PURE AND APPLIED MATHEMATICS

A Series of Monographs and Textbooks

AN INTRODUCTION TO FUNCTIONAL ANALYSIS

Charles Swartz

AN INTRODUCTION TO FUNCTIONAL ANALYSIS

Charles Swartz
New Mexico State University
Las Cruces, New Mexico

Marcel Dekker, Inc. New York • Basel • Hong Kong

Library of Congress Cataloging-in-Publication Data

Swartz, Charles.
Functional analysis / Charles Swartz.
p. cm.
Includes bibliographical references and index.
ISBN 0-8247-8643-2
1. Functional analysis. II. Title.
QA320.S9 1992
515'.7--dc20 91-44028
CIP

This book is printed on acid-free paper.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

Current printing (last digit):
10 9 8 7 6 5 4 3 2 1

PRINTED IN THE UNITED STATES OF AMERICA

TO

NITA RUTH

Preface

These notes evolved from the introductory functional analysis course given at New Mexico State University. In the course, we attempted to cover enough of the topics in functional analysis in depth and to give a sufficient number of applications to give the students a feel for the subject. Of course, the choice of topics reflects the interests and prejudices of the author and many important and interesting topics are not covered at all. We have given references for further study at many points. The students in this course were assumed to have had an introductory course in point set topology and a course in measure and integration. The text assumes a knowledge of elementary point set topology including Tychonoff's Theorem and the theory of nets and a background in real analysis equivalent to an introductory course from a text such as [Ro] or [AB]; a few basic results from complex analysis are required, especially in the sections on spectral theory. The reader is also assumed to be familiar with Hilbert space; an appendix supplies the basic properties of Hilbert spaces which are required.

Functional analysis is a subject which evolved from abstractions of situations which repeatedly occurred in concrete function spaces, differential and integral equations, calculus of variations, etc. The abstractions have repeatedly repaid their debts to its foundations with multitudes of applications. We have made an attempt to give at least enough of these applications to indicate the interplay between the abstract and the concrete and illustrate the beauty and usefulness of the subject.

The first section of the book introduces the basic properties of topological vector spaces which will subsequently be used. In the second section the three basic principles of functional analysis -- the Hahn-Banach Theorem, the Uniform Boundedness Principle and the Open Mapping/ Closed Graph Theorems -- are established. Applications to such topics as Fourier series, measure theory, summability and Schauder basis are given. Locally convex spaces are studied in the third section. Topics such as duality and polar topologies are discussed and applications to topics such as Liapounoff's Theorem, the Stone-Weierstrass Theorem and the Orlicz-Pettis Theorem are given. We do not present an exhaustive study of the enormously large subject of locally convex spaces, but rather try to present the most important and useful topics and particularly stress the important special case of normed linear spaces. In the fourth section we study the basic properties of linear operators between topological vector spaces. The Uniform Boundedness Principle for barrelled spaces, the basic properties of the transpose or adjoint operator, and the special classes of projection, compact, weakly compact and absolutely summing operators are discussed.

Applications to such topics as the integration of vector-valued functions, Schwartz distributions, the Fredholm Alternative and the Dvoretsky-Rogers Theorem are given. Finally, in section five we discuss spectral theory for continuous linear operators. We first establish the spectral theorem for compact symmetric operators and use this result to motivate the spectral theorem for continuous symmetric (Hermitian) operators. A simple straightforward proof of this version of the spectral theorem is given using the Gelfand mapping. We derive the spectral theorem for normal operators from basic results from Banach algebras; this affords the students an introduction to the subject of Banach algebras. Applications to such topics as Lomonosov's Theorem on the existence of invariant subspaces for compact operators, Sturm-Liouville differential equations, Hilbert-Schmidt operators and Wiener's Theorem are given.

These notes cover many of the basic topics presented in the classic texts on functional analysis such as [TL], [Ru1], [C1], etc., although some of the topics covered in the applications such as the Nikodym Boundedness and Convergence Theorems, the Pettis and Bochner integrals, summability and the Orlicz-Pettis Theorem are often not treated in introductory texts. We do, however, include some material that is not found in the standard introductory texts. We use the stronger notion of $\mathcal{K}$ convergence of sequences and $\mathcal{K}$ bounded sets developed by the members of the Katowice branch of the Mathematics Institute of the Polish Academy of Sciences under the direction of Piotr Antosik and Jan Mikusinski to establish versions of the Uniform Boundedness Principle which require no completeness or barrelledness assumptions on the domain space of the operators. It is shown that the general version of the Uniform Boundedness Principle yields the classic version of the Uniform Boundedness Principle for complete metric linear spaces as a corollary and the general Uniform Boundedness Principle is employed to obtain a version of the Nikodym Boundedness Theorem from measure theory where the classic version of the Uniform Boundedness Principle is not applicable. These results are obtained from a beautiful result due to Antosik and Mikusinski on the convergence of the elements on the diagonal of an infinite matrix with entries in a topological vector space. The matrix theorem can be viewed as an abstract "sliding hump" technique and is used in many proofs in the text. The notion of $\mathcal{K}$ bounded sets is also used to obtain hypocontinuity results for bilinear maps which require no completeness or barrelledness assumptions.

I would like to thank the many graduate students at New Mexico State University who took the introductory functional analysis course from me during the time that these notes were evolving into their final form and especially Chris Stuart, who read through the final manuscript. Special thanks go to my colleague and friend Piotr Antosik, who introduced me to the matrix methods that are employed throughout this text, and to Valerie Reed, who did such a superb job of translating my handwriting from the original class notes.

Charles Swartz

Contents

AN INTRODUCTION TO FUNCTIONAL ANALYSIS

Part I
Topological Vector Spaces (TVS)

1

Definition and Basic Properties

In this chapter we give the definition of a topological vector space and develop the basic properties of such spaces.

Let X be a vector space (VS) over the field $\mathbb{F}$, where $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$.

Definition 1. (X, τ) is a topological vector space (TVS) is τ is a topology on X such that:

(i) $(x, y) \longrightarrow x + y$ is continuous from $X \times X \to X$ and

(ii) $(t, x) \longrightarrow tx$ is continuous from $\mathbb{F} \times X \to X$,

that is, the topology τ is compatible with the algebraic structure of X. For brevity, we often say that X is a TVS and suppress the topology τ. Any topology τ on X such that (X, τ) is a TVS is called a vector topology.

Proposition 2. Let (X, τ) be a TVS.

(i) If $t_0 \in \mathbb{F}$, $x_0 \in X$ then the map $x \to t_0x + x_0$ is continuous from X into X; if $t_0 \neq 0$, the map is a homeomorphism from X onto X.

(ii) If $\mathcal{U}$ is a (open) neighborhood base of 0 in X, then $x_0 + \mathcal{U} = \{x_0 + U : U \in \mathcal{U}\}$ is a open neighborhood base at x_0.

(iii) If U is (open), tU is open $\forall\, t \in \mathbb{F}$, $t \neq 0$.

(iv) If $\mathcal{U}$ is a neighborhood base at 0 and $U \in \mathcal{U}$, then $\exists\, V \in \mathcal{U}$ such that $V + V \subseteq U$.

Proof: The first part of (i) follows from Definition 1 and the fact that joint continuity implies separate continuity. The second part follows from the first part and the fact that the map $x \to (x - x_0)/t_0$ is the inverse of the map in (i).

(ii) and (iii) follow from (i), and (iv) follows from the continuity of addition and the fact that $0 + 0 = 0$.

Definition 3. A subset $S \subseteq X$ is <u>symmetric</u> if $x \in S \Rightarrow -x \in S$ (i.e., $S = -S$).

Remark 4. Note that every neighborhood U of 0 in a TVS contains a symmetric neighborhood of 0 (take $U \cap (-U)$).

Definition 5. A subset $S \subseteq X$ is <u>balanced</u> if $x \in S$ and $t \in \mathbb{F}$, $|t| \leq 1$, $\Rightarrow$ $tx \in S$ (i.e., $S = DS$ where $D = \{t \in \mathbb{F} : |t| \leq 1\}$).

Remark 6. Note that a balanced set is symmetric and contains 0.

Definition 7. If $S \subseteq X$, we set $\text{bal}(S) = DS$. bal(S) is the smallest balanced set containing S and is called the balanced hull of S.

Proposition 8. Let X be a TVS.

(i) If U is a (open) neighborhood of 0, then bal(U) is a (open) neighborhood of 0.

(ii) If S is balanced, then $\bar{S}$ is balanced.

Proof: (i) follows from Proposition 2 (iii). For (ii) note that $D \times S$ is mapped into S by the continuous map, $(t, x) \to tx$ so $D \times \bar{S}$ is mapped into $\bar{S}$ and $\bar{S}$ is balanced.

Theorem 9. If $\mathcal{U}$ is a (open) neighborhood base of 0 in a TVS, X, and $\mathcal{V}$ is the family of balanced hulls of members of $\mathcal{U}$, then $\mathcal{V}$ is a (open) neighborhood base at 0. That is, $\exists$ a base at 0 consisting of balanced sets.

Proof: Let $U \in \mathcal{U}$. Since $0 \cdot 0 = 0$ and scalar multiplication is continuous, $\exists$ a (open) neighborhood of 0, W_1, whose balanced hull W_2 is contained in U. Now W_2 is a (open) neighborhood of 0 by Proposition 8 so $\exists\, U_1 \in \mathcal{U}$ such that $U_1 \subseteq W_2$. If $V = \text{bal}(U_1)$, then $V \subseteq W_2 \subseteq U$ and $V \in \mathcal{V}$.

Corollary 10. Let (X, τ) be a TVS. The following are equivalent:

(i) τ is Hausdorff, i.e., (X, τ) is a Hausdorff TVS.

(ii) $\overline{\{0\}} = \{0\}$

(iii) If $x_0 \neq 0$, then $\exists$ a neighborhood U of 0 such that $x_0 \notin U$.

Proof: (i) $\Rightarrow$ (ii) $\Rightarrow$ (iii) is clear. Suppose (iii) holds and $x, y \in X$ with $x \neq y$. There exists a neighborhood of 0, U, such that $x - y \notin U$. By Theorem 9 and Proposition 2, $\exists$ a balanced neighborhood V of 0 such that $V + V \subseteq U$. Thus, $(x + V) \cap (y + V) = \emptyset$ and (i) holds.

Proposition 11. If X is a TVS, then X has a neighborhood base at 0 which consists of closed sets.

Proof: Let U be a neighborhood of 0. Let V be a neighborhood of 0 such that $V - V \subseteq U$ (Theorem 9). Then $\bar{V} \subseteq U$; for if $x \in \bar{V}$, then every neighborhood of x meets V so $(x + V) \cap V \neq \emptyset$ which implies $\exists\, y, z \in V$ such that $y + x = z$ or $x = z - y \in V - V \subseteq U$.

Theorem 12. If X is a TVS, then there is a base at 0 which consists of closed, balanced sets.

Proof: Every neighborhood of 0, U, contains a closed neighborhood of 0, V, which contains a balanced neighborhood of 0, W.

Then $\bar{W} \subseteq V \subseteq U$ and $\bar{W}$ is closed and balanced by Proposition 7.

Definition 13. Let X be a TVS and $A, B \subseteq X$. Then B <u>absorbs</u> A (or A is <u>absorbed</u> by B) if there exists $t_0 > 0$ such that $|t| \geq t_0$ implies $tB \supseteq A$ (or, equivalently, if $\exists\ \varepsilon > 0$ such that $B \supseteq tA$ $\forall |t| \leq \varepsilon, t \neq 0$). A subset $S \subseteq X$ is absorbing if it absorbs all singletons.

Proposition 14. A balanced set S is absorbing if and only if $\forall\ x \in X$ $\exists\ t \neq 0$ such that $tx \in S$.

Proof: Exercise.

Proposition 15. Every neighborhood of 0 in a TVS is absorbing.

Proof: This follows from the continuity of the scalar product and the fact that $0 \cdot x = 0\ \forall x$.

Theorem 16. Let X be a TVS. $\exists$ an open neighborhood base at 0, $\mathcal{U}$ satisfying:

(i) each $U \in \mathcal{U}$ is balanced and absorbing,

(ii) if $U \in \mathcal{U}$, $t \neq 0$, then $tU \in \mathcal{U}$,

(iii) if $U \in \mathcal{U}$, $\exists\ V \in \mathcal{U}$ such that $V + V \subseteq U$.

Proof: Let $\mathcal{U}$ be the family of all balanced, open neighborhoods of

0. This is a neighborhood base by Theorem 9 and (i) holds by Proposition 15 while (ii) and (iii) hold by Proposition 2.

We now show conversely that if a family of subsets satisfies the conditions of Theorem 16, it forms a neighborhood base for a vector topology. This will give a convenient means of constructing vector topologies.

Theorem 17. Let X be a vector space. Let $\mathcal{U} \neq \emptyset$ be a family of non-void subsets of X such that:

(i) each $U \in \mathcal{U}$ is balanced and absorbing,

(ii) if $U \in \mathcal{U}$, $\exists\, V \in \mathcal{U}$ such that $V + V \subseteq U$,

(iii) if $U_1, U_2 \in \mathcal{U}$, then $\exists\, U_3 \in \mathcal{U}$ such that $U_3 \subseteq U_1 \cap U_2$,

(iv) if $U \in \mathcal{U}$ and $x \in U$, $\exists\, V \in \mathcal{U}$ such that $x + V \subseteq U$.

Then $\exists$ a unique topology τ on X such that (X, τ) is a TVS with $\mathcal{U}$ an open neighborhood base at 0 with respect to τ.

Proof: We define a topology τ be defining a non-void set V to be open if $\forall x \in V\ \exists\, U \in \mathcal{U}$ such that $x + U \subseteq V$. Condition (iii) implies that the intersection of open sets is open so τ is a topology on X, and condition (iv) implies that each element of $\mathcal{U}$ is open.

Condition (ii) implies that addition is continuous.

To show that scalar multiplication is continuous, first observe that: if $U \in \mathcal{U}$ and $t \neq 0$, $\exists\, V \in \mathcal{U}$ such that $tV \subseteq U$ [from (ii) and induction, if $U \in \mathcal{U}$ and n is a positive integer, $\exists\, V \in \mathcal{U}$ such that $2^n V \subseteq U$. Now choose n such that $|t| < 2^n$ and $V \in \mathcal{U}$ such that $2^n V \subseteq U$. Since U is balanced, $tV \subseteq U$].

Now suppose that $t_0 \in \mathbb{F}$, $x_0 \in X$ and let $U \in \mathcal{U}$. To show that scalar multiplication is continuous, it suffices to show $\exists\ \varepsilon > 0$, $V \in \mathcal{U}$ such that $tx \in t_0x_0 + U$ if $|t - t_0| < \varepsilon$, $x \in x_0 + V$. Choose $W \in \mathcal{U}$ such that $W + W + W \subseteq U$ [by (ii) and $0 \in W$]. If $t_0 = 0$, set $V = W$; if $t_0 \neq 0$, choose $W_1 \in \mathcal{U}$ such that $t_0W_1 \subseteq W$ [by the observation above] and then choose $V \in \mathcal{U}$ such that $V \subseteq W \cap W_1$ [(iii)]. By (i) $\exists\ \varepsilon > 0$ such that $sx_0 \in V$ if $0 < |s| < \varepsilon$ where we may assume $\varepsilon \leq 1$. Now suppose $|t - t_0| < \varepsilon$ and $x \in x_0 + V$. Then $(t - t_0)x_0 \in V$ and $(t - t_0)(x - x_0) \in V$ since V is balanced. Also $t_0(x - x_0) \in t_0V \subseteq W$. From

$$tx - t_0x_0 = (t - t_0)x_0 + t_0(x - x_0) + (t - t_0)(x - x_0),$$

it follows that $tx - t_0x_0 \in W + W + W \subseteq U$, and scalar multiplication is continuous.

The uniqueness of τ is clear.

A similar result is given by

Theorem 18. Let X be a vector space. Let $\mathcal{U}$ be a family of subsets of X satisfying:

(i) each $U \in \mathcal{U}$ is balanced and absorbing,

(ii) if $U_1, U_2 \in \mathcal{U}$, $\exists\ U_3 \in \mathcal{U}$ such that $U_3 \subseteq U_1 \cap U_2$,

(iii) if $U_1 \in \mathcal{U}$, $\exists\ U_2 \in \mathcal{U}$ such that $U_2 + U_2 \subseteq U_1$.

Then $\exists$ a unique topology τ on X such that (X, τ) is a TVS and $\mathcal{U}$ is a neighborhood base at 0 (not necessarily of open sets) for τ.

Proof: Define a subset $V \neq \emptyset$ to be τ-open if $\forall x \in V$, $\exists\, U \in \mathcal{U}$ such that $x + U \subseteq V$. As in Theorem 17 this defines a vector topology τ on X and $\mathcal{U}$ is a neighborhood base at 0 for τ. In general, the elements of $\mathcal{U}$ are not τ-open [assumption (iv) in Theorem 17 guarantees this].

We give several examples of TVS. More examples are given in §2.

Example 19. Let X be a vector space and $\mathcal{F}$ a family of linear maps $f : X \to Y_f$, where Y_f is a TVS $\forall f \in \mathcal{F}$. Let <u>$w(\mathcal{F})$</u> be the weakest topology on X such that each $f \in \mathcal{F}$ is continuous. A net $\{x_\delta\}$ in X converges to $x \in X$ if and only if $f(x_\delta) \to f(x)$ in Y_f $\forall f \in \mathcal{F}$. Thus, $w(\mathcal{F})$ is a vector topology on X.

As a special case we have,

Example 20. The product of TVS's is a TVS (under the product topology).

Example 21. Let Φ be a collection of vector topologies on the vector space X. The sup topology, $\vee \Phi$, or sup Φ, on X generated by Φ is the weakest topology on X which is stronger than each topology in Φ. A net $\{x_\delta\}$ in X converges to $x \in X$ in sup Φ if and only if $x_\delta \to x$ for each $\tau \in \Phi$. Thus, sup Φ is a vector topology on X.

Completeness:

A net $\{x_\delta\}$ in a TVS X is a Cauchy net if for every neighborhood of 0, U, in X $\exists\ \delta$ such that $\alpha, \beta \geq \delta$ implies $x_\alpha - x_\beta \in U$. A subset A of a TVS is said to be complete (sequentially complete) if every Cauchy net (sequence) in A converges to a point in A.

It is the case that every Hausdorff TVS X has a completion, i.e., there is a complete Hausdorff TVS $\hat{X}$ which contains X as a dense subspace. We do not need this general result and refer the reader to [H1] 2.9. See, however, 8.1.14 and §13.

√ Exercise 1. Show that if X is a Hausdorff TVS and $K \subseteq X$ is compact, then bal(K) is compact.

√ Exercise 2. Show that bal(S) can fail to be closed even when S is closed. (Hint: Consider $xy = 1$ in $\mathbb{R}^2$.)

Exercise 3. If Z is a TVS, show a map $g : Z \to X$ is continuous with respect to $w(\mathcal{F})$ if and only if $f \circ g : Z \to Y_f$ is continuous $\forall f \in \mathcal{F}$.

Exercise 4. If Z is a TVS, show that a map $f : Z \to X$ is continuous with respect to sup Φ if and only if it is continuous with respect to $\tau\ \forall\ \tau \in \Phi$.

Exercise 5. If L is a linear subspace of a TVS, show $\bar{L}$ is a linear subspace.

Exercise 6. If L is a linear subspace of a TVS X which contains an open set, show L = X.

Exercise 7. Show that any compact subset of a TVS is complete.

Exercise 8. Show that a complete subset of a TVS is closed, and a closed subset of a complete TVS is complete.

2

Quasi-normed and Normed Linear Spaces (NLS)

In this chapter we study two important examples of TVS, the quasi-normed and normed linear spaces.

Let X be a vector space.

Definition 1. A quasi-norm on X is a function $|\ |: X \to \mathbb{R}$ satisfying

(i) $|0| = 0$,

(ii) $|x| \geq 0 \ \forall x \in X$,

(iii) $|-x| = |x| \quad \forall x \in X$,

(iv) $|x + y| \leq |x| + |y| \ \forall x, y \in X$ (triangle inequality),

(v) if $t_k, t \in \mathbb{F}$, $|t_k - t| \to 0$ and $x_k, x \in X$, $|x_k - x| \to 0$, then $|t_k x_k - tx| \to 0$.

If the quasi-norm satisfies $|x| = 0$ if and only if $x = 0$, then it is said to be total.

We will show in 9.1.4 that axiom (V) can be replaced by a weaker condition.

Example 2. Let p be a quasi-norm on X. Then $q(x) = p(x)/(1 + p(x))$ also defines a quasi-norm on X. [The triangle inequality follows from the fact that the function $h(t) = t/(1 + t)$ is increasing for $t \geq 0$;

$$q(x + y) \leq (p(x) + p(y))/(1 + p(x) + p(y)) \leq q(x) + q(y).]$$

Example 3. Let $\{p_n\}$ be a sequence of quasi-norms on X. Then $p(x) = \sum_{n=1}^{\infty} p_n(x)/(2^n(1 + p_n(x)))$ defines a quasi-norm on X called the <u>Frechet quasi-norm</u> generated by $\{p_n\}$. [The triangle inequality follows from Example 2.] We show that a sequence $\{x_k\}$ in X has the property that $p(x_k) \to 0$ if and only if $\lim_k p_n(x_k) = 0$ for each n. First, if $p(x_k) \to 0$, then $p_n(x_k) \to 0$ as $k \to \infty$ since $p(x) \geq p_n(x)/2^n$ for each n. Conversely, suppose that $\lim_k p_n(x_k) = 0$ for each n. Let $\varepsilon > 0$. There exists N such that $\sum_{n=N}^{\infty} 1/2^n < \varepsilon/2$. There exists N_1 such that $k \geq N_1$ implies $p_n(x_k) < \varepsilon/2N$ for $n = 1, \ldots, N - 1$. Thus, $k \geq N_1$ implies

$$p(x_k) \leq \sum_{n=1}^{N-1} \varepsilon/2N + \sum_{n=N}^{\infty} 1/2^n < \varepsilon .$$

In particular, this shows that (v) holds.

If $|\ |$ is a quasi-norm on X, then $d(x, y) = |x - y|$ defines a semi- metric on X which is a metric if and only if $|\ |$ is total. The semi-metric d induced by $|\ |$ is also <u>translation invariant</u> in the sense that $d(x + z, y + z) = d(x, y)$ for all $x, y, z \in X$. A <u>quasi-normed space</u> (QNLS) is a pair $(X, |\ |)$ where $|\ |$ is a quasi-norm on X. If the quasi-norm is understood, we often say that X is a quasi-normed space. Note that a quasi-normed space is a TVS under the metric topology induced by the quasi- norm (the continuity of addition follows from the triangle inequality.) We always assume that a quasi-normed space carries this metric topology.

Proposition 4. Let (X, p) be a quasi-normed space. Then

(i) $|p(x) - p(y)| \leq p(x - y)\ \forall x, y \in X$,

(ii) the map $p : X \to \mathbb{R}$ is continuous.

Proof: $p(x) \leq p(x - y) + p(y)$ implies $p(x) - p(y) \leq p(x - y)$. By symmetry and 1.(iii), $p(y) - p(x) \leq p(y - x) = p(x - y)$, and (i) is established. (ii) follows from (i).

The convergence in a quasi-normed space has an interesting property which is useful and will now be described. A sequence $\{x_j\}$ in a TVS X is said to be <u>Mackey convergent</u> to 0 (<u>locally null</u>) if $\exists$ a sequence of positive scalars $t_j \to \infty$ such that $t_j x_j \to 0$. We now show that in a quasi-normed space every sequence which converges to 0 is Mackey convergent to 0.

Proposition 5. Let $(X, |\ |)$ be a quasi-normed space and $x_j \to 0$ in X.

Then $\exists$ an increasing sequence of non-negative integers $k_j \to \infty$ such that $k_j x_j \to 0$.

Proof: Set $n_1 = 0$. $\exists$ an increasing sequence of positive integers $\{n_k\}$ such that $|x_j| \le 1/k^2$ for $j \ge n_k$. For $n_k \le j < n_{k+1}$, put $k_j = k$ so $|k_j x_j| \le 1/k$ and $k_j x_j \to 0$.

For an example of a sequence which converges to 0 but is not Mackey convergent to 0 see Exercise 14.10.

Definition 6. A <u>semi-norm</u> on X is a function $\| \, \| : X \to \mathbb{R}$ satisfying

(i) $\|x\| \ge 0 \ \forall x \in X$,

(ii) $\|tx\| = |t| \|x\| \ \ \forall t \in \mathbb{F}, x \in X$

(iii) $\|x + y\| \le \|x\| + \|y\| \ \ \forall x, y \in X$ (triangle inequality).

A semi-norm is called a <u>norm</u> if it satisfies the further condition,

(iv) $\|x\| = 0$ if and only if $x = 0$.

Note from (ii) it follows that $\|0\| = 0$, and we show below in Proposition 7 that any semi-norm is a quasi-norm.

A <u>semi-normed linear space</u> (<u>semi-NLS</u>) is an ordered pair $(X, \| \, \|)$ where $\| \, \|$ is a semi-norm on X. Normed linear spaces (<u>NLS</u>) are defined similarly. If the semi-norm is understood, we often say that X is a semi-NLS (or NLS). If $(X, \| \, \|)$ is a semi-NLS (NLS), then

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

defines a translation invariant semi-metric (metric) on X, and we always assume that a semi-NLS is equipped with this metric topology. A NLS which is complete in this metric topology is called a <u>Banach space</u> or a <u>B-space</u> in honor of the Polish mathematician, Stefan Banach. The following proposition shows that a semi-NLS is a TVS under the metric topology.

Proposition 7. Let $(X, \| \|)$ be a semi-NLS. Then

(i) the map $(t, x) \to tx$ from $\mathbb{F} \times X \to X$ is continuous,

(ii) the map $(x, y) \to x + y$ from $X \times X \to x$ is continuous,

(iii) the map $x \to \|x\|$ from X to $\mathbb{R}$ is continuous.

Proof: Let $|t_k - t| \to 0$, $\|x_k - x\| \to 0$. Then

$$\|t_k x_k - tx\| \le |t_k - t| \|x_k\| + |t| \|x_k - x\|$$

so (i) holds.

(ii) follows from the triangle inequality.

From (i) it follows that a semi-norm norm is a quasi-norm so (iii) follows from Proposition 4.

We consider completeness for a QNLS. For such spaces the notions of completeness and sequential completeness coincide.

Proposition 8. A QNLS X is complete if and only if X is sequentially complete.

Proof: Suppose that X is sequentially complete and let $\{x_\delta\}$ be a Cauchy net in X. $\forall n$ choose δ_n such that $\delta_{n+1} > \delta_n$ and $\alpha, \beta \ge \delta_n$

implies $|x_\alpha - x_\beta| < 1/n$. Set $y_n = x_{\delta_n}$. Then $\{y_n\}$ is Cauchy and so converges to a point $y \in X$. The net $\{x_\delta\}$ also converges to y since if $\varepsilon > 0$ and $n > 2/\varepsilon$ is chosen such that $|y_n - y| < \varepsilon/2$, then $\alpha \geq \delta_n$ implies $|x_\alpha - y| \leq |x_\alpha - y_n| + |y_n - y| < 1/n + \varepsilon/2 < \varepsilon$.

Let (X, τ) be a TVS and $\{x_k\} \subseteq X$. The formal series $\sum_{k=1}^{\infty} x_k$ is said to converge in X if the sequence of partial sums $\{s_n\} = \{\sum_{k=1}^{n} x_k\}$ is τ-convergent to an element of X. In this case, we write $\sum_{k=1}^{\infty} x_k = \lim_n s_n$. If $(X, |\ |)$ is a quasi-normed space, the series $\sum_{k=1}^{\infty} x_k$ is said to be absolutely convergent if the series $\sum_{k=1}^{\infty} |x_k|$ converges. We have the following criterion for completeness in quasi-normed spaces.

Theorem 9. A quasi-normed space $(X, |\ |)$ is complete if and only if every absolutely convergent series in X is convergent.

Proof: Suppose $\sum x_k$ is absolutely convergent and set $s_n = \sum_{k=1}^{n} x_k$. If $n > m$, $|s_n - s_m| \leq \sum_{k=m+1}^{n} |x_k|$ so $\{s_n\}$ is Cauchy and, therefore, convergent.

