2 \geq \delta$ if either or both do not converge. The graph $\Gamma(A)$ is closed if and only if

$$\lim_{n\to\infty} \|\varphi - \varphi_n, \phi - A\varphi_n\|^2 = 0$$

for φ in the domain $D(A)$ and at the same time, $A\varphi_n$ converges to the vector ϕ. The original awkard definition of a closed operator is replaced by the simple concept of a closed subspace.

Let us repeat the proof of theorem (J.18.1) in terms of graphs.

Theorem J.18.2 *If a is a densely defined operator on $\mathcal{H}$, then $a^\dagger$ is closed.*

Proof: Define $U : \mathcal{H} \oplus \mathcal{H} \to \mathcal{H} \oplus \mathcal{H}$ by

$$U[\phi, \psi] = [-\psi, \phi].$$

U is unitary, with $U^{-1} = U^\dagger = -U$. Now, let $[\phi, \psi] \in \Gamma(a)^\perp$, the orthogonal complement of the graph $\Gamma(a)$. Then for each $\chi \in D(a)$,

$$< \phi, \chi > + < \phi, a\chi >= \left\langle \left(\begin{array}{c} \phi \\ \psi \end{array} \right), \left(\begin{array}{c} \chi \\ a\chi \end{array} \right) \right\rangle = 0.$$

Hence,

$$< \phi, \chi > = - < \phi, a\chi >$$

for each $\chi \in D(a)$, so, by definition, $\psi \in D(a^\dagger)$ and $a^\dagger\psi = -\phi$. Hence,

$$[-a^\dagger\psi, \psi] \in \Gamma(a)^\perp.$$

But

$$U\Gamma(a^\dagger) = \{[-a^\dagger\psi, \psi] : \psi \in D(a^\dagger)\}$$

so

$$U\Gamma(a^\dagger) = \Gamma(a)^\perp. \tag{J.345}$$

In a Hilbert space, the orthogonal complement of any subspace is closed. In particular, $\Gamma(a)^\perp$ is a closed subspace of $\mathcal{H} \oplus \mathcal{H}$. Unitary operators preserve closedness of subspaces, and

$$\Gamma(a^\dagger) = U^\dagger U\Gamma(a^\dagger) = U^\dagger[\Gamma(a)^\perp],$$

so $\Gamma(a^\dagger)$ is closed. Therefore $a^\dagger$ is a closed linear operator.

□

Theorem J.18.3 *If an operator T is closable, we can form its closure $\overline{T}$ via the following prescription:*

The domain $D(\overline{T})$ is taken to be the set of all $x \in \mathcal{H}$ for which there is a sequence $\{x_n\}$ in $D(T)$ and a $y \in \mathcal{H}$ such that

$$x_n \to x \qquad \textit{and} \qquad Tx_n \to y. \tag{J.346}$$

It is not difficult to see that $D(\overline{T})$ is a vector space. Clearly, $D(T) \subset D(\overline{T})$. On $D(\overline{T})$ we define $\overline{T}$ by setting

$$y = \overline{T}x \tag{J.347}$$

with y given by (J.346).

Proof: For $\overline{T}$ to be well defined there must correspond a unique y for every $x \in D(\overline{T})$. So in addition to $\{x_n\}$ let $\{\tilde{x}_n\}$ be another sequence in $D(T)$ such that

$$\tilde{x}_n \to x \qquad \text{and} \qquad T\tilde{x}_n \to \tilde{y},$$

we must have

$$T(x_n - \tilde{x}_n) \to y - \tilde{y} = 0.$$

Proof that, if it exists, $\overline{T}$ is a closed linear operator. Since T is linear, so is $\overline{T}$ by (J.346) and (J.347). Now consider any sequence $\{w_m\}$ in $D(\overline{T})$ such that

$$w_m \to x \qquad \text{and} \qquad \overline{T}w_m \to y, \tag{J.348}$$

we will prove that $x \in D(\overline{T})$ and $\overline{T}x = y$.

Every w_m (for fixed m) is in $D(\overline{T})$. By definition of $D(\overline{T})$ there is a sequence in $D(T)$ which converges to w_m and whose image under T converges to $\overline{T}w_m$. Hence for every fixed m there is a $v_m \in D(T)$ such that

$$\|w_m - v_m\| < \frac{1}{m} \qquad \text{and} \qquad \|\overline{T}w_m - \overline{T}v_m\| < \frac{1}{m}.$$

From this and (J.348) we conclude that

$$v_m \to x \qquad \text{and} \qquad \overline{T}v_m \to y.$$

This shows that $x \in D(\overline{T})$ and $y = \overline{T}v_m$. Recall that an operator is closed if and only if

$$x_n \to x \quad [x_n \in D(T)] \qquad \text{and} \qquad Tx_n \to y.$$

together imply that $x \in D(T)$ and $Tx = y$. From the defintion, we see that every point of $D(\overline{T})$ must also belong to the domain of every closed linear extension of T. This shows that $\overline{T}$is the closure of T, moreover, it implies that the closure is unique.

□

The concept of a closable operator is helpful to view from the graph perspective.

Since a closed linear operator is determined by its closed graph, each closable linear operator a has a minimal closed extension $\overline{a}$: its graph $\Gamma(\overline{a})$ is in the intersection of the graphs of all closed extensions of a.

Proposition J.18.4 *If T is closable, then* $\Gamma(\overline{T}) = \overline{\Gamma(T)}$.

Proof: Suppose that S is a closed extension of T. Then $\overline{\Gamma(T)} \subseteq \Gamma(S)$ so if

Define an operator R with domain

$$D(R) = \{\psi : [\psi, \phi] \in \overline{\Gamma(T)}\}$$

by

$$R\psi = \phi$$

where $\phi \in \mathcal{H}$ is the unique vector so that $[\psi, \phi] \in \overline{\Gamma(T)}$.

□

The simplest way of extending an operator A is to take the closure of its graph $\Gamma(A) = \{[\psi, A\psi] : \psi \in D(A)\} \subset \mathcal{H} \oplus \mathcal{H}$. That is, if

$$[\psi_n, A\psi_n] \to [\psi, \phi]$$

we might try to define

$$A\psi = \phi.$$

For $A\psi$ to be well-defined, we need if

$$\psi_n \to \psi \qquad \text{and} \qquad \tilde{\psi}_n \to \psi$$

then we should have $\lim_{n\to\infty} A\psi_n = \lim_{n\to\infty} A\tilde{\psi}_n$, that is,

$$[\psi_n - \tilde{\psi}_n, A(\psi_n - \tilde{\psi}_n)] \to [0, 0].$$

Stateted otherwise, for $A\psi$ to be well-defined, we need that

$$[\psi_n, A\psi_n] \to [0, \phi] \qquad \text{implies} \qquad \phi = 0.$$

In this case A is called closable and the unique operator $\overline{A}$ which satisfies $\Gamma(\overline{A}) = \overline{\Gamma(A)}$ is called the closure of A. Clearly, A is called closed if $\overline{A} = A$, which is the case if and only if the graph of A is closed.

Give a first easily proven simple criterion fo an operator to be closable.

Theorem J.18.5 *If a is a densely defined linear operator on a Hilbert space $\mathcal{H}$, then a is closable if and only if $D(a^\dagger)$ is dense in $\mathcal{H}$. In such a case, $\overline{a} = a^{\dagger\dagger}$*

Proof: In a Hilbert space the closure of a subspace is equal to the orthogonal complement of its orthogonal complement:

$$\overline{M} = M^{\perp\perp}.$$

Since $\Gamma(a)$ is a subspace of $\mathcal{H} \oplus \mathcal{H}$,

$$\begin{aligned}
\overline{\Gamma(a)} &= \Gamma(a)^{\perp\perp} \\
&= (U^2\Gamma(a)^{\perp})^{\perp} \\
&= (U(U\Gamma(a))^{\perp})^{\perp} \quad (U[M^{\perp}] = (U[M])^{\perp} \text{ for any subspace } M) \\
&= (U\Gamma(a^\dagger))^{\perp}
\end{aligned}$$

Now, by the proof of (J.345) in theroem J.18.2, repeated here for the densely defined operator $a^\dagger$, we have $U\Gamma(a^{\dagger\dagger}) = \Gamma(a^\dagger)^{\perp}$. Equivalently, $\Gamma(a^{\dagger\dagger}) = U(\Gamma(a^\dagger)^{\perp}) = (U\Gamma(a^\dagger))^{\perp}$. Thus, if $a^\dagger$ is densely defined, then $\overline{\Gamma(a)}$ is the graph of $a^{\dagger\dagger}$. a is closable with closure $\overline{a} = a^{\dagger\dagger}$.

Now suppose that $D(a^\dagger)$ is not dense in $\mathcal{H}$, then there is a non-zero $\psi \in D(a^\dagger)^{\perp}$ (see corollary J.4.9). Let $\phi \in D(a^\dagger)$,and note that

$$\left\langle \begin{pmatrix} \psi \\ 0 \end{pmatrix}, \begin{pmatrix} \phi \\ a^\dagger\phi \end{pmatrix} \right\rangle = <\psi, \phi> = 0,$$

so $[\psi, 0] \in \Gamma(a^\dagger)^{\perp}$. Thus

$$\begin{pmatrix} 0 \\ \psi \end{pmatrix} = U \begin{pmatrix} \psi \\ 0 \end{pmatrix} \in U[\Gamma(a^\dagger)^{\perp}].$$

is not the graph of a linear operator and so a is not closable.

□

Proposition J.18.6 *If a is closable, then $(\overline{a})^\dagger = a^\dagger$.*

Proof: Notice that if a is closable,

$$a^\dagger = \overline{(a^\dagger)} = a^{\dagger\dagger\dagger} = (a^{\dagger\dagger})^\dagger = (\overline{a})^\dagger.$$

□

Extending a symmetric operator

Every symmetric operator is closable.

Theorem J.18.7 *Let $T : D(T) \to \mathcal{H}$ be a linear operator, where $D(T)$ is dense in $\mathcal{H}$. Then if T is symmetric, its closure $\overline{T}$ exists and is unique. $\overline{T}$ is a symmetric linear extension of T.*

Proof. The domain $D(\overline{T})$ is taken to be the set of all $x \in \mathcal{H}$ for which there is a sequence $\{x_n\}$ in $D(T)$ and a $y \in \mathcal{H}$ such that $x_n \to x$ and $Tx_n \to y$. We define $\overline{T}$ by setting $y = \overline{T}x$. We must prove there corresponds a unique y for every $x \in D(\overline{T})$. In addition to $\{x_n\}$ let $\{\tilde{x}_n\}$ be another sequence in $D(T)$ such that

$$\tilde{x}_n \to x \qquad \text{and} \qquad T\tilde{x}_n \to \tilde{y}.$$

Since T is symmetric, we have for every $v \in D(T)$

$$(v, Tx_n - T\tilde{x}_n) = (Tv, x_n - \tilde{x}_n).$$

We let $n \to \infty$. By continuity of the inner product, we have

$$(v, y - \tilde{y}) = (Tv, x - x) = 0,$$

that is,

$$y - \tilde{y} \perp D(T).$$

Since $D(T)$ is dense in $\mathcal{H}$, we have $D(T)^\perp = \{0\}$ by lemma ??, and $y - \tilde{y} = 0$.

Proof that $\overline{T}$ is a symmetric linear extension of T. Since T is linear, so is $\overline{T}$ by (J.346) and (J.347). We show that the symmetry of $\overline{T}$ follows from the symmetry of T. By (J.346) and (J.347), for all $x, z \in D(\overline{T})$ there are sequences $\{x_n\}$ and $\{z_n\}$ in $D(T)$ such that

$$
\begin{aligned}
x_n &\to x, & \qquad Tx_n &\to y \\
z_n &\to z, & \qquad Tz_n &\to \overline{T}z.
\end{aligned}
$$

Since T is symmetric,

$$(z_n, Tx_n) = (Tz_n, x_n).$$

Letting $n \to \infty$, we obtain

$$(z, \overline{T}x) = (\overline{T}z, x)$$

by the continuity of the inner product. Since $x, z \in D(\overline{T})$ were arbitrary, this shows that $\overline{T}$ is symmetric.

□

The extension of a symmetric operator T can fail to be self-adjoint only if the domain of $A^\dagger$ is larger than that of T.

Definition If a symmetric operator cannot be extended to a symmetric operator on a larger domain, we call it a *maximal symmetric* operator.

□

Theorem J.18.8 *Every self-adjoint operator is a maximal symmetric operator.*

Proof. Let a be a self-adjoint operator and let b be a symmetric operator that is an extension of a. We proof the theorem by showing that b must be equal to a.

□

If A is a graph, then its null space $\mathrm{Ker}(A)$ consists of all $\psi \in \mathcal{H}$ such that $(\psi, 0)$ is in A.

If A is self-adjoint, then there is only one self-adjoint extension (if B is another one, we have $\overline{A} \subseteq B$ and hence $\overline{A} = B$ by theorem J.18.8. In this case A is called **essentially self-adjoint** and $D(A)$ is called a **core** for A. Otherwise there might be more than one self-adjoint extension or none at all.

Theorem J.18.9 *Let a be a symmetric operator on a Hilbert space $\mathcal{H}$. Then the following are equivalent:*

1) a is self-adjoint;

2) a is closed and Ker $(a^\dagger \pm i1_{\mathcal{H}}) = \{0\}$;

3) $[a \pm i1_{\mathcal{H}}]D(a) = \mathcal{H}$.

Proof. First we prove that 1) implies 2). Suppose that a is a self-adjoint operator and there is a $\varphi \in D(a^\dagger) = D(a)$ so that $a^\dagger\varphi = i\varphi$. Then $a\varphi = i\varphi$ and

$$-i(\varphi, \varphi) = (i\varphi, \varphi) = (a\varphi, \varphi) = (\varphi, a^\dagger\varphi) = i(\varphi, \varphi)$$

so $\varphi = 0$, that is, Ker $(a^\dagger - i1_{\mathcal{H}}) = \{0\}$. A similar proof shows that Ker $(a^\dagger + i1_{\mathcal{H}}) = \{0\}$.

We now prove how 2) implies 3). If $\psi \in$ Ran $(a - i)^\perp$

$$((a - i)\varphi, \psi) = 0$$

for all $\varphi \in D(a)$. Implying

$$(\varphi, (a^\dagger + i)\psi) = 0.$$

Recall the domain of $a^\dagger$ is the set of all vectors ψ for which there is a vector $a^\dagger\psi$ such that $(\varphi, a^\dagger\psi) = (a\varphi, \psi)$ for every vector φ in the domain of a. So $\psi \in D(a^\dagger)$ as $(a\varphi, \psi) = -i(\varphi, \psi) = (\varphi, a^\dagger\psi)$ for all $\varphi \in D(a)$. And $(a - i)^\dagger\psi = (a^\dagger + i)\psi = 0$ which is impossible since $a^\dagger\psi = -i\psi$ has no solutions.

Since Ran $(a-i)$ is dense, we need only prove it is closed to conclude that Ran $(a-i) = \mathcal{H}$. But for all $\varphi \in D(a)$

$$\begin{aligned}
\|(a - i)\varphi\|^2 &= ((a - i)\varphi, (a - i)\varphi) \\
&= \|a\varphi\|^2 - i(a\varphi, \varphi) + i(\varphi, a\varphi) + \|\varphi\|^2 \\
&= \|a\varphi\|^2 + \|\varphi\|^2.
\end{aligned}$$

Thus if $\varphi_n \in D(a)$ and $(a - i)\varphi_n \to \psi_0$,

$$\infty > \|\psi_0\|^2 = \lim_{n\to\infty} \|a\varphi_n\|^2 + \lim_{n\to\infty} \|\varphi_n\|^2$$

we conclude that φ_n convergences to some vector φ_0, and $a\varphi_n$ converges too. Since a is closed, $\varphi_0 \in D(a)$ and $(a-i)\varphi_0 = \psi_0$. Thus $\mathrm{Ran}(a-i)$ is closed, so $\mathrm{Ran}(a-i) = \mathcal{H}$. Similarly, $\mathrm{Ran}(a+i) = \mathcal{H}$.

Finally, we will show that 3) implies 1). Let $\varphi \in D(a^\dagger)$ then $(a^\dagger - i)\varphi \in \mathcal{H}$. Since $\mathrm{Ran}(a-i) = \mathcal{H}$, there is $\eta \in D(a)$ so that $(a-i)\eta = (a^\dagger - i)\varphi$. As $D(a) \subset D(a^\dagger)$, $\varphi - \eta \in D(a^\dagger)$ (because $\|a^\dagger(\varphi - \eta)\| \leq \|a^\dagger\varphi\|\ \|a^\dagger\eta\| = \|a^\dagger\varphi\|\ \|a\eta\|$) and

$$(a^\dagger - i)(\varphi - \eta) = (a - a^\dagger)\eta = 0.$$

Since $\mathrm{Ran}(a+i) = \mathcal{H}$, by lemma J.7.27, $\mathrm{Ker}(a^\dagger - i) = \{0\}$, so $\varphi = \eta \in D(a)$. This proves that $D(a^\dagger) = D(a)$, so that a is self-adjoint.

□

Computing $a^\dagger$ might be difficult, criterion 3) for doess not involve $a^\dagger$ and could be useful.

Lemma J.18.10 *If a is symmetric then*

$$Ran(a + i1_{\mathcal{H}})^\perp = Ker(a^\dagger - i1_{\mathcal{H}}), \qquad Ran(a - i1_{\mathcal{H}})^\perp = Ker(a^\dagger + i1_{\mathcal{H}})$$

Proof. In the proof of the previous theorem we established that if $\mathrm{Ker}(a^\dagger \pm i1_{\mathcal{H}}) = \{0\}$ then $\mathrm{Ran}(a \pm i1_{\mathcal{H}})$ is dense in $\mathcal{H}$.

□

Theorem J.18.11 *Let a be a symmetric operator on a Hilbert space $\mathcal{H}$. Then the following are equivalent:*

1) a is essentially self-adjoint

2) $Ker(a^\dagger \pm i1_{\mathcal{H}}) = \{0\}$;

3) $\overline{[a \pm i1_{\mathcal{H}}]D(a)} = \mathcal{H}$.

Proof. The equivalence of 2) and 3) follow from the lemma: 3) implies 2),

$$\begin{aligned} \overline{\mathrm{Ran}(a + i1_{\mathcal{H}})} &= \mathrm{Ran}(a + i1_{\mathcal{H}})^{\perp\perp} \\ &= \mathrm{Ker}(a^\dagger - i1_{\mathcal{H}})^\perp \\ &= \{0\}^\perp = \mathcal{H}. \end{aligned}$$

2) implies 3),

$$\begin{aligned}\mathrm{Ker}(a^{\dagger} - i1_{\mathcal{H}}) &= \mathrm{Ran}(a + i1_{\mathcal{H}})^{\perp} \\ &= (\mathrm{Ran}(a + i1_{\mathcal{H}})^{\perp})^{\perp\perp} \\ &= (\mathrm{Ran}(a + i1_{\mathcal{H}})^{\perp\perp})^{\perp} \\ &= \mathcal{H}^{\perp} = \{0\}.\end{aligned}$$

1) implies 2). If a is essentially self-adjoint then $\overline{a} = a^{\dagger}$ is self-adjoint.

Suppose that a is an essentially self-adjoint operator and there is a $\varphi \in D(a^{\dagger}) = D(\overline{a})$ so that $a^{\dagger}\varphi = i\varphi$. Then $\overline{a}\varphi = i\varphi$ and

$$-i(\varphi, \varphi) = (i\varphi, \varphi) = (\overline{a}\varphi, \varphi) = (\varphi, a^{\dagger}\varphi) = i(\varphi, \varphi)$$

so $\varphi = 0$, that is, Ker $(a^{\dagger} - i1_{\mathcal{H}}) = \{0\}$.

Hence $\mathrm{Ker}(a^{\dagger} \pm i1_{\mathcal{H}}) = \{0\}$.

A similar proof shows that Ker $(a^{\dagger} + i1_{\mathcal{H}}) = \{0\}$.

We prove that 2) implies 1).

....

$$\begin{aligned}\|(a - i)\psi\|^2 &= \|a\psi\|^2 + \|\psi\|^2 \\ &\leq \|\psi\|^2\end{aligned}$$

therefore $\mathrm{Ker}(a - i) = \{0\}$ and hence $(a - i)^{-1}$ exists. Moreover, setting $\psi = (a - i)^{-1}\varphi$ shows $\|(a-i)^{-1}\psi\| \leq |i|$. Hence $(a-i)^{-1}$ is bounded and closed. Since it is densely defined by assumption, its domain of dependence $\mathrm{Ran}(a+i)$ must be equal to $\mathcal{H}$. Replacing $(a-i)$ by $(a + i)$, we then apply theorem J.18.9.

□

J.18.3 Properties of Symmetric Operators

Theorem J.18.12 *Suppose an operator T has an inverse T^{-1} and that $D(T)$ and $D(T^{-1})$ are dense in $\mathcal{H}$, so that $T^{\dagger}$ and $(T^{-1})^{\dagger}$ exist. Then $T^{\dagger}$ also has an inverse, and*

$$(T^{\dagger})^{-1} = (T^{-1})^{\dagger} \tag{J.349}$$

Proof. We assume first that f transverses $D(T)$, and g transverses $D((T^{-1})^{\dagger})$; then

$$(f, g) = (T^{-1}Tf, g) = (Tf, (T^{-1})^{\dagger}g).$$

This shows that

$$(T^{-1})^{\dagger}g \in D(T^{\dagger}),$$

and

$$T^{\dagger}(T^{-1})^{\dagger}g = g. \tag{J.350}$$

On the other hand, if f transverses $D(T^{-1})$, and h transverses $D(T^{\dagger})$, then

$$(f, h) = (TT^{-1}f, h) = (T^{-1}f, T^{\dagger}h).$$

$$(T^{-1})^{\dagger}T^{\dagger}h = h. \tag{J.351}$$

The relations (J.350) and (J.351) show that

□

Theorem J.18.13 *If a self-adjoint operator A has an inverse A^{-1}, then A^{-1} is a self-adjoint operator.*

Proof. By () Ker(A) of an operator is

$$\mathrm{Ker}(A) = \mathcal{H} \ominus \overline{\mathrm{Ran}(A)}$$

but as the inverse exists, Ker$= 0$. Therefore $\overline{\mathrm{Ran}(A)} = \mathcal{H}$, i.e.,

$$\overline{D(A^{-1})} = \mathcal{H}.$$

The result to be proved then follows from theorem J.18.12.

□

J.18.4 Deficiency Indices and Self-adjoint Closed Extensions

For symmetric operators, there are always closed extensions. A smallest closed extension always exists (the double adjoint), but it is possible that none of these closed extensions is self-adjoint.

Definition The positive and negative deficiency indices of an operator a are defined respectively as

$$\begin{aligned} n_+(a) &= \dim(\mathrm{Ker}([a^\dagger + i1_{\mathcal{H}}])) \\ n_-(a) &= \dim(\mathrm{Ker}([a^\dagger - i1_{\mathcal{H}}])) \end{aligned}$$

□

Note a is essentially self-adjoint if and only if $n_+(a) = n_-(a) = 0$.

In general, symmetric operators may or may not have self-adjoint extension. One can show that this is possible if and only if the deficiency indices are equal to each other.

essentially self-adjoint operators their self-adjoint extension is unique and given by their closure.

Note if an operator a is closable then $\overline{a}^\dagger = a^\dagger$ so

$$n_+(a) = n_+(\overline{a}) \qquad \text{and} \qquad n_-(a) = n_-(\overline{a})$$

Definition A number λ (real or complex) is called a **point of regular type** for an operator T if there is a $k = k(\lambda) > 0$ such that, for all $f \in D(T)$,

$$\|(T - \lambda \mathbb{I})f\| \geq k\,\|f\|.$$

□

Therefore the eigenvalues of T are not points of regular type for T.

If λ is is a point of regular type for T, then the operator $(T - \lambda \mathbb{I})^{-1}$ exists...

If λ_0 is a point of regular type, then, for $|\lambda - \lambda_0| \leq \delta \leq \frac{1}{2}k(\lambda_0)$ and for any $f \in D(T)$,

$$\|(T - \lambda \mathbb{I})f\| \geq \|(T - \lambda \mathbb{I})f\| - |\lambda - \lambda_0|\,\|f\|. \tag{J.352}$$

This shows that the set of points of regular type is open.

Definition The set of points of regular type is called the **field of regularity** of the operator.

$\square$

If A is a symmetric operator, and if $z = x + iy$ (with $y \neq 0$), then () for any $f \in D(T)$,

$$\|(A - z\mathbb{I})f\| = \|(A - x\mathbb{I})f\| + y^2\|f\| \geq y^2\|f\|,$$

from which we see that each of the upper and the lower halves of the $z-$plane is a connected component of the field of regularity of any symmetric operator.

The field of regularity of any isometric operator V also consists of two connected components, the interior and exterior of the unit circle; because if $|\zeta| < 1$,

$$\|(V - \zeta\mathbb{I})f\| \geq \|Vf\| - |\zeta| \, \|f\| = (1 - |\zeta|)\|f\|,$$

and similarly, if $|\zeta| > 1$,

$$\|(V - \zeta\mathbb{I})f\| \geq |\zeta| - \|Vf\| \, \|f\| = (|\zeta| - 1)\|f\|.$$

Theorem J.18.14 *If Γ is a connected component of the field of regularity of a linear operator T, then the dimension of the subspace $\mathcal{H} \ominus Ran(T - \lambda\mathbb{I})$ is the same for all $\lambda \in \Gamma$.*

Proof: Let P_λ be the operator of orthogonal projection on to the subspace

$$\mathrm{Ker}(T - \lambda\mathbb{I})$$

If we can show that, for any $\lambda_0 \in \Gamma$, there is a $\delta = \delta(\lambda_0) > 0$ such that $|\lambda - \lambda_0| < \delta$ implies

$$\|P_\lambda - P_{\lambda_0}\| < 1, \tag{J.353}$$

then by theorem J.4.16

$$\dim\, N_\lambda = \dim\, N_{\lambda_0}.$$

Any two points λ_1, λ_2 in Γ can be joined by a path which is covered by a finite number of these $\delta-$neighbourhoods. This leads to the result

$$\dim N_{\lambda_1} = \dim N_{\lambda_2}.$$

Let λ_0 be a fixed point of the domain Γ and let

$$\delta = \delta(\lambda_0) \leq \frac{1}{3}k(\lambda_0)$$

As

$$\begin{aligned} k(\lambda_0)\|f\| &= \|(T-\lambda_0)f\| \\ &\leq \|(T-\lambda_0)f\| + |\lambda-\lambda_0|\ \|f\| \end{aligned}$$

if we assume that $|\lambda - \lambda_0| \leq \delta$, then

$$\|(T-\lambda_0)f\| \geq \frac{2}{3}k(\lambda_0)\|f\|.$$

Then for any $h \in N_\lambda$ (so that $(h,(T-\lambda)f) = 0$) with $\|h\| = 1$,

$$\begin{aligned} \|(\mathbb{I} - P_{\lambda_0})h\| &= \sup \frac{|(h,(T-\lambda_0)f)|}{\|(T-\lambda_0\mathbb{I})f\|} \\ &= \sup_{f\in D(T)} \frac{|(h,(T-\lambda)f + (\lambda-\lambda_0)f)|}{\|(T-\lambda_0\mathbb{I})f\|} \\ &= \sup_{f\in D(T)} \frac{|\lambda-\lambda_0|(h,f)}{\|(T-\lambda_0\mathbb{I})f\|} \\ &\leq \frac{1}{2} \end{aligned} \tag{J.354}$$

and similarly for any $h \in N_{\lambda_0}$ with $\|h\| = 1$,

$$\|(\mathbb{I} - P_\lambda)h\| \leq \frac{1}{2} \tag{J.355}$$

By the alternative defintion of the aperture (J.72), it follows from (J.354) and (J.355) that

$$\|P_\lambda - P_{\lambda_0}\| \leq 1/2.$$

□

The deficiency number of a linear subspace M is

$$\dim(\mathcal{H} \ominus M)$$

The deficiency number of the linear subspace

$$M_\lambda = \mathrm{Ran}(a - \lambda\mathbb{I})$$

for points λ belonging to a given connected component of the field of regularity of the operator a. Note that for any operator in Hilbert space

$$(\mathrm{Ran}\ a)^\perp = \mathrm{Ker}\ a^\dagger,$$

therefore the deficiency number of the manifold is

$$M_\lambda = \mathrm{Ker}(a^\dagger - \overline{\lambda}\mathbb{I}).$$

As a consequence of this and theorem J.18.14, a symmetric operator's deficiency numbers are equivalent to

$$\dim(\mathcal{H} \ominus (A - z\mathbb{I})) = \begin{cases} n_+(A), & (Im(z) > 0), \\ n_-(A), & (Im(z) < 0). \end{cases}$$

We will define deficiency numbers of an isometric operator V as:

$$\dim(\mathcal{H} \ominus (V - \zeta\mathbb{I})) = \begin{cases} n_+(V), & (|\zeta| < 1), \\ n_-(V), & (|\zeta| > 1). \end{cases}$$

Lemma J.18.15 *The deficiency numbers of an isometric operator V can be defined by*

$$m = dim(\mathcal{H} \ominus D(V)), \qquad n = dim(\mathcal{H} \ominus Ran(V))$$

Proof: Only the first equality needs proof. For any $\zeta \neq 0$,

$$
\begin{aligned}
\mathrm{Ran}(V-\zeta\mathbb{I}) &= (V-\zeta\mathbb{I})D(V) \\
&= \left(\frac{1}{\zeta}V-\mathbb{I}\right)D(V) \\
&= \left(\frac{1}{\zeta}\mathbb{I}-V^{-1}\right)VD(V) \\
&= \left(V^{-1}-\frac{1}{\zeta}\mathbb{I}\right)D(V^{-1}) \\
&= \mathrm{Ran}\left(V^{-1}-\frac{1}{\zeta}\mathbb{I}\right)
\end{aligned}
$$

and therefore, for $|\zeta| > 1$,

$$
\begin{aligned}
n_+(V) &= \dim(\mathcal{H}\ominus \mathrm{Ran}(V-\zeta\mathbb{I})) \\
&= \dim(\mathcal{H}\ominus \mathrm{Ran}\left(V^{-1}-\frac{1}{\zeta}\mathbb{I}\right)) \\
&= \dim(\mathcal{H}\ominus \mathrm{Ran}(V^{-1}-0\mathbb{I}) \\
&= \dim(\mathcal{H}\ominus \mathrm{Ran}(V^{-1})) \\
&= \dim(\mathcal{H}\ominus D(V)).
\end{aligned}
$$

□

Cayley transform

We know that every symmetric operator admits closure, so in what follows when discussing symmetric operators we shall assume them to be closed.

We have already introdued the Cayley transform V of a closed, symmetric operator A. The operator V is expressed in terms of the operator A by the formula

$$Vf = (A-z\mathbb{I})(A-\overline{z}\mathbb{I})^{-1}f$$

where the domain $D(V)$ is $\mathrm{Ran}(A-\overline{z}\mathbb{I})$.

In this section we assume $\mathrm{Im}(z) < 0$

Lemma J.18.16 *The deficiency index (m,n) of the operator A is the same as the deficiency index of the operator V.*

Proof: For, by definition,

$$m = \dim(\mathcal{H} \ominus \mathrm{Ran}(A - \overline{z}\mathbb{I})).$$

But $\mathrm{Ran}(A - \overline{z}\mathbb{I}) = D(V)$ and so

$$m = \dim(\mathcal{H} \ominus D(V))$$

Again, by definition,

$$n = \dim(\mathcal{H} \ominus \mathrm{Ran}(A - z\mathbb{I}))$$

and $\mathrm{Ran}(A - z\mathbb{I}) = \mathrm{Ran}(V)$ so

$$n = \dim(\mathcal{H} \ominus \mathrm{Ran}(V)).$$

□

Theorem J.18.17 *If V is an isometric operator, and if Ran$(V - \mathbb{I})$ is dense in $\mathcal{H}$, then A defined by*

$$Ah = (\overline{z}V - z\mathbb{I})(V - \mathbb{I})^{-1}h,$$

where $h \in D(A)$, is symmetric and V is its Cayley transform.

Proof: First we need to prove that since $\mathrm{Ran}(V - \mathbb{I})$ is dense in $\mathcal{H}$, the operator $(V - \mathbb{I})^{-1}$ exists. The operator $V - \mathbb{I}$ has an inverse if and only if

$$(V - \mathbb{I})g = 0$$

implies $g = 0$. Now for any $f \in D(V)$,

$$(Vf - f, g) = (Vf, g) - (f, g) = (Vf, Vg) - (f, g) = 0,$$

i.e., $g \perp \mathrm{Ran}(V - \mathbb{I})$. Since $\mathrm{Ran}(V - \mathbb{I})$ is dense, $g = 0$.

Since the operator $(V - \mathbb{I})^{-1}$ exists, the operator

$$A = (\overline{z}V - z\mathbb{I})(V - \mathbb{I})^{-1}$$

is well defined and its domain, $D(A) = \mathrm{Ran}(V - \mathbb{I})$, is dense in $\mathcal{H}$.

Let f, g be any elemnts of $D(A) = \mathrm{Ran}(V - \mathbb{I})$, then f, g are of the form

$$\begin{aligned} f &= V\phi - \phi \\ g &= V\psi - \psi \end{aligned}$$

for some $\phi, \psi \in D(V - \mathbb{I}) = D(V)$. Then

$$\begin{aligned} Af &= (\overline{z}V - z\mathbb{I})\phi = \overline{z}V\phi - z\phi \\ Ag &= (\overline{z}V - z\mathbb{I})\psi = \overline{z}V\psi - z\psi \end{aligned}$$

and so

$$\begin{aligned} (Af, g) &= (\overline{z}V\phi - z\phi, V\psi - \psi) \\ &= (z + \overline{z})(\phi, \psi) - \overline{z}(V\phi, \psi) - z(\phi, V\psi), \end{aligned}$$

$$\begin{aligned} (f, Ag) &= (V\phi - \phi, \overline{z}V\psi - z\psi) \\ &= (z + \overline{z})(\phi, \psi) - \overline{z}(V\phi, \psi) - z(\phi, V\psi). \end{aligned}$$

Hence

$$(Af, g) = (f, Ag).$$

It is easily shown that

$$V = \frac{A - z\mathbb{I}}{A - \overline{z}\mathbb{I}}$$

so V is the Cayley transform of A.

$\square$

Theorem J.18.18 *If A_1, A_2 are symmetric operators and if V_1, V_2 are their Cayley transforms, then A_2 is an extension of A_1 if and only if V_2 is an extension of V_1.*

Proof: Suppose A_2 is an extension of A_1. Obviously

$$V_1\varphi = \frac{A_1 - z\mathbb{I}}{A_1 - \overline{z}\mathbb{I}}\varphi = \frac{A_2 - z\mathbb{I}}{A_2 - \overline{z}\mathbb{I}}\varphi = V_2\varphi$$

for all $\varphi \in D(V_1)$. Now $D(A_1) \subset D(A_2)$ implies $\mathrm{Ran}(V_1 - \mathbb{I}) \subset \mathrm{Ran}(V_2 - \mathbb{I})$ which in turn implies $D(V_1 - \mathbb{I}) \subset D(V_2 - \mathbb{I})$ or $D(V_1) \subset D(V_2)$.

Suppose V_2 is an extension of V_1. Obviously

$$A_1 f = (\overline{z}V_1 - z\mathbb{I})(V_1 - \mathbb{I})^{-1} f = (\overline{z}V_2 - z\mathbb{I})(V_2 - \mathbb{I})^{-1} f = A_2 f$$

for all $f \in D(A_1)$. Now $D(V_1) \subset D(V_2)$ implies $D(V_1 - \mathbb{I}) \subset D(V_2 - \mathbb{I})$ which, as $(V_2 - \mathbb{I})$ is an injection, implies $\mathrm{Ran}(V_1 - \mathbb{I}) \subset \mathrm{Ran}(V_2 - \mathbb{I})$ which means $D(A_1) \subset D(A_2)$.

$\square$

This reduces the problem of symmetric extensions of a given operator A to the problem of isometric extensions of its Cayley transform V. This problem is much simplier than the original problem.

Two closed linear subspaces can be the domain and range of an isometric operator if and only if their dimensions are equal.

In the subspaces

$$\mathcal{H} \ominus D(V) \qquad \text{and} \qquad \mathcal{H} \ominus \mathrm{Ran}(V)$$

we choose two subspaces F and G of equal dimension, and construct an arbitrary isometric operator V_1 having F as its domain and G its range. We then define a linaer operator $\tilde{V}$, with

$$D(\tilde{V}) = D(V) \oplus F$$

as its domain and

$$\mathrm{Ran}(\tilde{V}) = \mathrm{Ran}(V) \oplus G$$

as its range, by

$$\tilde{V} f = \begin{cases} Vf & \text{for } f \in D(V) \\ V_1 f & \text{for } f \in F \end{cases}$$

$\tilde{V}$ is evidently an isometric extension of V_1, and by taking all possible choices of F, G, V_1, we obtain all possible isometric extensions $\tilde{V}$of the operator V and each extension once only.

To find some symmetric extension $\tilde{A}$ of an operator A, we have to take the Cayley transform V of A, find an isometric extension $\tilde{V}$ of V by the above procedure, and finally carry out a Cayley transform on $\tilde{V}$ to get $\tilde{A}$.

Recall an operator V from the subspace $D(V) \subset \mathcal{H}$ and mapping $D(V)$ to the subspace $\mathrm{Ran}(V) \subset \mathcal{H}$ is said to be isometric if, for all $f, g \in D(V)$,

$$(Vf, Vg) = (f, g).$$

A unitary operator is a particular case of an isometric operator where $D(V)$ and $\mathrm{Ran}(V)$ coincide. From the above considerations it follows that an operator A is maximal, symmetric (or self-adjoint) operator if and only if its Cayley transform V is a maximal, isometric (or unitary) operator.

Recall that the deficiency number of a not necessarily closed symmetric operator A is equal to the deficiency number of its closure $\overline{A}$.

Theorem J.18.19 *The closure $\overline{A}$ of a symmetric operator A is maximal if and only if one of its deficiency numbers is zero.*

The closure of a symmetric operator is self-adjoint if and only if both its deficiency numbers are zero, that is, a symmetric operator is essentially self-adjoint if and only if both its deficiency numbers are zero

Proof: An isometric operator is obviously maximal if and only if dimension of the deficiency subspace is zero.

□

Theorem J.18.20 *Any symmetric operator A can be extended into a maximal operator.*

If $n_+ \neq n_-$, then none of these extensions is self-adjoint;

if $n_+ = n_- < \infty$, then any maximal extension of A is self-adjoint;

if $n_+ = n_- = \infty$, then some of the maximal extensions of A are self-adjoint, and some are not.

Proof:

□

J.18.5 The Friedrichs Extension

Operators bounded below often occur in problems of mathematical physics. The importance of this class of operators is the remarkable porperty that each such operator has a self-adjoint extension (even if it is not essentially self-adjoint).

The Friedrichs extension is a self-adjoint extension of a non-negative densely defined symmetric operator.

characterising prooperty is it is the largest, self-adjoint extension with respect to the usual partial ordering of positive operators.

An operator T is said to be non-negative if and only if

$$(\xi, T\xi) \geq 0, \tag{J.356}$$

for all elements ξ in the domain of T, and a symmetric operator if

$$(T\xi, \eta) = (\xi, T\eta) \tag{J.357}$$

for all elements ξ and η in the domain of T.

Theorem J.18.21 *Every semi-bounded symmetric operator S can be extended to a semi-bounded self-adjoint operator A in such a way that A has the same bound as S.*

Proof: Let S be a symmetric operator such that

$$\|f\|_S^2 := (Sf, f) \geq \|f\|^2$$

for all $f \in D_S$. We introduce on D_S a new scalar product

If S is non-negative, then

$$q_S(\xi, \eta) = (\xi, S\eta) \tag{J.358}$$

is a sesquilinear for on $D(S)$, i.e.,

$$\begin{aligned}
q_S(\xi + \zeta, \eta) &= q_S(\xi, \eta) + q_S(\zeta, \eta) \\
q_S(\alpha\xi, \eta) &= \alpha\, q_S(\xi, \eta) \\
q_S(\xi, \eta + \delta) &= q_S(\xi, \eta) + q_S(\xi, \delta) \\
q_S(\xi, \alpha\eta) &= \overline{\alpha}\, q_S(\xi, \eta).
\end{aligned} \tag{J.359}$$

Also

$$q_S(\xi, \xi) = (\xi, S\xi) \ \geq \|\xi\|^2 \tag{J.360}$$

Thus q_S defines an inner product on the domain of S. Let $\mathcal{H}_0$ be the completion of the domain of A with respect to q_S. $\mathcal{H}_0$ is an abstractly defined space; its elements can be represented as equivalence classes Cauchy sequences of elements of $D(S)$. It is not obvious that all elements in $\mathcal{H}_0$ can be identified with elements of $\mathcal{H}$. However, it turns out that the ideal elements can be identified of the space $\mathcal{H}$:

$$D_S \subseteq \mathcal{H}_0 \subseteq \mathcal{H}.$$

In fact, it follows from the inequality (J.358) that if a sequence of elements of D_S is a Cauchy sequence in the new norm, it is also a Cauchy sequence in in the original norm, and that two Cauchy sequences equivalent in the new norm are also equivalent in the original norm, so that they converge in the original metric to a well-determined element of the space $\mathcal{H}$. We can thus asign to each element $\tilde{g}$ ideal or not a definite element g of $\mathcal{H}$. Obviously it will be such that $\tilde{g} = g$ for the elements $\tilde{g} \in D_S$.

We must prove that two different elements of $\mathcal{H}_0$ can not be represented by the same element g of $\mathcal{H}$. Since the equation

$$q_S(f, \tilde{g}) = (Sf, g),$$

where g denotes the element of $\mathcal{H}$ which we have just identified with the element $\tilde{g}$ of $\mathcal{H}_0$, is valid by defintion if f and $\tilde{g}$ both belong to $D(S)$, extends to the case where $\tilde{g}$ is an arbitrary element of $\mathcal{H}_0$ by continuity. Now, let $\tilde{g}_1, \tilde{g}_2 \in \mathcal{H}_0$, then if they correspond to the same element $g \in \mathcal{H}$, then

$$q_S(f, \tilde{g}_1 - \tilde{g}_2) = (Sf, g - g) = 0$$

for $f \in D(S)$. But $D(S)$ is dense in $\mathcal{H}_0$ (in the new norm), hence $\tilde{g}_1 - \tilde{g}_2 = 0$.

We regard $\mathcal{H}_0$ as a subspace of $\mathcal{H}$.

Inequality (J.358) carries over by continuity to all elements $f \in \mathcal{H}_0$.

Let us fix an element an arbitrary element h of $\mathcal{H}$ and consider the functional

$$L_h(f) := (f, h).$$

If f is in $\mathcal{H}_0$, then

$$|L_h(f)| \leq \|f\| \, \|h\| \leq \|f\|_S \, \|h\|$$

On the Hilbert space $\mathcal{H}_0$, $L_h(f)$ is therefore a linear functional whose norm does not exceed $\|h\|$ ($|L_h(f)| \leq \|h\|$). By the Riesz lemma there exists an element g of $\mathcal{H}_0$ such that

$$L_h(f) = q_S(f, g)$$

and whose norm $\|g\|_S$ equals that of the functional, hence

$$\|g\| \leq \|g\|_S = |L_h(f)| \leq \|h\|. \tag{J.361}$$

Moreover, this element g is uniquely determined by the functional, and hence by h, so that if we write

$$g = Bh$$

we define a unique aoperator whose domain is the entire space $\mathcal{H}$ and whose values are contained in $\mathcal{H}_0$. This oprator is obviously linear, and by defintion

$$(f, g) = q_S(f, Bh) \tag{J.362}$$

for $f \in \mathcal{H}_0$, $h \in \mathcal{H}$; by (J.361),

$$\|Bh\| \leq \|h\|.$$

Setting $f = Bh'$ in (J.362), where h' is again an arbitrary element of $\mathcal{H}$, we obtain

$$(Bh', h) = q_S(Bh', Bh) = \overline{q_S(Bh, Bh')} = \overline{(Bh, h')} = (h', Bh),$$

and in particular, for $h = h'$, we have

$$(Bh, h) = q_S(Bh, Bh) \geq (Bh, Bh) \geq 0, \tag{J.363}$$

hence B is a positive and symmetric operator in the space $\mathcal{H}$.