Conversely, suppose that $\{x_k\}$ is Cauchy in X. For each k choose n_k such that $n \geq n_k$ implies $|x_n - x_{n_k}| < 1/2^k$. We may assume that $n_{k+1} > n_k$. Set $z_k = x_{n_{k+1}} - x_{n_k}$. Then $\sum_{k=1}^{\infty} z_k$ is convergent since

$$\sum_{k=1}^{\infty} |z_k| < \sum_{k=1}^{\infty} 1/2^k < \infty.$$

Then $z = \sum_{k=1}^{\infty} z_k = \lim_j \sum_{k=1}^{j} (x_{n_{k+1}} - x_{n_k}) = \lim_j (x_{n_{j+1}} - x_{n_1})$ so the sequence $\{x_{n_k}\}$ converges to $z + x_{n_1}$. Since $\{x_k\}$ is Cauchy, $\{x_k\}$ converges to $z + x_{n_1}$ (Exer. 2).

We now introduce a weaker form of completeness for a TVS. A sequence $\{x_k\}$ in a TVS (X, τ) is said to be <u>τ-$\mathcal{K}$ convergent</u> to 0 if every subsequence of $\{x_k\}$ has a subsequence $\{x_{n_k}\}$ such that the series $\sum_{k=1}^{\infty} x_{n_k}$ is τ-convergent in X. If $\{x_k\}$ is τ-$\mathcal{K}$ convergent to 0, then $\{x_k\}$ is τ-convergent to 0 [every subsequence has a subsequence which converges to 0 (Exer. 3)]; the converse is false (see Example 15). A TVS (X, τ) is called a <u>$\mathcal{K}$-space</u> if every sequence which is τ convergent to 0 is τ-$\mathcal{K}$ convergent to 0. By the proof of Theorem 7, it follows that a complete quasi-normed space is a $\mathcal{K}$-space [if $x_k \to 0$, then any subsequence has a further subsequence $\{x_{n_k}\}$ satisfying $\sum |x_{n_k}| < \infty$]. However, there are normed $\mathcal{K}$-spaces which are not complete ([Kl]);

indeed, any complete quasi-normed space contains dense proper subspaces which are $\mathscr{K}$-spaces ([LL], [BKL]). It will be seen later that the notion of $\mathscr{K}$ convergence can often be used as an effective substitute for completeness.

We can give examples of some of the classical quasi-normed and NLS which we will often use later in examples. We begin with some of the classical sequence spaces.

Example 10. Let s be the space of all real or complex valued sequences. Addition and scalar multiplication are defined coordinatewise. If $x = \{t_k\} \in s$, $|x| = \sum_{k=1}^{\infty} |t_k|/(2^k(1 + |t_k|))$ defines a quasi-norm on s (Example 3). Since convergence in $|\ |$ is just coordinatewise convergence (Example 3), s is a complete quasi-normed space.

Example 11. ℓ^∞ is the subspace of s consisting of all the bounded sequences. $\|\{t_k\}\|_\infty = \sup_k |t_k|$ defines a norm on ℓ^∞ called the sup-norm. Under this norm, ℓ^∞ is a B-space (Exer. 8).

Example 12. c is the subspace of ℓ^∞ consisting of all the convergent sequences. We assume that c is equipped with the sup-norm.

We show that c is complete by showing that it is a closed subset of the complete space ℓ^∞.

Corollary 13. c is a B-space.

Proof: Let $x^n = \{t_j^n\}_{j=1}^{\infty} \in c$ converge to a point $x = \{t_j\} \in \ell^{\infty}$. Let $\lim_j t_j^n = t^n$. Since $\|x^n - x\|_{\infty} \geq |t_j^n - t_j| \ \forall j$, $\lim_n t_j^n = t_j$ uniformly in j. Thus, $\lim_n \lim_j t_j^n = \lim_n t^n = \lim_j \lim_n t_j^n = \lim_j t_j$, and $x \in c$.

Example 14. c_0 is the subspace of c consisting of the sequences which converge to 0. We assume that c_0 is equipped with the sup-norm, and just as in Corollary 13, c_0 is a B-space.

Example 15. c_{00} is the subspace of c_0 consisting of the sequences which are eventually 0, i.e., $x = \{t_j\} \in c_{00}$ if and only if $\exists$ N (depending on x) such that $t_j = 0$ for $j > N$. We assume that c_{00} is equipped with the sup-norm. c_{00} is not complete with respect to the sup-norm. Let e_j be the sequence with a 1 in the j^{th} coordinate and 0 in the other coordinates. If $\{t_j\}$ is a scalar sequence which converges to 0 with $t_j \neq 0$, then the sequence $s_n = \sum_{j=1}^{n} t_j e_j$ is Cauchy but does not converge to an element of c_{00}. The sequence $\{t_j e_j\}$ also furnishes an example of a sequence which converges to 0 but which is not $\mathscr{K}$ convergent.

Example 16. Let $1 \leq p < \infty$. ℓ^p consists of all scalar sequences $\{t_j\}$ satisfying $\|\{t_j\}\|_p = (\sum_{j=1}^{\infty} |t_j|^p)^{1/p} < \infty$. ℓ^p is a B-space under the norm

$\|\ \|_p$ (this is a special case of Example 22).

Example 17. Let $0 < p < 1$. ℓ^p consists of all scalar sequences $\{t_j\}$ satisfying $|\{t_j\}|_p = \sum_{j=1}^{\infty} |t_j|^p < \infty$. $|\ |_p$ defines a quasi-norm on ℓ^p (the triangle inequality follows from the inequality $|a + b|^p \le |a|^p + |b|^p$, $0 < p < 1$). Under $|\ |_p$, ℓ^p is a complete, Hausdorff quasi-normed space (Exer. 6).

Example 18. Let m_0 be the subspace of ℓ^∞ consisting of all sequences with finite range. If $E \subset \mathbb{N}$ and C_E denotes the characteristic function of E, then m_0 is just the span of $\{C_E : E \subseteq \mathbb{N}\}$. m_0 is a dense proper subspace of ℓ^∞ with respect to $\|\ \|_\infty$ (Exer. 11).

Example 19. Let $S \neq \emptyset$. B(S) consists of all bounded scalar valued functions defined on S and $\|f\|_\infty = \sup\{|f(t)| : t \in S\}$ defines a norm on B(S) called the sup-norm. Note ℓ^∞ is just $B(\mathbb{N})$. B(S) is a B-space (Exer. 8).

Example 20. Let S be a compact Hausdorff space. C(S) is the subspace of B(S) consisting of all continuous functions. We assume that C(S) is equipped with the sup-norm; C(S) is a B-space under the sup-norm since convergence in the sup-norm is just uniform convergence on S.

Example 21. Let (S, Σ, μ) be a finite measure space and let $L^0(\mu)$ be the vector space of all real-valued Σ-measurable functions defined on S. Then $|f| = \int_S \frac{|f|}{1+|f|}\,d\mu$ defines a quasi-norm on $L^0(\mu)$ such that convergence in $|\ |$ is exactly convergence in μ-measure ([TL], p. 117). $L^0(\mu)$ is, therefore, a complete quasi-normed space ([M], p. 223). Note $|\ |$ is not a semi-norm.

Example 22. Let (S, Σ, μ) be a measure space and $1 \le p < \infty$. $L^p(\mu)$ consists of all Σ-measurable functions such that $\|f\|_p = (\int_S |f|^p d\mu)^{1/p} < \infty$. $L^p(\mu)$ is complete with respect to the semi-norm $\|\ \|_p$, and if we agree to identify functions which are equal μ-a.e., then $L^p(\mu)$ is a B-space. (This is Riesz's Theorem, [Ro], p. 125.) Example 16 is a special case of this example where $S = \mathbb{N}$ and μ is counting measure ([Ro], p. 55).

Example 23. $L^\infty(\mu)$ consists of all Σ-measurable functions $f : S \to \mathbb{F}$ which are μ-essentially bounded, i.e., $\|f\|_\infty = \mu$ - essensup(f)$< \infty$. $L^\infty(\mu)$ is complete under $\|\ \|_\infty$, and if functions which are equal μ-a.e. are identified, then $L^\infty(\mu)$ is a B-space ([Ro], p. 125).

In Examples 21, 22 and 23, when $I = [a, b]$ we write <u>$L^p(I)$</u> for $L^p(m)$, where m is Lebesgue measure on I.

Example 24. Let $a, b \in \mathbb{R}$, $a < b$, and let $\mathcal{R}[a, b]$ be the space of all Riemann integrable functions defined on $[a, b]$. $\|f\|_{\mathcal{R}} = \int_a^b |f|$ defines a semi-norm on $\mathcal{R}[a, b]$ which is not complete ([M], p. 242).

Example 25. Let $D \subseteq \mathbb{C}$ be an open, connected set, and let $\mathscr{H}(D)$ be the space of all analytic functions $f : D \to \mathbb{C}$. Let $\{K_n\}$ be an increasing sequence of compact subsets of D each of which has non-void interior, $\bigcup_{n=1}^{\infty} K_n = D$ and each compact subset of D is contained in some K_n. For each n set $|f|_n = \sup\{|f(z)| : z \in K_n\}$. Each $|\ |_n$ is a norm on $\mathscr{H}(D)$, and if $|\ |$ is the Fréchet quasi-norm of Example 3 induced by $\{|\ |_n\}$, then $\mathscr{H}(D)$ is a complete quasi-normed space since convergence in $|\ |$ is just uniform convergence on compact subsets of D. Note that the topology of $\mathscr{H}(D)$ does not depend upon the particular sequence $\{K_n\}$ chosen.

Example 26. Let $a, b \in \mathbb{R}$, $a < b$, and $k \in \mathbb{N}$. Let $C^k[a, b]$ be the subspace of $C[a, b]$ which consists of all functions which have at least k continuous derivatives. Define a norm on $C^k[a, b]$ by

$$\|f\|_{\infty,k} = \sum_{j=0}^{k} \|f^{(j)}\|_{\infty},$$

where $f^{(0)} = f$. Then $C^k[a, b]$ is a B-space under $\|\ \|_{\infty,k}$ ([DeS], p. 130).

Example 27. Let $K \subseteq \mathbb{R}^n$ be compact. Let $\mathscr{D}_K$ be all scalar valued functions $f : \mathbb{R}^n \to \mathbb{F}$ which have continuous partial derivatives of all orders with support contained in K. For any multi-index $\alpha = (\alpha_1, \ldots, \alpha_n)$ with α_j a non-negative integer, let $|\alpha| = \sum_{j=1}^{n} \alpha_j$, and write

$$D^{\alpha}f = \frac{\partial^{|\alpha|} f}{\partial x_n^{\alpha_n} \cdots \partial x_1^{\alpha_1}}$$

for $f \in \mathscr{D}_K$. For $f \in \mathscr{D}_K$ set

$$\|f\|_k = \sum_{|\alpha| \le k} \|D^{\alpha}f\|_{\infty}$$

for $k = 0, 1, \ldots$. Then $\| \ \|_k$ is a norm on $\mathscr{D}_K$, and the sequence $\{\| \ \|_k\}$ induces a Fréchet quasi-norm on $\mathscr{D}_K$. A sequence $\{f_k\}$ converges in the quasi-norm if and only if $\{D^{\alpha}f_k\}$ converges uniformly $\forall$ multi-index α. Thus, $\mathscr{D}_K$ is complete under this quasi-norm ([DeS], p. 130).

It is not clear that $\mathscr{D}_K$ contains anything other than the zero function. We give an example of a non-zero function in $\mathscr{D}_K$. Define $\varphi : \mathbb{R} \to \mathbb{R}$ by

$$\varphi(t) = \begin{cases} \exp(1/(1 - t^2)) & |t| < 1 \\ 0 & |t| \ge 1 . \end{cases}$$

Then $\varphi \in \mathscr{D}_{[-1,1]}$, and $f : \mathbb{R}^n \to \mathbb{R}$ defined by

$$f(x_1, \ldots, x_n) = \varphi(x_1) \ldots \varphi(x_n)$$

is a function in $\mathscr{D}_K$, where $K = [-1, 1] \times \ldots \times [-1, 1]$.

Example 28. Let $\Omega \subseteq \mathbb{R}^n$ be open. $\mathscr{E}(\Omega)$ consists of all functions $f : \Omega \to \mathbb{R}$ which have continuous partial derivatives of all orders in Ω. Let $\{K_j\}$ be an increasing sequence of compact subsets of Ω with non-void interiors such that $\cup K_j = \Omega$ and every compact subset of Ω is contained in some K_j. For each j set $|f|_j = \sup\{|D^{\alpha}f(x)| : x \in K_j, |\alpha| \le j\}$. Let

$|\ |$ be the Fréchet quasi-norm induced by the sequence of norms $\{|\ |_j\}$ as in Example 3. A sequence $\{f_k\}$ converges to 0 if and only if the sequence $\{D^{\alpha}f_k\}$ converges uniformly on compact subsets of Ω $\forall$ multi-index α. Thus, $\mathscr{E}(\Omega)$ is complete ([DeS], p. 130). Note that the topology of $\mathscr{E}(\Omega)$ does not depend upon the particular sequence $\{K_j\}$ chosen.

Example 29. $\mathbb{R}^n$ and $\mathbb{C}^n$ are B-spaces under any of the norms

$$\|x\|_{\infty} = \sup_{i \le j \le n} |x_j|$$

or

$$\|x\|_p = (\sum_{j=1}^{n} |x_j|^p)^{1/p} \quad (1 \le p < \infty), \text{ where } x = (x_1, \ldots, x_n).$$

We will show later that all of these norms are, in some sense, equivalent.

Example 30. Let $S \neq \emptyset$ and let Σ be a σ-algebra of subsets of S. $ba(\Sigma)$ denotes the space of all bounded, finitely additive set functions $\nu : \Sigma \to \mathbb{R}$. Any bounded finitely additive set function $\nu : \Sigma \to \mathbb{R}$ has bounded variation, $\mathrm{var}(\nu) = |\nu|$, and satisfies the inequality

$$\sup\{|\nu(E)| : E \in \Sigma\} \le |\nu|(S) \le 2 \sup\{|\nu(E)| : E \in \Sigma\}$$

([DS], III.1.5). $ba(\Sigma)$ is a B-space under the variation norm $\|\nu\| = |\nu|(S)$ ([DS], III.7). $ca(\Sigma)$ denotes the subspace of $ba(\Sigma)$ which consists of all countably additive measures. $ca(\Sigma)$ is a closed subspace of $ba(\Sigma)$ and is, therefore, a B-space ([DS], III.7).

Example 31. Let $B(S, \Sigma)$ be the space of all bounded Σ-measurable functions $f : S \to \mathbb{R}$. Under the sup-norm, $\|f\|_\infty = \sup\{|f(t)| : t \in S\}$, $B(S, \Sigma)$ is a B-space (Exer. 17). Let $\mathscr{S}(\Sigma)$ be the linear subspace of $B(S, \Sigma)$ which consists of the Σ-simple function, i.e., functions of the form $f = \sum_{j=1}^{n} t_j C_{E_j}$, where $t_j \in \mathbb{R}$, $E_j \in \Sigma$ and C_E denotes the characteristic function of the set E. $\mathscr{S}(\Sigma)$ is a dense subspace of $B(S, \Sigma)$ ([HS], 11.35).

Finite Products:

Let X and Y be quasi-normed spaces. Then $X \times Y$ carries several natural quasi-norms. Namely, $|(x, y)|_1 = |x| + |y|$, $|(x, y)|_2 = \sqrt{|x|^2 + |y|^2}$ and $|(x, y)|_\infty = \max\{|x|, |y|\}$. Each of these quasi-norms induces the product topology on $X \times Y$ so any of them can be used on the product. Moreover, if X and Y are semi-NLS, then each of these quasi-norms is a semi-norm and is a norm if and only if X and Y are NLS.

A further important example of NLS are given by the inner product and Hilbert spaces. For the convenience of the reader not familiar with these spaces their basic properties are described in the appendix.

Exercise 1. Give an example of a function satisfying (i)-(iv) but not (v). [Hint: $p(x) = 1$ if $x \neq 0$, $p(0) = 0$.]

Exercise 2. If $\{x_k\}$ is a Cauchy sequence in a metric space and has a subsequence $\{x_{n_k}\}$ which converges to x, show $x_k \to x$.

Exercise 3. If Σx_k is a convergent series in a TVS, show $x_k \to 0$.

Exercise 4. Show c_{00} is dense in c_0 with respect to $\| \|_\infty$.

Exercise 5. Show c_{00} is dense in ℓ^p with respect to $\| \|_p$ for $1 \le p < \infty$.

Exercise 6. Show ℓ^p is complete for $0 < p < 1$ and c_{00} is a dense subspace.

Note that $| |_p$ is <u>not</u> a norm for $0 < p < 1$.

Exercise 7. Show $c_{00} \subsetneq m_0$ and $c_{00} \neq m_0$.

Exercise 8. Show $B(S)$ is complete under $\| \|_\infty$.

Exercise 9. Show $\| \|_{\mathscr{R}}$ is not a norm on $\mathscr{R}[a, b]$.

Exercise 10. Give an explicit example of an absolutely convergent series in c_{00} (m_0) which is not convergent.

Exercise 11. Show m_0 is dense in ℓ^∞ with $m_0 \neq \ell^\infty$.

Exercise 12. Give an example of a convergent series in ℓ^p $(1 < p < \infty)$ which is not absolutely convergent.

Exercise 13. Show $L^0[0, 1]$ is separable, where

$$L^0[0, 1] = L^0([0, 1], \mathcal{M}, m)$$

and m is Lebesgue measure on $[0, 1]$.

Exercise 14. If X is a TVS and p a semi-norm on X, show p is continuous if and only if p is continuous at 0.

Exercise 15. If the sequence $\{x_k\}$ is Mackey convergent to 0, show that it converges to 0.

Exercise 16. Show that c_{00}, c, c_0, ℓ^p $(1 \le p < \infty)$ are separable. Show ℓ^∞ is not separable.

Exercise 17. Show $B(S, \Sigma)$ of Example 31 is complete.

3
Metrizable TVS

If X is a TVS whose topology is semi-metrizable, then the topology is certainly first countable. This simple necessary condition turns out to be sufficient for the topology of a TVS to be semi-metrizable even by a quasi-norm. To establish this result, we require two lemmas.

Lemma 1. Let X be a vector space and q a non-negative function defined on X. For $x \in X$ set $|x| = \inf\{ \sum_{k=1}^{n} q(x_k) : x = \sum_{k=1}^{n} x_k, n \in \mathbb{N}\}$. Then $|x| \geq 0$ and $|x + y| \leq |x| + |y|$. If $q(0) = 0$, then $|0| = 0$; if $q(x) = q(-x)$ for all x, then $|x| = |-x|$.

Proof: Let $\varepsilon > 0$, $x, y \in X$. Pick $x = \sum_{k=1}^{n} x_k$, $y = \sum_{k=1}^{m} y_k$ such that $\sum_{k=1}^{n} q(x_k) < |x| + \varepsilon$ and $\sum_{k=1}^{m} q(y_k) < |y| + \varepsilon$. Then

$$x + y = \sum_{k=1}^{n} x_k + \sum_{k=1}^{m} y_k$$

and

$$|x + y| \le \sum_{k=1}^{n} q(x_k) + \sum_{k=1}^{m} q(y_k) < |x| + |y| + 2\varepsilon$$

so $|x + y| \le |x| + |y|$.

If $q(0) = 0$, then $|x| = 0$.

If $q(x) = q(-x)$ and $x = \sum_{k=1}^{n} x_k$, then $-x = \sum_{k=1}^{n} - x_k$ and

$$\sum_{k=1}^{n} q(x_k) = \sum_{k=1}^{n} q(-x_k)$$

so $|x| = |-x|$.

Lemma 2. Let X be a vector space and q a non-negative function of X such that $q(0) = 0$ and $q(x + y + z) \le 2 \max\{q(x), q(y), q(z)\}$. Then for $x_1, ..., x_n \in X$, $q(\sum_{i=1}^{n} x_i) \le 2 \sum_{i=1}^{n} q(x_i)$.

Proof: Set $a = \sum_{i=1}^{n} q(x_i)$, where we may assume $a > 0$. The proof is by induction on n. For $n = 1, 2, 3$, the result is trivial so assume that $n > 3$. Let m be the largest integer such that $\sum_{i=1}^{m} q(x_i) \le a/2$ [if this inequality fails to hold for $m = 1$, set $m = 0$ and ignore this inequality].

Then $0 \le m < n$ and $\sum_{i=1}^{m+1} q(x_i) > a/2$ so $\sum_{i=m+2}^{n} q(x_i) \le a/2$. The sums on the left hand side of the following inequalities have fewer than n terms or are 0 so the induction hypothesis gives

$$q(\sum_{i=1}^{m} x_i) \le 2 \sum_{i=1}^{m} q(x_i) \le a$$

$$q(\sum_{i=m+2}^{n} x_i) \le 2 \sum_{i=m+2}^{n} q(x_i) \le a$$

$$q(x_{m+1}) \le a .$$

These inequalities and the hypothesis give the result.

Theorem 3 (Kakutani). The topology of a first countable TVS X is given by a quasi-norm.

Proof: Put $U_0 = X$ and choose a neighborhood base at 0 of balanced sets, $\{U_n\}_{n=1}^{\infty}$, satisfying $U_{n+1} + U_{n+1} + U_{n+1} \subseteq U_n$. For $x \in \overline{\{0\}}$, set $q(x) = 0$, and for $x \notin \overline{\{0\}}$ set $q(x) = 2^{-k}$, where $k = k(x)$ is the largest integer such that $x \in U_k$ [thus, if $k = k(x)$, $x \in U_k \backslash U_{k+1}$]. Note $q(x) = q(-x)$ since the U_k's are balanced.

We claim that $x_n \to 0$ in X if and only if $q(x_n) \to 0$. Suppose $x_n \to 0$ and let m be a positive integer. Then $x_n \in U_m$ eventually and for such x_n either $x_n \in \overline{\{0\}}$ so $q(x_n) = 0$ or $k(x_n) \ge m$ so $q(x_n) \le 2^{-m}$. Thus, $q(x_n) \to 0$. Conversely, suppose that $q(x_n) \to 0$. Let m

be a positive integer. Then for sufficiently large n, $q(x_n) < 2^{-m}$. If $x_n \in \overline{\{0\}}$, $x_n \in U_m$ while if $x_n \notin \overline{\{0\}}$, we have with $k = k(x_n)$, $q(x_n) = 2^{-k} < 2^{-m}$ which implies $k > m$ and so $x_n \in U_k \subseteq U_m$. Thus, $x_n \in U_m$ eventually and $x_n \to 0$.

Next we show $q(x + y + z) \le 2 \max\{q(x), q(y), q(z)\}$. We may assume that not all $x, y, z \in \overline{\{0\}}$. Suppose, for definiteness, that

$$q(x) = 2^{-k} \ge q(y) \vee q(z).$$

Then $x, y, z \in U_k$ so by choice of $\{U_k\}$, $x + y + z \in U_{k-1}$. Thus, $q(x + y + z) \le 2^{-k+1} = 2 \cdot 2^{-k} = 2q(x)$.

We now have established the conditions of Lemma 2 and we define $|\ |$ by the formula in Lemma 1. We claim that $\frac{1}{2} q(x) \le |x| \le q(x)$. The second inequality follows from the definition of $|\ |$. Suppose $x = \sum_{k=1}^{n} x_k$. By Lemma 2

$$\sum_{k=1}^{n} q(x_k) \ge \frac{1}{2} q\Big(\sum_{k=1}^{n} x_k\Big) = \frac{1}{2} q(x).$$

This yields the first inequality.

It now follows from the above that $x_n \to 0$ in X if and only if $|x_n| \to 0$. Since the scalar multiplication on X is continuous, it follows from this that $|\ |$ is a quasi-norm on X, and induces the original topology of X.

A TVS X is called a <u>semi-metric linear space</u> (<u>metric linear space</u>) if its topology is semi-metrizable (metrizable). From Theorem 3 we see that a space is a semi-metric linear space if and only if it is first countable if and only if its topology is generated by a quasi-norm. A complete metric linear space or equivalently a complete quasi-normed space is often referred to as an F-space in honor of M. Fréchet.

Whereas the topology of a TVS is generated by a single quasi-norm if and only if the topology is semi-metrizable, Burzyk and P. Mikusinski ([BM]) have shown that the topology of any TVS is generated by a family of quasi-norms.

4

Bounded Sets in a TVS

We define the notion of boundedness in a general TVS. We use the notion of boundedness as originally introduced by Banach ([MO]).

Definition 1. A subset B of a TVS X is <u>bounded</u> if $\{x_k\} \subseteq B$ and $t_k \to 0$, then $t_k x_k \to 0$.

We have the following criteria for boundedness.

Theorem 2. Let B be a subset of a TVS X. The following are equivalent

(i) B is bounded,

(ii) $\forall \{x_k\} \subseteq B, \frac{1}{k} x_k \to 0$,

(iii) $\forall$ neighborhood U of 0 in X $\exists\, \varepsilon > 0$ such that $|t| \le \varepsilon \Rightarrow tB \subseteq U$ [we say that B is <u>absorbed</u> by every neighborhood of 0],

(iv) B is absorbed by every balanced neighborhood of 0,

(v) $\forall$ neighborhood U of 0 in X $\exists$ $\varepsilon > 0$ such that $\varepsilon B \subseteq U$

(vi) $\forall$ neighborhood U of 0 in X $\exists$ $r > 0$ such that $B \subseteq rU$

(viii) $\forall$ neighborhood U of 0 in X $\exists$ a positive integer n such that $B \subseteq nU$.

Proof: Clearly (i) $\Rightarrow$ (ii).

If (iii) fails, $\exists$ a neighborhood of 0, U, such that $tB \not\subseteq U$ for $|t| \le \varepsilon$ no matter what $\varepsilon > 0$. In particular, $\frac{1}{n} B \not\subseteq U$ so pick $x_n \in B \setminus nU$. Then $\frac{1}{n} x_n \notin U$ so $\{\frac{1}{n} x_n\}$ doesn't converge to 0, and (ii) fails. Hence, (ii) $\Rightarrow$ (iii).

Clearly (iii) $\Rightarrow$ (iv). (iv) $\Rightarrow$ (v) by Theorem 1.12. That (v) $\Rightarrow$ (vi) is clear. Suppose (vi) holds. Pick a balanced neighborhood V of 0 such that $V \subseteq U$ (1.12). $\exists$ $r > 0$ such that $B \subseteq rV$. Pick $n \in \mathbb{N}$ such that $n \ge r$. Then $B \subseteq nV \subseteq nU$ since V is balanced $[rV = \frac{r}{n}(nV) \subseteq nV]$. Thus, (vi) $\Rightarrow$ (vii).

Suppose (vii) holds. Let $\{x_k\} \subseteq B$ and $t_k \to 0$. Let U be a balanced neighborhood of 0. $\exists$ N such that $|t_k| \le \frac{1}{n}$ for $k \ge N$. Then $\frac{1}{n} x_k \in U$ for $k \ge N$, and since U is balanced, $t_k x_k = (nt_k)\frac{1}{n} x_k \in U$ for $k \ge N$. Hence, $t_k x_k \to 0$.

Remark 3. Condition (iii) was introduced by von Neumann and is often used for the definition of boundedness ([vN]).

Corollary 4. A linear subspace L of a TVS is bounded $\Leftrightarrow L \subseteq \overline{\{0\}}$. In particular, no non-trivial subspace of a Hausdorff TVS is bounded.

Proof: $\Leftarrow$: Every neighborhood of 0 contains $\overline{\{0\}}$ so $\overline{\{0\}}$ is bounded.

$\Rightarrow$: Suppose $\exists\, a \in L \backslash \overline{\{0\}}$. Then $na \in L\ \forall n$. But $(\frac{1}{n})na = a \not\to 0$ since $a \notin \overline{\{0\}}$. So L is not bounded.

Recall that a subset S of a metric space (X, d) is said to be <u>metrically bounded</u> if $\exists\, a \in X,\ r > 0$ such that $S \subseteq S(a, r)$, where $S(a, r) = \{x : d(a, x) < r\}$ is the <u>sphere</u> with center at a and radius r.

Proposition 5. In a quasi-normed space a bounded subset is metrically bounded, but the converse does not hold in general.

Proof: Let B be bounded. Then $\exists\, n \in \mathbb{N}$ such that $B \subseteq nS(0, 1)$. But $nS(0, 1) \subseteq S(0, n)$ so $B \subseteq S(0, n)$ is metrically bounded.

Note the space $L^0(\mu)$ of Example 2.21 is metrically bounded, but not bounded by Corollary 4.