By (J.362) $Bh = 0$ implies that $(f, h) = 0$ for all $f \in \mathcal{H}_0$, and since $\mathcal{H}_0$ is dense in $\mathcal{H}$, this implies $h = 0$. By ?? B has an inverse.

Since B is bounded and symmetric, and therefore self-adjoint, its inverse

$$A = B^{-1}$$

will also be self-adjoint (see theorem J.18.13). A is semi-bounded below by 1; in fact if $g = Bh$ is an element of $D(A)$, then by (J.363)

$$(g, Ag) \geq \|g\|^2.$$

Moreover, by (J.362)

$$(f, Ag) = q_S(f, g)$$

for $f \in \mathcal{H}_0$ and $g \in D(A)$. Furthermore, the domain $D(A)$, which is a linear subset of $\mathcal{H}_0$, is dense in $\mathcal{H}_0$ in the sense of the new norm as well as the original norm. For if we assumed the contrary case there would be an element $f_0 \neq 0$ in $\mathcal{H}_0$ such that

$$(f_0, Ag) = q_S(f_0, g) = 0$$

for all elements $g \in D(A)$. But the $\mathrm{Ran}(A) = \mathcal{H}$, and consequently $f_0 = 0$, in contradiction to the hypothesis that $f_0 \neq 0$.

It remains to show that the self-adjoint operator A is an extension of the given symmetric operator S.

Let f and g be any two arbitrary elemets of $D(S)$. We have on the one hand, by (J.362),

$$(f, Sg) = q_S(f, BSg),$$

and on the other hand, by defintion of the new norm,

$$(f, Sg) = q_S(f, g).$$

Since $D(S)$ is dense in $\mathcal{H}_0$ (in the sense of the new norm), the equation

$$q_S(f, BSg) = q_S(f, g)$$

is possible for all elements f of $D(S)$ only if

$$BSg = g.$$

This shows that g belongs to the domain of A and that

$$Ag = Sg.$$

Hence $A \supseteq S$.

□

Another definition:

Let $D(S)$, $D(S^*)$ denote the domains of S and its adjoint S^*, respectively. The domain of $D_F(S)$ of the Friedrichs extension S_F of S consists of all y in $D(S^*)$ for which there exists a sequence $y_k \in D(S)$ such that

1. $y_k \to y$ in $\mathcal{H}$ as $k \to \infty$,
2. $(S(y_k - y_l), y_k - y_l) \to 0$ as $k, l \to \infty$,

and S_F is the restriction of S^* to $D_F(S)$.

J.18.6 Quadratic Forms

In the construction of operators a in physics one often starts from its matrix elements $Q_a(\psi, \psi')$, which should equal $< \psi, a\psi' >$. However, this is not enough to define an operator in inifite dimensions because given an ortho-normal basis (b_n) we must have

$$\|a\psi\| = \sum_n |Q_a(b_n, \psi)|^2 < \infty$$

in order that $\psi \in D(a)$. Hence it may happen that the quadratic form $Q_a(\psi, \psi')$ exists for ψ, ψ' in a dense subset of $\mathcal{H}$ but on the other hand it could be that $D(a) = \{0\}$.

A bilinear Hermitian form.

Definition A **quadrartic form** is called is a map

$$q : Q(q) \times Q(q) \to \mathbb{C},$$

where $Q(q)$ is a dense linear subset of $\mathcal{H}$ calfunctionled the **form domain**, such that

$$
\begin{aligned}
q(\lambda\psi_1 + \mu\psi_2, \phi) &= \lambda q(\psi_1, \phi) + \mu q(\psi_2, \phi) \\
q(\psi, \lambda\phi_1 + \mu\phi_2) &= \overline{\lambda} q(\psi, \phi_1) + \overline{\mu} q(\psi, \phi_2)
\end{aligned} \tag{J.364}
$$

If

$$
q\{\psi, \phi\} = \overline{q\{\phi, \psi\}} \tag{J.365}
$$

then we say that q is **symmetric**.

□

We use the notation

$$
q(u) \equiv q(u, u). \tag{J.366}
$$

Definition A quadrartic form is called **semibounded** provided that $q(u, u) \geq -c\|u\|^2$ for some $c \geq 0$ and **positive** if $c = 0$.

□

Definition The numerical range $W(a)$ of a quadratic form $a(u, v)$

Definition A quadratic form $q(u, v)$ will be **closed** if for $\{u_n\} \subset D(q)$, $u_n \to u$ in $\mathcal{H}$, and $q(u_n - u_m) \to 0$ as $m, n \to \infty$ imply that

$$
u \in D(q) \ \text{ and } \ q(u_n - u) \to 0 \ \text{ as } \ n \to \infty.
$$

□

We can give a defintion of a closed quadratic form analogous to the defintion for a closed operator, an operator A is closed if and only if its graph is closed which is the same as saying that $D(A)$ is complete under the norm $\|\psi\|_A = \|A\psi\| + \|\psi\|$.

Definition Let q be a semibounded quadratic form, $q(\psi, \psi) \geq -M\|\psi\|$. q is closed if $Q(q)$ is complete under the norm

$$
\|\psi\|_{+1} = \sqrt{q(\psi, \psi) + (M+1)\|\psi\|^2}
$$

□

Notice that $\|\psi\|_{+1}$ comes from the inner product

$$(\psi, \varphi)_{+1} = q(\psi, \varphi) + (M+1)(\psi, \varphi).$$

Proposition J.18.22 *The above two defintions for a closed form are equivalent.*

Proof: Proof of the equivalence. Say for $\varphi_m, \varphi_n \in Q(q)$, $\|\varphi_m - \varphi_n\|_{+1} \to 0$ as $n, m \to \infty$ implies $\|\psi - \psi_n\|_{+1} \to 0$.

$$\|\varphi_m - \varphi_n\|_{+1}^2 = (q(\varphi_m - \varphi_n, \varphi_m - \varphi_n) + M\|\varphi_m - \varphi_n\|^2) + \|\varphi_m - \varphi_n\|^2 \to 0$$

as $n, m \to \infty$ implies $q(\varphi_m - \varphi_n, \varphi_m - \varphi_n) \to 0$ and $\{\varphi_n\}$ is Cauchy in the norm of $\mathcal{H}$. Now

$$\|\varphi - \varphi_n\|_{+1}^2 = (q(\varphi - \varphi_n, \varphi - \varphi_n) + M\|\varphi_m - \varphi_n\|^2) + \|\varphi - \varphi_n\|^2 \to 0$$

implies that if $\varphi_n \to \varphi$ in the norm of $\mathcal{H}$ then $q(\varphi_n - \varphi, \varphi_n - \varphi) \to 0$ and

$$\|0\|_{+1}^2 = (M+1)\|0\|^2 = 0.$$

This implies the inner product $(\cdot, \cdot)_{+1}$ is continuous (and as a concequence $q(\cdot, \cdot)$ is continuous). Also, by the Schwarz inequality

$$|(\phi, 0)_{+1}| \leq \|\phi\|_{+1} \, \|0\|_{+1} = 0.$$

From this we have

$$q(\phi, 0) = (M+1)(\phi, 0) - (\phi, 0)_{+1} = 0.$$

For $\psi_1, \psi_2 \in Q(q)$ we can write

$$q(\lambda\psi_1 + \mu\psi_2, \varphi_n) = \lambda q(\psi_1, \varphi_n) + \mu q(\psi_2, \varphi_n)$$

By the continuity of the quadratic form, in the limit $n \to \infty$ we obatin

$$q(\lambda\psi_1 + \mu\psi_2, \varphi) = \lambda q(\psi_1, \varphi) + \mu q(\psi_2, \varphi).$$

This and the similarly proven result for $q(\varphi, \cdot)$ proves that $\varphi \in Q(q)$.

Now the other way around. Say whenever $\varphi_n \in Q(q)$, $\varphi_n \to \varphi$ in the norm of $\mathcal{H}$ and $q(\varphi_n - \varphi_m, \varphi_n - \varphi_m) \to 0$, as $n, m \to \infty$, then

$$\|\varphi_m - \varphi_n\|_{+1} \to \sqrt{(M+1)}\, \|\varphi_m - \varphi_n\| \to 0$$

as $n, m \to \infty$. Now say that $\varphi \in Q(q)$ and $q(\varphi_n - \varphi, \varphi_n - \varphi) \to 0$, then

$$\|\varphi - \varphi_n\|_{+1} \to \sqrt{(M+1)}\, \|\varphi - \varphi_n\| \to 0,$$

that is, $Q(q)$ is complete in the norm $\| \cdot \|_{+1}$.

□

Definition A bilinear form $b(u, v)$ is called an *extension* of a bilinear form $a(u, v)$ if $D(b) \supseteq D(a)$ and $b(u, v) = a(u, v)$ for $u, v \in D(a)$.

□

Definition A set U will be called dense in $D(a)$ if for each $w \in D(a)$ and each $\epsilon > 0$ there is a $u \in U$ such that

$$a(w - u) < \epsilon \;\text{ and }\; \|w - u\| < \epsilon.$$

□

A symmetric operator can fail to be self-adjoint only if $A^\dagger$ is an extension of A to a larger domain.

Note on the distinction between symmetric operators and symmetric forms:

For symmetric operators, there are always closed extensions. A smallest closed extension always exists (the double adjoint), but it is possible that none of these closed extensions is self-adjoint. On the other hand, semibounded forms need not have any closed extensions (defintions for quadratic forms to be given below), but when such extensions exsit and are semibounded, they are the quadratic forms associated with self-adjoint operators.

Theorem J.18.23 *Let T be a semi-bounded quadratic form. Then T may not be closable, but if it is and the closure is semi-bounded, then T is the quadratic form of a unique self-adjoint operator T according to*

$$(\psi, T\phi) = (T\psi, \phi)??$$

Proof: When the quadratic form is real we have $T\{\psi, \psi\}$ is real for every vector ψ in the domain of T. In particular $T\{\psi+\phi, \psi+\phi\}$ and $T\{\psi+i\phi, \psi+i\phi\}$ are real. Therefore

$$(\psi+\phi, T(\psi+\phi)) = (T(\psi+\phi), \psi+\phi) \tag{J.367}$$

$$(\psi+i\phi, T(\psi+i\phi)) = (T(\psi+i\phi), \psi+i\phi) \tag{J.368}$$

Breaking up the inner product, we find that

$$\begin{aligned}(\psi, T\phi) + (\phi, T\psi) &= (T\psi, \phi) + (T\phi, \psi)\\ (\psi, T\phi) - (\phi, T\psi) &= (T\psi, \phi) - (T\phi, \psi)\end{aligned} \tag{J.369}$$

and adding we have

$$(\psi, T\phi) = (T\psi, \phi) \tag{J.370}$$

for arbitrary vectors ψ and ϕ in the domain of a_T, therefore T is self-adjoint.

Theorem J.18.24 *Let T be a positive, symmetric operator. Then the corresponding positive quadratic form a_T has a positive closure $\overline{a_T}$. The uniqe positive operator $\overline{T}$ corresponding to that closure via $a_{\overline{T}} = \overline{a_T}$ is called the* Friedrichs extension *of T. It may extend the closure $\overline{T}$ of T and is the only self-adjoint extension which contains $D(\overline{a_T})$.*

In an arbitrary orthonormal basis, we have

$$\langle x|H|x\rangle = x_i H^i_j x^j \equiv \sum_{i,j} H^i_j \bar{x}^i x^j \tag{J.371}$$

where H^i_j are elements of a Hermitian matrix $\mathbf{H}$. The last expression is called a Hemitian quadratic form.

Theorem J.18.25 *The numerical range $W(a)$ of a bilinear form is convex. That is, if $u, v \in D(a)$ satisfying $\|u\| = \|v\| = 1$, then for each $0 < \tau < 1$, we can find a $w \in D(a)$ such that $\|w\| = 1$, and*

$$a(w) = (1-\tau)a(u) + \tau a(v). \tag{J.372}$$

Proof: see details

J.19 Semianalytic Theory

J.19.1 Introduction

We introduce semianalytic functions, semianalytic manifolds,and semianalytic geometry. Used in proof of uniqueness theorem for the Kinematic Hilbert space of LQG.

Analytic difeomorphisms are non-local.

In particular, the group of anaytic diffeomorphisms has local degrees of freedom, in the sense that for every point $x \in \mathcal{M}$ and its neighbourhood $\mathcal{U}'$, there is a semianalytic diffeomorphism of $\mathbb{R}^n$ which moves x, but restricted to $\mathbb{R}^n/\mathcal{U}'$ is the identity map.

J.19.2 Semianalytic Functions in $\mathbb{R}^n$

Our non-analyticity surfaces will have an appropriate analytic structure. We introduce the idea of a semianalytic partition.

Let $\{h_1, \ldots, h_N\}$ be a set of analytic functions defined on a domain containing an open set $\mathcal{U}$. Consider the collection of all possible sequences of equalities/inequalities

$$\begin{aligned}
&h_1(x) > 0, \quad h_2(x) > 0, \ldots, h_N(x) > 0;\\
&h_1(x) < 0, \quad h_2(x) > 0, \ldots, h_N(x) > 0;\\
\\
&h_1(x) > 0, \quad h_2(x) = 0, \ldots, h_N(x) < 0;\\
\\
&h_1(x) = 0, \quad h_2(x) = 0, \ldots, h_N(x) = 0.
\end{aligned} \tag{J.373}$$

A particular member of this collection is identified by the $\{>, =, <\}$ associated to each h_I. Formally, there is a map

$$\sigma : h = \{h_1, \ldots, h_N\} \to \{>, =, <\} \tag{J.374}$$

and we can write

$$\sigma_I := \sigma(h_I), \tag{J.375}$$

where the integer I runs from 1 to N. A particular member of (J.373) can then be written

$$\begin{array}{ccc} h_1(x) & \sigma_1 & 0, \\ & \ldots & \\ h_N(x) & \sigma_N & 0. \end{array} \tag{J.376}$$

The set of conditions (J.376) determines the following subset of $\mathcal{U}$.

$$\mathcal{U}_{h,\sigma} = \{x \in \mathcal{U} : (J.376)\}. \tag{J.377}$$

Definition Given a finite set h of real valued analytic functions defined on a neighbourhood of an open subset $\mathcal{U}$ of $\mathbb{R}^n$, the corresponding semianalytic partition of $\mathcal{U}$ is the set of all the subsets $\mathcal{U}_{h,\sigma} \subset \mathcal{U}$ defined by (J.376 , J.377) such that σ is an arbitrary map (J.374). Given $\mathcal{U}$ and h as above, the partition will be denoted by $\mathcal{P}(\mathcal{U}, h)$.

Obviously, every semianalytic partition covers $\mathcal{U}$.

$$\mathcal{U} = \bigcup_{\sigma} \mathcal{U}_{h,\sigma} \tag{J.378}$$

where σ runs through all the maps (J.374). Also,

$$\sigma \neq \sigma' \quad \Rightarrow \quad \mathcal{U}_{h,\sigma} \cap \mathcal{U}_{h,\sigma'} = \emptyset, \tag{J.379}$$

and a set $\mathcal{U}_{h,\sigma}$ may be empty itself. Another obvious property is that given a semianalytic covering $\mathcal{P}(\mathcal{U}, h)$ and an open suset $\mathcal{V} \subset \mathcal{U}$, the family of functions h defines a semianalytic covering $\mathcal{P}(\mathcal{V}, h)$.

Example

The following subsets are empty sets:

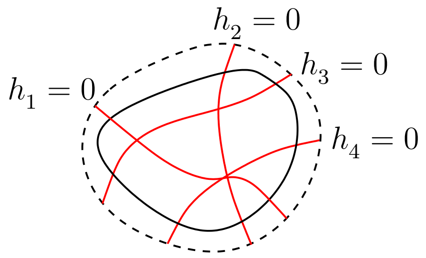

Figure J.20:

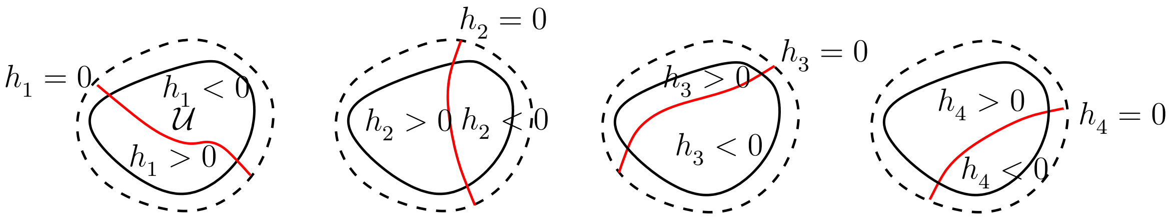

Figure J.21:

$$
\begin{aligned}
\mathcal{U}_{h_1=0,\ h_2=0,\ h_3=0,\ h_4=0} &= \emptyset \\
\mathcal{U}_{h_1>0,\ h_2=0,\ h_3=0,\ h_4=0} &= \emptyset \\
\mathcal{U}_{h_1<0,\ h_2=0,\ h_3=0,\ h_4=0} &= \emptyset \\
\mathcal{U}_{h_1=0,\ h_2=0,\ h_3=0,\ h_4>0} &= \emptyset \\
\mathcal{U}_{h_1=0,\ h_2=0,\ h_3=0,\ h_4<0} &= \emptyset \\
\mathcal{U}_{h_1=0,\ h_2=0,\ h_3=0,\ h_4<0} &= \emptyset \\
etc\ldots & .
\end{aligned}
\tag{J.380}
$$

Roughly speaking, a real valued function f defined on an open subset of $\mathcal{U} \subset \mathbb{R}^n$ will be called semianalytic if it is analytic on an open and dense subset of $\mathcal{U}$, and if the non-analyticity surfaces have also an appropriate analytic structure, and if the restrictions of f to the non-analyticity surfaces are again analytic in an appropriate sense.

Definition A function $f : \mathcal{U} \to \mathbb{R}^m$, where $\mathcal{U}$ is an open subset of $\mathbb{R}^n$, is called semianalytic if every $x \in \mathcal{U}$ has an open neighbourhood $\tilde{\mathcal{U}}$ equipped with a semianalytic partition $\mathcal{P}(\tilde{\mathcal{U}}, h)$, such that for every $\tilde{\mathcal{U}}_{h,\sigma} \in \mathcal{P}(\tilde{\mathcal{U}}, h)$ there is an analytic function $f_\sigma : \tilde{\mathcal{U}} \to \mathbb{R}^m$, such that

$$
f\big|_{\tilde{\mathcal{U}}_{h,\sigma}} = f_\sigma\big|_{\tilde{\mathcal{U}}_{h,\sigma}}, \tag{J.381}
$$

that is, such that f_σ coincides with f on $\mathcal{U}_{h,\sigma}$.

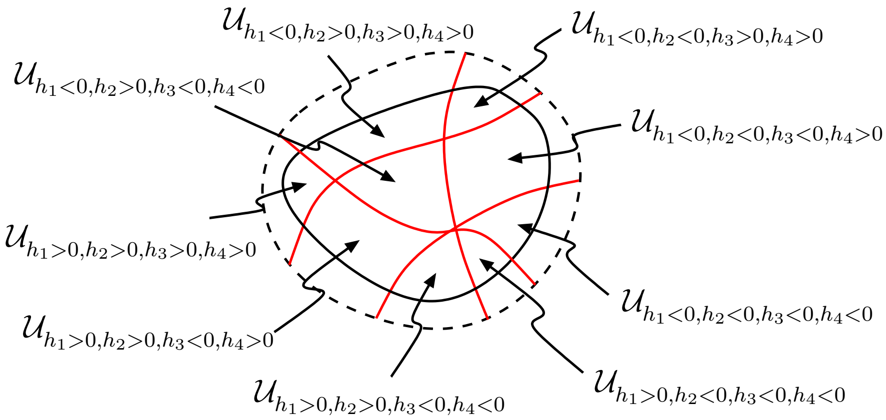

Figure J.22: Subsets, $\mathcal{U}_{h,\sigma}$, of dimension two.

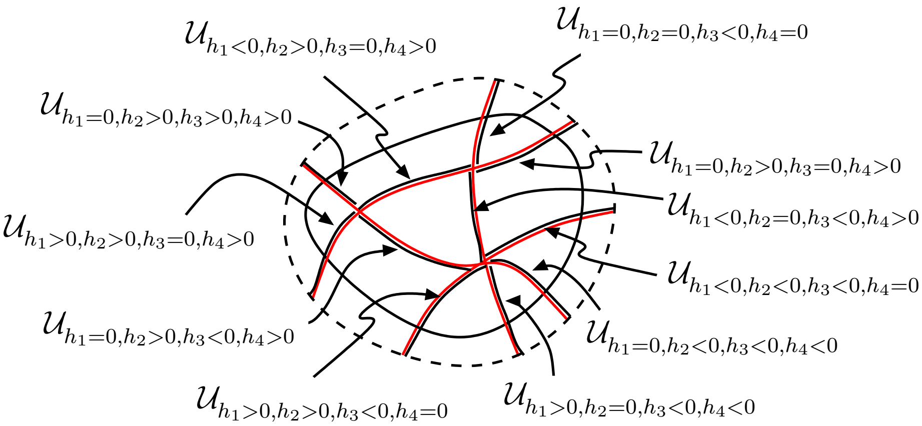

Figure J.23: Subsets, $\mathcal{U}_{h,\sigma}$, of dimension one.

Clearly, if $f : \mathcal{U} \to \mathbb{R}^n$ is semianalytic, and $\mathcal{V} \subset \mathcal{U}$ is open, then the restriction function $f_{|\mathcal{V}}$ is semianalytic.

Simple facts of semianalytic partitions

$$\mathcal{P}(\mathcal{U}, h^{(1)} \cup h^{(2)}) = \mathcal{P}(\mathcal{U}, h^{(1)}) \cap \mathcal{P}(\mathcal{U}, h^{(2)}).$$

Proposition J.19.1 *Let $f_1 : \mathcal{U} \to \mathbb{R}$, and $f_2 : \mathcal{U} \to \mathbb{R}^m$ be two semianalytic functions where $\mathcal{U}$ is an open subset of $\mathbb{R}^n$. Then the functions*

$$\mathcal{U} \ni x \mapsto f_1(x) f_2(x) \in \mathbb{R}^m, \qquad \mathcal{U} \ni x \mapsto (f_1(x), f_2(x)) \in \mathbb{R}^{m+1}, \tag{J.382}$$

are also semianalytic.