For semi-NLS, we do have

Proposition 6. A subset B of a semi-NLS is bounded if and only if it is metrically bounded. Therefore, B is bounded if and only if

$$\sup\{\|x\| : x \in B\} < \infty.$$

Proof: If B is metrically bounded, $B \subseteq S(0, r)$ for some $r > 0$. Let U be a neighborhood of 0. $\exists\, \varepsilon > 0$ such that $S(0, \varepsilon) \subseteq U$. Thus, $B \subseteq S(0, r) = \frac{r}{\varepsilon} S(0, \varepsilon) \subseteq \frac{r}{\varepsilon} U$.

Using the notion of $\mathcal{K}$ convergence discussed in §2, we introduce a stronger version of boundedness by using $\mathcal{K}$ convergence and the analogue of the sequential definition of boundedness given in Definition 1. The definition is due to P. Antosik ([A]).

Definition 7. A subset B of a TVS (X, τ) is $\underline{\tau\text{-}\mathcal{K} \text{ bounded}}$ (or $\mathcal{K}$ bounded if the topology τ is understood) if $\{x_k\} \subseteq B$ and $t_k \to 0$ implies that $\{t_k x_k\}$ is τ-$\mathcal{K}$ convergent to 0.

Clearly a $\mathcal{K}$ bounded set is bounded, but the following example shows that the converse is false.

Example 8. Consider $B = \{e_k : k \in \mathbb{N}\}$ in c_{00}. Clearly B is bounded with respect to $\| \|_\infty$, but is not $\mathcal{K}$ bounded as noted in Example 2.15.

As will be seen later, the notion of $\mathcal{K}$ boundedness will be very useful in formulating general results which usually require some sort of completeness conditions. However, there is one annoying property of $\mathcal{K}$ convergence; namely, a $\mathcal{K}$ convergent sequence needn't be $\mathcal{K}$ bounded as the following example shows.

Example 9. Let $\{t_k\} \in \ell^1$ with $t_k > 0\ \forall k$. Define a norm on m_0 (see Example 2.18) by $\|\{s_k\}\| = \sum_{k=1}^{\infty} |t_k s_k|$ ($\|\ \|$ clearly depends on $\{t_k\}$). Consider the sequence $\{e_k\}$ in m_0. For any subsequence $\{e_{n_k}\}$, the series $\sum e_{n_k}$ is $\|\ \|$-convergent in m_0

$$[\|\sum_{j=1}^{k} e_{n_j} - C_{\{n_j:j\in\mathbb{N}\}}\| = \sum_{j=k+1}^{\infty} |t_{n_j}|]$$

so $\{e_k\}$ is $\|\ \|$-$\mathscr{K}$ convergent to 0. However, no subseries of $\sum \frac{1}{k} e_k$ can converge to an element of m_0 with respect to $\|\ \|$ since convergence in $\|\ \|$ clearly implies coordinatewise convergence.

A TVS in which bounded sets are $\mathscr{K}$ bounded is called an <u>$\mathscr{A}$-space</u>. Clearly a $\mathscr{K}$-space (§2) is an $\mathscr{A}$-space, but there are examples of $\mathscr{A}$-spaces which are not $\mathscr{K}$-spaces (14.20). However, we show in Proposition 10 below that any quasi-normed $\mathscr{A}$-space is a $\mathscr{K}$-space. We will see later that $\mathscr{A}$-spaces are useful in obtaining uniform boundedness principles (§9).

Proposition 10. If X is a quasi-normed $\mathscr{A}$-space, then X is a $\mathscr{K}$-space.

Proof. Let $x_k \to 0$ in X. Pick $t_k \uparrow \infty$ such that $t_k x_k \to 0$ (2.5). Then $\{t_k x_k\}$ is bounded (Exer. 5) and, therefore, $\mathscr{K}$ bounded. Thus, $\{\frac{1}{t_k}(t_k x_k)\} = \{x_k\}$ is $\mathscr{K}$ convergent.

Exercise 1. Show that a finite set is bounded.

Exercise 2. Show a compact subset of a TVS is bounded.

Exercise 3. Show that a subset of a bounded set is bounded.

Exercise 4. Show that a finite union or (vector) sum of bounded sets is bounded.

Exercise 5. Show that a convergent sequence is bounded.

Exercise 6. If $\{x_k\}$ is a bounded sequence in a TVS and $\{t_i\} \in \ell^1$, show $\{t_i x_i\}$ is Cauchy.

5

Linear Operators and Linear Functionals

In this section we develop the basic properties of continuous linear mappings between TVS.

If X and Y are vector spaces and $T : X \to Y$ is linear, we denote the range of T by $\mathscr{R}T$ and denote the kernel or null manifold of T by ker T or $\mathscr{N}(T)$. We also follow the tradition of writing $T(x) = Tx$.

Proposition 1. Let E, F be TVS and $T : E \to F$ linear. The following are equivalent.

(i) T is continuous.

(ii) T is continuous at 0.

Proof: Assume (ii) holds and let $x \in E$. Let V be a neighborhood of Tx. Then $-Tx + V$ is a neighborhood of 0 so $\exists$ a neighborhood U of 0 in X such that $TU \subseteq -Tx + V$. Thus, $x + U$ is a neighborhood of x with $T(x + U) \subseteq V$, and (i) holds.

Definition 2. A linear map between TVS is said to be <u>bounded</u> if it carries bounded sets to bounded sets.

Proposition 3. Let E, F be TVS and $T : E \to F$ sequentially continuous. Then T is bounded.

Proof: Let $B \subseteq E$ be bounded. Let $y_k \in TB$ with $y_k = Tx_k$, $x_k \in B$. Then $\frac{1}{k} x_k \to 0$ so $T(\frac{1}{k} x_k) = \frac{1}{k} Tx_k = \frac{1}{k} y_k \to 0$ and TB is bounded by 4.2.

In particular, a continuous linear map between TVS is bounded. As we will see later (14.18), the converse of this statement is false, but it does hold for quasi-normed spaces.

Proposition 4. Let $(E, |\ |)$ be a quasi-normed space and F a TVS. If $T : E \to F$ is bounded, then T is continuous.

Proof: If T is not continuous at 0, $\exists$ a neighborhood V of 0 in F such that $U = T^{-1}V$ is not a neighborhood of 0 in E. Then nU is not a neighborhood of 0 $\forall n \in \mathbb{N}$ so $\exists\, x_n \in E$ with $|x_n| < \frac{1}{n}$ and $x_n \notin nU$. Then $\{x_n\}$ is bounded (Exer. 4.5), but since $Tx_n \notin nV$, $\{Tx_n\}$ is not bounded (4.2).

We will show later that the conclusion of Proposition 4 holds for a large class of spaces called bornological spaces (§21).

For semi-NLS, we have

Proposition 5. Let X, Y be semi-NLS and $T : X \to Y$ linear. The following are equivalent.

(i) T is uniformly continuous.

(ii) T is continuous at 0.

(iii) T is continuous.

(iv) T is bounded.

(v) $\exists M > 0$ such that (*) $\|Tx\| \le M\|x\|$ $\forall x \in X$.

Proof: (i) $\Rightarrow$ (ii) $\Rightarrow$ (iii) $\Rightarrow$ (iv) by Propositions 1 and 3.

Assume (iv) holds. Then $S = \{x : \|x\| \le 1\}$ is bounded so

$$M = \sup\{\|Tx\| : x \in S\} < \infty$$

(4.6). If $\|x\| \ne 0$, $x/\|x\| \in S$ so $\|Tx\|/\|x\| \le M$ and (*) holds; if $\|x\| = 0$, then (*) holds since $\|Tx\| = 0$ [$\{nx : n \in \mathbb{N}\}$ is bounded so

$$\sup\{n\|Tx\| : n \in \mathbb{N}\} < \infty].$$

Assume (v) holds. Let $\varepsilon > 0$. Set $\delta = \varepsilon/M$. If $\|x - y\| < \delta$, $\|T(x - y)\| = \|Tx - Ty\| \le M\|x - y\| < \varepsilon$ and (i) holds.

Definition 6. Let X, Y be TVS over $\mathbb{F}$. $L(X, Y)$ denotes the space of all linear, continuous mappings (called <u>operators</u>) from X into Y. $L(X, Y)$ is a vector space under the pointwise operations of addition and scalar multiplication. If $X = Y$, we often write $L(X) = L(X, X)$.

If X and Y are semi-NLS, we define a semi-norm on $L(X, Y)$ by $\|T\| = \sup\{\|Tx\| : \|x\| \le 1\}$; this semi-norm is called the <u>operator norm</u> or <u>uniform norm</u> on $L(X, Y)$. Note that if (*) $\|Tx\| \le M\|x\|$ $\forall x \in X$, then $\|T\| \le M$, and $\|T\|$ is the infimum of all such numbers M satisfying (*). We have the following easily checked properties of the operator norm.

Proposition 7. (i) $(L(X, Y), \|\ \|)$ is a semi-NLS and $\|Tx\| \le \|T\|\|x\|$ $\forall x \in X$.

(ii) $\|T\| = 0$ if and only if $\|Tx\| = 0$ $\forall x \in X$.

(iii) If Y is a NLS, then $L(X, Y)$ is a NLS under the operator norm.

(iv) If Z is a NLS, $T \in L(X, Y)$, $S \in L(Y, Z)$, then $ST \in L(X, Z)$ with $\|ST\| \le \|S\|\|T\|$.

Concerning completeness, we have

Proposition 8. If Y is a complete NLS, then $L(X, Y)$ is complete.

Proof: Suppose that $\{T_k\}$ is Cauchy with respect to the operator norm. If $x \in X$,

(1) $$\|T_k x - T_j x\| \le \|T_k - T_j\|\|x\|$$

so $\{T_k x\}$ is Cauchy in Y. Let $Tx = \lim T_k x$. Then $T : X \to Y$ is clearly linear. We show $T \in L(X, Y)$ and $T_k \to T$. Let $\varepsilon > 0$. $\exists N > 0$ such that $k, j \ge N \Rightarrow \|T_k - T_j\| < \varepsilon$. Let $j \to \infty$ in (1) giving

$\|T_k x - Tx\| \le \varepsilon\|x\|$ for $x \in X$. Therefore, $T_k - T \in L(X, Y)$ with $\|T_k - T\| \le \varepsilon$ for $k \ge N$, and $T \in L(X, Y)$.

Remark 9. We will establish the converse of Proposition 8 later (8.1.18).

Definition 10. Let X be a TVS. A linear map from X into the scalar field is called linear functional. The dual of X, denoted by X', is the space of all continuous linear functionals on X, i.e., $X' = L(X, \mathbb{F})$.

If x' is a linear functional on X, we often write $\langle x', x\rangle = x'(x)$ for the value of x' at $x \in X$.

If X is a semi-NLS, it follows from Proposition 8 that X' is a B-space under the norm $\|x'\| = \sup\{|\langle x', x\rangle| : \|x\| \le 1\}$; this norm is called the dual norm on X'.

If X and Y are semi-NLS, an isometry from X into Y is a map $U : X \to Y$ such that $\|Ux - Uy\| = \|x - y\|$ $\forall x, y \in X$. If X and Y are NLS, then X and Y are said to be linearly isometric if $\exists$ a linear isometry from X onto Y; if X and Y are linearly isometric, it is customary to write $X = Y$.

We give isometric descriptions of the duals of some of the classical function spaces. It is assumed that the reader is familiar with some of the classic dual spaces from real analysis; references to such duals are given.

Example 11. c_0' and ℓ^1 are linearly isometric.

Let $f \in c_0'$. Set $y_k = \langle f, e_k\rangle$ for $k \in \mathbb{N}$ and $y = \{y_k\}$. We claim that $y \in \ell^1$. If $x^N = \sum_{k=1}^{N} (\text{sign } y_k)e_k$, then

$$\langle f, x^N \rangle = \sum_{k=1}^{N} |y_k| \le \|f\| \|x^N\|_\infty \le \|f\|$$

so $y \in \ell^1$ and $\|y\|_1 \le \|f\|$. For $x = \{x_k\} \in c_0$, $x = \sum_{k=1}^{\infty} x_k e_k$ so

$\langle f, x \rangle = \sum_{k=1}^{\infty} x_k \langle f, e_k \rangle = \sum_{k=1}^{\infty} x_k y_k$ and

$$|\langle f, x \rangle| \le \|x\|_\infty \sum_{k=1}^{\infty} |y_k| = \|x\|_\infty \|y\|_1$$

which implies that $\|f\| \le \|y_1\|$. Hence, $\|f\| = \|y\|_1$.

Thus, the map $U : f \to y$ from c_0' to ℓ^1 is an isometry and is obviously linear. The map U is also onto since if $y = \{y_k\} \in \ell^1$, $\langle f_y, x \rangle = \sum_{k=1}^{\infty} y_k x_k$ defines a continuous linear functional with $U(f_y) = y$. Every $y = \{y_k\} \in \ell^1$ induces a continuous linear functional on c_0 via the mapping $x \to \sum_{k=1}^{\infty} y_k x_k$ and every continuous linear functional arises in this manner. We identify c_0' and ℓ^1 and simply write $c_0' = \ell^1$.

Example 12. Similarly, c' and ℓ^1 are linearly isometric under the correspondence which associates with each $y \in \{y_k\} \in \ell^1$ the linear functional $\langle f_y, \{x_k\} \rangle = \sum_{k=1}^{\infty} y_k x_{k-1}$, where $x_0 = \lim x_k$ (Exer. 16).

Let (S, Σ, μ) be a measure space.

Example 13. For $1 < p < \infty$, the dual of $L^p(S, \Sigma, \mu)$ is linearly isometric to $L^q(S, \Sigma, \mu)$, where $1/p + 1/q = 1$. The correspondence between $F \in (L^p)'$ and $f \in L^q$ is given by $\langle F, g\rangle = \int_S fg\, d\mu$, $g \in L^p$, and $\|F\| = \|f\|_q$ ([Ro]11.7).

If μ is a σ-finite measure, $L^\infty(S, \Sigma, \mu)$ is likewise the dual of $L^1(S, \Sigma, \mu)$ ([Ro]11.7).

In particular, if μ is counting measure on $\mathbb{N}$, we have

Example 14. For $1 \le p < \infty$, the dual of ℓ^p is ℓ^q, where $1/p + 1/q = 1$, $1 < p$, and if $p = 1$, $q = \infty$. The correspondence between $y = \{y_k\} \in \ell^q$ and the corresponding linear functional $f_y \in (\ell^p)'$ is $\langle f_y, x\rangle = \sum_{k=1}^{\infty} y_k x_k$, $\|f_y\| = \|y\|_q$.

Example 15. We now show that the dual of $B(S, \Sigma)$ is $ba(\Sigma)$ (Example 2.30, 2.31). For this we require the integration of bounded Σ-measurable functions with respect to bounded finitely additive measures belonging to $ba(\Sigma)$. Let $\mu \in ba(\Sigma)$. If $\varphi = \sum_{j=1}^{n} t_j C_{E_j}$ is a Σ-simple function with $\{E_j\}$ pairwise disjoint, we define the integral of φ with respect to μ to be $\int_E \varphi\, d\mu = \sum_{j=1}^{n} t_j \mu(E \cap E_j)$ for $E \in \Sigma$. It is easily checked from the finite additivity of μ that the integral is independent of the representation of φ. We also have

(2) $\left|\int_S \varphi\, d\mu\right| \le \|\varphi\|_\infty |\mu|(S)$, where $|\mu|$ is the variation of μ.

Every $f \in B(S, \Sigma)$ is the uniform limit of a sequence of simple functions, $\{\varphi_k\}$, so we can define the integral of f with respect to μ to be $\int_S f\, d\mu = \lim \int_S \varphi_k d\mu$ [the limit exists since by (2),

$$\left|\int_S \varphi_k d\mu - \int_S \varphi_j d\mu\right| \le \|\varphi_k - \varphi_j\|_\infty |\mu|(S)$$

and, moreover, the value of the limit is independent of the particular sequence $\{\varphi_k\}$]. The inequality (2) still holds if φ belongs to $B(S, \Sigma)$.

If $\mu \in ba(\Sigma)$, then by (2) μ induces a continuous linear functional f_μ on $B(S, \Sigma)$ with $\|f_\mu\| \le |\mu|(S)$. It is actually the case that $\|f_\mu\| = |\mu|(S) = \|\mu\|$. Let $\varepsilon > 0$. There is a partition $\{E_j : j = 1, \ldots, n\}$ of S such that $|\mu|(S) < \sum_{j=1}^{n} |\mu(E_j)| + \varepsilon$. Define $\varphi : S \to \mathbb{R}$ by $\varphi = \sum_{j=1}^{n} \text{sign}\, \mu(E_j) C_{E_j}$. Then $\|\varphi\| = 1$ and

$$\langle f_\mu, \varphi\rangle = \sum_{j=1}^{n} |\mu(E_j)| > \|\mu\| - \varepsilon.$$

Hence, $\|f_\mu\| = \|\mu\|$. Thus, $U : \mu \to f_\mu$ defines a linear isometry from $ba(\Sigma)$ into $B(S, \Sigma)'$. We show U is onto. Let $f \in B(S, \Sigma)'$. For $E \in \Sigma$ define $\mu(E) = \langle f, C_E\rangle$. Then μ is clearly finitely additive, and if $\{E_j : j = 1, \ldots, n\}$ is a partition of S, then

$$\sum_{j=1}^{n} |\mu(E_j)| = \langle f, \sum_{j=1}^{n} \text{sign}\, \mu(E_j) C_{E_j} \rangle \le \|f\|$$

so μ has bounded variation. Clearly $U\mu = f_\mu$ so U is onto and $ba(\Sigma)$ and $B(S, \Sigma)'$ are linearly isometric.

In particular, it follows that the dual of ℓ^∞ is ba, the space of all bounded, finitely additive set functions on the power set of the positive integers.

The dual of $L^\infty(S, \Sigma, \mu)$ for a general measure μ has a similar description. This requires the development of a general integration theory with respect to finitely additive measures which we will not give. The reader can consult [DS] or [HS].

Example 16. Let S be a compact Hausdorff space. The dual of $C(S)$ is the space of all regular, finite Borel measures μ on the Borel sets of S, rca(S). The norm of μ is the variation norm $\|\mu\| = var(\mu)(S)$. The correspondence between $\mu \in rca(S)$ and the corresponding continuous linear functional f_μ is given by $\langle f_\mu, \varphi \rangle = \int_S \varphi \, d\mu$,

$$\|f_\mu\| = var(\mu)(S) = \|\mu\|.$$

([DS]IV.6.3).

The dual of $C[a, b]$ can also be described as the space normalized functions of bounded variation on $[a, b]$; the correspondence between $F \in C[a, b]'$ and a normalized function of bounded variation f is given by the Riemann-Stieltjes integral $\langle F, \varphi\rangle = \int_a^b \varphi df$ (see [TL] III.5).

Finally we describe the dual of some of the quasi-normed spaces given in Chapter 2.

Example 17. The dual of s is c_{00}. If $y \in c_{00}$, y induces a continuous linear functional $f_y : s \to \mathbb{R}$ by $\langle f_y, x\rangle = \sum_{k=1}^{\infty} x_k y_k$.

Conversely, if $f \in s'$, set $y_k = \langle f, e_k\rangle$. For $x \in s$, $x = \sum_{k=1}^{\infty} x_k e_k$ so $\langle f, x\rangle = \sum_{k=1}^{\infty} x_k \langle f, e_k\rangle = \sum_{k=1}^{\infty} x_k y_k$. Thus, for any sequence $x = \{x_k\} \in s$, the series $\sum_{k=1}^{\infty} x_k y_k$ converges which implies that $y_k = 0$ eventually, or $y = \{y_k\} \in c_{00}$. Clearly, $f = f_y$ and the linear map $y \to f_y$ from c_{00} to s' is onto. We write $s' = c_{00}$; however, note that we do not have any topology defined on s', as the dual of s, at this point.

Example 18. Let $L^0[0, 1]$ be the vector space of all real-valued, Lebesgue measurable functions defined on $[0, 1]$ and equip $L^0[0, 1]$ with the quasi- norm $|f| = \int_0^1 \frac{|f|}{1+|f|}$, the integral being with respect to Lebesgue measure m on $[0, 1]$ (Example 2.21). We claim that the dual of $L^0[0, 1]$ is exactly $\{0\}$.

For suppose $0 \neq f \in L^0[0, 1]'$. Then $\exists$ $\varphi \in L^0[0, 1]$ such that

$<f, \varphi> \neq 0$. Let $J_1 = [0, 1/2]$, $J_2 = (1/2, 1]$ and set $g_i = C_{J_i}\varphi$. Since $\varphi = g_1 + g_2$, either $<f, g_1> \neq 0$ or $<f, g_2> \neq 0$; choose one such and label it φ_1. Note $m(\{t : \varphi_1(t) \neq 0\}) \leq 1/2$.

Continuing this bisection method produces a sequence

$$\{\varphi_j\} \subseteq L^0[0, 1]$$

such that $<f, \varphi_j> \neq 0$ and $m(\{t : \varphi_j(t) \neq 0\}) \leq 1/2^j$. Let $h_j = \varphi_j/<f, \varphi_j>$. Then $<f, h_j> = 1$ $\forall j$ and $h_j \to 0$ in m-measure since

$$m(\{t : h_j(t) \neq 0\} \leq 1/2^j.$$

Thus, f is not continuous at 0.

For $0 < p < 1$, the space $L^p[0, 1]$ is defined to be the vector space of all Lebesgue measurable functions on $[0, 1]$ such that

$$\int_0^1 |f|^p = |f|_p < \infty.$$

The functional $| \ |_p$ defines a quasi-norm on $L^p[0, 1]$ under which it is complete. This quasi-norm has a trivial dual space exactly like $L^0[0, 1]$ above. See [K1] p. 157 or [RR] p. 43 for details.

Definition 19. Let X and Y be TVS. A subset $\mathcal{F} \subseteq L(X, Y)$ is said to be <u>equicontinuous</u> if $\forall$ neighborhood V of 0 in Y, there is a neighborhood U of 0 in X such that $TU \subseteq V$ $\forall T \in \mathcal{F}$.

Theorem 20. Let $\mathcal{F} \subseteq L(X, Y)$ and consider the conditions:

(i) $\mathcal{F}$ is equicontinuous,

(ii) $x_k \to 0$ in X implies $\lim Tx_k = 0$ uniformly for $T \in \mathcal{F}$,

(iii) $x_k \to 0$ in X implies $\lim T_k x_k = 0 \ \forall \{T_k\} \subseteq \mathcal{F}$.

Always (i) ⇒ (ii) ⇒ (iii). If X is a QNLS, then (iii) ⇒ (i) so all 3 conditions are equivalent.

Proof: The first implications are clear. If (i) fails to hold and X is a QNLS, then there is a neighborhood of 0, V, in Y such that $\forall k \exists$ $x_k \in X$ and $T_k \in \mathcal{F}$ with $|x_k| < 1/k$ and $T_k x_k \notin V$. This violates (iii).

Without the quasi-norm assumption, the implication (iii) ⇒ (i) may fail (Exer. 16.16).

For semi-NLS, we have the following norm criterion for equicontinuity.

Theorem 21. Let X and Y be semi-NLS and $\mathcal{F} \subseteq L(X, Y)$. Then $\mathcal{F}$ is equicontinuous if and only if $\{\|T\| : T \in \mathcal{F}\}$ is bounded.

Proof: If $\mathcal{F}$ is equicontinuous, $\exists\ \delta > 0$ such that

$$\|x\| \le \delta \Rightarrow \|Tx\| \le 1 \quad \forall T \in \mathcal{F}$$

Therefore, $\|x\| \le 1 \Rightarrow \|Tx\| \le 1/\delta$ so $\|T\| \le 1/\delta\ \forall T \in \mathcal{F}$.

If $\|T\| \le M \ \ \forall T \in \mathcal{F}$, then $\|Tx\| \le M\|x\| \ \ \forall \ \ T \in \mathcal{F}$ and $\mathcal{F}$ is equicontinuous.

We give one further result on equicontinuity for future use.

Theorem 22. Let X and Y be TVS with Y sequentially complete and

Hausdorff. Assume $\{T_i\} \subseteq L(X, Y)$ is equicontinuous. If $\lim T_i x = Tx$ exists for each x in a dense linear subspace X_0 of X, then $\lim T_i x = Tx$ exists for each $x \in X$ and $T : X \to Y$ is linear and continuous.

Proof: Let V be a neighborhood of 0 in Y. Choose a closed balanced neighborhood of 0, W, in Y such that $W + W + W \subseteq V$. There is a neighborhood of 0, U, in X such that $T_i U \subseteq W$ $\forall i \in \mathbb{N}$. Given $x \in X$, $\exists\, x_0 \in X_0$ such that $x_0 \in x + U$. $\exists\, N$ such that $i, j \geq N$ implies $T_i x_0 - T_j x_0 \in W$. Therefore, if $i, j \geq N$,

$$T_i x - T_j x = T_i(x - x_0) + T_j(x_0 - x) + (T_i - T_j)x_0 \in W + W + W \subseteq V.$$

Hence, $\{T_i x\}$ is Cauchy in Y and converges to some $Tx \in Y$. T is clearly linear and $TU \subseteq W \subseteq V$ since W is closed so T is continuous.

Exercise 1. Define $f : B(S) \to \mathbb{R}$ by $\langle f, \varphi \rangle = \varphi(s_0)$, where s_0 is a fixed element of S. Show $f \in B(S)'$ and compute $\|f\|$.

Exercise 2. $(1 \leq p < \infty)$. Define $R, L : \ell^p \to \ell^p$ by

$$R(\{t_j\}) = (0, t_1, t_2, \ldots) \text{ (\underline{right shift})},$$

$$L(\{t_j\}) = (t_2, t_3, \ldots) \text{ (\underline{left shift})}.$$

Show R, L are linear, continuous and compute $\|R\|, \|L\|, \|R^k\|, \|L^k\|$.

Exercise 3. Let X_1 be a dense linear subspace of the NLS X; let Y be a B-space. If $T : X_1 \to Y$ is linear, continuous, show that T has a unique linear, continuous extension $\hat{T} : X \to Y$ with $\|T\| = \|\hat{T}\|$. [Thus, $L(X_1, Y)$ and $L(X, Y)$ are linearly isometric; in particular, $X_1' = X'$.]

Exercise 4. Describe the dual of c_{00} and m_0 (under $\| \|_\infty$).

Exercise 5. Let X, Y be semi-NLS and equip $X \times Y$ with the semi-norm $\|(x, y)\|_2 = \sqrt{\|x\|^2+\|y\|^2}$. Compute the dual norm of $\| \|_2$ in terms of the dual norms of X and Y.

Exercise 6. Let X be a B-space and $T \in L(X)$. Define e^T and show $e^T \in L(X)$. If T and S commute, show $e^{T+S} = e^T e^S$.

Exercise 7. Let X be a NLS over $\mathbb{F}$. Show X and $L(\mathbb{F}, X)$ are linearly isometric.

Exercise 8. Let $X = L^\infty(0, \infty)$ and define $T : X \to X$ by $Tf(t) = \frac{1}{t}\int_0^t f(s)ds$. Show $T \in L(X)$.

Exercise 9. Show that c and c_0 are linearly homeomorphic.

Exercise 10. Let X, Y be TVS. Describe $(X \times Y)'$ in terms of X' and Y'.

Exercise 11. Let X, Y be TVS and $T : X \to Y$ linear. Show T is bounded if and only if whenever $x_k \to 0$, $\{Tx_k\}$ is bounded if and only if T carries Mackey convergent sequences to Mackey convergent sequences if and only if T carries Mackey convergent sequences to bounded sets.

Exercise 12. Let X, Y be TVS and $T : X \to Y$ linear. If T is bounded on some neighborhood of 0, show T is continuous.

Exercise 13. If X, Y are TVS and $T : X \to Y$ is linear and sequentially continuous, show that T carries $\mathcal{K}$ convergent sequences ($\mathcal{K}$ bounded sets) to $\mathcal{K}$ convergent sequences ($\mathcal{K}$ bounded sets).

Exercise 14. Let X, Y be TVS and $\{T_\delta\}$ be a net in L(X, Y). If $\{T_\delta\}$ is equicontinuous and $\lim T_\delta x = Tx$ exists $\forall x \in X$, show $T \in L(X, Y)$.

Exercise 15. Define $L : c \to \mathbb{R}$ by $L\{t_j\} = \lim t_j$. Show L is continuous, linear and compute $\|L\|$.

Exercise 16. Complete Example 12.

Exercise 17. Let X be an infinite dimensional NLS and $Y \neq \{0\}$ a NLS. Show there exists a linear map $T : X \to Y$ which is not continuous. [Hint: Use a Hamel basis.]

6
Quotient Spaces

Let (X, τ) be a TVS and M a linear subspace of X. Let $\varphi : X \to X/M$ be the quotient map $x \to x + M = [x]$. We give X/M the quotient topology relative to the map φ, i.e., the quotient topology which is denoted by τ/M. The quotient topology is the strongest topology on X/M such that the map φ is continuous and open. It is routine to check that the neighborhoods of 0 defined above satisfy the conditions of 1.18 so τ/M is a vector topology on X/M.

Proposition 1. X/M is Hausdorff if and only if M is closed.

Proof: By 1.10, X/M is Hausdorff if and only if the complement of the origin in X/M is open. But the complement of the origin in X/M is the image under φ of the complement of M, and φ is open and continuous.

Corollary 2. The TVS $X/\{\bar{0}\}$ is Hausdorff.

Consider the case when X is a semi-NLS and M a linear subspace. We show the quotient topology is semi-normable in this case.

If $x \in X$, write $[x] = x + M$ for the coset determined by x. For $x \in X$, set $\|[x]\|' = \inf\{\|x + m\| : m \in M\} = \text{distance}(x, M)$.

Proposition 3. (i) $\| \|'$ defines a semi-norm on X/M.

(ii) $\| \|'$ is a norm if and only if M is closed.

(iii) The quotient map is norm reducing (with respect to $\| \|'$), continuous and open.

(iv) X/M is complete if X is complete.

(v) If X is complete and M is closed, X/M is a B-space.

Proof: (i):

$\|[x + y]\|' = \inf\{\|x + y + m\| : m \in M\}$

$= \inf\{\|x + y + m_1 + m_2\| : m_1, m_2 \in M\}$

$\leq \inf\{\|x + m_1\| : m_1 \in M\} + \inf\{\|y + m_2\| : m_2 \in M\} = \|[x]\|' + \|[y]\|'$:

$\|[tx]\|' = \inf\{\|tx + m\| : m \in M\} = |t| \inf\{\|x + m/t\| : m \in M\} = |t| \, \|[x]\|'$

if $t \neq 0$, and $\|[tx]\|' = 0$ if $t = 0$.

(ii): $\|[x]\|' = 0$ if and only if $x \in \overline{M}$.

(iii): Clearly $\|x\| \geq \|[x]\|'$ so φ is continuous. To show φ is open we show that $\{x : \|x\| < 1\}$ is mapped onto

$$\{[x] : \|[x]\|' < 1\}.$$

Let $\|[x]\|' = 1 - \delta < 1$. $\exists\, m \in M$ such that $\|x + m\| < 1$ and $[x + m] = [x]$.

(iv): Let $\sum_k [x_k]$ be an absolutely convergent series in X/M with respect to $\|\ \|'$. For each k choose $m_k \in M$ such that $\|x_k + m_k\| \le \|[x_k]\|' + 1/2^k$. Thus, $\sum_k (x_k + m_k)$ is absolutely convergent in X and, therefore, convergent (2.9). Set $z = \sum_{k=1}^{\infty} (x_k + m_k)$. By (iii),

$$[z] = \sum_{k=1}^{\infty} [x_k + m_k] = \sum_{k=1}^{\infty} [x_k]$$

and X/M is complete under $\|\ \|'$ by 2.9.

(v): Follows from (iv) and (ii).

It follows from (iii) that the topology induced by $\|\ \|'$ is exactly the quotient topology τ/M. For if U is open in τ/M, then $\varphi^{-1}(U)$ is open in X. But then by (iii), $\varphi(\varphi^{-1}(U)) = U$ is open with respect to $\|\ \|'$. Hence, $\tau/M \subseteq \|\ \|'$. But, τ/M is the strongest topology on X/M such that φ is continuous so $\tau/M = \|\ \|'$.

Let X be a semi-NLS and set $k(X) = \{x : \|x\| = 0\}$.

Proposition 4. (i) k(X) is a closed linear subspace of X.

(ii) The quotient map $\varphi : X \to X/k(X)$ is norm preserving (i.e., is an isometry).

(iii) $X/k(X)$ is a NLS and is a B-space if X is complete.

Proof: (i): If $x, y \in k(X)$, then $\|x + y\| \le \|x\| + \|y\| = 0$ and $\|tx\| = |t|\,\|x\| = 0$. $k(X)$ is closed by 2.4.

(ii): $\|x\| \ge \|[x]\|'$ by Proposition 3 (iii). For $m \in k(X)$, $\|x\| = \|x\| - \|m\| \le \|x + m\| \le \|x\|$ which implies $\|x\| = \|[x]\|'$.

(iii): Follows from Proposition 3.

Proposition 5. Let X be a semi-NLS and M a linear subspace. A subset $B \subseteq X/M$ is bounded if and only if $\exists\, A \subseteq X$ bounded with $[A] \supseteq B$.

Proof: If $A \subseteq X$ is bounded, $[A]$ is bounded by Proposition 3 (iii). If $B \subseteq X/M$ is bounded, $\forall [x] \in B$ $\exists\, m_x \in M$ such that

$$\|x + m_x\| \le \|[x]\|' + 1.$$

Since $\sup\{\|x + m_x\| : [x] \in B\} \le \sup\{\|[x]\|' : [x] \in B\} + 1$, set

$$A = \{x + m_x : [x] \in B\}.$$

Exercise 1. If X is metrizable and M is a closed subspace, show X/M is metrizable. (Hint: 3.3.)

Exercise 2. X, Y TVS, $T \in L(X, Y)$. Show the induced map $\dot{T} : X/\ker T \to Y$ is continuous.

Exercise 3. Let X and Y be semi-NLS and $T \in L(X, Y)$. Show the induced map $\dot{T} : X/\ker(T) \to Y$ is linear, continuous and $\|T\| = \|\dot{T}\|$.

Exercise 4. If X is a TVS and M a linear subspace, show that $(X/M)'$ and $M^{\perp} = \{x' \in X' : \langle x', m \rangle = 0 \ \forall m \in M\}$ are algebraically isomorphic. If X is a NLS, show that these spaces are isometric. ($M^{\perp}$ is called the <u>annihilator</u> of M.)

Exercise 5. If X is a NLS and M a linear subspace, show that M' and $X'/M^{\perp}$ are linearly isometric.

7

Finite Dimensional TVS

In this section we consider finite dimensional TVS. First we show that a Hausdorff vector topology on a finite dimensional TVS is unique.

Theorem 1. Let τ be a vector Hausdorff topology on $\mathbb{F}^n$. Then τ is equivalent to the topology induced by the norm $p(x) = \sum_{k=1}^{n} |x_k|$, $x = (x_1, ..., x_n) \in \mathbb{F}^n$.

Proof: Let p also denote the topology induced by p. Denote the vector with a 1 in the kth coordinate and 0 elsewhere by e_k.

First, we show $\tau \subseteq p$. Let V be a τ-neighborhood of 0. Choose a τ-neighborhood of 0, U, such that $U + ... + U \subseteq V$ (where there are n terms in the sum). Now U is absorbing so $\exists\ \varepsilon > 0$ such that $te_k \in U$ for $|t| \le \varepsilon$ and $k = 1, ..., n$. Hence, if $p(x) \le \varepsilon$, then $x = \sum_{k=1}^{n} x_k e_k \in V$ and

$\tau \subseteq p$.

Next, we show $p \subseteq \tau$. Let $\mathscr{V}$ be the family of τ-closed, balanced neighborhoods of 0. Since τ is Hausdorff, $\bigcap_{V \in \mathscr{V}} V = \{0\}$. Let $B = \{x : p(x) = 1\}$. Then B is p-compact and, hence, τ-compact since $\tau \subseteq p$. Since $B \cap (\bigcap_{V \in \mathscr{V}} V) = \emptyset$, $\mathscr{V}$ does not have the finite intersection property with respect to B. Therefore, $\exists\ V_1, \ldots, V_k \in \mathscr{V}$ such that if $V = \bigcap_{j=1}^{k} V_j$, $V \cap B = \emptyset$. Now V is a balanced τ-neighborhood of 0 and $V \cap B = \emptyset$ so $S(0, 1) = \{x : p(x) < 1\} \supseteq V$. Hence, $S(0, 1)$ is a τ-neighborhood of 0, and $p \subseteq \tau$.

Corollary 2. On a finite dimensional Hausdorff TVS every linear functional is continuous.

Proof: Theorem 1 and Exercise 1.

Corollary 3. If X is a Hausdorff TVS and M is a finite dimensional linear subspace, then M is closed.

Corollary 4. If X is a finite dimensional Hausdorff TVS, then closed, bounded subsets of X are compact.

Remark 5. The converse is false. Closed, bounded subsets of s are compact.

We next show that the converse of Corollary 4 is true for NLS. The proof utilizes a result of F. Riesz which is of interest in its own right and will be used in the study of compact operators in §28.

Lemma 6 (Riesz's Lemma). Let X be a NLS and X_0 a proper, closed subspace of X. Then $\forall\theta$, $0 < \theta < 1$, $\exists\, x_\theta \in X$ such that $\|x_\theta\| = 1$ and $\|x - x_\theta\| \geq \theta\ \forall x \in X_0$.

Proof: Let $x_1 \in X \backslash X_0$ and set $d = \inf\{\|x - x_1\| : x \in X_0\}$. Since X_0 is closed, $d > 0$. $\exists\, x_0 \in X_0$ such that $\|x_1 - x_0\| \leq d/\theta$ since $d/\theta > d$. Set $x_\theta = (x_1 - x_0)/\|x_1 - x_0\|$. Then $\|x_\theta\| = 1$, and if $x \in X_0$, $\|x_0 - x_1\|x + x_0 \in X_0$ so

$$\|x-x_\theta\| = \|x- \frac{x_1}{\|x_1-x_0\|} + \frac{x_0}{\|x_1-x_0\|}\| = \frac{1}{\|x_1-x_0\|}\|(\|x_1-x_0\|x+x_0) - x_1\|$$

$$\geq d/\|x_1-x_0\| \geq \theta.$$

Example 7. In general, x_θ cannot be chosen to be distance 1 from X_0 although this is the case when X_0 is finite dimensional.

Let $X \subseteq C[0, 1]$ be the subspace consisting of those functions x satisfying $x(0) = 0$. Set $X_0 = \{x \in X : \int_0^1 x(t)dt = 0\}$. Suppose $\exists\, x_1 \in X$ such that $\|x_1\| = 1$ and $\|x_1 - x\| \geq 1\ \ \forall x \in X_0$. For $y \in X \backslash X_0$, let $c = \int_0^1 x_1 / \int_0^1 y$. Then $x_1 - cy \in X_0$ and, therefore,

$$1 \leq \|x_1 - (x_1 - cy)\| = |c|\,\|y\|$$

which implies $|\int_0^1 y| \le \|y\| \, |\int_0^1 x_1|$. Now we can make $|\int_0^1 y|$ as close to 1 as we please and still have $\|y\| = 1$ [$y_k(t) = t^{1/k}$ as $k \to \infty$ will work]. Thus, $1 \le |\int_0^1 x_1|$. But since $\|x_1\| = 1$ and $x_1(0) = 0$, $|\int_0^1 x_1| < 1$.

Theorem 8. Let X be a NLS and suppose the unit ball $\{x : \|x\| \le 1\} = B$ is compact. Then X is finite dimensional.

Proof: Suppose that X is not finite dimensional. Let $0 \ne x_1 \in X$ and set $X_1 = \text{span}\{x_1\}$. Then $X_1 \subsetneq X$ and X_1 is closed by Corollary 3. By Lemma 6, $\exists\, x_2 \in B$ such that $\|x_2 - x_1\| \ge 1/2$. Let

$$X_2 = \text{span}\{x_1, x_2\}.$$

Then X_2 is a proper closed subspace of X so by Lemma 6 $\exists\, x_3 \in B$ such that $\|x_3 - x_2\| \ge 1/2$, $\|x_3 - x_1\| \ge 1/2$. Inductively, $\exists$ a sequence $\{x_k\} \subsetneq B$ with $\|x_i - x_j\| \ge 1/2$ for $i \ne j$. But, then B is not sequentially compact.

Corollary 9. Let X be a NLS. The following are equivalent.

(i) X is finite dimensional.

(ii) The closed, unit ball B of X is compact.

(iii) Closed, bounded subsets of X are compact.

(iv) X is locally compact.

It is actually the case that (i) and (iv) are equivalent for general

TVS.

Theorem 10. A locally compact Hausdorff TVS is finite dimensional.

Proof: Let K be a compact neighborhood of 0. Since K contains a closed, balanced neighborhood of 0, we may assume that K is balanced. Since K is compact and $(1/2)K$ is a neighborhood of 0, $\exists$ $x_1, ..., x_k$ such that $K \subseteq \bigcup_{j=1}^{k} (x_j + (1/2)K)$. Let $M = \text{span}\{x_1, ..., x_k\}$. Then M is closed (Corollary 3) and X/M is Hausdorff (6.1). Let $\varphi : X \to X/M$ be the quotient map. Since $K \subseteq M + (1/2)K$, $\varphi(K) \subseteq (1/2)\varphi(K)$ or $2\varphi(K) \subseteq \varphi(K)$. By induction, $\varphi(2^n K) \subseteq \varphi(K)$ $\forall n$. Since K is balanced, $X = \bigcup_{n=1}^{\infty} 2^n K$ and $\varphi(X) = X/M \subseteq \varphi(K)$. But K is compact and φ is continuous so $\varphi(K)$ is compact. Hence, X/M is a compact Hausdorff TVS so $X/M = \{0\}$ or $X = M$.

Exercise 1. If $\mathbb{F}^n$ is normed by p, show any linear functional on $\mathbb{F}^n$ is continuous with respect to p.

Exercise 2. Show that any infinite dimensional B-space has uncountable algebraic dimension.

Exercise 3. Let X be a Hausdorff TVS and $x_0 \in X$, $x_0 \neq 0$. Suppose $t_k \in \mathbb{F}$ is such that $t_k x_0 \to x \in X$. Show $\exists$ $t_0 \in \mathbb{F}$ such that $x = t_0 x_0$ and $t_k \to t_0$. [Hint: Apply Corollary 3 to the subspace spanned by x_0.]

Exercise 4. If X is a finite dimensional Hausdorff TVS and Y a TVS, show that any linear map $T : X \to Y$ is continuous.

Exercise 5. If X, Y are TVS with Y finite dimensional and if $T : X \to Y$ is linear and has a closed kernel, show that T is continuous.

Exercise 6. Two norms $\| \|_1, \| \|_2$ on a vector space X are said to be <u>equivalent</u> if $\exists$ a, b > 0 such that $a\| \|_1 \le \| \|_2 \le b\| \|_1$. Show any 2 norms on a finite dimensional space are equivalent. Give an example of 2 norms which are not equivalent.

Part II
The Three Basic Principles

8

The Hahn-Banach Theorem

In this section we establish the first of the three basic principles of functional analysis, the Hahn-Banach Theorem. For an interesting discussion of the history of the Hahn-Banach Theorem, see [Ho].

Definition 1. Let X be a vector space. A function $p : X \to \mathbb{R}$ is a sublinear functional if

(i) $p(x + y) \le p(x) + p(y)$ $\forall x, y \in X$,

(ii) $p(tx) = tp(x)$ $\forall t \ge 0, x \in X$.

A semi-norm is obviously sublinear but not conversely.

The Hahn-Banach Theorem guarantees that any linear functional defined on a subspace of a vector space which is dominated by a sublinear functional can be extended to a linear functional defined on the entire vector space and the extension is still dominated by the sublinear functional.

Theorem 2 (Hahn-Banach; real case). Let X be a real vector space and $p : X \to \mathbb{R}$ a sublinear functional. Let M be a linear subspace of X. If $f : M \to \mathbb{R}$ is a linear functional such that $f(x) \le p(x)$ $\forall x \in M$, then $\exists$ a linear functional $F : X \to \mathbb{R}$ such that $F(x) = f(x)$, $\forall x \in M$ and $F(x) \le p(x)$ $\forall x \in X$.

Proof: Let $\mathscr{E}$ be the class of all linear extensions g of f such that $g(x) \le p(x)$ $\forall x \in \mathscr{D}(g)$, the domain of g, with $\mathscr{D}(g) \supseteq M$. Note $\mathscr{E} \neq \emptyset$ since $f \in \mathscr{E}$. Partial order $\mathscr{E}$ by $g < h$ if and only if h is a linear extension of g. If $\mathscr{C}$ is a chain in $\mathscr{E}$, then $\bigcup_{g \in \mathscr{C}} g \in \mathscr{E}$ is clearly an upper bound for $\mathscr{C}$ so by Zorn's Lemma $\mathscr{E}$ has a maximal element F. The result follows if we can show $\mathscr{D}(F) = X$.

Suppose $\exists\, x_1 \in X \backslash \mathscr{D}(F)$. Let M_1 be the linear subspace spanned by $\mathscr{D}(F)$ and x_1. Thus, if $y \in M_1$, y has a unique representation in the form $y = m + tx_1$, where $m \in \mathscr{D}(F)$, $t \in \mathbb{R}$. If $z \in \mathbb{R}$, then

$$F_1(y) = F_1(m + tx_1) = F(m) + tz$$

defines a linear functional on M_1 which extends F. If we can show that it is possible to choose z such that $F_1(y) \le p(y)$ $\forall y \in M_1$, this will show $F_1 \in \mathscr{E}$ and contradict the maximality of F.

In order to have

(1) $F_1(y) = F_1(m + tx_1) = F(m) + tz \le p(y) = p(m + tx_1)$,

we must have for $t > 0$, $z \le -\frac{1}{t}F(m) + \frac{1}{t}p(y) = -F(\frac{m}{t}) + p(\frac{m}{t} + x_1)$, or since $m/t \in \mathscr{D}(F)$, if z satisfies

$$(2) \qquad z \le -F(m) + p(m + x_1) \quad \forall m \in \mathscr{D}(F),$$

then (1) holds for $t \ge 0$. For $t < 0$,

$$z \ge -\frac{1}{t}F(m) + \frac{1}{t}p(m + tx_1) = F(-m/t) - p(-m/t - x_1),$$

or since $-m/t \in \mathscr{D}(F)$, if z satisfies,

$$(3) \qquad z \ge F(m) - p(m - x_1) \quad \forall m \in \mathscr{D}(F),$$

then (1) holds. Thus, z must satisfy

$$F(m_1) - p(m_1 - x_1) \le z \le -F(m_2) + p(m_2 + x_1) \quad \forall m_1, m_2 \in \mathscr{D}(F),$$

i.e., we must have

$$(4) \qquad F(m_1) - p(m_1 - x_1) \le -F(m_2) + p(m_2 + x_1) \ \forall m_1,\ m_2 \in \mathscr{D}(F).$$

But,

$$F(m_1 + m_2) = F(m_1) + F(m_2) \le p(m_1 + m_2) \le p(m_1 - x_1) + p(m_2 + x_1)$$

so (4) does hold.

To obtain a complex form of the Hahn-Banach Theorem, we need the following interesting observation which shows how to write a complex linear functional in terms of its real part.

Lemma 3 (Bohnenblust-Sobczyk). Let X be a vector space over $\mathbb{C}$. Suppose $F = f + ig$ is a linear functional on X. Then for $x \in X$, $F(x) = f(x) - if(ix)$ and $f : X \to \mathbb{R}$ is $\mathbb{R}$-linear. Conversely, if $f : X \to \mathbb{R}$ is $\mathbb{R}$-linear, then $F(x) = f(x) - if(ix)$ defines a $\mathbb{C}$-linear functional on X.

Proof: f and g are clearly $\mathbb{R}$-linear. Now $F(ix) = iF(x)$ implies $f(ix) + ig(ix) = if(x) - g(x)$ so $f(ix) = -g(x)$ and $F(x) = f(x) - if(ix)$.

The converse is easily checked.

Theorem 4 (Hahn-Banach; complex case). Let X be a vector space and $p : X \to \mathbb{R}$ a semi-norm. Let M be a linear subspace of X and $f : M \to \mathbb{F}$ a linear functional. If $|f(x)| \le p(x)$ $\forall x \in M$, then f has a linear extension $F : X \to \mathbb{F}$ such that $|F(x)| \le p(x)$ $\forall x \in X$.

Proof: Suppose $\mathbb{F} = \mathbb{R}$. Then $f(x) \le |f(x)| \le p(x)$ $\forall x \in M$ so Theorem 2 implies $\exists$ a linear extension $F : X \to \mathbb{R}$ such that $F(x) \le p(x)$ $\forall x \in X$. But then $F(-x) = -F(x) \le p(-x) = p(x)$ so $|F(x)| \le p(x)$.

Suppose $\mathbb{F} = \mathbb{C}$. Then $\mathcal{R}f$, the real part of f, is an $\mathbb{R}$-linear functional on X such that $|\mathcal{R}f(x)| \le |f(x)| \le p(x)$ $\forall x \in M$. By the first part, $\exists$ a real linear functional $f_1 : X \to \mathbb{R}$ which extends $\mathcal{R}f$ and satisfies $|f_1(x)| \le p(x)$ $\forall x \in X$. Set $F(x) = f_1(x) - if_1(ix)$. Then F is $\mathbb{C}$-linear and extends f by Lemma 3.

For $x \in X$, write $F(x) = |F(x)|e^{i\theta}$. Then

$$|F(x)| = e^{-i\theta}F(x) = F(e^{-i\theta}x) = \mathcal{R}F(e^{-i\theta}x) = f_1(e^{-i\theta}x) \le p(e^{-i\theta}x) = p(x) .$$

Exercise 1. Give an example of a sublinear functional which is not a semi-norm.

8.1 Applications of the Hahn-Banach Theorem in NLS

Despite its somewhat esoteric appearance, we show in this section that the Hahn-Banach Theorem has many applications which show the existence of continuous linear functionals on NLS with important and useful properties.

Theorem 1. Let X be a semi-NLS and M a linear subspace. If $y' \in M'$, then $\exists\, x' \in X'$ such that x' extends y' and

$$\|x'\| = \sup\{|\langle x', x\rangle| : x \in X, \|x\| \le 1\} = \|y'\|$$
$$= \sup\{|\langle y', x\rangle| : x \in M, \|x\| \le 1\},$$

i.e., the norm of the extension is equal to the norm of the original linear functional.

Proof: Define a semi-norm p on X by $\|x\|\|y'\| = p(x)$. Then $|\langle y', x\rangle| \le p(x)\ \forall\, x \in M$. By Theorem 8.3, $\exists$ a linear functional x' extending y' such that $|\langle x', x\rangle| \le p(x) = \|x\|\|y'\|\ \forall\, x \in X$. Thus, $x' \in X'$ and $\|x'\| \le \|y'\|$. But, certainly, $\|y'\| \le \|x'\|$.

There is no analogue of Theorem 1 for linear operators between NLS; for example, $\ell^\infty \supseteq c_0$ but there is no continuous linear extension of the identity operator on c_0 to an operator from ℓ^∞ to c_0 (27.5).

Theorem 2. Let M be a linear subspace of a NLS X. Suppose $x_0 \in X$ is such that distance$(x_0, M) = d > 0$. Then $\exists\, x_0' \in X'$ such that $\|x_0'\| = 1$, $\langle x_0', M\rangle = 0$ and $\langle x_0', x_0\rangle = d$.

Proof: Let M_0 be the linear subspace spanned by M and x_0. Define a linear functional f on M_0 by $f(m_1) = f(m + tx_0) = td$, where $m \in M$, $t \in \mathbb{F}$. Then $\langle f, M \rangle = 0$ and $f(x_0) = d$. Also $f \in M_0'$ with $\|f\| \le 1$ since if $t \neq 0$, $m \in M$, then

$$\|m + tx_0\| = |t|\,\|m/t + x_0\| \ge |t|d = |f(m + tx_0)|.$$

Actually, $\|f\| = 1$ since $\exists\ \{m_k\} \subseteq M$ with $\|m_k - x_0\| \downarrow d$ and, therefore, $d = |f(m_k - x_0)| \le \|f\|\|m_k - x_0\| \downarrow \|f\|d$ so $\|f\| \ge 1$.

Now apply Theorem 1 to f.

Remark 3. Note in particular that Theorem 2 is applicable if M is closed and $x_0 \notin M$. Therefore, if M is a linear subspace of X and if $\langle x', M \rangle = 0 \Rightarrow x' = 0$ for all $x' \in X'$, then M is dense in X.

As a corollary of Theorem 2 we show that the dual of a NLS always separates the points of the NLS, i.e., a NLS always has a non-trivial dual (recall Example 5.18).

Corollary 4. Let X be a NLS and $x_0 \in X$, $x_0 \neq 0$. Then $\exists\ x_0' \in X'$ such that $\|x_0'\| = 1$ and $\langle x_0', x_0 \rangle = \|x_0\|$. (In particular, if $x \neq y$, $x, y \in X$, then $\exists\ x_0' \in X'$ such that $\langle x_0', x - y \rangle \neq 0$, i.e., the dual X' separates the points of X.)

Proof: Put $M = \{0\}$ in Theorem 2.

If X is a NLS, then the dual norm on X';

$$\|x'\| = \sup\{ |\langle x', x\rangle| : \|x\| \le 1\} ,$$

is found by computing the sup of $|x'|$ over the unit sphere in X. We establish a dual result; we show that the norm of an element $x \in X$ can be found by computing the sup of x over the unit sphere in X'.

Corollary 5. For each x in a NLS X, $\|x\| = \sup\{ |\langle x', x\rangle| : \|x'\| \le 1\}$ and the sup is attained.

Proof: If $x' \in X'$, $\|x'\| \le 1$, then $|\langle x', x\rangle| \le \|x\|$. On the other hand, by Corollary 4 $\exists\, x' \in X'$, $\|x'\| = 1$, such that $\langle x', x\rangle = \|x\|$.

As an interesting consequence of Corollary 5, we can show that any NLS can be isometrically imbedded in the space B(S) for some S.

Corollary 6. Let X be a NLS. Then $\exists\, S \neq \emptyset$ such that X is linearly isometric to a linear subspace of B(S).

Proof: Let $S = \{x' \in X' : \|x'\| \le 1\}$. For $x \in X$ define $Ux = \hat{x} \in B(S)$ by $\hat{x}(x') = \langle x', x\rangle$. By Corollary 5, $\|\hat{x}\|_\infty = \|x\|$ so $U : X \to B(S)$ is an isometry which is obviously linear.

Later we will show that S carries a natural topology under which it is a compact Hausdorff space and the space B(S) can be replaced by C(S).

We next consider some separability results for a NLS and its dual.

Proposition 7. If the dual of a NLS X is separable, then X is separable.

Proof: Let $\{x_k'\}$ be dense in X'. For each k choose $x_k \in X$ such that $\|x_k\| = 1$ and $|<x_k', x_k>| \geq \|x_k'\|/2$. The subspace X_1 spanned by $\{x_k : k \in \mathbb{N}\}$ is dense in X; for if this were not the case, by Theorem 2, $\exists\ x' \in X'$, $\|x'\| = 1$, such that $<x', X_1> = 0$. Since $\{x_k'\}$ is dense in X', $\exists$ a subsequence $\{x_{n_k}'\}$ such that $x_{n_k}' \to x'$ and since

$$\|x_{n_k}' - x'\| \geq |<x_{n_k}' - x', x_{n_k}>| = |<x_{n_k}', x_{n_k}>| \geq \|x_{n_k}'\|/2\ ,$$

$\|x_{n_k}'\| \to 0$. But, $\|x_{n_k}'\| \to \|x'\| = 1$. This contradiction shows that X_1 is dense in X.

Example 8. The converse is false; consider ℓ^1 and its dual ℓ^∞. (Exer. 2.16).

Even though a separable space needn't have a separable dual, it does have a countability property that can sometimes be substituted for separability.

Definition 9. If X is a NLS, a subset $\Gamma \subseteq X'$ is said to be a <u>norming set</u> (for X) if $\|x\| = \sup\{|<x', x>| : x' \in \Gamma\}$ $\forall\ x \in X$.

Example 10. By Corollary 5, $\{x' \in X' : \|x'\| \leq 1\}$ is always a norming set. However, some spaces have much smaller norming sets. For example, in c_0 (c_{00}, c), the set $\{e_k : k \in \mathbb{N}\} \subseteq \ell^1$ is a norming set. In C(S), the set of Dirac measures, $\{\delta_t : t \in S\}$, is a norming set [here, $<\delta_t, f> = f(t)$].