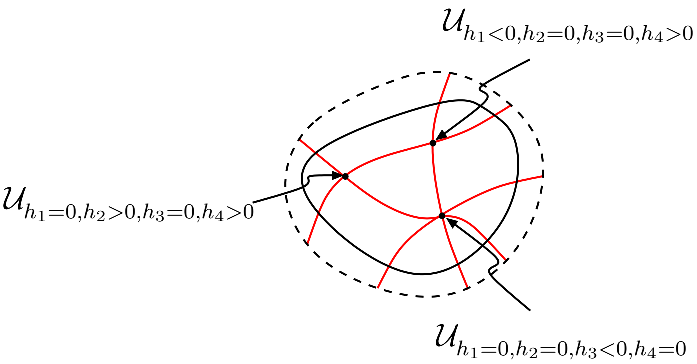

Figure J.24: Subsets, $\mathcal{U}_{h,\sigma}$, of dimension zero.

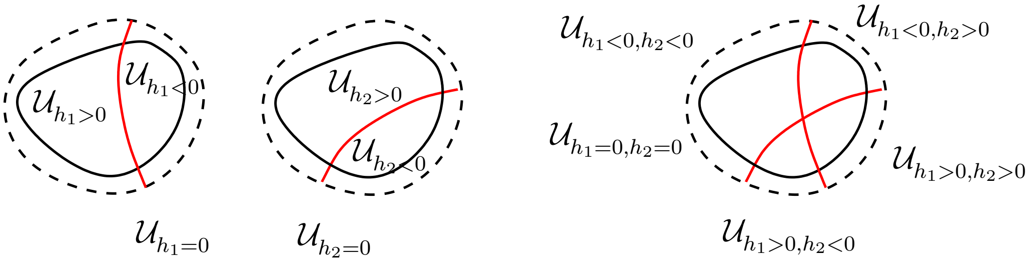

Figure J.25: An example of $\mathcal{P}(\mathcal{U}, h^{(1)} \cup h^{(2)}) = \mathcal{P}(\mathcal{U}, h^{(1)}) \cap \mathcal{P}(\mathcal{U}, h^{(2)})$. Also, for every $\mathcal{U}_{h,\sigma} \in \mathcal{P}(\mathcal{U}, h)$ there are some $\mathcal{U}^{(1)}_{h^{(1)},\sigma^{(1)}} \in \mathcal{P}(\mathcal{U}^{(1)}, h^{(1)})$ and $\mathcal{U}^{(2)}_{h^{(2)},\sigma^{(2)}} \in \mathcal{P}(\mathcal{U}^{(2)}, h^{(2)})$ such that $\mathcal{U}_{h^{(1)}\cup h^{(2)},\sigma} \subset \mathcal{U}^{(1)}_{h^{(1)},\sigma^{(1)}}$ and $\mathcal{U}_{h^{(1)}\cup h^{(2)},\sigma} \subset \mathcal{U}^{(2)}_{h^{(2)},\sigma^{(2)}}$.

Proof: Let $x \in \mathcal{U}$. Let $\mathcal{P}(\tilde{\mathcal{U}}^{(1)}, h^{(1)})$ be a semianalytic partition compatible with f_1 at x, and $\mathcal{P}(\tilde{\mathcal{U}}^{(2)}, h^{(2)})$ be a semianalytic partition compatible with f_2 at x. The proof becomes obvious if we construct a single semianalytic partition compatible at x with both functions. The natural choice is just the semianalytic partition of the intersection

$$\tilde{\mathcal{U}} := \tilde{\mathcal{U}}^{(1)} \cap \tilde{\mathcal{U}}^{(2)} \tag{J.383}$$

defined by the set of functions

$$h := h^{(1)} \cup h^{(2)}. \tag{J.384}$$

It is sufficient to notice, that for every $\tilde{\mathcal{U}}_{h,\sigma} \in \mathcal{P}(\mathcal{U}, h)$ there are some $\tilde{\mathcal{U}}^{(1)}_{h^{(1)},\sigma^{(1)}} \in \mathcal{P}(\tilde{\mathcal{U}}^{(1)}, h^{(1)})$ and $\tilde{\mathcal{U}}^{(2)}_{h^{(2)},\sigma^{(2)}} \in \mathcal{P}(\tilde{\mathcal{U}}^{(2)}, h^{(2)})$ such that

$$\tilde{\mathcal{U}}_{h,\sigma} \subset \tilde{\mathcal{U}}^{(1)}_{h^{(1)},\sigma^{(1)}} \qquad \text{and} \qquad \tilde{\mathcal{U}}_{h,\sigma} \subset \tilde{\mathcal{U}}^{(2)}_{h^{(2)},\sigma^{(2)}}. \tag{J.385}$$

□

J.19.3 Morphisms of Semianalytic Functions

It is obvious, that every analytic map $\phi : \mathcal{U} \to \mathcal{U}'$ between two open subsets $\mathcal{U} \subset \mathbb{R}^n$ and $\mathcal{U}' \subset \mathbb{R}^{n'}$ pullbacks all the semianalytic functions defined on $\mathcal{U}'$ into semianalytic functions defined on $\mathcal{U}$.

Lemma J.19.2 *Let $\mathcal{P}(\tilde{\mathcal{U}}', h')$ be a semianalytic partition. Let $\phi : \mathcal{U} \to \tilde{\mathcal{U}}'$ be a semianalytic function, where the subset $\mathcal{U} \subset \mathbb{R}^n$ is open. For every $x_0 \in \mathcal{U}$ there exists an open neighbourhood $\tilde{\mathcal{U}}$ and a semianalytic partition $\mathcal{P}(\tilde{\mathcal{U}}, \tilde{h})$ such that for every element of $\mathcal{P}(\tilde{\mathcal{U}}, \tilde{h})$, say $\tilde{\mathcal{U}}_{\tilde{h},\tilde{\sigma}}$, there is an element of $\mathcal{P}(\tilde{\mathcal{U}}', h')$, say $\tilde{\mathcal{U}}'_{\tilde{h}',\tilde{\sigma}'}$, such that*

$$\tilde{\mathcal{U}}_{\tilde{h},\tilde{\sigma}} \subset \phi^{-1}\big(\tilde{\mathcal{U}}'_{\tilde{h}',\tilde{\sigma}'}\big). \tag{J.386}$$

Proof:

If ϕ is analytic, then the set of pullbacks of all the functions $h'_I \in h'$ defines a suitable partition of the whole of $\mathcal{U}$.

$$\tilde{h} := \{h'_I \circ \phi_\sigma\} \cup \{h_I\}$$

We now prove the corresponding semianalytic partition $\mathcal{P}(\tilde{\mathcal{U}}, \tilde{h})$ satisfies the conclusion.

Denote the restriction of $\tilde{\sigma}$ to the set of functions h by σ

$$\sigma := \tilde{\sigma}\big|_h$$

so

$$\begin{matrix} h_1 & \sigma_1 & 0 \\ & \dots & \\ h_N & \sigma_N & 0 \end{matrix} \tag{J.387}$$

Denote the restriction of $\tilde{\sigma}$ to the set of functions $\phi^*_\sigma(h')$ by σ'

$$\sigma' := \tilde{\sigma}\big|_h$$

so

$$
\begin{array}{ccc}
\phi_\sigma^*(h_1') & \sigma_1' & 0 \\
 & \dots & \\
\phi_\sigma^*(h_M') & \sigma_M' & 0
\end{array}
\tag{J.388}
$$

$$
\tilde{\mathcal{U}}_{\tilde{h},\tilde{\sigma}} \subset \tilde{\mathcal{U}}_{h,\sigma} \tag{J.389}
$$

$$
\phi_\sigma(\tilde{\mathcal{U}}_{\tilde{h},\tilde{\sigma}}) \subset \tilde{\mathcal{U}}'_{h',\sigma'} \tag{J.390}
$$

By the definition of semianalytic functions, we have

$$
\phi_\sigma\big|_{\tilde{\mathcal{U}}_{h,\sigma}} = \phi\big|_{\tilde{\mathcal{U}}_{h,\sigma}},
$$

then by (J.389) we see that ϕ is equal to every one of analytic functions ϕ_σ when restricted to the set $\tilde{\mathcal{U}}_{\tilde{h},\tilde{\sigma}}$. This combied with (J.390) allows us to define a suitable partition of the of $\tilde{\mathcal{U}}$ and concludes the proof.

□

Proposition J.19.3 *Let $\mathcal{U} \subset \mathbb{R}^n$ and $\mathcal{U} \subset \mathbb{R}^{n'}$ be open subsets. Suppose the functions $f' : \mathcal{U}' \to \mathbb{R}^m$ and $\phi : \mathcal{U} \to \mathcal{U}'$ are semianalytic. Then, the composition function $f' \circ \phi : \mathcal{U} \to \mathbb{R}^m$ is semianalytic.*

Proof:

□

In general, the inverse of an invertible semianalytic function is not necessarilty semianalytic. However, a carefully formulated set of assumptions ensures the semianalyticity of the inverse.

Proposition J.19.4 *Let $\phi : \mathcal{U} \to \mathcal{U}'$ be a semianalytic and bijective function, where $\mathcal{U},\mathcal{U}' \subset \mathbb{R}^n$ are open. Suppose, that for every $x_0 \in \mathcal{U}$ there exists a semianalytic partition $\mathcal{P}(\mathcal{U}, h)$ compatible with ϕ at x_0, and such that for every $\tilde{\mathcal{U}}_{h,\sigma} \in \mathcal{P}(\mathcal{U}, h)$ the restriction $\phi\big|_{\tilde{\mathcal{U}}_{h,\sigma}}$ is extendable to an analytic, injective function $\phi_\sigma : \tilde{\mathcal{U}} \to \mathcal{U}'$, such that:*

(i) $\phi(\tilde{\mathcal{U}})$ is an open subset of $\mathbb{R}^n$, and

(ii) the invrese $\phi_\sigma^{-1} : \phi_\sigma(\tilde{\mathcal{U}}) \to \tilde{\mathcal{U}}$ is analytic. Then, ϕ^{-1} is semianalytic.

Proof:

□

Corollary J.19.5 *Suppose $\phi : \mathcal{U} \to \mathcal{U}'$ is a diffeomorphisms of the differentiability class C^m, where $\mathcal{U}, \mathcal{U}' \subset \mathbb{R}^n$ are open and $m > 0$. If ϕ is semianalytic, then so is $\phi^{-1} : \mathcal{U}' \to \mathcal{U}$*

Proof:

□

Lemma J.19.6 *Suppose $\mathcal{P}(\mathcal{U}, h)$ is a semi-semianalytic partition of an open $\mathcal{U} \subset \mathbb{R}^n$. Then, every $x \in \mathcal{U}$ has a neighbourhood $\tilde{\mathcal{U}}$ which admits a semianalytic partition finer than $\mathcal{P}(\tilde{\mathcal{U}}, h)$.*

Proof:

□

Proposition J.19.7 *For every semiaanlytic partition $\mathcal{P}(\mathcal{U}, h)$ of an open $\mathcal{U} \subset \mathbb{R}^n$, every point $x \in \mathcal{U}$ has a neighbourhood $\tilde{\mathcal{U}}$ which admits an analytic partitio finer than $\mathcal{P}(\tilde{\mathcal{U}}, h)$.*

Proof:

□

J.19.4 Semianalytic Manifolds and Submanifolds

By analogy with the definitions of an analytic structure, analytic function, and analytic submanifolds, we introduce now natural semianalytic generalizations. This is possible by results of the previous two subsections. The main result of this subsection is the proof of the finite intersection property of semianalytic submanifolds, given in the form of proposition J.19.8. Then in the next subsection we prove the second important property of semianalytic structures - a local character.

In this subsection, Σ is an n-dimensional differential manifold. Henceforth we will be assuming that Σ and all considered functions are of a differentiable class C^m, where $m > 0$.

Recall a collection of related charts such that every point of Σ lies in the domain of at least one chart forms an atlas.

We denote an atlas of Σ by $\{(\mathcal{U}_I, \chi_I)\}_{I \in \mathcal{I}}$, where $\mathcal{I}$ is some labeling set, $\{\mathcal{U}_I\}$ is an open covering of Σ, and $\{\chi_I\}_{I \in \mathcal{I}}$ is a family of diffeomorphisms $\chi_I : \mathcal{U}_I \to \mathcal{U}'_I \subset \mathbb{R}^n$.

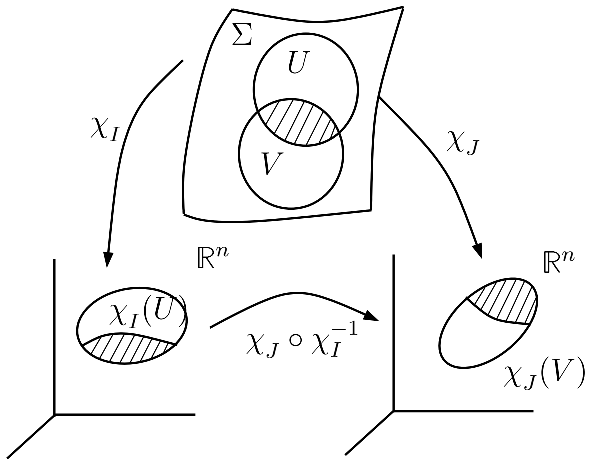

Figure J.26: semiana. The neighbourhoods U and V in Σ overlap. Their respective maps to $\mathbb{R}^n$, χ_I and χ_J, give two different coordinate systems to the overlap region. The differentiability class of the manifold is semianalytic when $\chi_J \circ \chi_I^{-1}$ is semianalytic.

Definition An atlas $\{(\mathcal{U}_I, \chi_I)\}$ of Σ is called semianalytic if for every pair $I, J \in \mathcal{I}$ the map

$$\chi_J \circ \chi_I^{-1} : \chi_I(\mathcal{U}_I \cap \mathcal{U}_J) \to \chi_J(\mathcal{U}_I \cap \mathcal{U}_J) \tag{J.391}$$

The collection of all semianalytic charts forms a maximal semianalytic atlas.

$\square$

Definition A semianalytic manifold is a differential manifold endowed with a maximal semianalytic atlas.

$\square$

Definition Given two semianalytic manifolds Σ and Σ', a map $f : \Sigma \to \Sigma'$ is called semianalytic if fo every semianalytic chart χ_I of Σ, and every semianalytic chart $\chi'_{I'}$ of Σ' the function $\chi'_{I'} \circ \chi_I^{-1}$ (whenever the composition can be applied) is semianalytic. .2 In particular, if

$$\Sigma' = \mathbb{R}^{n'} \tag{J.392}$$

and the semianalytic structure is the natural one defined by the atlas $\{(\mathbb{R}^{n'}, \mathrm{id})\}$, then the map f is a semianalytic function defined on Σ.

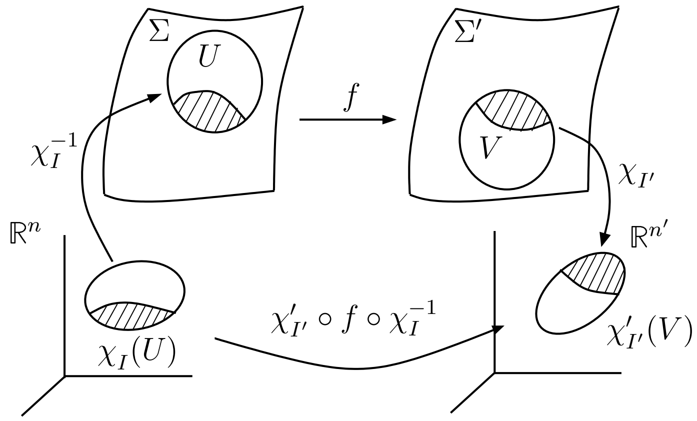

Figure J.27: Definition of a semianalytic function f between two semianalytic manifolds Σ and Σ'. Composition can be applied only when the range of $\chi_I(U) \subset \mathbb{R}^n$ under $\chi'_{I'} \circ f \circ \chi_I^{-1}$ overlaps with the subset $\chi'_{I'}(V) \subset \mathbb{R}^{n'}$ for subsets $U \subset \Sigma$ and $V \subset \Sigma'$.

$\square$

Definition A semianalytic submanifold of a semianalytic manifold Σ is a subset $S \subset \Sigma$ such that for very $x \in S$, there is a semianalytic chart χ_I defined in a neighbourhood $\mathcal{U}_I$ of x, such that

$$\begin{aligned} \chi_I(S \cap \mathcal{U}_I) &= \{(x^1, \ldots, x^n) \in \mathbb{R}^n : x^1 = \cdots = x^{n-n'} = 0, \\ &\quad 0 < x^{n-n'+1} < 1, \ldots, 0 < x^n < 1\}, \end{aligned} \tag{J.393}$$

where n' is a non-negative integer, $n' \leq n$, and n' is called the dimension of S.

$\square$

Definition An n' dimensional semianalytic submanifold with boundary of Σ is a subset $S \subset \Sigma$ such that for every $x \in S$, there is a semianalytic chart χ_I defined in a neighbourhood $\mathcal{U}_I$ of x, such that either (J.393) or

$$\begin{aligned} \chi_I(S \cap \mathcal{U}_I) &= \{(x^1, \ldots, x^n) \in \mathbb{R}^n : x^1 = \cdots = x^{n-n'} = 0, \\ &\quad 0 \leq x^{n-n'+1} < 1, 0 \leq x^{n-n'+2} < 1, \ldots, 0 < x^n < 1\}, \end{aligned} \tag{J.394}$$

$\square$

We are using extensively two particular classes of submanifolds: edges and faces.

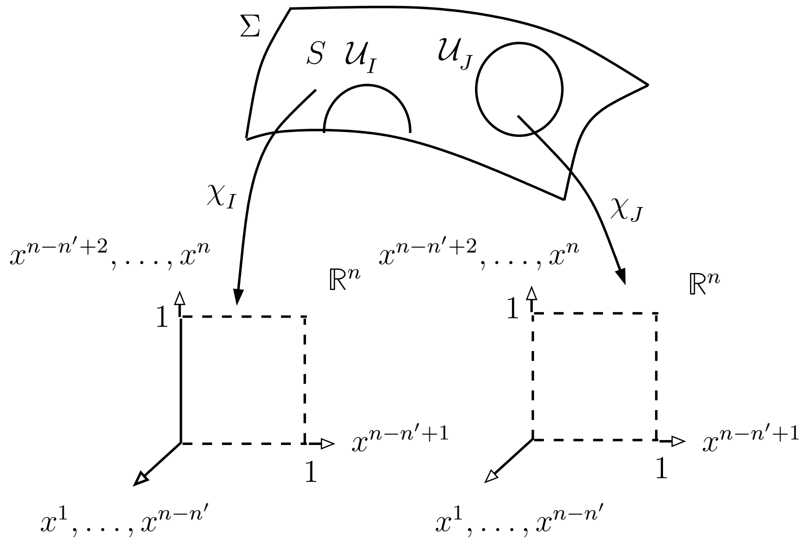

Figure J.28: .

Definition A semianalytic edge is a connected, one-dimensional semianalytic submanifold of Σ with two-point boundary.

$\square$

Definition A semianalytic face is a connected, codimensional one-dimensional semianalytic submanifold of Σ whose normal bundle is equipped with an orientation.

$\square$

All these definitions and results were to prepare for the following key property of semianalytic submanifolds crucial to the proof of the LOST uniqueness theorem. An infinite set of disjoint, embedded intervals may also form a single submanifold, disconnected though. For example the intersection of two edges where one of them is allowed to "oscillate" arbitarily rapidly. In theorem L.5.1 it is proved that for semianalytic edges this cannot occur: either the intersection of two edges results in a finite number of isolated intersections or a common connected finite segment. What is important in the conclusion of the following proposition, a generalization of this result for edges, is the finiteness of the partition and the connectedness of its elements. The following proposition apllied to the edges and faces contained in compact sets shows the intersection of edges is a finite collection of edges and isolated points, the intersection of edges and faces is a finite collection of edges and isolated points and the intersection of faces is a finite collection of faces, s.a. submanifolds of co-dimension two,..., edges and isolated points.

Proposition J.19.8 *Let S_1 and S_2 be two semianalytic submanifolds of a semianalytic manifold Σ. Suppose $x \in S_1 \cap S_2$. Then, there is an open neighbourhood $\mathcal{W}$ of x in Σ, such that $\mathcal{W} \cap S_1 \cap S_2$ is finite, distjoint union of connected semianalytic submanifolds.*

Proof: For every point $x \in S_1 \cap S_2$, there is a neighbourhood $\mathcal{W}$ which can be mapped by a semianalytic chart into an open subset $\mathcal{U} \subset \mathbb{R}^n$.

$\square$

J.19.5 Local Character of Semianalytic Structures

The property of semianalytic structures which distinguishes them from analytic ones is the local chararcter of the spaces of semianalytic functions and semianalytic diffeomorphisms. That feature is guaranteed by the existence of a partition of unity compatible with an arbitrary open covering.

Proposition J.19.9 *Suppose $\mathcal{W} \subset \Sigma$ is a compact subset. Let $\mathcal{U}_I \subset \Sigma$, $I = 1, \ldots, N$, be a family of open sets which covers $\mathcal{W}$. There exists a family of C^m semianalytic functions $\phi_I : \Sigma \to \mathbb{R}$, $I = 1, \ldots, N$ such that for every I,*

$$\textit{supp}\ \phi_I \subset \mathcal{U}_I \tag{J.395}$$

and

$$\sum_I \phi_I\big|_{\mathcal{W}} = 1. \tag{J.396}$$

Proof: The proof relies on the following two properties of semianalytic functions:

(i) For every open ball in $\mathbb{R}^D$, there is a C^m semianalytic function greater than zero at every point inside the ball and identically zero everywhere else.

(ii) If f is a nowhere vanishing C^m semianalytic function then so is $1/f$.

One employs smooth functions of the form

$$\exp(-1/(x_1^2 + \cdots + x_n^2)).$$

These functions are actually analytic except at the point $x = 0$ where they are C^∞. First we prove that there exists a smooth function $h : \mathbb{R}^n \to \mathbb{R}$ with $h \geq 0$ such that $h(x) = 1$ for $x_1^2 + \cdots + x_n^2 \leq 1$ and $h(x) = 0$ for $x_1^2 + \cdots + x_n^2 \geq 2$. Let

$$f(r) := \begin{cases} e^{-1/r^2} & r > 0 \\ 0 & r \leq 0. \end{cases}$$

Then $f : \mathbb{R} \to \mathbb{R}$ is smooth and the function

$$g(r) := \frac{f(2-r)}{f(2-r) + f(r-1)}$$

is C^∞ with values in $[0,1]$ such that $g(r) = 1$ for $r \le 1$ and $g(r) = 0$ for $r \ge 2$. We then just have to set $h(x) = g(\sqrt{x_1^2 + \cdots + x_n^2})$.

Now $\mathcal{W}$ is covered by finitely many open sets of the form $\mathcal{U}_I = \varphi_{x_I}^{-1}(B_1(0))$.

we now define functions $\phi_I : \Sigma \to \mathbb{R}$ by

$$\phi_I(y) := \begin{cases} h(\varphi_J(y)) & y \in \mathcal{U}_J \\ 0 & y \notin \mathcal{U}_J. \end{cases}$$

It follows from property () and the properties of h that

we put

$$\phi_I = \sum \frac{g_n}{g}$$

to get the required partion of unity subordinate to $\{\mathcal{U}_I\}$.

$\square$

J.20 Limits of Families of Measurable Spaces

The idea is actually quite simple.

Follows [409]

J.20.1 Hopf Theorem

$$E_1 \subset E_2 \;\Rightarrow\; \mu(E_1) \le \mu(E_2) \quad \text{monotonic} \tag{J.397}$$

$$\mu(\cup_i^\infty E_i) \;\le\; \sum_i^\infty \mu(E_i) \quad \text{countably subadditive.} \tag{J.398}$$

A set E is said to be $\mu-$measurable if

$$\mu(Y) = \mu(Y \cap E) + \mu(Y \cap E^c) \quad \text{for all } Y \subset X. \tag{J.399}$$

Theorem J.20.1 *For a Caratheodory's outer measure μ, the family of all $\mu-$measurable sets is a $\sigma-$algebra and the restriction of μ on this $\sigma-$algebra is a measure.*

$\mathcal{B}$ is closed under complement; $E \in \mathcal{B} \Leftrightarrow E^c \in \mathcal{B}$. Then we have

$$\begin{aligned} \mu(Y) &= \mu(Y \cap E_1) + \mu(Y \cap E_1^c) = \mu(Y \cap E_1) + \mu(Y \cap E_1^c \cap E_2) + \mu(Y \cap E_1^c \cap E_2^c) \\ &= \sum_{i=1}^{n} \mu(Y \cap E_1^c \cap \cdots \cap E_{i-1}^c \cap E_i) + \mu(Y \cap E_1^c \cap \cdots \cap E_n^c). \end{aligned} \tag{J.400}$$

By monotonicity of the measure μ, that $Y \cap \left(\bigcup_{i=1}^{\infty} E_i\right)^c \subset Y \cap E_1^c \cap \cdots \cap E_n^c$ implies $\mu(Y \cap E_1^c \cap \cdots \cap E_n^c) > \mu\big(Y \cap \bigcup_{i=1}^{\infty} E_i)^c\big)$. Therefore by (J.400) in the limit $n \to \infty$ we have

$$\mu(Y) \geq \sum_{i=1}^{\infty} \mu(Y \cap E_1^c \cap \cdots \cap E_{i-1}^c \cap E_i) + \mu\big(Y \cap \bigcup_{i=1}^{\infty} E_i)^c\big) \tag{J.401}$$

$$\mu(Y) \geq \mu\big(Y \cap (\bigcup_{i=1}^{\infty} E_i)\big) + \mu\big(Y \cap (\bigcup_{i=1}^{\infty} E_i)^c\big) \tag{J.402}$$

This implies $\bigcup_{i=1}^{\infty} E_i \in \mathcal{B}$, so the $\mu-$measurable sets form a $\sigma-$algebra. Now we establish that the measure μ is $\sigma-$additive. Assume $\{E_i\}$ are pairwise disjoint. If we choose as $Y = \cup_{i=1}^{\infty} E_i$ then E.q.(J.401) shows that

$$\begin{aligned} \mu\Big(\bigcup_{i=1}^{\infty} E_i\Big) &\geq \sum_{i=1}^{\infty} \mu\Big(\bigcup_{n=1} E_n \cap E_1^c \cap \cdots \cap E_{i-1}^c \cap E_i\Big) + \mu(\emptyset) \\ &= \mu\Big(\bigcup_{n=1} E_n \cap E_1\Big) + \mu\Big(\bigcup_{n=1} E_n \cap E_1^c \cap E_2\Big) + \mu\Big(\bigcup_{n=1} E_n \cap E_1^c \cap E_2^c \cap E_3\Big) + \ldots \\ &= \mu(E_1) + \mu\Big(\bigcup_{n=2}^{\infty} E_n \cap E_2\Big) + \mu\Big(\bigcup_{n=3}^{\infty} E_n \cap E_3\Big) + \ldots \\ &= \sum_{i=1}^{\infty} \mu(E_i) \end{aligned} \tag{J.403}$$

where we used pairwise disjointness in going from line 2 to line 3. This result means

$$\mu\Big(\bigcup_{i=1} E_i\Big) = \sum_{i=1}^{\infty} \mu(E_i) \tag{J.404}$$

because the converse inequality is true from (J.398).

□

Theorem J.20.2 (Hopf Theorem)

Let $\mathcal{B}(\mathcal{F})$ is the smallest $\sigma-$algebra containing $\mathcal{F}$. A finitely additive measure μ on $\mathcal{F}$ can be extended to a $\sigma-$additive measure on $\mathcal{B}(\mathcal{F})$, if and only if for

$$\begin{aligned} & F_i \in \mathcal{F} \;\; ; \;\; F_1 \supset F_2 \supset \cdots \supset F_n \supset \dots \\ & \lim_{n\to\infty} \mu(F_n) > 0 \implies \bigcap_{n=1}^{\infty} F_n \neq \emptyset \end{aligned} \tag{J.405}$$

or as in [105] "if and only if for

$F_i \in \mathcal{F}$; $F_1 \supset F_2 \supset \cdots \supset F_n \supset \dots$ with $\bigcap_{n=1}^{\infty} F_n = \emptyset$, we have

$$\lim_{n\to\infty} \mu(F_n) = 0. \tag{J.406}$$

Essentially, the condition (J.406) allows an extension $\tilde{\mu}$ to be consistently defined on elements of $\mathcal{B}(\mathcal{F})$ as limits of $\mu-$measures of sets in $\mathcal{F}$."

Proof of Theorem J.20.2:

The condition is necessary because if μ is extended to a $\sigma-$additive measure $\tilde{\mu}$ on $\mathcal{B}(\mathcal{F})$, we have

$$\lim_{n\to\infty} \mu(F_n) = \tilde{\mu}\big(\bigcap_{n=1}^{\infty} F_n\big). \tag{J.407}$$

Conversely, assume the condition (J.20.2). Caratheodory's outer measure

$$\mu^*(E) = \inf\{\sum_{i=1}^{\infty} \mu(F_n) \mid E \subset \bigcup_{i=1}^{\infty} F_i\} \tag{J.408}$$

Suppose $Y \subset X$ and $Y \subset \cup_{i=1}^{\infty} F_i$ are given. Since μ is finitely additive, we have

$$\sum_{i=1}^{\infty} \mu(F_i) = \sum_{i=1}^{\infty} \big(\mu(F_i \cap F) + \mu(F_i \cap F^c)\big) \geq \mu^*(Y \cap F) + \mu^*(Y \cap F^c) \tag{J.409}$$

The infimum of a smaller set cannot be less than the infimum of a larger set, so $\mu(Y) \geq \inf \sum_{i=1}^{\infty} \mu(F_i)$. Taking the infimum of the left-hand side of (J.409) and noting that $\mu(Y) = \mu^*(Y)$, we have

$$\mu^*(Y) \geq \mu^*(Y \cap F) + \mu^*(Y \cap F^c) \tag{J.410}$$

This means that F is μ^*-measurable.

Next, with $F_i = 0$ for $i > 1$ and setting $F_1 = F$ and $E = F$, it is clear that (J.408) implies $\mu^*(F) \leq \mu(F)$. We shall prove $\mu^*(F) \geq \mu(F)$, namely prove that

$$F \subset \bigcup_{i=1}^{\infty} F_i \implies \mu\,(F) \leq \sum_{i=1}^{\infty} \mu\,(F_i). \tag{J.411}$$

since

$$\mu(F) = \inf \mu(F) \leq \inf \sum_{n=1}^{\infty} \mu(F_n) \tag{J.412}$$

and (J.408) implies, (with $E = F$), $\mu(F) \leq \mu^*(F)$. Thus to this end we put

$$F_n' = \bigcap_{i=1}^{n} F_i^c \cap F.$$

Under the assumption of (J.411) then $\{F_n'\}$ is decreasing and $\bigcap_{n=1}^{\infty} F_n' = \emptyset$. From (J.20.2), we have $\mu(F_n') \to 0$, therefore for all $\epsilon > 0$, there exists an n such that $\mu(\bigcap_{i=1}^{n} F_i^c \cap F) < \epsilon$. Since

$$\mu(F) = \mu(\bigcap_{i=1}^{n} F_i \cap F) + \mu(\bigcap_{i=1}^{n} F_i^c \cap F)$$

we get

$$
\begin{aligned}
\mu(F) - \epsilon &< \mu(\bigcup_{i=1}^{n} F_i \cap F) \\
&\leq \sum_{i=1}^{n} \mu(F_i \cap F) \leq \sum_{i=1}^{n} \mu(F_i) \\
&\leq \sum_{i=1}^{\infty} \mu(F_i) \qquad \text{(J.413)}
\end{aligned}
$$

This holds for every $\epsilon > 0$, so that we have (J.411).

□

Theorem J.20.3 *A $\sigma-$additive extension of a finitely additive measure μ on $\mathcal{F}$ is unique on $\mathcal{B}(\mathcal{F})$, if possible.*

Proof of Theorem J.20.3.

□

J.21 Regular Borel Measures on the Projective Limit: The Uniform Measure

Here we describe how we can, based on the Riesz representatioon theorem, how to construct $\sigma-$additive measures on the porjective limit $\overline{X}$ starting from a self-consistent family of (so-called cylindrical) maeusres μ_α on the various factor spaces X_α.

The spaces X_α are compact Hausdorff spaces and inparticular topological spaces and are therefore naturally equipped with the $\sigma-$algebra $\mathcal{B}_\alpha$ of Borel sets (recall, the smallest $\sigma-$algebra containing all open subsets of X_α).

J.22 Almost Periodic Functions and Bohr Compactification

$$
\mathcal{N}_l(\overline{A}) = e^{il_j c} \qquad \text{(J.414)}
$$

the space of square integrable fucntions on a suitable completion $\overline{\mathcal{A}}_S$ of the classical configuration space.

is given as the spectrum of the C$\star$-algebra of almost periodic functions on H.

Periodic Functions

Defintion A **trignometric polynomial** on T is an expression of the form

$$P \sim \sum_{n=-N}^{N} a_n e^{int}. \tag{J.415}$$

The numbers n are called the frequencies of P. The largest integer n such that $|a_n| + |a_{-n}| \neq 0$ is called the **degree of** P.

$$P(t) = \sum_{n=-N}^{N} a_n e^{int}. \tag{J.416}$$

Knowing the function P we can compute the coefficients a_n by the formula

$$a_n = \frac{1}{2\pi} \int P(t) e^{-int} dt \tag{J.417}$$

which follows from

$$\frac{1}{2\pi} \int e^{i(m-n)t} dt = \delta_{mn} \tag{J.418}$$

Definition Let f be a comlex-valued function on R and let $\epsilon > 0$. An almost periodic of f is a number τ such that

$$\sup_x |f(x - \tau) - f(x)| < \epsilon. \tag{J.419}$$

Definition A function is almost periodic on $\mathbb{R}$ if it is continuous and if for every $\epsilon > 0$ there exists a number $\Sigma = \Sigma(\epsilon, f)$ such that every interval of length Σ on $\mathbb{R}$ is contains an $\epsilon-$almost period of f.

Examples

(1)

(2) $f = \cos x + \cos \pi x$

f is not continuous $f(x) = 2$ for $x = 0$ only.

(i) Every continuous a.p. function is bounded and uniformaly continuous on tha x-axis.

(ii) continuous periodic functions are also a.p. ,

(iii) The sums and products of a.p. functions and the limits of uniformaly convergent sequences of a.p. functions are also a.p.

J.22.1 Bohr Compactification

[157]

A function $f \in L^\infty(G)$ is, by definition, almost periodic if he set of all translates of f, $\{f_y\}_{y \in G}$ is precompact in the norm topology of $L^\infty(G)$.

As the dual of a discrete group, $\overline{G}$ it is compact; this is known as the *Borh compactification* of G.

A function f in sup norm L^∞ is, by definition, almost-periodic if the set of all translates of f, $\{f_y\}_{y \in G}$ precompact in the norm topology of L^∞. One proves that almost-periodic functions are uniformaly contiuous and are uniform limits of trignometric polynomials on G (i.e., of finite linear combinations of characters)

Bohr compacitification of the real line is the dual group of the discrete line and is usually called the Bohr group.

Assume $f \in AP(G)$; let $\{P_j\}$ be a sequence of trignometric polynomials which converges to f uniformally. Then, since G is dense in $\overline{G}$, $\{P_j\}$ converges uniformally on $\overline{G}$ (every character on G extends by continuity to a character on $\overline{G}$). It follows that f is the restriction to G of the $\lim P_i = F \in C(\overline{G})$. Conversely, since every continuous function F on $\overline{G}$ can be approximated uniformaly by trignometric polynomials, it follows that AP(G) is simply the restriction to G of $C(\overline{G})$.

J.22.2 Characters and the Dual Group

A *character* on a LCA group G is a continuous complex-valued function $\chi(x)$ on G with modulas 1 that preserves mulipilication, i.e.,

$$\chi(xy) = \chi(x)\chi(y) \text{ , with } |\chi(x)| = 1 \text{ for all } x, y \in G. \tag{J.420}$$

That is, a character is a continuous homomorphism of G into the multiplicative group of complex numbers of modulus 1.

J.23 Introduction to Ergodic Theory

The original ergodic hypothesis concerned the long run average behavior of individual trajectories in the phase space Σ of a dynamical system. Such trajectories define a measure preserving flow φ

J.23.1 Motivation

For a measure $d\hat{\mu}$ to correspond to a physically interesting field theory, it has to satisfy (some version) of the Osterwalder-Schrader axioms []. These axioms guarantee that from the measure it is possible to construct the physical Hilbert space with a well defined Hamiltonian and Green's functions with appropriate proerties.

One of the most important of these axioms is the Euclidean invariance that requires that the measure $d\hat{\mu}$ be invariant under the action of theEuclidean group on $\mathbb{R}^{d+1}$.

J.23.2 Introduction

statistical mechanics concerning phase space. If a system of N particles is enclosed in a box, the state of the system is given by a point in 6-dimensional phase space with q_i representing coordinates and p_i reprsenting momenta.

An absractization of measure-preserving character of flow.; the basic problem of statistical mechanics is to study the asymptotic ptoperties of measure-preserving transformations.

The state of the system

a one parameter group of transformation, $T_{s+t} = T_s T_t$ Liouville's theorem the flow of phase space all volumes are invariant. In

Ergodic theory an absractization of such physical considerations.

So we would like to show that the limit

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T f(T_t x) dt \tag{J.421}$$

exists, at least for continuous functions.

The map $f \mapsto f(x)$

(i) $\mu(1) = 1$

(ii) μ is linear

(iii) $\mu(f) \geq 0$ if $f \geq 0$.

Such a μ is always associated with a measure $\hat{\mu}$ on Ω_E with $\hat{\mu}(\Omega_E) = 1$, so that

$$\mu(f) = \int_{\Omega_E} f(w) d\hat{\mu}(w) \tag{J.422}$$

Lioville's theorem in classical mechanics.

We have an isolated system of N particles whoes state is specified by N generalized coordinates and momentum $\{q_1, \cdots, q_N; p_1, \cdots, p_N\}$ in phase space.

$$\dot{q}_i = \frac{\partial \mathcal{H}}{\partial p_i}, \quad \dot{p}_i = -\frac{\partial \mathcal{H}}{\partial q_i} \tag{J.423}$$

$$\rho \dot{q}_1 \, dt dq_2 \cdots dp_N - \left[\rho \dot{q}_1 + \frac{\partial(\rho \dot{q}_1)}{\partial q_1} dq_1 \right] dt dq_2 \cdots dp_N = -\frac{\partial}{\partial q_1}(\rho \dot{q}_1) dt \, dq_1 dq_2 \cdots \tag{J.424}$$

$$\frac{\partial \rho}{\partial t} dt dq_1 \cdots dp_N = \left[-\sum_{i=1}^{N} \frac{\partial}{\partial q_i}(\rho \dot{q}_i) - \sum_{i=1}^{N} \frac{\partial}{\partial p_i}(\rho \dot{p}_i) \right] dt \, dq_1 \cdots dp_N \tag{J.425}$$

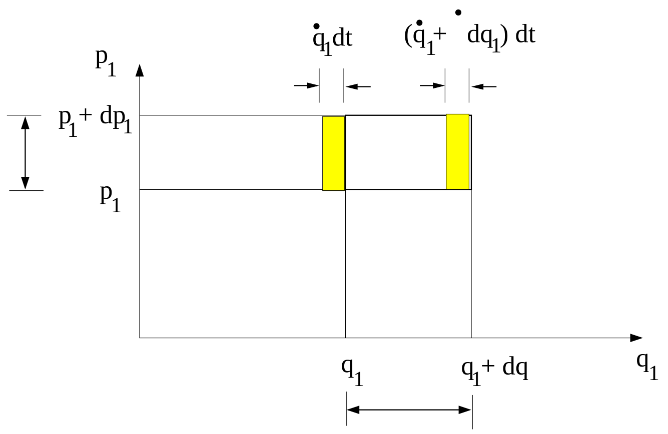

Figure J.29: Surface corresponding to first order term.

$$\mu_E(F) := \int_F \delta(\mathcal{H}(p, q) - E) d^{3N} p d^{3N} q \tag{J.426}$$

If we pich a set of local coordinates at $x \in \Omega_E$, say $\mathcal{Q}_1, \ldots, \mathcal{Q}_{6N-1}$, which are orthogonal and normalized, then

$$d\mu_E = Cd^{6N-1}\mathcal{Q}/|\text{grad}\ \mathcal{H}| \tag{J.427}$$

J.23.3 An Ergodic Theorem

we want to study

$$\frac{1}{T}\int_0^T (U_t f)(w) d\mu(w) \tag{J.428}$$

$$\frac{1}{N}\sum_{n=0}^{N-1} U^n f \tag{J.429}$$

statistical ergodic theorom of J. Von Neumann concerning motions $P \rightarrow P_t$ of a measurably set Ω onto itself which preserve the measure. The "discrete" counterpart of this theorem concerns iterates

$\int_\Omega f(P)d\sigma = \int_\Omega f(P')d\sigma$

$$\psi_{m,n} = \frac{1}{n-m}\sum_{k=m}^{n-1} f(P^{(k)}) \tag{J.430}$$

which tend in the mean, when $n - m \rightarrow \infty$.

Theorem (the mean ergodic theorem, or von Neumann's ergodic theorem) Let U be a unitary operator on a Hilbert space $\mathcal{H}$.

$$\lim_{N\rightarrow\infty} \frac{1}{N}\sum_{n=0}^{N-1} U^n f = Pf \tag{J.431}$$

in justifying the statement that phase-space average and time averages are equal.

[**?**] page 166

where μ is a finite measure, and U is given by $Ux = x \circ f$ for some measure-preserving map f. A map f is said to be *measure-preserving* if it is measurable and $\mu(f^{-1}(B)) = \mu(B)$ for all measurable sets $B \subset \Omega$. In other words the transport μ_f of the measure μ via f is the same as μ. To prove that such a U is indeed unitary we calculate as follows:

$$(Ux, Uy) = \int_\Sigma x(f(\omega))y(f(\omega))\ d\mu(\omega) = \int_\Omega xy\ d\mu_f = \int_\Omega xy\ d\mu = (x, y). \tag{J.432}$$

Theorem J.23.1 *(von Neumann mean ergodic theorem). Let $\mathbb{R} \to G$; $t \mapsto g_t$ be a one-parameter group and $\hat{U} : G \to \mathcal{B}(L_2(X, d\mu))$ be a unit representaation of G. Let $\hat{P}$ be the projection on the closure of the set of a.e. invariant vectors under $\hat{U}(g_t)$, $t \in \mathbb{R}$. Then*

$$(\hat{P}f)(x) = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} dt(\hat{U}(g_t)f)(x) \quad \mu - a.e. \tag{J.433}$$

J.24 The Infinite Tensor Product

We first consider the tensor product of a finite number of Hilbrt spaces. Say $f_k, g_k \in \mathcal{H}_k$ with inner product $(\cdot,\cdot)_k$ on $\mathcal{H}_k$. If an element of $\otimes_k \mathcal{H}_k$ is f, the inner product of the tensor product is defined as

$$(f, g) = \prod_{k=1}^{n} (f_k, g_k)_k$$

and the norm

$$\|f\| = \sqrt{\prod_{k=1}^{n} (f_k, f_k)_k} = \sqrt{\prod_{k=1}^{n} \|f_k\|_k^2} = \prod_{k=1}^{n} \|f_k\|_k.$$

We are lead to the consideration of the mathematics of products of arbitrary complex numbers.

Now when one forms the infinite tensor product of a collection of Hilbert spaces, a physical requirement is that this product must not depened on the order of the individual Hilbert spaces (whether the collection is countably or uncountably infinite).

Hence we are interested in the convergence properties of a countable or uncountable product of complex numbers which are independent of the ordering of the product. This was developed in the paper by von Neumann (availble at http://www.numdam.org/item?id=1939 0).