Proposition 11. Let X be a separable NLS. Then X' contains a countable norming set for X.

Proof: Let $\{x_k\}$ be dense in X. By Corollary 4, $\forall\, k\ \exists\, x_k' \in X'$, $\|x_k'\| = 1$, such that $|\langle x_k', x_k\rangle| = \|x_k\|$. Set $\Gamma = \{x_k' : k \in \mathbb{N}\}$. If $x \in X$ and $\varepsilon > 0$, then $\exists\, x_k$ such that $|\,\|x\| - \|x_k\|\,| \le \|x - x_k\| < \varepsilon$. Then $\|x\| \ge |\langle x_k', x\rangle| \ge |\langle x_k', x_k\rangle| - |\langle x_k', x - x_k\rangle| \ge \|x_k\| - \varepsilon \ge \|x\| - 2\varepsilon$ and $\|x\| = \sup\{|\langle x_k', x\rangle| : k\}$.

The Canonical Map and Reflexivity:

Let X be a NLS. Let X'' be the dual of X' (with the dual norm) and assume that X'' carries its dual norm from X'. X'' is called the <u>second dual</u> or <u>bidual</u> of X. Each $x \in X$ induces an element $\hat{x} \in X''$ defined by $\langle \hat{x}, x'\rangle = \langle x', x\rangle$ for $x' \in X'$. $\hat{x}$ is obviously linear and by Corollary 5, $\|\hat{x}\| = \sup\{|\langle x', x\rangle| : \|x'\| \le 1\} = \|x\|$, so $\hat{x} \in X''$. Thus, the map $J_X : X \to X''$, $J_X x = \hat{x}$, defines a linear, isometric imbedding of X into its second dual. When there is no possibility of misunderstanding, we write $J_X = J$. A NLS is said to be <u>reflexive</u> if the canonical imbedding J is onto. Note from 5.8 that any reflexive NLS must be complete, i.e., must be a B-space. Also note that for a B-space X to be reflexive, X and X'' must be linearly isometric <u>under</u> the canonical imbedding J; R.C. James has given an example of a non-reflexive B-space X such that X and X'' are linearly isometric [J1].

Example 12. If $1 < p < \infty$, $L^p(S, \Sigma, \mu)$ is reflexive; in particular, ℓ^p is reflexive for $1 < p < \infty$.

Example 13. ℓ^1 is not reflexive; more generally, $L^1(S, \Sigma, \mu)$ is not reflexive unless it is finite dimensional. Also, c_0 and c are not reflexive.

Theorem 14. Every NLS X is a dense subspace of a B-space $\tilde{X}$ (i.e., every NLS has a B-space completion).

Proof: Set $\tilde{X} = \overline{JX} \subseteq X''$.

Here we are following the common procedure of identifying X and JX under the linear isometry J.

We now establish several properties of reflexive spaces.

Theorem 15. A B-space X is reflexive if and only if X' is reflexive.

Proof: Suppose X is reflexive and let $x''' \in X'''$. Define $x' \in X'$ by $x' = x'''J_X$. For any $J_X x \in X''$,

$$<x''', J_X x> = <x', x> = <J_X x, x'> = <J_{X'}x', J_x x>$$

so $x''' = J_{X'}x'$.

Suppose X' is reflexive. By Remark 3, if we show $<x''', J_X X> = 0$ implies $x''' = 0$, then $J_X X$ is dense in X'' and must be equal to X'' since it is complete and closed. Let $x''' \in X'''$ and let $x' \in X'$ be such that $J_{X'}x' = x'''$. If

$$<x''', J_X x> = <J_{X'}x', J_X x> = <x', x> = 0$$

for all $x \in X$, then $x' = 0$ so $x''' = 0$.

Theorem 16. A closed linear subspace M of a reflexive space X is reflexive.

Proof: Let $m'' \in M''$. For $x' \in X'$ let $x'_M \in M'$ be the restriction of x' to M. Define $x'' \in X''$ by $\langle x'', x' \rangle = \langle m'', x'_M \rangle$. $\exists\, x \in X$ such that $x'' = \hat{x}$. But $x \in M$ since if $x' \in X'$ vanishes on M, $\langle x', x \rangle = \langle x'', x' \rangle = \langle m'', x'_M \rangle = \langle m'', 0 \rangle = 0$ and $x \in \bar{M}$ by Theorem 2. Also $\hat{x} = m''$ since if $m' \in M'$, we can extend m' to $x' \in X'$ (Theorem 1) and then

$$\langle x'', x' \rangle = \langle m'', m' \rangle = \langle x', x \rangle = \langle m', x \rangle = \langle \hat{x}, m' \rangle .$$

Another interesting property of reflexive spaces is given by

Theorem 17. Let X be a reflexive B-space. Then every continuous linear functional $x' \in X'$ attains its maximum on the closed unit ball of X.

Proof: By Corollary 5 $\exists\, x'' \in X''$ such that, $\|x''\| = 1$, $\langle x'', x' \rangle = \|x'\|$. But $\exists\, x \in X$, $\|x\| = 1$, such that $Jx = x''$.

It is an interesting result of R.C. James that the converse of Theorem 17 holds ([J2]). To illustrate the content of Theorem 17 consider the map $h : c_0 \to \mathbb{R}$ defined by $h(\{t_j\}) = \sum_{j=1}^{\infty} t_j/j!$. Then $h \in c_0'$ and $\|h\| = \sum_{j=1}^{\infty} 1/j!$. However, there is no $\{t_j\} \in c_0$ such that $\|h\| = \langle h, \{t_j\}\rangle$. [This also shows that c_0 is not reflexive.]

Finally, we establish the converse of Proposition 5.8.

Theorem 18. Let X and Y be NLS with $X \neq \{0\}$. If $L(X, Y)$ is a B-space, then Y is a B-space.

Proof: Let $\{y_k\}$ be Cauchy in Y. Choose $x_0 \in X$, $\|x_0\| = 1$. By Corollary 4 $\exists\, x_0' \in X'$ such that $\|x_0'\| = 1$ and $\langle x_0', x_0\rangle = 1$. Define $T_k \in L(X, Y)$ by $T_k x = \langle x_0', x\rangle y_k$. Then $\|(T_k - T_j)x\| \le \|x\|\|y_k - y_j\|$ so $\{T_k\}$ is Cauchy in $L(X, Y)$ and, therefore, convergent to, say, T. Then $\|y_k - Tx_0\| = \|T_k x_0 - Tx_0\| \le \|T_k - T\|$ implies $y_k \to Tx_0$.

Exercise 1. If X is a real NLS, show $\|x\| = \sup\{\langle x', x\rangle : \|x'\| \le 1\}$ and the sup is attained (compare Corollary 5).

Exercise 2. Show the dual of a separable reflexive space is separable.

Exercise 3. If X is reflexive and X' contains a countable set which separates the points of X, show X' is separable.

Exercise 4. If X is a NLS, show JX separates the points of X'.

Exercise 5. If X is reflexive, show X' has no proper closed subspaces which separates the points of X.

Exercise 6. Show there is no norm on $L^0[0, 1]$ (Example 5.18) which generates its topology.

Exercise 7. Give an example of a countable norming set for ℓ^1.

Exercise 8. Show that the quotient of a reflexive space by a closed linear subspace is reflexive.

8.2 Banach Limits

Define the linear functional ℓ on c (with real coefficients) by $\ell(\{t_j\}) = \lim_j t_j$. ℓ is a continuous linear functional on c with $\|\ell\| = 1$. We show that ℓ has an extension to ℓ^∞ which still retains many of the desirable properties of the limit operation. For $x = (x_1, x_2, ...) \in \ell^\infty$ define $\tau x = (x_2, x_3, ...)$, i.e., $\tau : \ell^\infty \to \ell^\infty$ is the shift operator which shifts the coordinates of a sequence one place to the left.

A <u>Banach limit</u> is a continuous linear functional L on ℓ^∞ which satisfies

(i) $L(x) \geq 0$ if $x \geq 0$ [i.e., if $x_k \geq 0$ $\forall k$],

(ii) $L(x) = L(\tau x)$ $\forall$ x, and

(iii) $L((1, 1, ...)) = 1$.

Any Banach limit agrees with ℓ on convergent sequences. More generally, we have

Proposition 1. If L is a Banach limit, then $\underline{\lim}\, x_k \leq L(x) \leq \overline{\lim}\, x_k$ $\forall\, x \in \ell^\infty$.

Proof: Let $\varepsilon > 0$. Choose N such that $\inf_k x_k \leq x_N < \inf_k x_k + \varepsilon$. Thus, $x_k + \varepsilon - x_N > 0$ $\forall$ k. By (i) and (iii), we have $L(x) + \varepsilon - x_N \geq 0$ so $L(x) \geq \inf_k x_k - \varepsilon$ and $L(x) \geq \inf_k x_k$. Similarly, $L(x) \leq \sup_k x_k$.

Since $\overline{\lim}\, x_k = \inf_n \sup_{k \geq n} x_k$ and $\underline{\lim}\, x_k = \sup_n \inf_{k \geq n} x_k$, the result follows from (ii) and the observation above.

We now show that the Hahn-Banach Theorem can be used to show the existence of Banach limits.

Theorem 2. Banach limits exist.

Proof: For $x = (x_1, x_2, \ldots) \in \ell^\infty$, set $p(x) = \overline{\lim} \frac{x_1+\ldots+x_n}{n}$. Then $p(x + y) \le p(x) + p(y)$ and $p(tx) = tp(x)$ $\forall t \ge 0$ so p is a sublinear functional with $p(x) \le \|x\|_\infty$. For $x \in c$, $\ell(x) = p(x)$ so by Theorem 8.2, $\exists$ an extension $L : \ell^\infty \to \mathbb{R}$ of ℓ satisfying $L(x) \le p(x)$ $\forall x \in \ell^\infty$.

For $x \in \ell^\infty$, $L(-x) = -L(x) \le p(-x)$ so $L(x) \ge -p(-x) = \underline{\lim} \frac{x_1+\ldots+x_n}{n}$ and (i) follows. (iii) is clear so it remains to show (ii). For $x \in \ell^\infty$, $p(x - \tau x) = \overline{\lim} \frac{x_1-x_{n+1}}{n} = 0$ so $L(x - \tau x) \le 0$ $\forall$ x. Hence, $L(x) = L(\tau x)$ $\forall$ x.

Note from the construction, we also have that $\|L\| = 1$.

Banach limits have applications in the theory of finitely additive set functions. See Exer. 1 where the existence of a finitely additive set function which is not a measure is sketched. See also [Jor] for further such applications.

Banach also used the Hahn-Banach Theorem to show the existence of a translation invariant, positive, finitely additive set function which is defined on all subsets of $\mathbb{R}$ and assigns unit measure to $[0, 1]$. (Recall that no such countably additive set function exists.) [See [BN], p. 188, for details.]

Exercise 1. Let $\mathcal{P}$ be the power set of $\mathbb{N}$. Show that there exists a positive finitely additive set function μ on $\mathcal{P}$ which is not countably additive. [Hint: For each n construct a probability measure on μ_n on $\mathcal{P}$ such that $\mu_n(\{1, ..., n\}) = 0$ and then set $\mu(E) = L(\{\mu_n(E)\})$.]

8.3 The Moment Problem

A classical moment problem is the following: given a sequence $c_0, c_1, \ldots$ of real numbers, when does there exist a function of bounded variation g such that $c_n = \int_0^1 t^n dg(t)$ for all $n \geq 0$? That is, in probability terms, when do the moments of a probability distribution determine the distribution? Since the dual of $C[0, 1]$ can be identified with (normalized) functions of bounded variation (Example 5.16), this suggests an abstract formulation of the moment problem. Given a NLS X, a subset $\{x_a : a \in A\}$ of X and a subset $\{c_a : a \in A\}$ of scalars, when does there exist an $x' \in X'$ such that $\langle x', x_a \rangle = c_a$ $\forall\, a \in A$? We use the Hahn-Banach Theorem to give a complete answer to this question.

Theorem 1. The following are equivalent:

(i) $\exists\, x' \in X'$ such that $\langle x', x_a \rangle = c_a$ $\forall a \in A$,

(ii) $\exists\, M > 0$ such that $\left| \sum_{a \in \sigma} t_a c_a \right| \leq M \left\| \sum_{a \in \sigma} t_a x_a \right\|$ for all finite subsets $\sigma \subseteq A$ and scalars t_a.

Proof: (i) implies (ii): If σ is a finite subset of A,

$$\left| \sum_{a \in \sigma} t_a c_a \right| = \left| \sum_{a \in \sigma} t_a \langle x', x_a \rangle \right| \leq \|x'\| \left\| \sum_{a \in \sigma} t_a x_a \right\| .$$

(ii) implies (i): Let X_0 be the linear subspace spanned by $\{x_a : a \in A\}$. If $x = \sum_{a \in \sigma} t_a x_a$, $\sigma \subseteq A$ finite, belongs to X_0, define

$x' : X_0 \to \mathbb{F}$ by $\langle x', x\rangle = \sum_{a\in\sigma} t_a c_a$. First, note that x' is well-defined. For if $x = \sum_{a\in\sigma} t_a x_a = \sum_{b\in\tau} s_b x_b$, set $r_i = t_i - s_i$ if $i \in \sigma \cap \tau$, $r_i = t_i$ if $i \in \sigma\backslash\tau$ and $r_i = -s_i$ if $i \in \tau\backslash\sigma$. Then

$$|\sum_{a\in\sigma} t_a c_a - \sum_{b\in\tau} s_b c_b| = |\sum_{i\in\sigma\cup\tau} r_i c_i| \le M \|\sum_{i\in\sigma\cup\tau} r_i c_i\|$$

$$= M\|\sum_{a\in\sigma} t_a c_a - \sum_{b\in\tau} s_b x_b\| = 0$$

so x' is well-defined. Since

$$|\langle x', \sum_{a\in\sigma} t_a x_a\rangle| = |\sum_{a\in\sigma} t_a c_a| \le M\|\sum_{a\in\sigma} t_a x_a\| ,$$

x' is a continuous linear functional on X_0 with norm $\le M$. Now extend x' to X by the Hahn-Banach Corollary 8.1.1.

Whereas we have used the Hahn-Banach Theorem to solve the abstract moment problem above, it is interesting that a moment problem in $C[0, 1]$ motivated the first early version of the Hahn-Banach Theroem. E. Helly solved a moment problem of F. Riesz by developing a form of the Hahn-Banach Theorem in $C[0, 1]$. Although his theorem was proven in the special case of $C[0, 1]$, it is entirely modern in spirit and, in fact, is essentially the proof given in §8. For an interesting discussion of Helly's paper, which also contains an early version of the Uniform Boundedness Principle, see [Ho].

9
The Uniform Boundedness Principle (UBP)

The Uniform Boundedness Principle (UBP) was one of the earliest abstract results in functional analysis ([B], [H]). These early versions of the UBP were established by Hahn and Banach by employing "sliding hump" arguments. In this section we use a theorem about infinite matrices to establish a very general version of the UBP which requires no completeness assumptions on the domain space and then give several applications to illustrate its utility. We begin by establishing the matrix theorem, called the Basic Matrix Theorem, due to Antosik and Mikusinski; the Basic Matrix Theorem can be viewed as an abstract "sliding hump" result.

First, a simple lemma.

Lemma 1. Let X be a TVS and $x_{ij} \in X$ for $i, j \in \mathbb{N}$. If $\lim_i x_{ij} = 0 \ \forall j$ and $\lim_j x_{ij} = 0 \ \forall i$ and if $\{U_k\}$ is a sequence of neighborhoods of 0 in X, then $\exists$ an increasing sequence of positive integers $\{p_i\}$ such that

$x_{p_i p_j}, x_{p_j p_i} \in U_j$ for $j > i$.

Proof: Set $p_1 = 1$. $\exists\, p_2 > p_1$ such that $x_{ip_1} \in U_2$, $x_{p_1 j} \in U_2$ for $i, j \geq p_2$.

Then $\exists\, p_3 > p_2$ such that $x_{ip_1}, x_{ip_2}, x_{p_1 j}, x_{p_2 j} \in U_3$ for $i, j \geq p_3$. Continue.

We now establish the <u>Basic Matrix Theorem</u>. As one of its consequences the result gives sufficient conditions for the diagonal of an infinite matrix with values in a TVS to converge to 0. For this reason the result is sometimes referred to as a <u>Diagonal Theorem</u>. The result for metric linear spaces is due to Antosik and Mikusinski (see [AS] for the metric case and many applications).

Theorem 2 (Basic Matrix Theorem). Let $x_{ij} \in X$ for $i, j \in \mathbb{N}$. Suppose

(I) $\lim_i x_{ij} = x_j$ exists for each j and

(II) for each increasing sequence of positive integers $\{m_j\}$ there is a subsequence $\{n_j\}$ of $\{m_j\}$ such that $\{\sum_{j=1}^{\infty} x_{in_j}\}_{i=1}^{\infty}$ is Cauchy.

Then $\lim_i x_{ij} = x_j$ uniformly for $j \in \mathbb{N}$. In particular,

$$\lim_i \lim_j x_{ij} = \lim_j \lim_i x_{ij} = 0 \text{ and } \lim_i x_{ii} = 0 .$$

Proof: If the conclusion fails, there is a closed, symmetric neighborhood U_0 of 0 and increasing sequences of positive integers $\{m_k\}$ and $\{n_k\}$ such that $x_{m_k n_k} - x_{n_k} \notin U_0$ for all k. Pick a closed, symmetric neighborhood U_1 of 0 such that $U_1 + U_1 \subseteq U_0$ and set $i_1 = m_1$, $j_1 = n_1$. Since $x_{i_1 j_1} - x_{j_1} = (x_{i_1 j_1} - x_{ij_1}) + (x_{ij_1} - x_{j_1})$, there exists i_0 such that $x_{i_1 j_1} - x_{ij_1} \notin U_1$ for $i \geq i_0$. Choose k_0 such that

$$m_{k_0} > \max\{i_1, i_0\}, n_{k_0} > j_1 \text{ and set } i_2 = m_{k_0}, j_2 = n_{k_0}.$$

Then $x_{i_1 j_1} - x_{i_2 j_1} \notin U_1$ and $x_{i_2 j_2} - x_{j_2} \notin U_0$. Proceeding in this manner produces increasing sequences $\{i_k\}$ and $\{j_k\}$ such that $x_{i_k j_k} - x_{j_k} \notin U_0$ and $x_{i_k j_k} - x_{i_{k+1} j_k} \notin U_1$. For convenience, set $z_{k,\ell} = x_{i_k j_\ell} - x_{i_{k+1} j_\ell}$ so $z_{k,k} \notin U_1$.

Choose a sequence of closed, symmetric neighborhoods of 0, $\{U_n\}$, such that $U_n + U_n \subseteq U_{n-1}$ for $n \geq 1$. Note that

$$U_3 + U_4 + \cdots + U_m = \sum_{j=3}^{m} U_j \subseteq U_2$$

for each $m \geq 3$. By (I) and (II), $\lim_k z_{k\ell} = 0$ for each ℓ and $\lim_\ell z_{k\ell} = 0$ for each k so by Lemma 1 there is an increasing sequence of positive integers $\{p_k\}$ such that $z_{p_k p_\ell}, z_{p_\ell p_k} \in U_{k+2}$ for $k > \ell$. By (II) $\{p_k\}$ has a subsequence $\{q_k\}$ such that $\{\sum_{k=1}^{\infty} x_{iq_k}\}_{i=1}^{\infty}$ is Cauchy so

$$\lim_k \sum_{\ell=1}^{\infty} z_{q_k q_\ell} = 0 .$$

Thus, there exists k_0 such that $\sum_{\ell=1}^{\infty} z_{q_{k_0} q_\ell} \in U_2$. Then for $m > k_0$

$$\sum_{\substack{\ell=1 \\ \ell \neq k_0}}^{m} z_{q_{k_0} q_\ell} = \sum_{\ell=1}^{k_0 - 1} z_{q_{k_0} q_\ell} + \sum_{\ell=k_0+1}^{m} z_{q_{k_0} q_\ell} \in \sum_{\ell=1}^{k_0-1} U_{k_0+2}$$

$$+ \sum_{\ell=k_0+1}^{m} U_{\ell+2} \subseteq \sum_{\ell=3}^{m+2} U_\ell \subseteq U_2$$

so $z_{k_0} = \sum_{\substack{\ell=1 \\ \ell \neq k_0}}^{\infty} z_{q_{k_0} q_\ell} \in U_2$. Thus,

$$z_{q_{k_0} q_{k_0}} = \sum_{\ell=1}^{\infty} z_{q_{k_0} q_\ell} - z_{k_0} \in U_2 + U_2 \subseteq U_1 .$$

This is a contradiction and establishes the result.

A matrix which satisfies conditions (I) and (II) is called a <u>$\mathcal{K}$-matrix</u>.

Let X and Y be TVS. A Uniform Boundedness Principle (UBP) asserts that a family $\mathcal{F} \subseteq L(X, Y)$ which is pointwise bounded on X [i.e., $\{Tx : T \in \mathcal{F}\}$ is bounded in Y $\forall x \in X$] is uniformly bounded on some family of subsets of the domain space [$\mathcal{F}$ is <u>uniformly bounded</u> on a family $\mathcal{A}$ of subsets of X if $\{Tx : x \in A, T \in \mathcal{F}\}$ is bounded $\forall A \in \mathcal{A}$]. The classical UBP for NLS asserts that any pointwise bounded family of continuous linear operators on a B-space X is uniformly bounded on the family of all bounded subsets of X (see Corollaries 6 and 7 below). If the domain space X does not satisfy some type of completeness condition, then as the following example illustrates the family of bounded subsets of the domain space is in general too large to draw such a conclusion.

Example 3. Consider the sequence $\{ke_k\}$ in ℓ^1, the dual of c_{00} (Exer. 5.4). This sequence of continuous linear functionals on c_{00} is pointwise bounded on c_{00} but is not uniformly bounded on the bounded subset $\{e_k : k \in \mathbb{N}\}$ of c_{00}.

Whereas the family of all bounded subsets of the domain space is too large a family to draw the conclusion that a pointwise bounded family of continuous linear operators is uniformly bounded on each member of this family, we show that the families of $\mathcal{K}$ convergent sequences and $\mathcal{K}$ bounded sets do allow such a conclusion to be drawn. We then show that this general UBP yields the classical B-space version of the UBP as an immediate corollary.

If $\mathcal{F} \subseteq L(X, Y)$, let $w(\mathcal{F})$ be the weakest topology on X such that all members of $\mathcal{F}$ are continuous (1.19). Note that $w(\mathcal{F})$ is weaker than $w(L(X, Y))$ which is weaker than the original topology of X.

Theorem 4 (General UBP). Suppose that $\mathcal{F}$ is pointwise bounded on X, i.e., $\{Tx : T \in \mathcal{F}\}$ is bounded $\forall\, x \in X$. Then

(i) $\mathcal{S}$ is uniformly bounded on $w(\mathcal{S})$ - $\mathcal{K}$ convergent sequences and

(ii) $\mathcal{S}$ is uniformly bounded on $w(\mathcal{S})$ - $\mathcal{K}$ bounded subsets of X.

Proof: If (i) fails, there is a balanced neighborhood of 0, U, in Y and a $w(\mathcal{S})$ - $\mathcal{K}$ convergent sequence $\{x_j\}$ such that

$$\{Tx_j : T \in \mathcal{S}, j \in \mathbb{N}\}$$

is not absorbed by U. $\exists\, T_1 \in \mathcal{S}$, x_{n_1} such that $T_1x_{n_1} \notin U$. Put $k_1 = 1$. By the pointwise boundedness, $\exists\, k_2 > k_1$ such that

$$\{Tx_j : T \in \mathcal{S}, 1 \le j \le n_1\} \subseteq k_2U.$$

$\exists\, T_2 \in \mathcal{S}$, x_{n_2} such that $T_2x_{n_2} \notin k_2U$. Thus, $n_2 > n_1$. Continuing produces increasing sequences $\{n_i\}$, $\{k_i\}$ and a sequence $\{T_i\} \subseteq \mathcal{S}$ such that $T_ix_{n_i} \notin k_iU$. Set $t_i = 1/k_i$ and note $t_i \to 0$.

Now consider the matrix $M = [t_iT_ix_{n_j}]$. By the pointwise boundedness, $\lim_i t_iT_ix_{n_j} = 0$ for each j. By the $w(\mathcal{S})$ - $\mathcal{K}$ convergence of $\{x_j\}$, for any subsequence of $\{x_{n_j}\}$, there is a further subsequence $\{x_{p_j}\}$ such that the series $\sum_{j=1}^{\infty} x_{p_j}$ is $w(\mathcal{S})$ convergent to some $x \in X$. Thus,

$$\sum_{j=1}^{\infty} t_iT_ix_{p_j} = t_iT_i(x) \to 0$$

by the pointwise boundedness. Hence, M is a $\mathcal{K}$-matrix and Theorem 2 implies $t_iT_ix_i \to 0$. In particular, $t_iT_ix_i \in U$ for large i which contradicts the construction above.

For (ii), let $B \subseteq X$ be $w(\mathscr{F})$ - $\mathscr{K}$ bounded. To show that $\{Tx : T \in \mathscr{F}, x \in B\}$ is bounded, it suffices to show that $\{T_jx_j\}$ is bounded for each $\{x_j\} \subseteq B$, $\{T_j\} \subseteq \mathscr{F}$. Now $\{x_j/\sqrt{j}\}$ is $w(\mathscr{F})$ - $\mathscr{K}$ convergent so part (i) implies that $\{T_j(x_j/\sqrt{j})\}$ is bounded. Hence

$$1/\sqrt{j}\; T_j(x_j/\sqrt{j}) = \frac{1}{j} T_jx_j \to 0$$

and $\{T_jx_j\}$ is bounded (4.2).

Recall (4.9) that a $\mathscr{K}$ convergent sequence needn't be $\mathscr{K}$ bounded so this is the reason for the dual conclusions in (i) and (ii).

Note that the families of bounded sets in (i) and (ii) depend upon the family of operators $\mathscr{F}$. In order to obtain a family of bounded subsets, $\mathscr{A}$, which has the property that <u>any</u> pointwise bounded family of continuous linear operators is uniformly bounded on the members of $\mathscr{A}$, we can take the topology $w(L(X, Y))$ which is stronger than $w(\mathscr{F})$. Since the original topology, τ, of X is stronger than $w(L(X, Y))$, we have

Corollary 5. If $\mathscr{F} \subseteq L(X, Y)$ is pointwise bounded, then $\mathscr{F}$ is uniformly bounded on

(i) $w(L(X, Y))$ - $\mathscr{K}$ [τ - $\mathscr{K}$] convergent sequences and

(ii) $w(L(X, Y))$ - $\mathscr{K}$ [τ - $\mathscr{K}$] bounded subsets of X.

Recall that an $\mathscr{A}$-space is a TVS in which bounded sets are $\mathscr{K}$ bounded (§4) so for $\mathscr{A}$-spaces we have the following version of the UBP.

Corollary 6. If X is an $\mathcal{A}$-space, then any pointwise bounded family $\mathcal{F} \subseteq L(X, Y)$ is uniformly bounded on bounded subsets of X.

Without some form of completeness assumption, such as the $\mathcal{A}$-space assumption in Corollary 6, the conclusion in Corollary 6 is false (see Example 3). For semi-NLS, Corollary 6 takes the following form.

Corollary 7. Let X be a semi-NLS which is a $\mathcal{K}$-space and Y a semi-NLS. If $\mathcal{F} \subseteq L(X, Y)$ is pointwise bounded, then

(i) $\{\|T\| : T \in \mathcal{F}\}$ is bounded and

(ii) $\mathcal{F}$ is equicontinuous.

Proof: (i) follows from Corollary 6 since $\mathcal{F}$ is uniformly bounded on the unit ball $\{x \in X : \|x\| \le 1\}$. (ii) follows from (i) and 5.21.

Example 3 shows that some sort of completeness-type assumption on the domain space is necessary in Corollary 7.

Recall that a complete semi-NLS is a $\mathcal{K}$-space (§2) so Corollary 7 is applicable to such spaces. This is the classical version of the UBP for NLS. Mazur and Orlicz ([MO]) generalized conclusion (ii) of Corollary 7 to complete quasi-normed spaces. We can also obtain this form of the UBP from Corollary 6.

Corollary 8. Let X be a quasi-normed $\mathcal{K}$-space. If $\mathcal{F} \subseteq L(X, Y)$ is pointwise bounded, then $\mathcal{F}$ is equicontinuous.