As we will see, convergence of products is related to convergence of corresponding summations. Now, it is a remarkable fact that whether an infinite series converges or not can depend on the ordering of the terms of that series. From which it follows that whether or not the coresponding product of complex numbers converges depends on the ordering of the terms of that product.

Absolutely and conditionally convergent series have completely different behaviours under rearrangement

Theorem J.24.1 *(***Riemann's Rearrangement Theorem***) Suppose that $\sum_{n=1}^{\infty} a_n$ is a conditionally convergent real series. For each real number s, there is a rearrangement of $\sum_{n=1}^{\infty} a_n$ that converges and has sum s.*

Proof:

The nonnegative series $\sum_{n=1}^{\infty} p_n$ and $\sum_{n=1}^{\infty}(-q_n)$ diverge. In fact, if both were to converge, it would follow that $\sum_{n=1}^{\infty} |a_n|$ converges, that is, $\sum_{n=1}^{\infty} a_n$ would be absolutely convergent.

On the other hand, if one of these series converged and the other diverged, it would follow that the partial sums of $\sum_{n=1}^{\infty} a_n$ diverge to either $+\infty$ or to $-\infty$. The convergence of $\sum_{n=1}^{\infty} a_n$ itself implies that both $\{p_n\}$ and $\{q_n\}$ have limit zero.

Now we construct the rearrangement. Choose terms $p_1, p_2, \ldots$ up to the first index k_1 such that

$$p_1 + p_2 + \cdots + p_{k_1} > s.$$

This will occur , because $\sum_{n=1}^{\infty}$. Note

$$|p_1 + p_2 + \cdots + p_{k_1} - s| < p_{k_1}.$$

Next, we choose $q_1, q_2, \ldots$ up to the first index l_1

such that

$$(p_1 + p_2 + \cdots + p_{k_1}) + (q_1 + q_2 + \cdots + q_{l_1}) < s.$$

Note

$$|(p_1 + p_2 + \cdots + p_{k_1}) + (q_1 + q_2 + \cdots + q_{l_1}) - s| < \max\{p_{k_1}, |q_{l_1}|\}.$$

We then add just enough new p's to make the left hand side greater than s, followed by just enough q's to make it less than s, and continue. At each phase of the $2n-$th step, the difference between s and the partial sum of the new series has absolute value smaller than $\max\{p_{k_n}, |q_{l_n}|\}$, and at each phase of the $2n+1-$th step, the difference between s and the partial sum of the new series has absolute value smaller than $\max\{p_{k_{n+1}}, |q_{l_n}|\}$. As these have the limit 0, the rearranged series has sum s.

□

Corollary J.24.2 *Suppose that $\sum_{n=1}^{\infty} a_n$ is a conditionally convergent real series. It has a rearrangement that diverges to $+\infty$.*

Proof:

□

Theorem J.24.3 *Suppose that $\sum_{n=1}^{\infty} a_n$ is absolutely convergent and that $\sum_{n=1}^{\infty} b_n$ is a rearrangement. Then $\sum_{n=1}^{\infty} b_n$ converges and has the same sum as $\sum_{n=1}^{\infty} a_n$.*

Proof: Let $\{s_n\}$ be the sequence of partial sums of $\sum_{k=1}^{\infty} a_k$, i.e,

$$s_n = \sum_{k=1}^{n} a_k,$$

and let s be the limit. Let $\{t_n\}$ be the sequence of partial sums of $\sum_{k=1}^{\infty} b_k$. Given $\epsilon > 0$, choose M so large that

$$\sum_{k=M+1}^{\infty} |a_k| < \epsilon/2. \tag{J.434}$$

It follows from this that $|s - s_M| < \epsilon/2$,

$$|s - s_M| = |\sum_{k=M+1}^{\infty} a_k| \leq \sum_{k=M+1}^{\infty} |a_k| < \epsilon/2.$$

Choose N so large that every one of the first M terms of $\{a_k\}$ occurs among the first N terms of $\{b_k\}$. So that for any $n \geq N$

$$|t_n - s_M| = |t_n - \sum_{k=1}^{M} a_k| \leq \sum_{k=M+1}^{\infty} |a_k|.$$

Therefore

$$|t_n - s| \leq |t_n - s_M| + |s_M - s| < \epsilon/2 + \epsilon/2 = \epsilon.$$

□

Theorem J.24.4 *Any bounded nondecreasing sequence of reals if convergent.*

Proof:

Let a be the least upper bound of the set of numbers $\{a_1, a_2, \dots\}$. If ϵ is positive, then the number $a - \epsilon$ is less than a and therefore not an upper bound for this set. As such, there is some N such that $a_N > a - \epsilon$. This, and the inequalities

$$a_1 \leq a_2 \leq a_3 \leq \cdots \leq a_n, \qquad \text{all } n,$$

and the fact that a is an upper bound imply that for any integer $n \geq N$

$$a - \epsilon < a_N \leq a_n \leq a.$$

Therefore $n \geq N$ implies $|a_n - a| < \epsilon$, and a is the limit.

□

Theorem J.24.5 *If $b_n \geq 0$ for all n, then $\sum_{n=1}^{\infty} b_n$ converges if and only if the sequence of partial sums is a bounded sequence.*

Proof:

□

J.24.1 Infinite Products of Complex Numbers

A sequence of numbers $a_1, a_2, a_3, \ldots$ the infinite product

$$\prod_{n=1}^{\infty} a_n = a_1\ a_2\ a_3 \cdots$$

is defined to be the limit of the partial products $a_1 a_2 \ldots a_n$ as $n \to \infty$. That is, let P_n be the partial product

$$P_n = \prod_{k=1}^{n} a_k$$

then

$$\prod_{n=1}^{\infty} a_n := \lim_{n \to \infty} P_n.$$

The product is said to converge when the limit exists. If the product converges, then the limit of the sequence a_n as $n \to \infty$ must be 1. Proof is easy. Assume that an infinite product $\prod a_n$ is convergent. Then for each a_n of the series we have $a_n = P_n / P_{n-1}$. Since the product is convergent, there exists a such that as $n \to \infty$, $P_n = a$ and $P_{n-1} = a$. Therefore

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} P_n/P_{n-1} = a/a = 1.$$

The contrary to this is in general not true. Therefore, the logarithm a_n will be defined for all but a finite number of n, and for those we have ($a_n \neq 0$)

$$\ln \prod_{n=1}^{\infty} a_n = \sum_{n=1}^{\infty} \ln a_n$$

with the product on the left converging if and only if the sum on the right converges. This allows the translation of convergence criteria for infinite sums into convergence criteria for infinite products.

Lemma J.24.6 *By the defintion of convergence , if $\prod_n^\infty a_n$, $\prod_n^\infty a'_n$ converge to a, a' respectively then $\prod_n^\infty a_n a'_n$ converges to aa'.*

Proof: Note that if $\{C_n\}$ and $\{C'_n\}$ are convergent sequences, converging to C, C' respectively, then

$$\lim_{n\to\infty}(C_n + C'_n) = \lim_{n\to\infty} C_n + \lim_{n\to\infty} C'_n$$

as

$$\begin{aligned} |(C_n + C'_n) - (C + C')| &= |(C_n - C) + (C'_n - C')| \\ &\leq |(C_n - C)| + |(C'_n - C')| \end{aligned}$$

and because we can always choose N large enough so that $|(C_n - C)| < \epsilon/2$ and $|(C'_n - C')| < \epsilon/2$ for all $n \geq N$.

Therefore, with $C_n = \sum_k^n \ln a_k$, $C'_n = \sum_k^n \ln a'_k$, we have

$$\begin{aligned} \lim_{n\to\infty} \sum_k^n (\ln a_k + \ln a'_k) &= \lim_{n\to\infty} (\sum_k^n \ln a_k + \sum_k^n \ln a'_k) \\ &= \lim_{n\to\infty} \sum_k^n \ln a_k + \lim_{n\to\infty} \sum_k^n \ln a'_k \end{aligned}$$

We can the use this to prove the lemma,

$$
\begin{aligned}
\ln \prod_{n}^{\infty}(a_n a'_n) &= \sum_{n}^{\infty} \ln(a_n a'_n) \\
&= \sum_{n}^{\infty} (\ln a_n + \ln a'_n) \\
&= \sum_{n}^{\infty} \ln a_n + \sum_{n}^{\infty} \ln a'_n \\
&= \ln a + \ln a' \\
&= \ln(aa').
\end{aligned}
$$

$\square$

A criterion for a product to converge.

Theorem J.24.7 *For $\rho_k > 0$, if $\prod_{n=1}^{\infty} |\rho_n - 1|$ converges then $\prod_{n=1}^{\infty} \rho_n$ converges.*

Proof: Set

$$
P_n = \prod_{k=1}^{n}(1 + a_k), \qquad \tilde{P}_n = \prod_{k=1}^{n}(1 + |a_k|)
$$

Note that

$$
1 + \sum_{k=1}^{n} |a_k| \leq \tilde{P}_n \leq \exp(\sum_{k=1}^{n} |a_k|).
$$

If the summation $\sum_{n=1}^{\infty} |a_n|$ converges then the sequence $\{B_n\}$ defined by $B_n := \exp(\sum_{k=1}^{n} |a_k|)$ is a bounded and a nondecreasing sequence and so converges. This implies that $\tilde{P}_n$ converges if the summation $\sum_{n=1}^{\infty} |a_n|$ convergences.

If $P_n - 1 < 0$ then, expanding the product, we see

$$
-(\tilde{P}_n - 1) < P_n - 1
$$

and if $P_n - 1 \geq 0$ then

$$
P_n - 1 \leq (\tilde{P}_n - 1),
$$

therefore

$$|P_n - 1| \leq \tilde{P}_n - 1$$

Hence if the product

$$\prod_{n=1}^{\infty}(1 + |a_n|)$$

converges, then so does

$$\prod_{n=1}^{\infty}(1 + a_n).$$

If $\sum_{n=1}^{\infty} |a_n| < \infty$ then

$$\prod_{n=1}^{\infty}(1 + a_n)$$

converges. Putting $\rho_n = 1 + a_n$ ($\rho_n > 0$) we have $\sum_{n=1}^{\infty} |\rho_n - 1|$ converges then $\prod_{n=1}^{\infty} \rho_n$ converges.

□

In fact the implication can be reversed. To prove this we need to first establish a couple of lemmas.

Lemma J.24.8 *For $r_n > 0$, $\sum_{n=0}^{\infty} r_n$ converges if and only if $\prod_{n=0}^{\infty}(1 + r_n)$ converges.*

Proof:

From part of the proof given in the previous theorem we have that $\prod_{n=0}^{\infty}(1+r_n)$ converges if $\sum_{n=0}^{\infty} r_n$ converges, noting that $r_n > 0$.

□

Lemma J.24.9 *If $0 < r_n < 1$, then $\sum_{n=0}^{\infty} r_n$ converges if and only if $\prod_{n=0}^{\infty}(1 - r_n)$ converges.*

Proof: Note the product obviously is finite, it only diverges if it equals zero. First say the product $\prod_n^\infty r_n$ converges. As $0 < x <$ then

$$\ln(1-x) = -x - x^2/2 - \cdots < -x,$$

so

$$x < -\ln(1-x).$$

This implies

$$\sum_n^\infty r_n < -\sum_n^\infty \ln(1-r_n) = \ln\left(1/\prod_{n=0}^\infty (1-r_n)\right) < \infty,$$

i.e., the summation converges.

Now say the summation $\sum_{n=0}^\infty r_n$ converges. Then there exists finite N such that $r_n < \delta$ for all $n \geq N$. Since $\ln(1-x)$ is convex, it follows that for $x < \delta$

$$\ln(1-x) > -kx$$

where $k = -(\ln(1-\delta))/\delta$. To see this more clearly draw the function $\ln(1-x)$. Replacing x by r_n, summing over n and then exponentiating gives

$$\prod_{n=N}^\infty (1-r_n) > \exp(-k\sum_{n=N}^\infty r_n) > 0,$$

i.e., the product is non-zero, and so by definition convergent.

□

Theorem J.24.10 *For $\rho_n > 0$, $\sum_{n=0}^\infty |\rho_n - 1|$ converges if and only if $\prod_{n=0}^\infty \rho_n$ converges.*

Proof: Consider the product

$$\prod_{n=0}^\infty \rho_n = \prod_{n=0}^\infty [(1 + (\rho_n - 1)]$$

for $\rho_n > 1$. It follows from lemma (J.24.8) that $\sum_{n=0}^{\infty} |\rho_n - 1|$ converges if and only if $\prod_{n=0}^{\infty}(1 + |\rho_n - 1|)$ converges. This proves the theorem for the case of $\rho_n > 1$ for all n.

Consider the product

$$\prod_{n=0}^{\infty} \rho_n = \prod_{n=0}^{\infty}[(1 - (1 - \rho_n)]$$

for $0 < \rho < 1$. It follows from lemma (J.24.9), by replacing r_n by $1 - \rho_n$, the theorem is proved for the case of $\rho_n < 1$ for all n.

Now we consider the general case. We factor the partial product

$$P_N = \prod_{n=1}^{N} \rho_n$$

into the product of terms for which $0 < \rho < 1$ and into the product of terms for which $\rho > 1$, which we denote respectively as Q_N and R_N (obviously we can ignore terms for which $\rho_n = 1$). Then

$$P_N = Q_N \cdot R_N.$$

We then use the fact that if $\prod_n^{\infty} \rho_n$, $\prod_n^{\infty} \rho'_n$ converge to ρ, ρ' respectively then $\prod_n^{\infty} \rho_n \, \rho'_n$ converges to $\rho\rho'$ to complete the proof of the theorem.

□

Now we consider products of complex numbers.

Definition The infinite product of complex numbers

$$\prod_{n=0}^{\infty} z_n \tag{J.435}$$

is said to **converge** to the number z provided that for each positive number $\delta > 0$ there exists $N < \infty$ such that for all $n \geq N$

$$|z - \prod_{k=0}^{n} z_k| < \delta.$$

□

Lemma J.24.11 *If z_n are arbitrary complex numbers, then $\sum_n^\infty z_n$ converges if and only if $\sum_n^\infty |z_n|$ converges.*

Proof: Convergence of $\sum_n^\infty z_n$ is equivalent to the convegrence of both $\sum_n^\infty \mathrm{Re} z_n$ and $\sum_n^\infty \mathrm{Im} z_n$. Similarly, as

$$|\mathrm{Re} z_n|,\ |\mathrm{Im} z_n| \leq |z_n| \leq |\mathrm{Re} z_n| + |\mathrm{Im} z_n|.$$

the convergence of $\sum_n^\infty |z_n|$ is equivalent to the convegrence of both $\sum_n^\infty |\mathrm{Re} z_n|$ and $\sum_n^\infty |\mathrm{Im} z_n|$.

Sufficiency. As we have for real series

$$\left| \sum_k^n a_k \right| \leq \sum_k^n |a_k|,$$

$|\sum_k^n a_k|$ will be bounded as $\sum_k^n |a_k|$ is bounded by hypothesis, this proves sufficiency.

Now we prove necessity. Suppose $\sum_n^\infty z_n$ converges, and let a be its value.

□

in general it is not guatanteed we can write

$$\lim_{n\to\infty} (\prod_k^n |z_k|\, e^{i\varphi_k}) = \lim_{n\to\infty} \prod_k^n |z_k| \ \lim_{n\to\infty} \prod_k^n e^{i\varphi_k}$$

However we have: if $\{\xi_n\}$ and $\{\zeta_n\}$ are two complex convergent sequences with limits a and b, then

$$\lim_{n\to\infty} (\xi_n \zeta_n) = \xi\zeta$$

Proof:

$$\begin{aligned} |\xi_n\zeta_n - \xi\zeta| &= |(\xi_n\zeta_n - \xi\zeta_n) + (\xi\zeta_n - \xi\zeta)| \\ &\leq |(\xi_n\zeta_n - \xi\zeta_n)| + |(\xi\zeta_n - \xi\zeta)| \\ &= |\xi_n - \xi|\,|\zeta_n| + |\xi|\,|\zeta_n - \zeta| \end{aligned}$$

For $n \geq$ to some N, $|\zeta_n| = |\zeta + (\zeta_n - \zeta)| \leq |\zeta| + |(\zeta_n - \zeta)| \leq |\zeta| + 1$. Given $\epsilon > 0$ we can choose N so large that $n \geq N$ implies $|\zeta_n| \leq |\zeta| + 1$ and also

$$|\xi_n - \xi| < \frac{\epsilon}{2|\zeta|+2}, \qquad |\zeta_n - \zeta| < \frac{\epsilon}{2|\xi|}.$$

It follows that $n \geq N$ implies $|\xi_n\zeta_n - \xi\zeta| < \epsilon$.

Theorem J.24.12 *Let $z_n = \rho_n e^{i\varphi_n} \in \mathbb{C}$ where $\rho_n = |z_n|$, $\varphi_n \in [-\pi, \pi]$. Then $\prod_\alpha z_n$ converges if and only if*

i) either $\prod_n \rho_n$ converges to zero in which case $\prod_n z_n = 0$, prod ii) or $\prod_n \rho_n$ converges to $\rho > 0$ and $\sum_n |\varphi_n|$ converges in which case

$$\prod_n z_n = \rho e^{i\sum_n \varphi_n}.$$

Proof: the partial product can be written

$$P_n = \prod_{k=1}^n a_k = (\prod_{k=1}^n \rho_k)\ e^{i\sum_{k=1}^n \varphi_k}$$

$$\lim_{n\to\infty} e^{i\sum_{k=1}^n \varphi_k}$$

$$|z - e^{i\sum_{k=1}^n \varphi_k}| < \delta.$$

We require

$$\lim_{n\to\infty} P_n/P_{n-1} = 1$$

that is

$$\varphi_n \to 0$$

We show that when $\prod_n^\infty |z_n| > 0$, then it implies the convergence of $\sum_n^\infty |\varphi_n|$ also.

As $\prod_n^\infty z_n$, $\prod_n^\infty |z_n|$ converge, with the latter > 0

$$\prod_n^\infty \frac{z_n}{|z_n|} = \prod_n^\infty e^{i\varphi_n}$$

converges also.

□

Corollary J.24.13 *Quasi-convergence of $\prod_n z_n$ to a non-vanishing value $\neq 0$ is equivalent with convergence with such a value (this is by definition of quasi-convergence). It holds if and only if all $z_n \neq 0$, and $\sum_n^{\infty} |z_n - 1|$*

Proof: As it is quasi-convergent we have that $\sum_n^{\infty} |\,|z_n| - 1|$

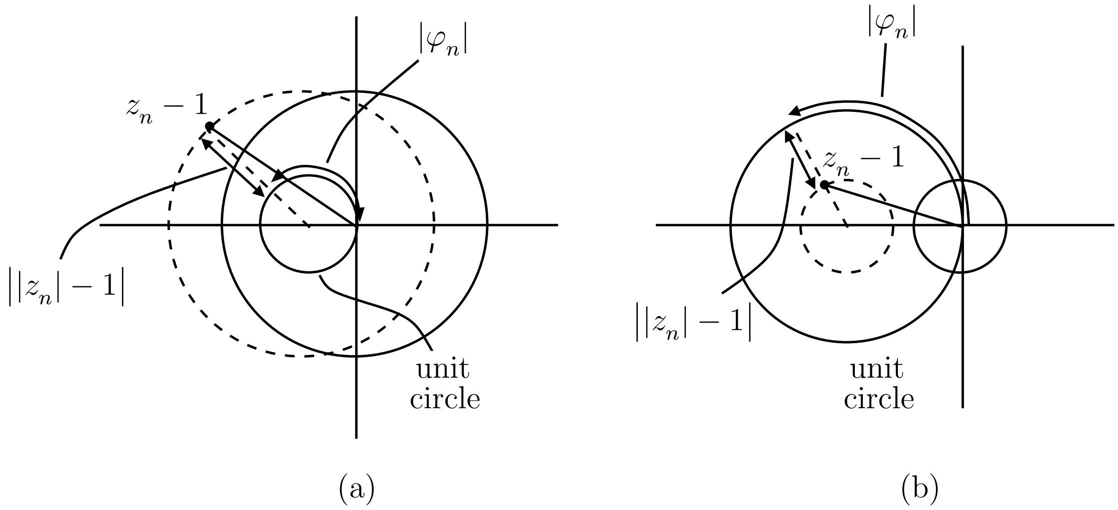

Figure J.30: Proof of the inequalities $\big||z_n| - 1\big| \leq |z_n - 1| \leq \big||z_n| - 1\big| + |\varphi_n|$. (a) Case where $|z_n| > 1$. (b) Case where $|z_n| < 1$.

$$\big||z_n| - 1\big|,\ \frac{1}{\pi}|\varphi_n| \leq |z_n - 1| \leq \big||z_n| - 1\big| + |\varphi_n|$$

Hence the convergence of the sum $\sum_n^{\infty} |z_n - 1|$ is equivalent to the convergence of both $\sum_n^{\infty} \big||z_n| - 1\big|$ and $\sum_n^{\infty} |\varphi_n|$.

Lemma J.24.14 *By the defintion of convergence, if*

1) $\prod_n^{\infty} z_n$, $\prod_n^{\infty} z'_n$ converge to z, z' respectively then $\prod_n z_n z'_n$ converges to zz',

2) if $\prod_n^{\infty} z_n$ converges to z, then $\prod_n^{\infty} z_n^*$ converges to z^*.

Proof: 1)

$$\begin{aligned}
\ln \prod_n^\infty (z_n z'_n) &= \sum_n^\infty (\ln \rho_n + \ln \rho'_n + i\varphi_n + i\varphi'_n) \\
&= \sum_n^\infty (\ln \rho_n + i\varphi_n) + \sum_n^\infty (\ln \rho'_n + i\varphi'_n) \\
&= \ln \prod_n^\infty z_n + \ln \prod_n^\infty z'_n \\
&= \ln(zz')
\end{aligned}$$

□

Lemma J.24.15 *For any complex numbers, the convergence of one of* $\sum_{n=0}^\infty | \, |z_n| - 1|$, $\sum_{n=0}^\infty | \, |z_n|^2 - 1|$ *implies the convergence of the other.*

Proof: If one or other of the series converges, then $\lim_{n\to\infty} |z_n| = 1$. So

$$\begin{aligned}
\lim_{n\to\infty} \frac{| \, |z_n|^2 - 1|}{| \, |z_n| - 1|} &= \lim_{n\to\infty} | \, |z_n| + 1| \\
&= 2 > 0.
\end{aligned}$$

Hence there exists finite N such that

$$1 < \frac{| \, |z_n|^2 - 1|}{| \, |z_n| - 1|} < 3 \qquad \text{for all } n \geq N.$$

Equivalently

$$| \, |z_n| - 1| < | \, |z_n|^2 - 1| < 3| \, |z_n| - 1| \qquad \text{for all } n \geq N.$$

If the series $\sum_{n=0}^\infty | \, |z_n|^2 - 1|$ converges then $\sum_{n=0}^\infty | \, |z_n| - 1|$ also converges as

$$\sum_{n=N}^\infty | \, |z_n| - 1| < \sum_{n=N}^\infty | \, |z_n|^2 - 1|.$$

If the series $\sum_{n=0}^\infty | \, |z_n| - 1|$ converges then $\sum_{n=0}^\infty | \, |z_n|^2 - 1|$ also converges as

$$\sum_{n=N}^{\infty} |\, |z_n|^2 - 1| < \sum_{n=N}^{\infty} 3|\, |z_n| - 1|.$$

The result also follows by applying theorem J.24.10: $\sum_{n=0}^{\infty} |\, |z_n| - 1|$ converges if and only if $\prod_{n=0}^{\infty} |z_n|$ converges, which converges if and only if $\prod_{n=0}^{\infty} |z_n|^2$ converges, which in turn converges if and only if $\sum_{n=0}^{\infty} |\, |z_n|^2 - 1|$ converges.

□

It is understood that the sum is meaniful only if at most a countable number of terms are different from zero.

Recall that for a series $\sum_{\alpha} z_{\alpha}$ to converge absolutely it is necessary that $z_{\alpha} = 0$ for all but countably infinitely many $\alpha \in \mathcal{I}$.

Definition Let $\{z_{\alpha}\}_{\alpha \in \mathcal{I}}$ be a collection of complex numbers. The infinite product

$$\prod_{\alpha \in \mathcal{I}} z_{\alpha} \tag{J.436}$$

is said to **converge** to the number z provided that for each positive number $\delta > 0$ there exists a finite set

$$\mathcal{I}_0 \subset \mathcal{I}$$

such that for any other finite $\mathcal{J}$ with

$$\mathcal{I}_0(\delta) \subset \mathcal{J} \subset \mathcal{I}$$

it holds that

$$|z - \prod_{\alpha \in \mathcal{J}} z_{\alpha}| < \delta.$$

□

In contrast to the case of an infinite series, absolute converge of an infinite product does not imply converge, the phases of the factors could fluctuate wildly. This motivates the following definition.

Definition We say that $\prod_{\alpha\in\mathcal{I}} z_\alpha$ is **quasi-convergent** if $\prod_{\alpha\in\mathcal{I}} |z_\alpha|$ converges. If $\prod_{\alpha\in\mathcal{I}} z_\alpha$ is quasi-convergent but not convergent we define $\prod_{\alpha\in\mathcal{I}} z_\alpha := 0$.

$\square$

This definition assigns a value to the infinite product of numbers which converge absolutely but not necessarily non-absolutely.

Theorem J.24.16 *1) Let $\rho_\alpha \geq 0$.*

i) If for every $\alpha_0 \in \mathcal{I}$ it holds that $\rho_{\alpha_0} = 0$ then $\prod_\alpha \rho_\alpha = 0$.

ii) If for $\rho_\alpha > 0$ for all α then $\prod_\alpha \rho_\alpha$ converges if and only if

$$\sum_\alpha \max(\rho_\alpha - 1, 0)$$

converges.

iii) If for $\rho_\alpha > 0$ for all α then $\prod_\alpha \rho_\alpha$ converges if and only if

$$\sum_\alpha |\rho_\alpha - 1|$$

converges.

2) Let $z_\alpha = \rho_\alpha e^{i\varphi_\alpha} \in \mathbb{C}$ where $\rho_\alpha = |z_\alpha|$, $\varphi_\alpha \in [-\pi, \pi]$. Then $\prod_\alpha z_\alpha$ converges if and only if

i) either $\prod_\alpha \rho_\alpha$ converges to zero in which case $\prod_\alpha z_\alpha = 0$,

ii) or $\prod_\alpha \rho_\alpha$ converges to $\rho > 0$ and $\sum_\alpha |\varphi_\alpha|$ converges in which case

$$\prod_\alpha z_\alpha = \rho e^{i\sum_\alpha \varphi_\alpha}.$$

Proof:

$\square$

Corollary J.24.17 *Quasi-convergence*

Proof:

$\square$

After having introduced convergence for infinite products of complex numbers we can now turn to ITP Hilbert spaces.

J.24.2 $C-$vectors of the ITP

Definition Let $\mathcal{H}_\alpha$, $\alpha \in \mathcal{I}$ be an arbitrary collection of Hilbert spaces. For a sequence $f := \{f_\alpha\}_{\alpha\in\mathcal{I}}$, $f_\alpha \in \mathcal{H}_\alpha$ the object

$$\otimes_f := \otimes_\alpha f_\alpha \tag{J.437}$$

is called a $C-$**vector** provided that $\prod_\alpha \|f_\alpha\|_\alpha$ converges, where $\|\cdot\|_\alpha$ denotes the Hilbert norm of $\mathcal{H}_\alpha$. The set of vectors will be denoted V_C.

$\square$

Lemma J.24.18 *For two $C-$vectors $\otimes_f = \otimes_\alpha f_\alpha$, $\otimes_g = \otimes_\alpha g_\alpha$ the inner product*

$$< \otimes_f, \otimes_g > := \prod_\alpha < f_\alpha, g_\alpha >_\alpha \tag{J.438}$$

is a quasi-convergent product of the individual inner products $< f_\alpha, g_\alpha >$ on $\mathcal{H}_\alpha$.

Proof: We prove that if $f_\alpha, g_\alpha \in \mathcal{H}_\alpha$ for all $\alpha \in \mathcal{I}$, and if $\prod_\alpha \|f_\alpha\|_\alpha$, $\prod_\alpha \|g_\alpha\|_\alpha$ are convergent, then so is $\prod_a |(f_\alpha, g_\alpha)_\alpha|$, that is, $\prod_a (f_\alpha, g_\alpha)_\alpha$ is quasi-convergent.

Since $\|f_\alpha\|_\alpha = 0$ or $\|g_\alpha\|_\alpha = 0$ implies $(f_\alpha, g_\alpha)_\alpha = 0$

by theorem J.24.16 we need only show that $\sum_\alpha \max(|(f_\alpha, g_\alpha)_\alpha| - 1, 0)$ converges as a consequence of the convergence of $\sum_\alpha \max(\|f_\alpha\|_\alpha^2 - 1, 0)$, $\sum_\alpha \max(\|g_\alpha\|_\alpha^2 - 1, 0)$.

Now as $|(f_\alpha, g_\alpha)_\alpha| \leq \frac{1}{2}\|f_\alpha\|_\alpha^2 + \frac{1}{2}\|g_\alpha\|_\alpha^2$,

$$|(f_\alpha, g_\alpha)_\alpha| - 1 \leq \frac{1}{2}(\|f_\alpha\|_\alpha^2 - 1) + \frac{1}{2}(\|g_\alpha\|_\alpha^2 - 1)$$

and hence

$$\max(|(f_\alpha, g_\alpha)_\alpha| - 1, 0) \leq \frac{1}{2}\max(\|f_\alpha\|_\alpha^2 - 1, 0) + \frac{1}{2}\max(\|g_\alpha\|_\alpha^2 - 1, 0).$$

□

There are $C-$vectors $\otimes_f$ such that $\prod_\alpha \|f_\alpha\|_\alpha = 0$ although $\|f_\alpha\|_\alpha > 0$ for all $\alpha \in \mathcal{I}$. Thus, it is conceivable that it happens that $< \otimes_f, \otimes_g > \neq 0$ for some $C-$vector $\otimes_g$. If that were the case, the Schwarz inequality is violated for the inner product (J.438) for $C-$vectors.

Definition We say that for $f_\alpha \in \mathcal{H}_\alpha$ the ITP $\otimes_f := \otimes_\alpha f_\alpha$ is a C_0-**vector** (and $f = \{f_\alpha\}$ a C_0 sequence) if $\| \otimes_f \| := \prod_{\alpha \in \mathcal{I}} \|f_\alpha\|_\alpha$ converges to a non-vanishing number. The set of C_0-vectors will be denoted V_0.

□

To distinguish trivial $C-$vectors from non-trivial ones we define

Definition A sequence $\{f_\alpha\}$ defines a C_0-**vector** $\otimes_f = \otimes_\alpha f_\alpha$ if and only if

$$\sum_\alpha | \; \|f_\alpha\|_\alpha - 1| \tag{J.439}$$

converges. The set of C_0-vectors will be denoted V_0.

□

Lemma J.24.19 *For any complex numbers,* $\sum_\alpha | \; |z_\alpha| - 1|$ *converges if and only if* $\sum_\alpha | \; |z_\alpha|^2 - 1|$.

Proof:

□

For a C_0-vector $\|f_\alpha\|_\alpha \neq 0$.

The norm of a C_0-vector does not zanish. By the previous lemma C_0-vector can be defined equivalently as a $\otimes_f = \otimes_\alpha f_\alpha$ such that

$$\sum_\alpha | \; \|f_\alpha\|_\alpha^2 - 1|$$

converges. By theorem J.24.16 this implies that

$$\prod_\alpha \|f_\alpha\|_\alpha^2 = \ (\otimes_f, \otimes_f) \ = \| \otimes_f \|^2$$

converges to a non-zero number, i.e. $\| \otimes_f \| > 0$.

J.24.3 Strong Equivalence Classes

Definition We will call two C_0-sequences $f = \{f_\alpha\}, g = \{g_\alpha\}$ **strongly equivalent**, denoted $f \approx g$, provided that

$$\sum_\alpha |(f_\alpha, g_\alpha)_\alpha - 1| \tag{J.440}$$

converges.

□

Lemma J.24.20 *Strong equivalence of C_0-sequences is an equivalence relation, i.e., a relation such that*

1) (Reflexivity) for any $\otimes_f \in V_0$, we have $\otimes_f \approx \otimes_f$.

2) (Symmetry) for all $\otimes_f, \otimes_g \in V_0$, if $\otimes_f \approx \otimes_g$ then $\otimes_g \approx \otimes_f$.

3) (Transitivity) for all $\otimes_f, \otimes_g, \otimes_h \in V_0$, if $\otimes_f \approx \otimes_g$ and $\otimes_g \approx \otimes_h$ then $\otimes_g \approx \otimes_h$.

Proof: Condition 1) follows from the fact that a C_0-vector is defined by the condition that $\sum_\alpha | \ \|f_\alpha\|_\alpha^2 - 1|$ converges.

$$|(f_\alpha, g_\alpha)_\alpha - 1| = |\overline{(f_\alpha, g_\alpha)_\alpha} - 1| = |\overline{(f_\alpha, g_\alpha)_\alpha - 1}| = |(g_\alpha, f_\alpha)_\alpha - 1|.$$

Condition 3) We prove that, apart from a possible finite number of exceptions, the following

$$|(f_\alpha, h_\alpha)_\alpha - 1| \le D \left(\left| \|f_\alpha\|_\alpha - 1 \right| + \left| \|g_\alpha\|_\alpha - 1 \right| + |(f_\alpha, g_\alpha)_\alpha - 1| + |(g_\alpha, h_\alpha)_\alpha - 1| \right)$$

holds for some constant D.

We put

$$\|f_\alpha\|_\alpha = 1+\lambda_f, \quad \|g_\alpha\|_\alpha = 1+\lambda_g, \quad \|h_\alpha\|_\alpha = 1+\lambda_h,$$
$$(f_\alpha, g_\alpha) = 1+\lambda_{fg}, \quad (g_\alpha, h_\alpha) = 1+\lambda_{gh}$$

So $|\lambda_f|$, $|\lambda_g|$, $|\lambda_h|$, $|\lambda_{fg}|$, $|\lambda_{gh}| \le C$, and except for a finite number of α's, $|\lambda_g| \le \frac{1}{2}$ (or $-1/2 \le \|g_\alpha - 1\|_\alpha \le 1/2$).

We express $g_\alpha, f_\alpha, h_\alpha \in \mathcal{H}_\alpha$ as:

$$\begin{aligned} g_\alpha &= a_{11}x_\alpha, \\ f_\alpha &= a_{21}x_\alpha + a_{22}x'_\alpha, \\ h_\alpha &= a_{31}x_\alpha + a_{32}x''_\alpha + a_{33}x'''_\alpha, \end{aligned}$$

where

$$\|x_\alpha\|_\alpha = \|x'_\alpha\|_\alpha = \|x''_\alpha\|_\alpha = 1, \qquad (x_\alpha, x'_\alpha) = (x_\alpha, x''_\alpha) = (x'_\alpha, x''_\alpha) = 0.$$

Then

$$\begin{aligned} \|g_\alpha\|_\alpha^2 &= |a_{11}|^2, \\ \|f_\alpha\|_\alpha^2 &= |a_{21}|^2 + |a_{22}|^2, \\ \|h_\alpha\|_\alpha^2 &= |a_{31}|^2 + |a_{32}|^2 + |a_{33}|^2, \\ \\ (f_\alpha, g_\alpha)_\alpha &= a_{21}\overline{a_{11}} \\ (g_\alpha, h_\alpha)_\alpha &= a_{11}\overline{a_{31}} \\ (f_\alpha, h_\alpha)_\alpha &= a_{21}\overline{a_{31}} + a_{22}\overline{a_{32}} \end{aligned}$$

Now

$$\begin{aligned} |(f_\alpha, g_\alpha)_\alpha - 1| &= |a_{21}\overline{a_{31}} + a_{22}\overline{a_{32}} - 1| \\ &= \left|(a_{21}\overline{a_{11}} \cdot a_{11}\overline{a_{31}}|a_{11}|^{-2} - 1) + a_{22}\overline{a_{32}}\right| \\ &\le \left|a_{21}\overline{a_{11}}a_{11}\overline{a_{31}}|a_{11}|^{-2} - 1\right| + |a_{22}\overline{a_{32}}|. \end{aligned}$$

$$
\begin{aligned}
\left|a_{21}\overline{a_{11}}a_{11}\overline{a_{31}}|a_{11}|^{-2}-1\right| &\leq (1+|\lambda_{fg}|)\,(1+|\lambda_{gh}|)\,(1-|\lambda_g|)^{-2}-1 \\
&\leq (1+|\lambda_{fg}|+|\lambda_{gh}|+|\lambda_{fg}|\,|\lambda_{gh}|)\,(1+2|\lambda_g|)-1 \\
&\leq \big(1+|\lambda_{fg}|+|\lambda_{gh}|+\frac{C}{2}(|\lambda_{gh}|+|\lambda_{fg}|)\big)\,(1+2|\lambda_g|)-1 \\
&\leq D_1(|\lambda_g|+|\lambda_{fg}|+|\lambda_{gh}|)
\end{aligned}
$$

where in the second line we used that $1/(1-x)^2$ is convex.

$$
|a_{22}|^2 = (|a_{21}|^2+|a_{22}|^2) - \frac{|a_{21}\overline{a_{11}}|^2}{|a_{11}|^2} \leq (1+|\lambda_f|)^2 - \frac{(1-|\lambda_{fg}|)^2}{(1+|\lambda_g|)^2}
$$

$$
|a_{32}|^2 = (|a_{31}|^2+|a_{32}|^2+|a_{33}|^2) - \frac{|a_{11}\overline{a_{31}}|^2}{|a_{11}|^2} \leq (1+|\lambda_h|)^2 - \frac{(1-|\lambda_{gh}|)^2}{(1+|\lambda_g|)^2}
$$

$$
\begin{aligned}
(1+|\lambda_f|)^2 - \frac{(1-|\lambda_{fg}|)^2}{(1+|\lambda_g|)^2} &= 1+2|\lambda_f|+|\lambda_f|^2 - \frac{1-2|\lambda_{fg}|}{(1+|\lambda_g|)^2} - \frac{|\lambda_{fg}|^2}{(1+|\lambda_g|)^2} \\
&\leq (2+C)|\lambda_f| - (1-2|\lambda_{fg}|)[1-2|\lambda_g|] \\
&\leq (2+C)|\lambda_f| + 2|\lambda_{fg}| + 2|\lambda_g| - 4|\lambda_{fg}|\,|\lambda_g| \\
&\leq (2+C)|\lambda_f| + 2|\lambda_{fg}| + 2|\lambda_g| \\
&\leq (2+C)(|\lambda_f|+|\lambda_g|+|\lambda_{fg}|)
\end{aligned}
$$

where in the second line we used $1/(1+x)^2 \geq 1-2x$ for $x \geq 0$. Obviously we also have

$$
|a_{32}|^2 \leq (2+C)(|\lambda_g|+|\lambda_h|+|\lambda_{gh}|)
$$

Thus

$$
|a_{22}a_{32}| \leq D_2(|\lambda_f|+|\lambda_g|+|\lambda_h|+|\lambda_{fg}|+|\lambda_{gh}|).
$$

Combining all the inequalities, we obtain

$$
|(f_\alpha, h_\alpha)-1| \leq D(|\lambda_f|+|\lambda_g|+|\lambda_h|+|\lambda_{fg}|+|\lambda_{gh}|).
$$

with a finite number of exceptions.

$\square$

An equivalence relation will split the set V_0 into disjoint subsets, the equivalence classes.

Definition The strong equivalence class of a C_0-sequence f is denoted by $[f]$. The set of strong equivalence classes of C_0-sequences will be called $\mathcal{S}$.

$\square$

Theorem J.24.21 *If two C_0-sequences f_α, g_α belong to two different equivalence classes, then*

$$(\prod_\alpha f_\alpha, \prod_\alpha g_\alpha) = 0.$$

If they belong to the same equivalnce class, then $(\prod_\alpha f_\alpha, \prod_\alpha g_\alpha) = 0$ if and only if some $(f_\alpha, g_\alpha) = 0$.

Proof: Clearly

$$(\prod_\alpha f_\alpha, \prod_\alpha g_\alpha) = \prod_\alpha (f_\alpha, g_\alpha)$$

in the sence of quasi-convergence,

$\square$

C_0-vectors from different strong equivalence classes are always orthogonal and those from the same strong equivalence class are orthogonal if and only if they are orthogonal in at least one tensor product factor.

Theorem J.24.22 $[f] = [g]$ *if and only if both* $\sum_\alpha \|f_\alpha^0 - g_\alpha^0\|_\alpha^2$ *and* $\sum_\alpha |Im(f_\alpha^0, g_\alpha^0)_\alpha|$ *converge*

Proof: In othere words, the combined convergence is equivalent to that of $\sum_\alpha |(f_\alpha^0, g_\alpha^0) - 1|$. Note that if $|z_{\alpha_1}| + |z_{\alpha_2}| + \cdots + |z_{\alpha_n}|$ is bounded then so is $|\zeta_{\alpha_1}| + |\zeta_{\alpha_2}| + \cdots + |\zeta_{\alpha_n}|$ if $|\zeta_{\alpha_1} - z_{\alpha_1}| + |\zeta_{\alpha_2} - z_{\alpha_2}| + \cdots + |\zeta_{\alpha_n} - z_{\alpha_n}|$ is bounded.

Now as $\sum_\alpha \big| \|f_\alpha\| - 1 \big|$ converges if and only if $\sum_\alpha \big| \|f_\alpha\|^2 - 1 \big|$, we can prove the theorem by demonstrating the convergence of

$$\sum_\alpha |(f_\alpha^0, g_\alpha^0) - \frac{1}{2}\|f_\alpha^0\|^2 + \frac{1}{2}\|g_\alpha^0\|^2|.$$

Now

$$\begin{aligned}\mathrm{Re}\big((f^0_\alpha,g^0_\alpha)-\frac{1}{2}\|f^0_\alpha\|^2+\frac{1}{2}\|g^0_\alpha\|^2\big) &= -\frac{1}{2}\big(\|f^0_\alpha\|^2+\|g^0_\alpha\|^2-2\mathrm{Re}(f^0_\alpha,g^0_\alpha)\big)\\ &= -\frac{1}{2}\|f^0_\alpha-g^0_\alpha\|^2_\alpha\end{aligned}$$

and

$$\mathrm{Im}\big((f^0_\alpha,g^0_\alpha)-\frac{1}{2}\|f^0_\alpha\|^2+\frac{1}{2}\|g^0_\alpha\|^2\big)=\mathrm{Im}(f^0_\alpha,g^0_\alpha)$$

Recall that the convergence of $\sum_\alpha z_\alpha$ is equivalent to the combined convergence of $\sum_\alpha |\mathrm{Re} z_\alpha|$, $\sum_\alpha |\mathrm{Im} z_\alpha|$. This completes the proof.

□

Lemma J.24.23 *for each $[f]\in\mathcal{S}$ there exists $f^0\approx f$ such that $\|f^0_\alpha\|_\alpha=1$ for all $\alpha\in\mathcal{I}$.*

Proof: Choose $f\in\mathcal{S}$. As $\sum_\alpha \big|\|f'_\alpha\|_\alpha-1\big|$ converges, apart from a finite number of α's, we have

$$\big|\|f_\alpha\|_\alpha-1\big|\leq\frac{1}{2},\qquad \|f_\alpha\|_\alpha\geq\frac{1}{2}.$$

Recall that $f\approx f'$ if $f_\alpha\neq f'_\alpha$ for a finite number of α's. So we replace f_α by any f'_α such that $1/2\leq\|f'_\alpha\|_\alpha\leq 3/2$ for the exceptional α's.