Proof: Let $x_i \to 0$ in X and $\{T_i\} \subseteq \mathcal{F}$. Pick $t_i \to \infty$ such that $t_i x_i \to 0$ (2.5). Then $\{t_i x_i\}$ is bounded so $\{T_i t_i x_i\}$ is bounded by Corollary 6. Hence, $\frac{1}{t_i} T_i t_i x_i = T_i x_i \to 0$, and $\mathcal{F}$ is equicontinuous (5.20).

Again recall that a complete quasi-normed space is a $\mathcal{K}$-space (§2) so Corollary 8 is applicable to such spaces.

In general, any equicontinuous family of continuous linear operators is uniformly bounded on bounded sets (Exer. 2) so the conclusion in Corollary 8 is sharper than that in Corollary 6. However, it is not generally the case that a pointwise bounded family of continuous linear operators (or even a family of operators which is uniformly bounded on bounded sets) is equicontinuous as the following example shows.

Example 9. Let ℓ^2 have the topology $w(\ell^2)$, where we consider each element of ℓ^2 to be a member of the dual of ℓ^2 (Example 5.14). Then $\{e_k\}$ is pointwise bounded on ℓ^2 but is not equicontinuous with respect to $w(\ell^2)$ since $e_k \to 0$ in $w(\ell^2)$ but $\langle e_k, e_k \rangle = 1$ (5.20). As we will see later $(\ell^2, w(\ell^2))$ is an $\mathcal{A}$-space and the family $\{e_k\}$ is actually uniformly bounded on bounded subsets of $(\ell^2, w(\ell^2))$ so the quasi-norm assumption in Corollary 8 is important.

Despite this example which shows that a pointwise bounded family of continuous linear operators needn't be equicontinuous, we will show later in §25 that in the locally convex case there is always a stronger topology on the domain space under which any pointwise bounded family of continuous linear operators is equicontinuous.

A result which is closely related to the UBP is a result often referred to as the Banach-Steinhaus Theorem. This result gives sufficient conditions for the pointwise limit of a sequence of continuous linear operators to be continuous. We first use the Basic Matrix Theorem to establish a general form of this result.

Theorem 10. Suppose that $\{T_i\} \subseteq L(X, Y)$ and $\lim T_i x = Tx$ exists for each $x \in X$ (we do not assume that $T \in L(X, Y)$). Then $\lim_i T_i x_j = Tx_j$ converges uniformly for $j \in \mathbb{N}$ for each $w(\{T_i\})$ - $\mathcal{K}$ convergence sequence $\{x_j\}$ in X.

Proof: Consider the matrix $M = [T_i x_j]$ in Y. From the definition of $w(\{T_i\})$ - $\mathcal{K}$ convergence, it follows immediately that M is a $\mathcal{K}$-matrix so Theorem 2 gives the result.

Again the topology $w(\{T_i\})$ depends upon the particular sequence $\{T_i\}$, but if either of the stronger topologies, $w(L(X, Y))$ or the original topology of X is used, then the same conclusion holds for these topologies. In particular, we have

Corollary 11 (Banach-Steinhaus). Let X be a quasi-normed $\mathcal{K}$-space and $\{T_i\} \subseteq L(X, Y)$. If $\lim T_i x = Tx$ exists $\forall x \in X$, then

(i) T is continuous, and

(ii) $\forall$ compact subset $K \subseteq X$, $\lim_i T_i x = Tx$ uniformly for

$x \in K$.

Proof: (i): If $x_j \to 0$ in X, then $\{x_j\}$ is $\mathcal{K}$ convergent so by Theorem 10, $\lim_i T_i x_j = T x_j$ uniformly for $j \in \mathbb{N}$. Therefore, $\lim_j T x_j = \lim_j \lim_i T_i x_j = \lim_i \lim_j T_i x_j = 0$.

(ii): From the sequential compactness of K, it suffices to show that $\lim_i T_i x_j = T x_j$ uniformly for $j \in \mathbb{N}$ for any convergent sequence $\{x_j\}$ in X. But, if $x_j \to x$ in X, then $\{x_j - x\}$ is $\mathcal{K}$ convergent in X so $\lim_i T_i(x_j - x) = T(x_j - x)$ uniformly for $j \in \mathbb{N}$ by Theorem 10. Since $\lim_i T_i x = Tx$, it follows that $\lim_i T_i x_j = T x_j$ uniformly for $j \in \mathbb{N}$.

The classical form of the Banach-Steinhaus Theorem assumes that X is a complete quasi-normed space. Since any such space is a $\mathcal{K}$-space, Corollary 11 gives a generalization of the classical Banach-Steinhaus Theorem to $\mathcal{K}$-spaces.

Finally, we employ the UBP to obtain two interesting results which will be employed later. First, we establish a simple and useful test for determining the boundedness of a subset in a NLS.

Proposition 12. Let X be a semi-NLS. Then $B \subseteq X$ is (norm) bounded if and only if $\langle x', B \rangle$ is bounded $\forall x' \in X'$.

Proof: $\Rightarrow$: clear.

$\Leftarrow$: The family $\{Jb : b \in B\}$ is pointwise bounded on X' and X' is complete (5.8) so the result is an immediate consequence of Corollary 7 and the fact that the canonical imbedding J is an isometry.

We will establish a more general form of this result later (14.15).

Finally, we give an application of the UBP to sequence spaces (see also Exercises 4 and 5).

Proposition 13. Let $\{t_i\} \subseteq \mathbb{R}$. Suppose that $\sum_{i=1}^{\infty} t_i s_i$ converges $\forall \{s_i\} \in c_0$. Then $\{t_i\} \in \ell^1$.

Proof: For each n define $f_n : c_0 \to \mathbb{R}$ by $\langle f_n, \{s_i\}\rangle = \sum_{i=1}^{n} t_i s_i$. Then f_n is linear, continuous and $\lim_n \langle f_n, \{s_i\}\rangle = \langle f, \{s_i\}\rangle = \sum_{i=1}^{\infty} s_i t_i$ exists $\forall \{s_i\} \in c_0$. By Corollary 11, f is continuous so $\{t_i\} \in \ell^1$ since $c_0' = \ell^1$ (5.11).

Sargent Spaces:

Although the early versions of the UBP were established by using sliding hump arguments, one of the most common methods currently used to establish the UBP for NLS or QNLS is by use of the Baire Category Theorem ([B]; see [Sw] for a historical description of the evolution of the UBP). This method of proof was refined by W.L.C. Sargent and used to obtain an equicontinuous version of the UBP as in Corollary 8. We present her results here.

Definition 14. A TVS X is called a Sargent space (a β-space by Sargent) ([S]) if there exists no sequence of subsets $\{E_k\}$ satisfying

(S1) $0 \in E_k$, $E_k - E_k \subseteq E_{k+1}$

(S2) $X = \bigcup_{k=1}^{\infty} E_k$

(S3) Every E_k is nowhere dense.

A QNLS of second category (such a space is usually called a Baire space) is obviously a Sargent space, but, remarkably, there are Sargent spaces which are first category. See [S] for examples of such spaces.

For Sargent spaces we have the following UBP.

Theorem 15. Let X be a Sargent space, Y a semi-NLS and let $\mathcal{F} \subseteq (X, Y)$ be pointwise bounded on X. Then $\mathcal{F}$ is equicontinuous.

Proof: Let $\varepsilon > 0$ and set

$$E_k = \{x \in X : \|Tx\| \le 2^k \varepsilon \text{ for all } T \in \mathcal{F}\}.$$

Then $\{E_k\}$ clearly satisfies (S1) and also satisfies (S2) since $\mathcal{F}$ is pointwise bounded. Hence, some E_k must be somewhere dense, i.e., there exists x_0 and a neighborhood of 0, U, such that $x_0 + U \subseteq \overline{E_k}$. If $x \in U$, then $x_0 + x \in \overline{E_k}$ and $x_0 \in \overline{E_k}$ so by (S1) $x = (x + x_0) - x_0 \in \overline{E_{k+1}}$ and $\|Tx\| \le 2^{k+1}\varepsilon$ for all $T \in \mathcal{F}$. Thus, $W = (1/2^{k+1})U$ is a neighborhood of

0 with $\|Tx\| \le \varepsilon$ for all $x \in W$ and $T \in \mathcal{F}$ so $\mathcal{F}$ is equicontinuous.

A more general form of the UBP for Sargent spaces is given in Exer. 24.15.

Exercise 1. Let $\{t_{ij}\} \subseteq \mathbb{R}$ and $\varepsilon_{ij} > 0$. If $\lim_i t_{ij} = 0 \ \forall j$ and $\lim_j t_{ij} = 0$ $\forall i$, show there is an increasing sequence of positive integers $\{p_k\}$ such that $|t_{p_i p_j}| \le \varepsilon_{ij}$ for $i \ne j$.

Exercise 2. If $\mathcal{F} \subseteq L(X, Y)$ is equicontinuous, show that $\mathcal{F}$ is uniformly bounded on bounded subsets of X.

Exercise 3. If X is a B-space and Y a semi-NLS and $T_k \in L(X, Y)$ satisfies $\lim T_k x = Tx$ exists $\forall x \in X$, show $\|T\| \le \underline{\lim}\|T_k\| < \infty$.

Exercise 4. Establish the analogue of Proposition 13 with the space c_0 replaced by c and ℓ^p for $1 \le p < \infty$.

Exercise 5. Let (S, Σ, μ) be a σ-finite measure space and $1 \le p < \infty$. Suppose $f : S \to \mathbb{R}$ is Σ-measurable and fg is μ-integrable $\forall g \in L^p(\mu)$. Show $f \in L^q(\mu)$, $1/p + 1/q = 1$.

Exercise 6. Show Corollary 11 is false if the condition on X is dropped. [Hint: Consider $f(\{t_j\}) = \sum_{j=1}^{\infty} t_j$ on c_{00}.]

9.1 Bilinear Maps

Let X, Y, Z be TVS. If $f : X \times Y \to Z$, for $x \in X$ $(y \in Y)$ we write $f(x, \cdot)$ $(f(\cdot, y))$ for the map $f(x, \cdot)(y) = f(x, y)$ $(f(\cdot, y)(x) = f(x, y))$. A map $B : X \times Y \to Z$ is <u>bilinear</u> if $B(x, \cdot)$ $(B(\cdot, y))$ is linear $\forall\, x \in X$ $(\forall\, y \in Y)$. We first establish the analogue of 5.1 for bilinear maps.

Proposition 1. A bilinear map $B : X \times Y \to Z$ is (jointly) continuous (with respect to the product topology) if and only if B is continuous at (0, 0).

Proof: Let $(x_0, y_0) \in X \times Y$ and W be a neighborhood of 0 in Z. Choose a neighborhood of 0, W_1, in Z such that $W_1 + W_1 + W_1 \subseteq W$. $\exists$ balanced neighborhoods U_1 of 0 in X and V_1 of 0 in Y such that $B(U_1, V_1) \subseteq W_1$. Since U_1 is absorbing, $\exists\, t > 1$ such that $x_0 \in tU_1$ and $\exists\, s > 1$ such that $y_0 \in sV_1$. Set $U = (1/s)U_1$, $V = (1/t)V_1$.

If $y \in V$, then $B(x_0, y) = B(x_0/t, ty) \in W_1$; if $x \in U$, then $B(x, y_0) = B(sx, y_0/s) \in W_1$. Thus, if $(x, y) \in U \times V$,

$$B(x_0+x, y_0+y) - B(x_0,y_0)$$
$$= B(x_0,y) + B(x,y_0) + B(x,y) \in W_1 + W_1 + W_1 \subseteq W .$$

A map $f : X \times Y \to Z$ is said to be <u>separately continuous</u> if $f(x, \cdot)$ and $f(\cdot, y)$ are continuous $\forall\, x \in X$, $y \in Y$. Even for bilinear maps separate continuity does not imply joint continuity.

Example 2. Define $B : c_{00} \times c_{00} \to \mathbb{R}$ by $B(\{s_j\}, \{t_j\}) = \sum_{j=1}^{\infty} s_j t_j$ (note this is a finite sum). Note B is separately continuous but is not continuous $[1/\sqrt{k} \sum_{j=1}^{k} e_j = x_k \to 0$ in c_{00} but $B(x_k, x_k) = 1 \ \forall \ k]$.

For $\mathcal{K}$-convergent sequences we do have a form of sequential joint continuity for separately continuous bilinear maps. Recall that $w(L(Y, Z))$ is the weakest topology on Y such that each element of $L(Y, Z)$ is continuous (§9).

Theorem 3. If $x_j \to 0$ in X and $\{y_j\}$ is $w(L(Y, Z))$ - $\mathcal{K}$ convergent to 0 in Y and if $B : X \times Y \to Z$ is a separately continuous bilinear map, then $\lim_i B(x_i, y_j) = 0$ uniformly for $j \in \mathbb{N}$. In particular, $\lim B(x_i, y_i) = 0$.

Proof: Since $B(x_k, y) \to 0 \ \forall \ y \in Y$, the sequence of linear operators $\{B(x_k, \cdot)\}$ in $L(Y, Z)$ is pointwise convergent to 0 on Y. The result now follows from 9.10.

Since the original topology of Y is stronger than $w(L(Y, Z))$, Theorem 3 also holds for the original topology of Y. As an immediate consequence of Theorem 3, we obtain a classic of Mazur and Orlicz on joint continuity.

Corollary 4 (Mazur-Orlicz). Let X and Y be QNLS with Y a $\mathcal{K}$-space.

If $B : X \times Y \to Z$ is separately continuous and bilinear, then B is continuous.

The first result of this type was established by Mazur and Orlicz when Y is a complete QNLS ([MO]). Corollary 4 gives a generalization of their result to $\mathscr{K}$-spaces.

Using Corollary 4 we can weaken axiom (v) in Definition 2.1 of a quasi-normed space. If we assume that scalar multiplication is only separately continuous (with respect to the metric topology induced by the function $|\ |$), then it follows that scalar multiplication is actually continuous since the scalar field is complete.

Bourbaki introduced a concept, called hypocontinuity, for bilinear maps which is intermediate between separate continuity and continuity.

Definition 5. Let $B : X \times Y \to Z$ be bilinear and separately continuous, and let $\mathscr{Y}$ be a family of bounded subsets of Y. B is <u>$\mathscr{Y}$-hypocontinuous</u> if for every neighborhood of 0, W, in Z and every $A \in \mathscr{Y}$, $\exists$ a neighborhood of 0, U, in X such that $B(U, A) \subseteq W$.

Thus, B is $\mathscr{Y}$-hypocontinuous if and only if whenever $\{x_\delta\}$ is a net which converges to 0 in X and $A \in \mathscr{Y}$, $\lim B(x_\delta, y) = 0$ uniformly for $y \in A$. If X is metrizable, the same statement holds with sequences replacing nets.

If $\mathscr{B}_x$ ($\mathscr{B}_y$) is the family of all bounded subsets of X (Y) and B is both $\mathscr{B}_x$ and $\mathscr{B}_y$-hypocontinuous, then B is said to be <u>hypocontinuous</u>.

Example 1 shows that a separately continuous bilinear map needn't be hypocontinuous [if $x_k = \sum_{j=1}^{k} e_j/k$ and $A = \{y_k = \sum_{j=1}^{k} e_j : k \in \mathbb{N}\}$, then $x_k \to 0$ but $B(x_k, y_k) = 1$]. However, we show that any separately continuous bilinear map is hypocontinuous with respect to the family of $\mathcal{K}$ bounded sets.

Theorem 6. Let X be a QNLS and $B : X \times Y \to Z$ bilinear, separately continuous. Then B is $\mathcal{K}_y$-hypocontinuous, where $\mathcal{K}_y$ is the family of $w(L(Y, Z))$ - $\mathcal{K}$ bounded subsets of Y.

Proof: It suffices to show that $B(x_k, y_k) \to 0$ whenever $x_k \to 0$ in X and $\{y_k\}$ is $w(L(Y, Z))$ - $\mathcal{K}$ bounded. Pick $t_k \to \infty$ such that $t_k x_k \to 0$ (2.5). Then $\{y_k/t_k\}$ is $w(L(Y, Z))$ - $\mathcal{K}$ convergent to 0 so $B(x_k, y_k) = B(t_k x_k, y_k/t_k) \to 0$ by Theorem 3.

The conclusion of Theorem 6 is also valid for the family of subsets of Y which are $\mathcal{K}$ bounded with respect to the original topology of Y.

Any continuous bilinear map is hypocontinuous (Exer. 1), but there are hypocontinuous bilinear maps which are not continuous. Examples of such bilinear mappings are given in 23.15, 23.16.

We consider results analogous to Theorem 3 and Corollary 4 for families of bilinear mappings. Let $\mathcal{F}$ be a family of separately continuous bilinear mappings from $X \times Y$ into Z. The family $\mathcal{F}$ is <u>left</u> (<u>right</u>) <u>equicontinuous</u> if $\forall\, y \in Y$ ($\forall\, x \in X$), the family $\mathcal{F}_y = \{B(\cdot, y) : B \in \mathcal{F}\}$ ($\{B(x, \cdot) : B \in \mathcal{F}\}$) is equicontinuous in $L(X, Z)$ ($L(Y, Z)$). The sequence in Example 9 below shows that a family can be left equicontinuous but not right equicontinuous. The family $\mathcal{F}$ is equicontinuous if it is equicontinuous as a family of mappings from $X \times Y$ into Z. As in Proposition 1, $\mathcal{F}$ is equicontinuous if and only if $\mathcal{F}$ is equicontinuous at $(0, 0)$ (Exer. 4). Analogous to Theorem 3 we have

Theorem 7. Let $\mathcal{F}$ be left equicontinuous. If $x_i \to 0$ in X and $\{y_i\}$ is $\mathcal{K}$ convergent in Y, then $\lim B(x_i, y_i) = 0$ uniformly for $B \in \mathcal{F}$.

Proof: It suffices to show that $\lim B_i(x_i, y_i) = 0$ for each $\{B_i\} \subseteq \mathcal{F}$. Consider the matrix $M = [B_i(x_i, y_j)]$. Since $\mathcal{F}$ is left equicontinuous, the columns of M converge to 0. For any increasing sequence of positive integers $\{m_j\}$, there is a subsequence $\{n_j\}$ of $\{m_j\}$ such that the series $\sum_{j=1}^{\infty} y_{n_j}$ converges to a point $y \in Y$. Then $\sum_{j=1}^{\infty} B_i(x_i, y_{n_j}) = B_i(x_i, y) \to 0$ by the separate continuity and the left equicontinuity of $\mathcal{F}$. Thus, M is a $\mathcal{K}$-matrix and by the Basic Matrix Theorem $\lim B_i(x_i, y_i) = 0$.

We have the immediate corollary analogous to Corollary 4.

Corollary 8. If X and Y are QNLS with Y a $\mathcal{K}$-space, then any left equicontinuous family of bilinear maps is equicontinuous.

Without some type of completeness assumption, Corollary 8 is false.

Example 9. For each $i \in \mathbb{N}$ define $B_i : \ell^\infty \times c_{00} \to \mathbb{R}$ by

$$B_i(\{s_j\}, \{t_j\}) = \sum_{j=1}^{i} s_j t_j.$$

Each B_i is separately continuous and the sequence $\{B_i\}$ is left equicontinuous [if $(t_1, \ldots, t_n, 0, \ldots) \in c_{00}$, then for $i \geq n$ $B_i(\{s_j\}, \{t_j\}) = \sum_{j=1}^{n} s_j t_j$]. However, $\{B_i\}$ is not even right equicontinuous; for if $e = (1, 1, \ldots) \in \ell^\infty$ and $y_i = \sum_{j=1}^{i} e_j/i$, then $B_i(e, y_i) = 1$ while $y_i \to 0$.

From Corollary 8 we obtain an analogue of Corollary 9.7 for bilinear maps.

Corollary 10. Let X, Y be QNLS which are $\mathcal{K}$-spaces. If $\mathcal{F}$ is a family of separately continuous bilinear maps which is pointwise bounded on $X \times Y$, then $\mathcal{F}$ is equicontinuous.

Proof: For $y \in Y$ the family $\{B(\cdot, y) : B \in \mathcal{F}\} \subset L(X, Z)$ is pointwise bounded on X and, therefore, equicontinuous by 9.7. That is, $\mathcal{F}$ is left equicontinuous. The result follows from Corollary 8.

Corollary 10 gives a generalization of Theorem 15.12 of [Gr]; see also 14(3) of [K1]. The proofs in these references use the theory of uniform spaces while the proof above uses only basic properties of convergent sequences and the Basic Matrix Theorem.

Exercise 1. Show any continuous bilinear map is hypocontinuous.

Exercise 2. Show a bilinear map $B : X \times Y \to Z$ between semi-NLS X, Y and Z is continuous if and only if $\sup\{\|B(x, y)\| : \|x\| \le 1, \|y\| \le 1\} < \infty$. (Compare with 5.5.)

Exercise 3. Let $\mathcal{A}$ be a family of bounded subsets of X. Show a separately continuous bilinear map $B : X \times Y \to Z$ is $\mathcal{A}$-hypocontinuous if and only if $\forall A \in \mathcal{A}$ the family $\{B(x, \cdot) : x \in A\} \subseteq L(Y, Z)$ is equicontinuous.

Exercise 4. Show that a family of bilinear maps is equicontinuous if and only if it is equicontinuous at $(0, 0)$.

Exercise 5. Show that if $\mathcal{F}$ is left equicontinuous, then $\mathcal{F}$ is pointwise bounded on $X \times Y$. Thus, the converse of Corollary 10 holds.

Exercise 6. Show that if $\mathcal{F}$ is equicontinuous, then $\mathcal{F}$ is uniformly bounded on sets of the form $C \times D$, where C is bounded in X and D is bounded in Y. (Compare with Exer. 9.2.)

9.2 The Nikodym Boundedness Theorem

The Nikodym Boundedness Theorem is a "striking improvement" in the UBP for the space of countably additive measures defined on a σ-algebra ([DS] IV.9.8). The theorem has been extended to the case of bounded finitely additive measures and we consider this more general version. We begin by establishing a remarkable lemma of Drewnowski ([Dr]) which shows that in some sense a bounded finitely additive set function is "not too far" from being countably additive.

Let Σ be a σ-algebra of subsets of a set S. Recall that $ba(\Sigma)$ is the B-space of all bounded, finitely additive set functions $\mu : \Sigma \to \mathbb{R}$ equipped with the variation norm, $\|\mu\| = |\mu|(S)$ (2.30). If $\mu \in ba(\Sigma)$ and $\{E_j\}$ is a pairwise disjoint sequence from Σ, then

$$\sum_{j=1}^{n} |\mu(E_j)| \le |\mu|(\bigcup_{j=1}^{n} E_j) \le |\mu|(S) \quad \forall\, n$$

so the series $\sum_{j=1}^{\infty} \mu(E_j)$ is absolutely convergent. Thus, if a finitely additive set function, $\mu \in ba(\Sigma)$, fails to be countably additive, it is not because the series $\sum_{j=1}^{\infty} \mu(E_j)$ fails to converge but rather that it fails to converge to the "proper" value, $\mu(\bigcup_{j=1}^{\infty} E_j)$.

Lemma 1 (Drewnowski [Dr]). Let $\mu \in ba(\Sigma)$. If $\{E_j\}$ is a pairwise disjoint sequence from Σ, then $\exists$ a subsequence $\{E_{n_j}\}$ such that μ is countably additive on the σ-algebra generated by $\{E_{n_j}\}$.

Proof: Partition $\mathbb{N}$ into a pairwise disjoint sequence of infinite sets $\{K_j^1\}_{j=1}^\infty$. By the observation above, $|\mu|(\bigcup_{j\in K_i^1} E_j) \to 0$ as $i \to \infty$. So $\exists$ i such that $|\mu|(\bigcup_{j\in K_i^1} E_j) < 1/2$. Let $N_1 = K_i^1$ and $n_1 = \inf N_1$. Now partition $N_1\backslash\{n_1\}$ into a pairwise disjoint sequence of infinite sets $\{K_j^2\}_{j=1}^\infty$. As before $\exists$ i such that $|\mu|(\bigcup_{j\in K_i^2} E_j) < 1/2^2$. Let $N_2 = K_i^2$ and $n_2 = \inf N_2$. Note $n_2 > n_1$ and $N_2 \subseteq N_1$. Continuing produces a subsequence $n_j \uparrow \infty$ and a sequence of infinite subsets of $\mathbb{N}$, $\{N_j\}$, such that $N_{j+1} \subseteq N_j$ and $|\mu|\,(\bigcup_{i\in N_j} E_i) < 1/2^j$. Let Σ_0 be the σ-algebra generated by $\{E_{n_j}\}$.

We claim that μ is countably additive on Σ_0. If $\{H_k\} \subseteq \Sigma_0$ and $H_k \downarrow \emptyset$, then $|\mu(H_k)| \le |\mu|(H_k)\downarrow 0$ so μ is countably additive [given j, $\exists H_i$ such that min $H_i > n_j$ so $|\mu|(H_i) \le |\mu|(\bigcup_{k\ge j} E_{n_k}) < 1/2^j$].

Corollary 2 (Drewnowski). Let $\mu_i \in ba(\Sigma)$ for $i \in \mathbb{N}$. If $\{E_j\}$ is a pairwise disjoint sequence from Σ, $\exists$ a subsequence $\{E_{n_j}\}$ such that each μ_i is countably additive on the σ-algebra generated by $\{E_{n_j}\}$.

Proof: Set $\mu(E) = \sum_{i=1}^\infty \frac{1}{2^i} \frac{|\mu_i|(E)}{1+|\mu_i|(S)}$. Then $\mu \in ba(\Sigma)$ (2.30, 2.9) so by Lemma 1, $\exists$ a subsequence $\{E_{n_j}\}$ such that μ is countably additive on the σ-algebra, Σ_0, generated by $\{E_{n_j}\}$. If $\{H_j\} \subseteq \Sigma_0$ and $H_j \downarrow \emptyset$, then $\lim_j \mu_i(H_j) = 0$ for each i, and μ_i is countably additive on Σ_0.

Finally, we require the following technical lemma.

Lemma 3. Let $\mathcal{A}$ be an algebra of subsets of set S, and let $\mu_i : \mathcal{A} \to \mathbb{R}$ be bounded and finitely additive for $i \in \mathbb{N}$. Then $\{\mu_i(A) : i \in \mathbb{N}, A \in \mathcal{A}\}$ is bounded if and only if $\{\mu_i(A_j) : i, j \in \mathbb{N}\}$ is bounded for each pairwise disjoint sequence $\{A_j\}$ from $\mathcal{A}$.

Proof: Suppose $\sup\{|\mu_i(A)| : i \in \mathbb{N}, A \in \mathcal{A}\} = \infty$. Note that for each $M > 0$ $\exists$ a partition (E, F) of S and an integer i such that $\min\{|\mu_i(E)|, |\mu_i(F)|\} > M$. [This follows since

$$|\mu_i(E)| > M + \sup\{|\mu_i(S)| : i \in \mathbb{N}\}$$

implies $|\mu_i(S \setminus E)| \geq |\mu_i(E)| - |\mu_i(S)| > M$.] Hence, $\exists\ i_1$ and a partition (E_1, F_1) of S such that $\min\{|\mu_{i_1}(E_1)|, |\mu_{i_1}(F_1)|\} > 1$. Now either $\sup\{|\mu_i(H \cap E_1)| : H \in \mathcal{A}, i \in \mathbb{N}\} = \infty$ or

$$\sup\{|\mu_i(H \cap F_1)| : H \in \mathcal{A}, i \in \mathbb{N}\} = \infty.$$

Pick whichever of E_1 or F_1 satisfies this condition and label it B_1 and set $A_1 = S \setminus B_1$. Now treat B_1 as S above to obtain a partition (A_2, B_2) of B_1 and an $i_2 > i_1$ satisfying $|\mu_{i_2}(A_2)| > 2$ and

$$\sup\{|\mu_i(H \cap B_2)| : H \in \mathcal{A}, i \in \mathbb{N}\} = \infty .$$

Proceeding produces a subsequence $\{i_j\}$ and a disjoint sequence $\{A_j\}$

such that $|\mu_{i_j}(A_j)| > j$. This establishes the sufficiency; the necessity is clear.

We now establish the Nikodym Boundedness Theorem for bounded, finitely additive set functions.

Theorem 4. Let $\{\mu_i\} \subseteq ba(\Sigma)$. If $\{\mu_i\}$ is pointwise bounded on Σ, then $\{\mu_i(E) : i \in \mathbb{N}, E \in \Sigma\}$ is bounded, i.e., $\{\mu_i\}$ is uniformly bounded on Σ.

Proof: By Lemma 3 it suffices to show that $\{\mu_i(E_j) : i, j \in \mathbb{N}\}$ is bounded for each pairwise disjoint sequence $\{E_j\}$ from Σ.

Let $\mathscr{S}(\Sigma)$ be the space of Σ-simple, real-valued functions equipped with the sup-norm (2.31). The dual of $\mathscr{S}(\Sigma)$ is $ba(\Sigma)$ (5.15 and Exer. 5.3) so each μ_i induces a continuous linear function on $\mathscr{S}(\Sigma)$ via integration, $<\mu_i, \varphi> = \int_S \varphi \, d\mu_i$, and $\|\mu_i\| = |\mu_i|(S)$. Consider the sequence $\{C_{E_i}\}$ in $\mathscr{S}(\Sigma)$. By Corollary 2, any subsequence of $\{E_j\}$ has a subsequence $\{E_{n_j}\}$ such that each μ_i is countably additive on the σ-algebra generated by $\{E_{n_j}\}$. This means that $\{C_{E_j}\}$ is $w(\{\mu_i\})$ - $\mathscr{K}$ convergent in $\mathscr{S}(\Sigma)$ $[\sum_{j=1}^{\infty} \mu_i(E_{n_j}) = \mu_i(\bigcup_{j=1}^{\infty} E_{n_j}) \; \forall \; i]$. The hypothesis of the theorem implies that the sequence $\{\mu_i\}$ is pointwise bounded on $\mathscr{S}(\Sigma)$ so the General UBP 9.4(i) yields that $\{\mu_i(E_j) : i, j\}$ is bounded and the proof is complete.

Note that the space $\mathscr{S}(\Sigma)$ is not complete (in general, it is not even second category [Sa]) so the classical form of the UBP is not applicable. Also note that the sequence $\{C_{E_j}\}$ above is not norm - $\mathscr{K}$ convergent so 9.7 is not applicable; the general form of the UBP given in 9.4 seems to be necessary.

From the inequality $|\mu|(S) \le 2 \sup\{|\mu(E)| : E \in \Sigma\}$ (Example 2.30), it follows that the norms of the linear functionals, $\{\|\mu_i\|\}$, in Theorem 4 are bounded. If the conclusion is written in this form, the theorem has the flavor of the classical UBP for B-spaces (9.7).

If Σ is the σ-algebra of all subsets of $\mathbb{N}$, then $\mathscr{S}(\Sigma)$ is just the space m_0 of Example 2.18 which has for its dual ba (Exer. 5.3 and Example 5.15). For later use, we give a statement of this special case.

Corollary 5. If $\{\mu_i\} \subseteq ba$ is pointwise bounded on m_0, then $\{\|\mu_i\|\}$ is bounded.

We can now obtain an analogue of Proposition 9.13 for m_0.

Proposition 6. If $\{r_i\} \in \mathbb{R}$ is such that $\sum_{i=1}^{\infty} r_i t_i$ converges $\forall \{t_i\} \in m_0$, then $\langle R, \{t_i\}\rangle = \sum_{i=1}^{\infty} r_i t_i$ defines a continuous linear functional on m_0 with $\|R\| = \sum_{i=1}^{\infty} |r_i| < \infty$.

Proof: For each n define $R_n : m_0 \to \mathbb{R}$ by $\langle R_n, \{t_j\}\rangle = \sum_{j=1}^{n} r_j t_j$. Then R_n is continuous on m_0 with $\|R_n\| = \sum_{j=1}^{n} |r_j|$. Since

$$\langle R_n, \{t_j\}\rangle \to \langle R, \{t_j\}\rangle,$$

it follows from Corollary 5 that R is continuous and

$$\|R\| \le \underline{\lim}\|R_n\| = \sum_{j=1}^{\infty} |r_j| < \infty$$

(see Exer. 9.3). For each n, set $x_n = \sum_{j=1}^{n} (\text{sign } r_j)e_j$. Then $\|x_n\| = 1$ and $\|R\| \ge \langle R, x_n\rangle = \sum_{j=1}^{n} |r_j|$ so $\|R\| \ge \sum_{j=1}^{\infty} |r_j|$.

The Nikodym Boundedness Theorem is such a striking result that it has been generalized in many different directions (see, however, Exer. 1). For example, the σ-algebra assumption has been relaxed and the range of the measures has been replaced by TVS's. For the history of the evolution of these generalizations and references, see [DU].

Exercise 1. Let $\mathcal{A}$ be the family of all subsets of $\mathbb{N}$ which are either finite or have finite complements. Let $\delta_n : \mathcal{A} \to \mathbb{R}$ be given by $\delta_n(E) = 1$ if $n \in E$ and $\delta_n(E) = 0$ otherwise. Define $\mu_n : \mathcal{A} \to \mathbb{R}$ by $\mu_n(E) = n(\delta_{n+1}(E) - \delta_n(E))$ if E is finite and

$$\mu_n(E) = -n(\delta_{n+1}(E) - \delta_n(E))$$

otherwise. Show that $\mathcal{A}$ is an algebra, $\{\mu_n\}$ is pointwise bounded on $\mathcal{A}$ but is not uniformly bounded on $\mathcal{A}$. Thus, the σ-algebra assumption in Theorem 4 cannot be replaced with Σ being an algebra.

Exercise 2. Let X be a NLS and Σ a σ-algebra. Let $\mu_n : \Sigma \to X'$ be finitely additive and have bounded range $\forall\, n \in \mathbb{N}$. If $\{\mu_n(E)\}$ is bounded $\forall\, E \in \Sigma$, show $\sup\{\|\mu_n(E)\| : n \in \mathbb{N}, E \in \Sigma\} < \infty$.

9.3 Fourier Series

Given a periodic function $f : \mathbb{R} \to \mathbb{R}$ of period 2π belonging to $L^1[0, 2\pi]$, its Fourier coefficients are given by $c_n(f) = \frac{1}{2\pi}\int_0^{2\pi} f(t)e^{-int}dt$ and the nth partial sum of the Fourier series for f is

$$s_n(f)(t) = \sum_{j=-n}^{n} c_j(f)e^{ijt} = \int_0^{2\pi} f(s)\frac{1}{2\pi}\sum_{j=-n}^{n} e^{ij(t-s)}ds = \frac{1}{2\pi}\int_0^{2\pi} f(s)D_n(t-s)ds,$$

where $D_n(t) = \frac{\sin(n+1/2)t}{\sin(t/2)}$ is the Dirichlet kernel ([HS] 18.27).

One of the problems of Fourier analysis is to determine the convergence of the Fourier series of a function f. We show that the UBP can be used to show the existence of a continuous periodic function whose Fourier series diverges at 0. For this, the following lemma is useful.

Lemma 1. Let $g : [a, b] \to \mathbb{R}$ be continuous. Define $G : C[a, b] \to \mathbb{R}$ by $G(f) = \int_a^b g(t)f(t)dt$. Then G is linear and continuous with $\|G\| = \int_a^b |g|$.

Proof: Since $|G(f)| \le \|f\|_\infty \int_a^b |g|$, G is linear, continuous and $\|G\| \le \int_a^b |g|$. Fix $n \in \mathbb{N}$. Then

$$\int_a^b |g| = \int_a^b |g| \frac{1+n|g|}{1+n|g|} = \int_a^b \frac{|g|}{1+n|g|}$$

$$+ \int_a^b g\frac{ng}{1+n|g|} \le \int_a^b \frac{1}{n} + G\left(\frac{ng}{1+n|g|}\right)$$

$$\le \frac{b-a}{n} + \|G\|\left\|\frac{ng}{1+n|g|}\right\|_\infty \le \frac{b-a}{n} + \|G\|$$

so $\int_a^b |g| \le \|G\|$.

Now let $f_n : C[0, 2\pi] \to \mathbb{R}$ be defined by

$$f_n(\varphi) = \frac{1}{2\pi}\int_0^{2\pi} \varphi(s)D_n(s)ds,$$

i.e., $f_n(\varphi)$ is the nth partial sum of the Fourier series of φ evaluated at $t = 0$. By Lemma 1, $\|f_n\| = \frac{1}{2\pi}\int_0^{2\pi} |D_n(s)|ds$. We claim that $\{\int_0^{2\pi} |D_n(s)|ds\}$ is unbounded. Since $\sin u \le u$ for $0 \le u \le \pi$, this will follow if $\{\int_0^{\pi} |\frac{\sin(2n+1)u}{u}|du\}$ is unbounded. But,

$$\int_0^{\pi} \left|\frac{\sin(2n+1)u}{u}\right|du = \sum_{k=0}^{2n} \int_{\frac{k\pi}{2n+1}}^{\frac{(k+1)\pi}{2n+1}} \left|\frac{\sin(2n+1)u}{u}\right|du$$

$$\ge \sum_{k=0}^{2n} \frac{2n+1}{(k+1)\pi}\int_{\frac{k\pi}{2n+1}}^{\frac{(k+1)\pi}{2n+1}} |\sin(2n+1)u|du$$

$$= \sum_{k=0}^{2n} \frac{1}{(k+1)\pi}\int_{k\pi}^{(k+1)\pi} |\sin v|dv$$

$$= \frac{2}{\pi} \sum_{k=0}^{2n} \frac{1}{k+1}.$$

It now follows from Corollary 9.7 that there must exist a continuous, periodic function φ whose Fourier series diverges at 0 because otherwise the sequence $\{\|f_n\|\}$ would be bounded. Of course, the same argument applies to any point in $[0, 2\pi]$.

We can use a construction of Banach and Steinhaus, called condensation of singularities, to show the existence of a continuous, periodic function whose Fourier series diverges at a sequence of points in $[0, 2\pi]$.

Theorem 2 (Banach, Steinhaus). Let X be a B-space, Y_k a NLS and $T_k \in L(X, Y_k)$ $\forall k \in \mathbb{N}$. Then $B = \{x \in X : \overline{\lim}\|T_k x\| < \infty\}$ either coincides with X or is first category in X.

Proof: Suppose B is second category in X. By the definition of B, $\forall x \in B$ $\lim_k \sup_{n\geq 1}\|\frac{1}{k} T_n x\| = 0$. Let $\varepsilon > 0$. Then $B \subseteq \bigcup_{k=1}^{\infty} B_k$, where $B_k = \{x \in X : \sup_{n\geq 1}\|\frac{1}{k} T_n x\| \leq \varepsilon\}$. Each B_k is closed so some B_k contains a sphere, i.e., $\exists\, x_0 \in X \cap B_k$, $r > 0$ such that

$$\|x - x_0\| \leq r \Rightarrow \sup_{n\geq 1}\|\frac{1}{k} T_n x\| \leq \varepsilon.$$

Thus, for $\|z\| \leq r$, if $x = x_0 + z$, then $\|\frac{1}{k} T_n z\| \leq \|\frac{1}{k} T_n x\| + \|\frac{1}{k} T_n x_0\| \leq 2\varepsilon$ and $\sup_n \|T_n z\| \leq 2\varepsilon k$. Hence, $X = B$ since B is a linear subspace.

Corollary 3 (Condensation of Singularities). Let $T_{p,q}$, $p, q \in \mathbb{N}$, be a sequence of continuous linear operators from a B-space X into a NLS Y_q. Suppose $\forall p\ \exists\, x_p \in X$ such that $\overline{\lim_q}\|T_{p,q} x_p\| = \infty$. Then

$$B = \{x \in X : \overline{\lim_q}\|T_{p,q} x\| = \infty\ \forall p \in \mathbb{N}\}$$

is second category in X.

Proof: $\forall\, p$ $B_p = \{x \in X : \overline{\lim_q} \|T_{p,q}x\| < \infty\}$ is first category in X by Theorem 2 and the hypothesis. Thus, $B = X \setminus \bigcup_{p=1}^{\infty} B_p$ is second category.

Now let $\{s_j\}$ be a sequence of distinct points from $[0, 2\pi]$. Let $S_{n,j}(\varphi)$ be the nth partial sum of the Fourier series of the continuous periodic function φ evaluated at s_j. By the example above $\forall\, j\ \exists\ \varphi_j$ such that $\overline{\lim_n} |S_{n,j}(\varphi_j)| = \infty$. By Corollary 3, $\exists\ \varphi$ such that $\overline{\lim_n} S_{n,j}(\varphi) = \infty$ $\forall\, j$, i.e., $\exists$ a continuous, periodic function φ whose Fourier series diverges at each t_j. Actually, Rudin shows that the set of divergent points is a dense G_δ set, and then shows by a simple argument that in a complete metric space every dense G_δ set is uncountable ([Ru1]).

9.4 Vector-Valued Analytic Functions

In order to extend a property of scalar-valued functions to a function with values in a Banach space there are often two natural approaches. One is to simply replace the absolute values in the scalar field by norms; this leads to what is generally referred to as a strong property. The second approach is to compose the Banach space valued function with an arbitrary continuous linear functional from the dual of the Banach space and require the resulting scalar-valued function to have the property being considered; this leads to what is called a weak property. We now consider such a situation for analytic functions; we will later consider integration of vector-valued functions.

Let D be an open subset of the complex plane. Let X be a complex B-space and $f : D \to X$. If $z_0 \in D$, we say that f is differentiable at z_0 with derivative $f'(z_0)$ provided the limit $\lim_{z \to z_0} \frac{f(z) - f(z_0)}{z - z_0} = f'(z_0)$ exists with respect to the norm topology. We say that f is (strongly) analytic on D if f is differentiable at every point in D. We say that f is weakly analytic on D if the function $x'f : D \to \mathbb{C}$ is analytic on D $\forall x' \in X'$. We now show that, surprisingly, these two notions of analyticity coincide.

Theorem 1 (Dunford). If $f : D \to X$ is weakly analytic on D, then f is analytic on D.

Proof: Let $z_0 \in D$. Since X is complete, it suffices to show that

$\frac{f(z)-f(z_0)}{z-z_0} - \frac{f(w)-f(z_0)}{w-z_0} \to 0$ as $z, w \to z_0$. Let C be a positively oriented circle with center at z_0 and radius r such that C and its interior lies inside D. For $x' \in X'$, $x'f$ is continuous on C so the UBP implies that $\exists\, M > 0$ such that $\|f(z)\| \le M \;\forall\, z \in C$. For $x' \in X'$, we have

$$x'f(z_0) = \frac{1}{2\pi i} \int_C \frac{x'f(z)}{z-z_0}\, dz$$

by Cauchy's Integral Formula. If $0 < |z - z_0| < r/2$ and $0 < |w - z_0| < r/2$, for $\|x'\| \le 1$, we have

$$\left|\left\langle x', \frac{f(z)-f(z_0)}{z-z_0} - \frac{f(w)-f(w_0)}{w-w_0}\right\rangle\right|$$

$$= \left|\frac{1}{2\pi i} \int_C x'f(\lambda) \frac{z-w}{(\lambda-z)(\lambda-w)(\lambda-z_0)}\, d\lambda\right| \le \frac{4M|z-w|}{r^2}$$

since $|\lambda - z| \ge r/2$, $|\lambda - w| \ge r/2$ for $\lambda \in C$. By 8.1.5

$$\left\|\frac{f(z)-f(z_0)}{z-z_0} - \frac{f(w)-f(z_0)}{w-z_0}\right\| \le \frac{4M|z-w|}{r^2},$$

and the result follows.

For a history of vector-valued analytic functions, see [Ta].

Exercise 1. Define Riemann integrability for a function $f : [a, b] \to X$ with values in a B-space.

Exercise 2. Prove a vector version of Liouville's Theorem. That is, if $f : \mathbb{C} \to X$ is analytic and bounded, then f is a constant.

9.5 Summability

A summability method for sequences is an attempt to attach a notion of convergence to certain divergent sequences. For example, the classical Cesàro summability method takes a sequence $\{t_k\}$ and associates with it the sequence of averages $s_k = \sum_{j=1}^{k} t_j/k$. If the original sequence $\{t_k\}$ converges to t, then also $s_k \to t$. However, there are divergent sequences $\{t_k\}$ such that the sequence $\{s_k\}$ converges; for example, take $t_k = (-1)^k$. The Cesàro method of summability can be described by employing infinite matrices.

Let $A = [a_{ij}]$ be an infinite matrix of real numbers. If X and Y are vector spaces of real-valued sequences, we say that the matrix A belongs to the class (X, Y) if for each $x = \{x_j\} \in X$, the sequence

$$\{\sum_{j=1}^{\infty} a_{ij}x_j\}_{i=1}^{\infty} = Ax$$

belongs to Y (note this requires that the series $\sum_{j=1}^{\infty} a_{ij}x_j$ converges for each i). Thus, to say that a matrix belongs to the class (c, c) means that it transforms convergent sequences into convergent sequences. A matrix of class (c, c) is said to be <u>regular</u> if it preserves limits, i.e., if $\lim_i x_i = \lim_i \sum_{j=1}^{\infty} a_{ij}x_j$ for each $x \in c$. Cesàro summability can be described in this manner by using the matrix

$$A = \begin{bmatrix} 1 & 0 & 0 & \ldots\ldots \\ 1/2 & 1/2 & 0 & \ldots\ldots \\ 1/3 & 1/3 & 1/3 & 0 \ \ldots \\ \vdots & & & \end{bmatrix},$$

and from the convergence properties discussed above, this matrix is regular. We use the results on the UBP to establish several classic results in summability theory. We first require several preliminary observations.

Lemma 1. Let $t_j \in \mathbb{R}, j \in \mathbb{N}$. If $\exists\, M > 0$ such that $|\sum_{j\in\sigma} t_j| \le M$ for every finite set $\sigma \subseteq \mathbb{N}$, then $\sum_{j=1}^{\infty} |t_j| \le 2M$.

Proof: For $\sigma \subseteq \mathbb{N}$ finite, let $\sigma_+ = \{k \in \sigma : t_k \ge 0\}$, $\sigma_- = \{k \in \sigma : t_k < 0\}$. Then $\sum_{k\in\sigma_+} |t_k| = \sum_{k\in\sigma_+} t_k \le M$ and

$$\sum_{k\in\sigma_-} |t_k| = -\sum_{k\in\sigma_-} t_k \le M$$

so $\sum_{k\in\sigma} |t_k| \le 2M$. Since σ is arbitrary, $\sum_{k=1}^{\infty} |t_k| \le 2M$.

The complex case of Lemma 1 is given in Exercise 1.

Lemma 2. Assume that $\sum_{j=1}^{\infty} |a_{ij}| < \infty\ \forall\, i$ and $\lim_i a_{ij} = a_j$ exists $\forall\, j$. Let $\mathcal{F} = \{\sigma \subseteq \mathbb{N} : \sigma \text{ finite}\}$. If $\lim_i \sum_{j\in\sigma} a_{ij} = \sum_{j\in\sigma} a_j$ is not uniform for $\sigma \in \mathcal{F}$, then $\exists\, \varepsilon > 0$, an increasing sequence, $\{i_j\}$, of positive integers and a pairwise disjoint sequence $\{\tau_j\} \subseteq \mathcal{F}$ such that $|\sum_{k\in\tau_j} (a_{i_j k} - a_k)| \geq \varepsilon \;\forall\, j$.

Proof: Suppose the limit is not uniform for $\sigma \in \mathcal{F}$. Then $\exists\, \varepsilon > 0$ such that $\forall\, i \;\exists\, k_i > i$ and a finite set σ_i such that $|\sum_{k\in\sigma_i} (a_{k_i k} - a_k)| \geq 2\varepsilon$.

Put $i_1 = 1$ and let $\sigma_1 \in \mathcal{F}$ be such that $|\sum_{k\in\sigma_1} (a_{i_1 k} - a_k)| \geq 2\varepsilon$.

Set $\tau_1 = \sigma_1$ and $M_1 = \max \sigma_1$. $\exists\, n_1$ such that $\sum_{j=1}^{M_1} |a_{ij} - a_j| < \varepsilon$ for $i \geq n_1$. $\exists\, i_2 > \max\{i_1, n_1\}$ and $\sigma_2 \in \mathcal{F}$ such that $|\sum_{k\in\sigma_2} (a_{i_2 k} - a_k)| \geq 2\varepsilon$.

Set $\tau_2 = \sigma_2 \backslash \sigma_1$ and note

$$|\sum_{k\in\tau_2} (a_{i_2 k} - a_k)| \geq |\sum_{k\in\sigma_2} (a_{i_2 k} - a_k)| - \sum_{k\in\sigma_1} |a_{i_2 k} - a_k| \geq 2\varepsilon - \varepsilon = \varepsilon .$$

Continuing by induction produces the desired sequences.

We now establish a result which contains two summability results of Hahn and Schur. Recall that m_0 is the subspace of ℓ^∞ which consists of the sequences with finite range.

Theorem 3. For a real infinite matrix $A = [a_{ij}]$, the following are equivalent.

(i) $A \in (\ell^\infty, c)$,

(ii) $A \in (m_0, c)$,

(iii) (I) $\lim_i a_{ij} = a_j$ exists $\forall j$

(II) $\{a_{ij}\}_j$ and $\{a_j\}$ belong to ℓ^1 $\forall i$

(III) $\lim_i \sum_{j=1}^{\infty} |a_{ij} - a_j| = 0$.

(iv) (I) and

(IV) $\{a_{ij}\}_j \in \ell^1$ $\forall i$ and $\sum_j |a_{ij}|$ converges uniformly for $i \in \mathbb{N}$.

Proof: Clearly (i) implies (ii). Suppose (ii) holds. Then (I) follows by setting $x = e_j \in m_0$. Each row $\{a_{ij}\}_j$ induces a continuous linear functional R_i on m_0 defined by $\langle R_i, x\rangle = \sum_{j=1}^{\infty} a_{ij}x_j$, $x = \{x_j\} \in m_0$, with $\|R_i\| = \sum_{j=1}^{\infty} |a_{ij}|$ (9.2.6). We claim that $\lim_i \sum_{j\in\sigma} a_{ij} = \sum_{j\in\sigma} a_j$ uniformly for $\sigma \in \mathcal{F}$, the finite subsets of $\mathbb{N}$. For assume that this is not the case and let the notation be as in Lemma 2. The sequence $\{C_{\tau_j}\}$ is $w(\ell^1)$ - $\mathcal{K}$ convergent in m_0 so in particular it is $w(\{R_i\})$ - $\mathcal{K}$ convergent. Since $\lim R_i(x) = R(x)$ exists $\forall x \in m_0$, the general Banach-Steinhaus Theorem (9.10) implies that

$$\lim_i R_i(C_{\tau_j}) = \lim_i \sum_{k\in\tau_j} a_{ik} = R(C_{\tau_j}) = \sum_{k\in\tau_j} a_k$$

converges uniformly for $j \in \mathbb{N}$. But, this contradicts the conclusion of

Lemma 2 and establishes the claim. If $\varepsilon > 0$ is given, then $|\sum_{j\in\sigma} (a_{ij} - a_j)| < \varepsilon$ for i large and all finite sets σ. By Lemma 1, $\sum_{j=1}^{\infty} |a_{ij} - a_j| \le 2\varepsilon$ for i large and (II), (III) follow.

Assume that (iii) holds. Let $\varepsilon > 0$. $\exists$ N such that $i \ge N$ implies $\sum_{j=1}^{\infty} |a_{ij} - a_j| < \varepsilon/2$. $\exists$ M such that $\sum_{j=M}^{\infty} |a_{ij}| < \varepsilon$ for $1 \le i \le N - 1$ and $\sum_{j=M}^{\infty} |a_j| < \varepsilon/2$. If $i \ge N$,

$$\sum_{j=M}^{\infty} |a_{ij}| \le \sum_{j=1}^{\infty} |a_{ij} - a_j| + \sum_{j=M}^{\infty} |a_j| < \varepsilon$$

so (IV) holds.

Assume (iv) holds. Let $\varepsilon > 0$. $\exists$ M such that $\sum_{j=M}^{\infty} |a_{ij}| < \varepsilon/4 \ \forall\, i$.

Let $\{t_j\} \in \ell^{\infty}$ and assume that $\|\{t_j\}\|_{\infty} \le 1$. The series $\sum_{j=1}^{\infty} a_{ij}t_j$ converges $\forall\, i$ since $|\sum_{j=M}^{M+P} a_{ij}t_j| \le \sum_{j=M}^{\infty} |a_{ij}| < \varepsilon/4 \ \forall\, P > 0$. $\exists$ N such that $i, k \ge N$ implies $|a_{ij} - a_{kj}| < \varepsilon/2M$ for $j = 1, \dots, M - 1$. If $i, k \ge N$, then

$$|\sum_{j=1}^{\infty} a_{ij}t_j - \sum_{j=1}^{\infty} a_{kj}t_j| \le \sum_{j=1}^{M-1} |a_{ij} - a_{kj}| + \sum_{j=M}^{\infty} |a_{ij}| + \sum_{j=M}^{\infty} |a_{kj}|$$

$$< M\varepsilon/2M + \varepsilon/4 + \varepsilon/4 = \varepsilon$$

so $\lim_i \sum_{j=1}^{\infty} a_{ij}t_j$ exists and (i) holds.

The equivalence of (i) and (iv) is a classic result of summability due to Schur ([Sc]). Schur's result was generalized by Hahn who established the equivalence of (i) and (ii) ([Ha]). It is interesting that Hahn obtained his result as a consequence of an abstract UBP, one of the earliest abstract forms of the UBP. We will use this result later in §16 to give an interesting example of weak convergence in ℓ^1.

In 9.2 we established a result of Nikodym on the boundedness of families of measures. There is another result of Nikodym, called the Nikodym Convergence Theorem, on measures which is an immediate consequence of Theorem 3. Let Σ be a σ-algebra of subsets of a set S, and let $\mu_i : \Sigma \to \mathbb{R}$ be a sequence of countably additive set functions on Σ. The sequence $\{\mu_i\}$ is said to be uniformly countably additive if for every pairwise disjoint sequence $\{E_j\}$ from Σ, the series $\sum_{j=1}^{\infty} \mu_i(E_j)$ are uniformly convergent for $i \in \mathbb{N}$.

Theorem 4 (Nikodym Convergence Theorem). If $\lim \mu_i(E) = \mu(E)$ exists for each $E \in \Sigma$, then $\{\mu_i\}$ is uniformly countably additive and μ is countably additive.

Proof: Let $\{E_j\}$ be a pairwise disjoint sequence from Σ. Set $a_{ij} = \mu_i(E_j)$ and $A = [a_{ij}]$. Then $A \in (m_0, c)$ by hypothesis so the conclusion follows from Theorem 3.

The Nikodym Convergence Theorem has generalizations to finitely additive, vector-valued set functions. For a discussion of these generalizations and the historical development of the theorem, see [DU].

We use the UBP to establish another result of summability theory due to Silvermann and Toeplitz. This result gives necessary and sufficient conditions for a matrix A to belong to the class (c, c) and to be regular.

Theorem 5. The matrix A belongs to (c, c) and is regular if and only if

(i) $\sup_i \sum_{j=1}^{\infty} |a_{ij}| < \infty$,

(ii) $\lim_i a_{ij} = 0$ for each $j \in \mathbb{N}$, and

(iii) $\lim_i \sum_{j=1}^{\infty} a_{ij} = 1$.

Proof: If $A \in (c, c)$ is regular, then (ii) and (iii) follow by setting $x = e_j$ and $x = (1, 1, \ldots)$. For (i) note that each row $\{a_{ij}\}_{j=1}^{\infty}$ induces a continuous linear functional R_i on c by $R_i x = \sum_{j=1}^{\infty} a_{ij} x_j$, $x = \{x_j\} \in c$, with norm $\|R_i\| = \sum_{j=1}^{\infty} |a_{ij}|$ (Exer. 9.4). Since $A \in (c, c)$, $\lim_i R_i(x)$ exists $\forall\, x \in c$ so by the UBP (9.7) $\sup_i \|R_i\| = \sup_i \sum_{j=1}^{\infty} |a_{ij}| < \infty$ and (i) holds.

Suppose that (i), (ii) and (iii) hold. Then (i) implies that $\sum_{j=1}^{\infty} a_{ij}x_j$ converges $\forall \{x_j\} \in c$. Let $x = \{x_j\} \in c$ with $\ell = \lim x_j$. Let $\varepsilon > 0$. Set $M = \max\{\sup_i \sum_{j=1}^{\infty} |a_{ij}|, \sup_j |x_j - \ell|\}$. Choose N such that $|x_j - \ell| < \varepsilon$ for $j \geq N$. Write

$$(1) \qquad |\sum_{j=1}^{\infty} a_{ij}x_j - \ell| \leq \sum_{j=1}^{N} |a_{ij}||x_j - \ell| + \sum_{j=N+1}^{\infty} |a_{ij}||x_j - \ell| + |\ell||\sum_{j=1}^{\infty} a_{ij} - 1|.$$

The second term on the right hand side of (1) is less than $M\varepsilon$ by (i). With N fixed by (ii) the first term is less than $M \cdot N\varepsilon$ for i large and the last term is less than $|\ell|\varepsilon$ for i large by (iii). Thus, $Ax \in c$ and $\lim Ax = \ell$.

The sufficiency of (i), (ii) and (iii) is due to Silvermann; the necessity is due to Toeplitz.

Another interesting result of summability related to the Silvermann-Toeplitz Theorem is a theorem due to Steinhaus which asserts that a regular matrix A belonging to the class (c, c) cannot "sum" all bounded sequences, i.e., cannot belong to the class (ℓ^{∞}, c). A more general form of this theorem asserts that a regular matrix $A \in (c, c)$ cannot belong to (m_0, c), i.e., cannot "sum" all sequences of 0's and 1's. We can give a simple proof of Steinhaus' Theorem based on Theorem 3.

Theorem 6 (Steinhaus). If $A \in (c, c)$ is regular, then $A \notin (m_0, c)$.

Proof: Suppose that $A = [a_{ij}] \in (m_0, c)$. By Theorem 3, $\lim_i a_{ij} = a_j$ exists for each j, $\{a_j\} \in \ell^1$ and

$$\lim_i \sum_{j=1}^{\infty} a_{ij}x_j = R(\{x_j\}) = \sum_{j=1}^{\infty} a_jx_j$$

for each $\{x_j\} \in m_0$. For each n let x^n be the sequence with 0 in the first n coordinates and 1 in the other coordinates. Then $Rx^n = 1$ for each n since A is regular while $Rx^n = \sum_{j=n+1}^{\infty} a_j \to 0$. This contradiction establishes the result.

Finally, we use the UBP 9.7 to establish an interesting result of Hellinger and Toeplitz on matrices which map ℓ^2 into itself. Hellinger and Toeplitz showed that if

(2) $$\sum_i y_i \sum_j a_{ij}x_j \text{ converges } \forall\, x = \{x_i\},\ y = \{y_i\} \in \ell^2,$$

then there is a constant M such that

(3) $$\left|\sum_i y_i \sum_j a_{ij}x_j\right| \le M \text{ for } \|x\|_2, \|y\|_2 \le 1 \quad ([HT]).$$

If (2) holds $\forall\, y \in \ell^2$, then $Ax \in \ell^2$ $\forall\, x \in \ell^2$ (Exer. 9.4), i.e., $A \in (\ell^2, \ell^2)$, and (3) implies that $\|Ax\|_2 \le M$ $\forall\, x \in \ell^2$ with $\|x\|_2 \le 1$, i.e., $A : \ell^2 \to \ell^2$ is a continuous linear operator. We use 9.7 to give a simple proof of this result.

Theorem 7. If $A \in (\ell^2, \ell^2)$, then $A : \ell^2 \to \ell^2$ is continuous.

Proof: If $x, y \in \ell^2$, we write $x \cdot y = \sum_i x_i y_i$. Let a_i be the ith row of the matrix $A = [a_{ij}]$; since $A \in (\ell^2, \ell^2)$, each $a_i \in \ell^2$ (Exer. 9.4). Define $A_n : \ell^2 \to \ell^2$ by $A_n x = \sum_{i=1}^{n} (a_i \cdot x)e_i$. Then each A_n is linear and continuous and $\lim y \cdot A_n x = y \cdot Ax$ for each $x, y \in \ell^2$. Thus, $\{A_n x : n\}$ is norm bounded in ℓ^2 $\forall x$ (9.12), and by the UBP 9.7, $\{\|A_n\|\}$ is bounded, say, by M. Since $|y \cdot A_n x| \le M$ for $\|x\|_2, \|y\|_2 \le 1$, and $y \cdot A_n x \to y \cdot Ax$, (3) follows.

Exercise 1. If $t_k \in \mathbb{C}$ and $|\sum_{k \in \sigma} t_k| \le M$ for every finite σ, show $\sum_{k=1}^{\infty} |t_k| \le 4M$.

Exercise 2. Show $A = [a_{ij}] \in (c_0, \ell^\infty)$ if and only if (α) $\sup_i \sum_{j=1}^{\infty} |a_{ij}| < \infty$.

Exercise 3. Show $A = [a_{ij}] \in (c_0, c)$ if and only if (α) and (β) $\lim_i a_{ij} = a_j$ exists $\forall j$.

Exercise 4. Show $A = [a_{ij}] \in (c_0, c_0)$ if and only if (α), (β) and $a_j = 0$ $\forall j$.

Exercise 5. Show $A \in (c, c)$ if and only if (α) and $\forall p \lim_i \sum_{j=p}^{\infty} a_{ij} = a_p$ exists.

Exercise 6. Check conditions (i), (ii), (iii) of Theorem 5 for the Cesàro matrix.

10

The Open Mapping and Closed Graph Theorems

In this section we establish two important results of Banach, the Open Mapping and Closed Graph Theorems ([B]).

Let X and Y be quasi-normed spaces.

Lemma 1. Let $T \in L(X, Y)$ be such that $\mathcal{R}T$ is second category in Y. If U is a neighborhood of 0 in X, then $\overline{TU}$ contains a neighborhood of 0 [mappings with this property are called almost open].

Proof: Let V be a balanced neighborhood of 0 in X such that $V + V \subseteq U$. Now $X = \bigcup_{k=1}^{\infty} kV$ so $TX = \bigcup_{k=1}^{\infty} kTV$. By the second category assumption some $\overline{kTV} = k(\overline{TV})$ contains an open set W. Therefore, $\overline{TU} \supseteq \overline{TV - TV} \supseteq \overline{TV} - \overline{TV} \supseteq \frac{1}{k}(W - W)$ and $\frac{1}{k}(W - W)$ is an open neighborhood of 0 in Y.

Lemma 2. Let X be complete. If $T \in L(X, Y)$ is almost open, then T is open (and, hence, onto).

Proof: Let $S(r) = \{x \in X : |x| < r\}$. Since every neighborhood of 0 in X contains some $S(r)$, it suffices to show that $T(S(r))$ is a neighborhood of 0 for each $r > 0$. For $n = 0, 1, 2, \ldots$, let $G_n = S(r/2^n)$. By hypothesis, $\overline{TG_n}$ is a neighborhood of 0 in Y so $\exists\ \varepsilon_n > 0$ such that

$$(1) \qquad \overline{TG_n} \supseteq H_n = \{y \in Y : |y| < \varepsilon_n\},$$

and we may assume that $\varepsilon_n \to 0$.

Then $H_1 \subseteq TG_0$; for if $y \in H_1$, then by (1) $\exists\ x_1 \in G_1$ such that $|y - Tx_1| < \varepsilon_2$ or $y - Tx_1 \in H_2$. By (1) $\exists\ x_2 \in G_2$ such that $|y - Tx_1 - Tx_2| < \varepsilon_3$ or $y - Tx_1 - Tx_2 \in H_3$. By induction, $\exists\ x_n \in G_n$ such that $|y - Tx_1 - Tx_2 - \ldots - Tx_n| < \varepsilon_{n+1}$. Thus, $y = \sum_{k=1}^{\infty} Tx_k$ and since $|x_k| < r/2^k$, the series $\sum_{k=1}^{\infty} x_k$ converges to some $x \in X$. Since

$$|x| \le \sum_{k=1}^{\infty} |x_k| < \sum_{k=1}^{\infty} r/2^k = r,$$

$x \in G_0$ and continuity implies that $Tx = y \in TG_0$ so $H_1 \subseteq TG_0$. Thus, TG_0 is a neighborhood of 0.

Theorem 3 (Open Mapping Theorem, OMT). Let X be complete. If $T \in L(X, Y)$ is such that $\mathcal{R}T$ is second category in Y, then T is open (and, hence, onto).

Proof: Lemmas 1 and 2.

Theorem 3 holds in particular if both X and Y are complete and T is onto, but there are examples of NLS which are second category but not complete ([Bo] Exer. III.4.4).

Corollary 4. If X and Y are complete and $T \in L(X, Y)$ is 1-1 and onto, then T is a homeomorphism.

Corollary 5. Suppose the vector space X is a quasi-normed space under the two complete quasi-norms $|\ |_1$ and $|\ |_2$. If $|\ |_1$ is stronger than $|\ |_2$, then $|\ |_1$ and $|\ |_2$ are equivalent.

Proof: The identity map from $(X, |\ |_1)$ to $(X, |\ |_2)$ is continuous so Corollary 4 gives the result.

Example 6. The completeness of both quasi-norms in Corollary 5 is important. The identity map $I : (C[0, 1], \|\ \|_\infty) \to (C[0, 1], \|\ \|_1)$ is continuous but does not have a continuous inverse. This example also shows that the completeness of the range space in the OMT is important even when the domain space is complete.

See also Exer. 9.

Example 7. The completeness of the domain space in the OMT is also important even when the range space is complete. Let X be a separable B-space and $\{x_k : k \in H\}$ a Hamel basis for X with $\|x_k\| = 1 \ \forall k \in H$. Define a norm, $\| \ \|'$, on X by $\|x\|' = \sum_{h \in H} |t_h|$, where $x = \sum_{h \in H} t_h x_h$. Since $\|x_h\| = 1 \ \forall h \in H$, $\|x\| \le \sum_{h \in H} |t_h| = \|x\|'$ and the identity $I : (X, \| \ \|') \to (X, \| \ \|)$ is continuous. However, since $(X, \| \ \|')$ is not separable ($\|x_h - x_k\|' = 1$ for $h \neq k$), $I : (X, \| \ \|) \to (X, \| \ \|')$ is not continuous.

We now consider the Closed Graph Theorem.

Let X and Y be topological spaces and f a mapping with domain, $\mathscr{D}(f) \subseteq X$, and range, $\mathscr{R}f$, in Y. The graph of f is $G(f) = \{(x, f(x)) : x \in \mathscr{D}(f)\} \subseteq X \times Y$. f is said to be a closed mapping if $G(f)$ is closed in $X \times Y$ (note this is not in agreement with the usual topological definition).

In semi-metric spaces we have the following simple test for closedness.

Proposition 8. If X and Y are semi-metric spaces and $f : \mathscr{D}(f) \subseteq X \to Y$, then f is closed if and only if whenever $\{x_k\} \subseteq \mathscr{D}(f)$ is such that $x_k \to x \in X$ and $f(x_k) \to y \in Y$, then $x \in \mathscr{D}(f)$ and $y = f(x)$.

Corollary 9. If X and Y are semi-metric spaces and $f : X \to Y$ is continuous, then f is closed.

Example 10. The converse of Corollary 9 is false. Let $X = Y = C[0, 1]$ and $X_0 = C^1[0, 1] = \{x \in X : x' \text{ exists on } [0, 1] \text{ and } x' \in X\}$. Define $D : X_0 \to Y$ by $Dx = x'$. Then D is linear and is closed with respect to $\| \|_\infty$. But D is not continuous since if $x_k(t) = t^k$, then $\|x_k\|_\infty = 1$ but $\|Dx_k\|_\infty = k$.

Note X_0 is not closed in X and is, therefore, not complete with respect to $\| \|_\infty$.

Example 11. It is important in Corollary 9 that the domain of the function be all of the space X. Let $X = Y = C[a, b]$ and $\mathscr{P}$ the polynomials in X. The identity operator $I : \mathscr{P} \subseteq X \to Y$ is continuous with respect to $\| \|_\infty$, linear but not closed.

Proposition 12. If $f : \mathscr{D}(f) \subseteq X \to Y$ is closed and 1-1, then f^{-1} is closed.

Proof: $G(f^{-1}) = \{(f(x), x) : x \in \mathscr{D}(f)\}$.

Proposition 13. Let X and Y be semi-NLS with Y complete. Let $T : \mathscr{D}(T) \subseteq X \to Y$ be linear. If T is closed and continuous, then $\mathscr{D}(T)$ is closed.

Proof: Let $x \in \overline{\mathscr{D}(T)}$. Choose $x_k \in \mathscr{D}(T)$ such that $x_k \to x$. Then $\{Tx_k\}$ is Cauchy $[\|Tx_k - Tx_j\| \le \|T\|\|x_k - x_j\|]$ so $\exists\, y \in Y$ such that $Tx_k \to y$. Since T is closed, $x \in \mathscr{D}(T)$.

Corollary 14. Let X and Y be semi-NLS with X complete. Let $T : \mathscr{D}(T) \subseteq X \to Y$ be a closed linear operator. If T^{-1} exists and is continuous, then $\mathscr{R}T$ is closed.

Proof: T^{-1} is closed by Proposition 12 and $\mathscr{R}T = \mathscr{D}(T^{-1})$ so the result follows from Proposition 13.

We now establish the Closed Graph Theorem which under appropriate completeness assumptions asserts that a closed linear operator is continuous.

Theorem 15 (Closed Graph Theorem, CGT). Let X and Y be complete quasi-normed spaces and $T : X \to Y$ a closed linear operator. Then T is continuous.

Proof: $X \times Y$ is a complete quasi-normed space and $G(T) \subseteq X \times Y$ is closed so it is complete. Let $P_1 : G(T) \to X$ be $P_1(x, Tx) = x$. Then P_1 is linear, continuous, 1-1 and onto X so P_1^{-1} is continuous by the OMT. If $P_2 : G(T) \to Y$ is $P_2(x, Tx) = Tx$, then P_2 is continuous. But then $T = P_2P_1^{-1}$ is continuous.

Remark 16. Examples 6 and 7 show that completeness in both the domain and range spaces is important.

The great utility of the CGT is that it is often easy to check the closedness of an operator.

To illustrate the use of the CGT, suppose that (S, Σ, μ) is a measure space and $f : S \to \mathbb{R}$ is Σ-measurable. Let $1 \le p < \infty$ and suppose that $fg \in L^p(\mu)$ $\forall\, g \in L^p(\mu)$. We use the CGT to show that $f \in L^\infty(\mu)$. Define a linear map $F : L^p(\mu) \to L^p(\mu)$ by $Fg = fg$. We claim that F is closed; for suppose that $g_k \to g$ and $Fg_k \to h$ in $L^p(\mu)$. Then there is a subsequence $\{g_{n_k}\}$ which converges μ-a.e. to g so $Fg_{n_k} = fg_{n_k} \to fg$ μ-a.e. so $fg = Fg = h$ μ-a.e. Thus, F is continuous, and we may assume that $\|F\| \le 1$. Let $\delta > 0$ and set $E = \{t \in S : |f(t)| \ge 1 + \delta\}$. Since $\|F^n g\| \le \|g\|$ $\forall\, n \ge 1$, $\int |g|^p d\mu \ge \int |f^n g|^p d\mu \ge \int_E (1+\delta)^{np} |g|^p d\mu$ and since $(1+\delta)^{np} \to \infty$, $\int_E |g|^p d\mu = 0$ $\forall\, g \in L^p(\mu)$ so $\mu(E) = 0$. Hence, $f \in L^\infty(\mu)$ with $\|f\|_\infty \le 1$.

We consider another application of the OMT. For $f \in L^1[0, 2\pi]$ and $n \in \mathbb{Z}$, set $c_n(f) = \frac{1}{2\pi}\int_0^{2\pi} f(t)e^{-int}dt$ (the nth Fourier coefficient of f; see 9.3). Let $c_0(\mathbb{Z})$ be the space of all complex-valued sequences defined on $\mathbb{Z}$ such that $\lim_{n \to \pm\infty} t_n = 0$ equipped with the sup-norm. By the Riemann-Lebesgue Lemma ([HS] 16.35), $\{c_n(f)\} \in c_0(\mathbb{Z})$ $\forall\, f \in L^1[0, 2\pi]$ so the map $F : f \to \{c_n(f)\}$ carries $L^1[0, 2\pi]$ into $c_0(\mathbb{Z})$ and clearly $\|F\| \le 1/2\pi$. Is the map F onto $c_0(\mathbb{Z})$? Since the map F is 1-1 ([HS], 16.34), if F is onto, then it has a bounded inverse by Corollary 4. However, if D_k is the Dirichlet kernel (see 9.3), then $\|F(D_k)\|_\infty = 1/2\pi$

while $\|D_k\|_1 \to \infty$ so F cannot have a bounded inverse.

See Exer. 6 for another application of the CGT.

Exercise 1. Let X, Y be complete quasi-normed spaces and $T : X \to Y$ linear. Suppose $A \subseteq Y'$ separates the points of Y. If $y'T$ is continuous $\forall\, y' \in A$, show that T is continuous.

Exercise 2. Show there is no sequence, $\{t_k\}$, in $\mathbb{R}$ such that $\sum_{k=1}^{\infty} a_k$ converges absolutely if and only if $\{t_k a_k\}$ is bounded. [Assume such a sequence exists with $t_k \neq 0 \;\forall\, k$. Consider the map $T : \ell^\infty \to \ell^1$, $\{a_k\} \to \{a_k/t_k\}$.]

Exercise 3. Show the CGT $\Rightarrow$ Corollary 4. [If $T : X \to Y$, consider X/ker T.]

Exercise 4. Show that Corollary 4 $\Rightarrow$ Theorem 3.

Exercise 5. Let X, Y be complete semi-NLS and $T : X \to Y$ linear. The graph norm on X is $\|x\|' = \|x\| + \|Tx\|$. Show that $\|\;\|'$ is complete if and only if T is closed.

Exercise 6. Let X be a B-space and (S, Σ, μ) a measure space. Suppose that $f : S \to X$ is such that $x' \circ f = x'f \in L^1(\mu)\; \forall\, x' \in X'$. Show the map $T : X' \to L^1(\mu)$, $Tx' = x'f$, is continuous (with respect to the dual norm).

Exercise 7. Let X, Y be complete and $T \in L(X, Y)$. Then either $\mathcal{R}T$ is closed or is first category in Y.

Exercise 8. Suppose that $\| \ \|$ is a complete norm on $C[0, 1]$ such that $\|f_k - f\| \to 0$ implies $f_k(t) \to f(t)$ $\forall\, t \in [0, 1]$. Show $\| \ \|$ is equivalent to $\| \ \|_\infty$.

Exercise 9. Let X be a B-space and f a linear functional on X which is not continuous. Set $p(x) = \|x\| + |\langle f, x\rangle|$. Show p is strictly stronger than $\| \ \|$ and p is not complete.

Exercise 10. Show that $L^2[0, 1]$ is first category in $L^1[0, 1]$.

10.1 Schauder Basis

Definition 1. Let X be a TVS. A sequence $\{b_k\}_{k=1}^{\infty} \subseteq X$ is a <u>Schauder basis</u> for X if each $x \in X$ has a unique representation $x = \sum_{k=1}^{\infty} t_k b_k$.

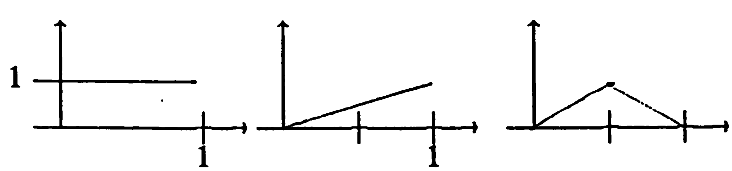
If $\{b_k\}$ is a Schauder basis and $x = \sum_{k=1}^{\infty} t_k b_k \in X$, then the linear functionals $f_k : X \to \mathbb{F}$ defined by $\langle f_k, x\rangle = t_k$ are called the <u>coordinate functionals</u> relative to $\{b_k\}$. Note $\langle f_i, b_j\rangle = \delta_{ij}$ and $x = \sum_{k=1}^{\infty} \langle f_k, x\rangle b_k$.

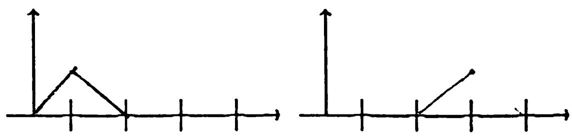
Example 2. $\{e_k\}_{k=1}^{\infty}$ form a Schauder basis in c_0, ℓ^p $(1 \le p < \infty)$ and s.

Example 3. c also has a Schauder basis, namely, $\{e\} \cup \{e_k\}_{k=1}^{\infty}$ where $e = (1, 1, \ldots)$. For if $x = \{x_k\} \in c$ and $x_0 = \lim x_k$, then

$$x = x_0 e + \sum_{k=1}^{\infty} (x_k - x)e_k.$$

Example 4. $C[0, 1]$ has a Schauder basis (the original Schauder basis!) consisting of the functions $f_0 = 1$, $f_1(t) = t$, $f_2(t) = \begin{cases} 2t & 0 \le t < 1/2 \\ 2 - 2t & 1/2 \le t \le 1 \end{cases}$ and $f_{2^n+i}(t) = f_2(2^n t - i + 1)$ for $n \ge 1$, $i = 1, \ldots, 2^n$.





See [Si] for the verification that $\{f_k\}$ forms a Schauder basis.

Example 5. The Haar system forms a Schauder basis in $L^p[0, 1]$ $(1 \le p < \infty)$ (see [Si]).

Example 6. ℓ^∞ has no Schauder basis (Exer. 1).

Schauder conjecture. In the 1920's Schauder introduced the notion of a basis and advanced the conjecture that every separable B-space has a Schauder basis. This problem was only solved in the negative by Per Enflo in 1972 ($[E_1]$). Enflo gave an example of a separable reflexive B-space with no Schauder basis.

As an application of the OMT we show that the coordinate functionals relative to the Schauder basis of a complete Hausdorff quasi-normed space are always continuous. First an example where the coordinate functionals are not continuous, the NLS in the example is, of course, not complete.

Example 7. If $\{b_k\}$ is a Schauder basis and $\{f_k\}$ are the coordinate functionals relative to $\{b_k\}$, then f_1 cannot be continuous if $b_k \to b_1$ for in this case $\langle f_1, b_k \rangle = 0$ for $k > 1$ and $\langle f_1, b_1 \rangle = 1$. We construct a

Schauder basis with this property in c_{00}. Put $b_1 = e_1$, $b_k = e_1 + \frac{1}{k}e_k$ $(k \geq 2)$. Clearly $b_k \to b_1$. Moreover, $\{b_k\}_{k=1}^{\infty}$ is a Schauder basis; for if $x = (x_1, \ldots, x_k, 0 \ldots) \in c_{00}$, then set $t_1 = x_1 - \sum_{j=1}^{k} jx_j$, $t_2 = 2x_2, \ldots, t_k = kx_k$ so that $x = \sum_{j=1}^{k} t_j b_j$.

Definition 8. A Schauder basis $\{b_k\}$ in a quasi-normed space $(X, |\ |)$ is monotone if for each $x = \sum_{k=1}^{\infty} t_k b_k$, the sequence $\{|\sum_{k=1}^{n} t_k b_k|\}$ is monotone increasing. Note in this case, $|\sum_{k=1}^{n} t_k b_k| \uparrow |x|$ (2.4).

The Schauder bases in Examples 2 and 3 are monotone; the one in Example 7 is not monotone.

We first observe that the coordinate functionals relative to a monotone basis are continuous.

Lemma 9. Let X be a quasi-normed space and $x_0 \in X$, $x_0 \neq 0$. Let f be a linear functional on X and $g : X \to X$ be defined by $g(x) = \langle f, x \rangle x_0$. If g is continuous, then f is continuous.

Proof. If f is not continuous at 0, $\exists\, x_k \to 0$ with $\{1/\langle f, x_k \rangle\}$ bounded. Then $g(x_k) \to 0$ in X so $g(x_k)/\langle f, x_k \rangle = x_0 \to 0$. This contradiction establishes the result.

Theorem 10. Let X be a quasi-normed space with a monotone Schauder basis $\{b_k\}$. Then the coordinate functionals $\{f_k\}$ with respect to $\{b_k\}$ are continuous.

Proof. Fix k. Define $g_k : X \to X$ by $g_k(x) = \langle f_k, x\rangle b_k$. Then

$$(1) \qquad |g_k(x)| = |\sum_{j=1}^{k} \langle f_j, x\rangle b_j - \sum_{j=1}^{k-1} \langle f_j, x\rangle b_j|$$

$$\leq |\sum_{j=1}^{k} \langle f_j, x\rangle b_j| + |\sum_{j=1}^{k-1} \langle f_j, x\rangle b_j| \leq 2|x|$$

by monotonicity. Hence, g_k is continuous and f_k is continuous by Lemma 9.

We next show that a complete quasi-normed space with a Schauder basis can be renormed so that the basis is monotone with respect to the new norm. If $\{b_k\}$ is a Schauder basis, we set $P_n x = \sum_{k=1}^{n} t_k b_k$ where $x = \sum_{k=1}^{\infty} t_k b_k$.

Lemma 11. Let $(X, |\ |)$ be a quasi-normed space with a Schauder basis $\{b_k\}$. Then $|x|' = |\sum_{k=1}^{\infty} t_k b_k|' = \sup_n |\sum_{k=1}^{n} t_k b_k|$ defines a quasi-norm on X under which $\{b_k\}$ is a monotone basis.

Proof. Everything is clear except that scalar multiplication is continuous. Let $\varepsilon > 0$.

Suppose $a_k \to 0$, $|x_k|' \to 0$. Note as in (1), we have

(2) $$\left|\sum_{k=p}^{q} t_k b_k\right| \le 2|x|' \text{ for all } q \ge p \text{ where } x = \sum_{k=1}^{\infty} t_k b_k .$$

Let U be a balanced neighborhood of 0 such that $|x| \le \varepsilon \; \forall x \in U$. (2) implies $\lim_k |P_n x_k| = 0$ uniformly for $n \in \mathbb{N}$ so $\exists k_0$ such that $k \ge k_0$ implies $|a_k| < 1$ and $P_n x_k \in U \; \forall n \in \mathbb{N}$. Then $a_k P_n x_k \in U$ and $|a_k x_k|' \le \varepsilon$ for $k \ge k_0$, i.e. $|a_k x_k|' \to 0$. Apply 9.1.1.

Theorem 12. If $(X, |\ |)$ is a complete quasi-norm space, then $|\ |'$ is complete.

Proof. Let $\{x_n\}$ be Cauchy in $|\ |'$ and $x_n = \sum_{k=1}^{\infty} t_{nk} b_k$. For each k $\{t_{nk} b_k\}_{n=1}^{\infty}$ is $|\ |$-Cauchy since $|t_{pk} b_k - t_{qk} b_k| \le 2|x_p - x_q|'$ by (2). Put $y_k = |\ |\text{-}\lim_n t_{nk} b_k$. By Exer. 7.3, $\exists t_k \in \mathbb{F}$ such that $y_k = t_k b_k$ and

(3) $$\lim_n t_{nk} = t_k .$$

We show that $\sum_{k=1}^{\infty} t_k b_k = \sum_{k=1}^{\infty} y_k$ converges with respect to $|\ |$.

We have

$$\left|\sum_{k=p}^{q} y_k\right| \le \left|\sum_{k=p}^{q} (t_k - t_{nk})b_k\right| + \left|\sum_{k=p}^{q} (t_{nk} - t_{mk})b_k\right| + \left|\sum_{k=p}^{q} t_{mk}b_k\right| = T_1 + T_2 + T_3.$$

Let $\varepsilon > 0$. Choose and fix m such that $n \ge m$ implies $T_2 < \varepsilon/3$; this is possible since by (2) $T_2 \le 2|x_n - x_m|'$. Chooose N such that $p, q \ge N$ implies $T_3 < \varepsilon/3$ (note $\sum_{k=1}^{\infty} t_{mk}b_k$ is $|\ |$-convergent). Fix $p, q \ge N$ and choose n so large that $T_1 < \varepsilon/3$ (by (3)). Thus, $\sum_{k=1}^{\infty} y_k$ is $|\ |$-convergent by the completeness of $|\ |$. Set $x = \sum_{k=1}^{\infty} y_k$.

Next we show that $|x_n - x|' \to 0$. Choose N such that $p, q \ge N$ implies $|x_p - x_q|' < \varepsilon/2$. For $p, q \ge N$ and $m \in \mathbb{N}$ we have

$$|P_m x_p - P_m x_q| \le 2|x_p - x_q|' < \varepsilon$$

by (2) so letting $q \to \infty$ implies $\left|\sum_{k=1}^{m} t_{pk}b_k - \sum_