As $1/x$ is convex, for $1/2\leq x\leq 1$ we have $(1/x)-1\leq 2(1-x)$ and for $1\leq x\leq 3/2$ we have $1-(1/x)\leq 2(x-1)$, that is,

$$\left|\frac{1}{x}-1\right|\leq 2|x-1|$$

for $1/2\leq x\leq 3/2$. So we have

$$\left|\frac{1}{\|f'_\alpha\|_\alpha}-1\right|\leq 2\big|\|f'_\alpha\|_\alpha-1\big|,$$

therefore

$$\sum_\alpha\left|\frac{1}{\|f'_\alpha\|_\alpha}-1\right|$$

converges. By theorem J.24.16 ii), replace $|z_\alpha|$ by $1/\|f'_\alpha\|_\alpha$, as f'_α is a C_0−sequence so is

$$f^0_\alpha = \frac{1}{\|f'_\alpha\|_\alpha} f'_\alpha.$$

By theorem J.24.16 iv), replace z_α by $1/\|f'_\alpha\|_\alpha$, the C_0−sequences f'_α and f^0_α are equivalent. $\|f^0_\alpha\|_\alpha = 1$ is obvious, therefore the proof is complete.

□

By $\otimes_\alpha f_\alpha$ we denote the functional on X defined by

$$\otimes_\alpha f_\alpha(\{x_\alpha\}) = \prod_\alpha (f_\alpha, x_\alpha)_\alpha$$

We need to prove

$$\otimes_\alpha f_\alpha \neq 0 \quad \Leftrightarrow \quad \prod_\alpha \|f_\alpha\|_\alpha \neq 0$$

But $\otimes_\alpha f_\alpha \neq 0$ implies $\prod_\alpha \|f_\alpha\|_\alpha \neq 0$.

J.24.4 ITP Hilbert Space

1) All the $\mathcal{H}^\otimes_{[f]}$ are isomorphic and mutually orthogonal.

2) Every $\mathcal{H}^\otimes_{(f)}$ is the closed direct sum of all the $\mathcal{H}^\otimes_{[f']}$ with $[f'] \in \mathcal{S} \cap (f)$.

3) The ITP $\mathcal{H}^\otimes$ is the closed direct sum of all the $\mathcal{H}^\otimes_{(f)}$ with $(f) \in \mathcal{W}$.

4) Every $\mathcal{H}^\otimes_{[f]}$ has an explicitly known orthonormal von Neumann basis.

5) If s, s' are two different strong equivalence classes in the same weak equivalence class then there exists a unitary operator on $\mathcal{H}^\otimes$ that maps $\mathcal{H}^\otimes_s$ to $\mathcal{H}^\otimes_{s'}$, otherwise such an operator does not exist, the two Hilbert spaces are unitarily inequivalent subspaces of $\mathcal{H}^\otimes$.

Notice that two isomorphic Hilbert spaces can always be mapped into each other such that scalar products are preserved (just map some orthonormal bases) but here the question is whether this map can be extended unitarily to all of $\mathcal{H}^\otimes$. Intuitively then, strong classes within the same weak classes describe the same physics, those in different weak classes describe different physics such as an infinite difference in energy, magnetization, volume etc.

Definition By $\mathcal{H}_C$ we denote the completion of the complex vector space of finite linear combinations of elements from V_C equipped with the sesquilinear form $< \cdot, \cdot >$ obtained by extending (J.438) from V_C to $\mathcal{H}_C$ by sesquilinearity,

$$\begin{aligned} < \alpha\xi + \beta\chi, \zeta > &= \overline{\alpha} < \xi, \zeta > + \overline{\beta} < \chi, \zeta > \\ < \xi, \alpha\chi + \beta\zeta > &= \alpha < \xi, \chi > + \beta < \xi, \zeta > \end{aligned}$$

for all $\xi, \chi, \zeta \in \mathcal{H}_C$ and $\alpha, \beta \in \mathbb{C}$.

□

We can now give the defintion of the ITP.

Definition We denote by

$$\mathcal{H}^{\otimes} := \otimes_\alpha \mathcal{H}_\alpha \tag{J.441}$$

the Cauchy completion of the pre-Hilbert space $\mathcal{H}_C$. It is called the ITP of the $\mathcal{H}_\alpha$.

□

The strong equivalence classes provide the basic tool to analyze the structure of $\mathcal{H}^{\otimes}$.

Definition For a strong equivalence class $[f] \in \mathcal{S}$ we define the closed subspace $\mathcal{H}_{[f]}$ of $\mathcal{H}^{\otimes}$ by the closure of the finite linear combinations of $\otimes_{f'}$'s with $f' \in [f]$, i.e., the closure of

$$\left\{ \sum_{k=1}^{N} z_k \otimes_{f^k} : z_k \in \mathbb{C},\ f^k \in [f],\ N < \infty \right\}. \tag{J.442}$$

It is called the $[f]-$**adic incomplete ITP** of the $\mathcal{H}_\alpha$'s.

□

Theorem J.24.24 *The complete ITP decomposes as the direct sum over strong equivalence classes $[f]$ of $[f]-$adic ITP's,*

$$\mathcal{H}^{\otimes} = \overline{\otimes_{[f] \in \mathcal{S}} \mathcal{H}_{[f]}} \tag{J.443}$$

Proof:

□

J.24.5 Non-Associativity of ITPs

The associative law of tensor products is false. By this we mean the following: Let us subdivide the index set $\mathcal{I}$ into mutually disjoint index sets

$$\mathcal{I} = \cup_\beta \mathcal{I}_\beta$$

where β runs over some other index set $\mathcal{L}$. One can now form the different ITP

$$\mathcal{H}^{\otimes\prime} = \otimes_\beta \mathcal{H}_\beta^\otimes, \qquad \mathcal{H}_\beta^\otimes = \otimes_{\gamma \in \mathcal{I}_\beta} \mathcal{H}_\gamma.$$

Unless the index set $\mathcal{L}$ is finite, a generic C_0-vector of $\mathcal{H}'^\otimes$ is orthogonal to all of $\mathcal{H}^\otimes$.

Scalar multiplication is not multi-linear. That is, if f and $z \cdot f$ are C_0-sequences where

$$(z \cdot f)_\alpha = z_\alpha f_\alpha$$

for some complex numbers z_α then $\otimes_{z \cdot f} = (\prod_\alpha z_\alpha) \otimes_f$.

Lemma J.24.25 *Let $\prod_\alpha z_\alpha$ be quasi-convergent. Then*

i) If f is a $C-$sequence, so is $z \cdot f$ with $(z \cdot f)_\alpha := z_\alpha f_\alpha$.

ii) If moreover $\sum_\alpha \big||z_\alpha| - 1\big|$ converges and f is a C_0-sequence, so is $z \cdot f$.

iii) The product formula

$$\otimes_{z \cdot f} = [\prod_\alpha z_\alpha] \otimes_f \tag{J.444}$$

fails to hold only if

1) $\prod_\alpha z_\alpha$ is (quasi-convergent but) not convergent and

2) $(\otimes_f, \cdot\,) \neq 0$ considered as a linear functional over $C-$vectors.

In this case, $\{z_\alpha\}$, f satisfy the assumptions of ii), and all $z_\alpha \neq 0$.

iv) If $\{z_\alpha\}$, f satisfy the assumptions of ii), then $[z \cdot f] = [f]$ if and only if $\sum_\alpha |z_\alpha - 1|$ converges. If all $z_\alpha \neq 0$, then this is equivalent to the convergence of $\prod_\alpha z_\alpha$ (beyond mere quasi-convergence).

Proof: i): As both $\prod_\alpha |z_\alpha|$, $\prod_\alpha \|f_\alpha\|_\alpha$ converge, so does

$$\prod_{\alpha} \|z_\alpha f_\alpha\|_\alpha = \prod_{\alpha} |z_\alpha| \cdot \prod_{\alpha} \|f_\alpha\|_\alpha$$

ii) $|z_\alpha| \leq C$

$$\begin{aligned} \big|\|z_\alpha f_\alpha\|_\alpha - 1\big| &= \big||z_\alpha| \ \|f_\alpha\|_\alpha - 1\big| \\ &= \big|(|z_\alpha| - 1) + |z_\alpha|(\|f_\alpha\|_\alpha - 1)\big| \\ &\leq \big|(|z_\alpha| - 1) + C(\|f_\alpha\|_\alpha - 1)\big| \end{aligned}$$

and thus $\sum_\alpha \big|\|z_\alpha f_\alpha\|_\alpha - 1\big|$ converges.

$\square$

J.24.6 Von-Neumann Algebras on ITPs

A von Neumann algebra over a Hilbert space is weakly (equivalently strongly) closed sub$-^*$ algebra of the algebra of bounded operators on that Hilbert space.

Given a bounded operator a_α on $\mathcal{H}_\alpha$ (notice that closed unbounded operators have a polar decomposition into a unitary and self-adjoint piece and that the self-adjoint operator is completely determined by its bounded spectral projections so that restriction to bounded operators is no loss of generality) we can extend it in the natural way to $\mathcal{H}^\otimes$ by defining $\hat{a}_\alpha$ densely on C_0-vectors through

$$\hat{a}_\alpha \otimes_f = \otimes_{f'}$$

with

$$f'_{\alpha'} = f_{\alpha'} \quad \text{for} \quad \alpha' \neq \alpha$$

and

$$f'_\alpha = a_\alpha f_\alpha$$

As we will see, it turns out that the algebra of these extended operators for a given label is automatically a von Neumann algebra for $\mathcal{H}^\otimes$ and we call the weak closure of all these algebras the von Neumann algebra $\mathcal{R}^\otimes$ of local operators.

Definition We denote by $\mathcal{B}(\mathcal{H}_\alpha)$ the set of bounded operators on $\mathcal{H}_\alpha$ and by $\mathcal{B}^\otimes := \mathcal{B}(\mathcal{H}^\otimes)$ the set of bounded operators on the ITP $\mathcal{H}^\otimes$.

□

Definition We denote by $\mathcal{B}_\alpha$ the extension of $\mathcal{B}(\mathcal{H}_\alpha)$ to the ITP, that is,

$$\mathcal{B}_\alpha = \{\hat{A}_\alpha : A_\alpha \in \mathcal{B}(\mathcal{H}_\alpha)\} \tag{J.445}$$

□

Lemma J.24.26 *For all $\alpha \in \mathcal{I}$, the algebra $\mathcal{B}_\alpha$ is a von Neumann algebra over $\mathcal{H}^\otimes$.*

Proof: To prove the theorem one uses the von Neumann density theorem, proved in the next chapter. This states that a set $\mathcal{U}$ of the space of bounded operators of a Hilbert space $\mathcal{H}$ is a von Neuman algrba if it is equal to its double commutant, the commutant of a set $\mathcal{U} \subset \mathcal{B}(\mathcal{H})$ being

$$\mathcal{U}' = \{B \in \mathcal{B}(\mathcal{H}) : [A, B] = 0 \text{ for all } A \in \mathcal{U}\},$$

Let us write

$$\mathcal{B}^\otimes = \mathcal{B}(\mathcal{H}_\alpha \otimes \mathcal{H}_{\overline{\alpha}})$$

where $\overline{\alpha} = \mathcal{I} - \alpha$.

$$\mathcal{B}'_{\overline{\alpha}} = \{\hat{B} \in \mathcal{B}^\otimes : [\hat{A}, \hat{B}] = 0 \text{ for all } \hat{A} \in \mathcal{B}_{\overline{\alpha}}\},$$

the commutant of $\mathcal{B}_{\overline{\alpha}}$.

□

Definition Two C_0-sequences f, g are said to be weakly equivalent, denoted $f \sim g$ provided that there are complex numbers z_α such that $z \cdot f$ and g are strongly equivalent, that is, $z \cdot f \approx g$.

□

Lemma J.24.27 *The definition of weak equivalenceis unchanged if we restrict to complex numbers with* $|z_\alpha| = 1$.

Proof: We prove that if f and $z \cdot f$ are C_0-sequences we can find z'_α with $|z'_\alpha| = 1$ such that $(z \cdot f) \approx (z' \cdot f)$.

As $\sum_\alpha \left| \|z_\alpha f_\alpha\|_\alpha - 1 \right|$ converges, $z_\alpha f_\alpha = 0$ can occur only for a finite number of α's. For these we may replace z_α by 1 and f_α by non-zero f^0_α. As this so constructed C_0-sequence differes from f by only a finite number of terms, it is srtongly equivalent to it (this follows from theorem J.24.22). So we may assume that $z_\alpha f_\alpha \neq 0$ for all α.

As $\sum_\alpha \left| \|f_\alpha\|_\alpha - 1 \right|$, $\sum_\alpha \left| \|z_\alpha f_\alpha\|_\alpha - 1 \right|$ converge, and all $\|f_\alpha\|_\alpha, \|z_\alpha f_\alpha\|_\alpha \neq 0$, by theorem J.24.16, $\prod_\alpha \|f_\alpha\|_\alpha$, $\prod_\alpha \|z_\alpha f_\alpha\|_\alpha$ converge and have non-zero values. Thus

$$\prod_\alpha \frac{1}{|z_\alpha|}$$

converges too, and has a non-zero value, as

$$\frac{\|f_\alpha\|_\alpha}{\|z_\alpha f_\alpha\|_\alpha} = \frac{1}{|z_\alpha|}.$$

Therefore

$$\sum_\alpha \left| \frac{1}{|z_\alpha|} - 1 \right|$$

converges by theorem J.24.16. Now it follows from lemma, (ii), (iv), that $(z \cdot f)$ and $(\frac{z}{|z|} \cdot f)$ are equivalent. Thus

$$z'_\alpha = \frac{z_\alpha}{|z_\alpha|}.$$

Lemma J.24.28 *Weak equivalence is an equivalence relation.*

Proof: Reflextivity: $f \sim f$ as $(1 \cdot f) \approx f$.

Symmetry: If $|z_\alpha| = 1$ we have

$$(f_\alpha, (1/z_\alpha) g_\alpha)_\alpha = z_\alpha (f_\alpha, g_\alpha)_\alpha = (z_\alpha f_\alpha, g_\alpha)_\alpha$$

and $\|(1/z_\alpha)g_\alpha\|_\alpha = \|g_\alpha\|_\alpha$. These together with the symmetry property of strong equivalence implies that if $(z \cdot f) \approx g$ with $|z_\alpha| = 1$ then $(1/z) \cdot g \approx f$ with $(1/z_\alpha)g_\alpha$ a C_0-sequence, that is, $g \sim f$.

Transitivity: If $z \cdot f \approx g$ and $z' \cdot g \approx h$, $|z_\alpha|, |z'_\alpha| = 1$, then $z \cdot f \approx (1/z') \cdot h$ and $z' \cdot z \cdot f \approx h$. So $f \sim g$.

□

Lemma J.24.29 *f and g are weakly equivalent if and only if*

$$\sum_\alpha \big||(f_\alpha, g_\alpha)_\alpha| - 1\big|$$

converges.

Proof: By definition, $f \sim g$ if we can find complex numbers z_α with $|z_\alpha| = 1$ such that

$$\sum_\alpha \big|(z_\alpha f_\alpha, g_\alpha)_\alpha - 1\big|$$

converges. If

$$\sum_\alpha \big||(f_\alpha, g_\alpha)_\alpha| - 1\big|$$

converges, then $f \sim g$ with $z_\alpha = (f_\alpha, g_\alpha)^*_\alpha / |(f_\alpha, g_\alpha)_\alpha|$.

Now, from fig (), it is easy to see that the minimum value of $|(z_\alpha f_\alpha, g_\alpha) - 1| = |z_\alpha (f_\alpha, g_\alpha)_\alpha - 1|$ is $\big||(f_\alpha, g_\alpha)_\alpha| - 1\big|$. Thus $\sum_\alpha \big|(z_\alpha f_\alpha, g_\alpha)_\alpha - 1\big|$ converges for some complex numbers z_α, when $f \sim g$, only if

$$\sum_\alpha \big||(f_\alpha, g_\alpha)_\alpha| - 1\big|$$

converges.

□

Lemma J.24.30 *Assume that a z_α with $|z_\alpha| = 1$ is given for each $\alpha \in \mathcal{I}$. Then there exists one and only one closed, linear operator U, such that*

$$U \otimes_f = \otimes_{z \cdot f}$$

for every C_0-vector f. The operator U is unitary.

Proof:

Proof of uniqueness. If U' is another closed, linear operator meeting the requirements, then

$\square$

J.25 Biblioliographical notes

In this chapter I have relied on the following references:

Linear operators for Quantum Mechanics by Thomas F. Jordan.

Functional Analysis by Reed and Simon.

In proof of Tynchonov's Theorem I used J.L. Bell General topology [406].

An introduction to Hilbert space by N. Young.

To do with nets very good introduction that can be downloaded called "Topological spaces", King's College, London. http://www.mth.kcl.ac.uk/ iwilde/notes/fa2/fa2.pdf. [408].

T. Thiemann's notes "Introduction to Modern Canonical Quantum General Relativity" [32] and "Lectures of Loop Quantum Gravity" [33].

J.26 Worked Exercises and Details

Details: Numerical range of a form is convex.

Comming from Martin Schechter "princples of functional analysis".

Proof of Theorem J.18.25 Any linear combination of u and v is also in $D(a)$. We will show that for $w = \alpha e^{i\varphi}u + \beta$, $a(w)$ satisfies Eq.(J.372), where α, β, φ are real numbers to be determined.

$$\text{Im}\gamma a(u) = \text{Im}\gamma a(v) = d. \tag{J.446}$$

$$\begin{aligned} g(\alpha,\beta) &= \gamma a(\alpha e^{i\varphi}u + \beta v) - id\,\|\alpha e^{i\varphi}u + \beta v\|^2 && \text{(J.447)} \\ &= \alpha^2[\gamma a(u) - id] + \beta^2[\gamma a(v) - id] \\ &+ \alpha\beta e^{i\varphi}[\gamma a(u,v) - (u,v)id] + \alpha\beta e^{-i\varphi}[\gamma a(v,u) - (v,u)id] && \text{(J.448)} \end{aligned}$$

Now given any two complex number z and z', we can find φ such that $\mathrm{Im}(ze^{i\varphi} + e^{-i\varphi}z') = 0$.[1] We will pick φ so that

$$\mathrm{Im}\{e^{i\varphi}[\gamma a(u,v) - (u,v)id] + e^{-i\varphi}[\gamma a(v,u) - (v,u)id]\} = 0$$

We see form $g(\alpha, \beta)$ is real. Now if

$$\alpha e^{i\varphi}u + \beta v = 0, \tag{J.449}$$

then we have

$$\|\alpha e^{i\varphi}u\|^2 = \beta v\|^2$$

and

$$a(\alpha e^{i\varphi}u) = a(\beta v).$$

The first implies $\alpha^2 = \beta^2$, while the second gives

$$\alpha^2 a(u) = \beta^2 a(v).$$

Since we are assuming $a(u) \neq a(v)$, the only way (J.449) can hold is if $\alpha = \beta = 0$. From this, we see that the function

$$h(t) = \frac{g(t, 1-t)}{\|te^{i\varphi}u + (1-t)v\|^2} \tag{J.450}$$

is contiuous and real valued in $0 \leq t \leq 1$. Moreover,

$$h(0) = \gamma a(v) - id, \qquad h(1) = \gamma a(u) - id.$$

Hence, there is a value t_1 satisfying $0 < t_1 < 1$ such that (see Fig(J.26))

$$\begin{aligned} h(t_1) &= \tau[\gamma a(v) - id] + (1-\tau)[\gamma a(u) - id] \\ &= \gamma[\tau a(v) + (1-\tau)a(u)] - id \end{aligned} \tag{J.451}$$

Now equation (J.447) tells us

[1] $\mathrm{Im}(|z|e^{i(\theta+\varphi)} + |z'|e^{i(\theta'-\varphi)}) = |z|\sin(\theta+\varphi) + |z'|\sin(\theta'-\varphi) = 0$ so that $-\frac{|z'|}{|z|} = \frac{\sin(\theta+\varphi)}{\sin(\theta'-\varphi)}$

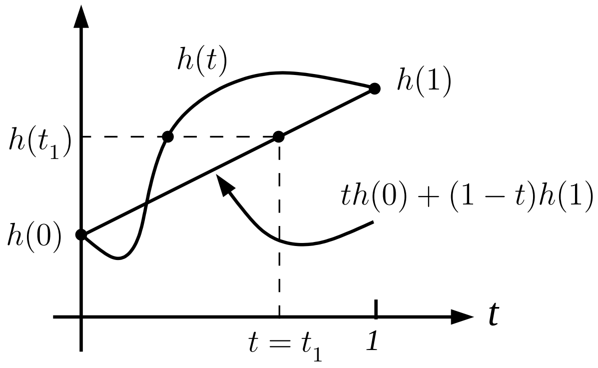

Figure J.31: NumConvex.

$$g(t_1, 1-t_1) = \gamma a(t_1 e^{i\varphi} u + (1-t_1)v) - id\ \|t_1 e^{i\varphi} u + (1-t_1)v\|^2$$

Dividing both sides by $\|t_1 e^{i\varphi} u + (1-t_1)v\|^2$

$$\begin{aligned} h(t_1) &= \frac{a(t_1 e^{i\varphi} u + (1-t_1)v)}{\|t_1 e^{i\varphi} u + (1-t_1)v\|^2} - id \\ &= \gamma a(w) - id. \end{aligned} \tag{J.452}$$

where

$$w = \frac{t_1 e^{i\varphi} u + (1-t_1)v)}{\|t_1 e^{i\varphi} u + (1-t_1)v\|^2}. \tag{J.453}$$

Equating (J.452) with (J.451) gives (J.372).

□

Details: Caratheodory's outer measure.

Prove that, under condition (N.-19), that $\mu(E) = \inf\{\sum_{j=1}^{\infty} \mu(F_j) \mid E \subset \bigcup_{j=1}^{\infty} F_j\}$ becomes Caratheodory's outer measure.

Proof:

A non-negative function on the power set. Monoticity: Let $E_1 \subset E_2$,

the infimum of a smaller set cannot be less than the infimum of a larger set $\Rightarrow \mu(E_1) \leq \mu(E_2)$.

subadditivity

$$\mu(\cup_i E_i) = \inf\{\sum_{j=1}^{\infty} \mu(F_j) \mid \cup_i E_i \subset \bigcup_{i=1}^{\infty} F_i\} \tag{J.454}$$

$$\sum_i \mu(E_i) = \sum_i \inf\{\sum_{i=1}^{\infty} \mu(F_j) \mid E_i \subset \bigcup_{j=1}^{\infty} F_j\} \tag{J.455}$$

Details: Axiom of choice and Zorn's Lemma.

Proof: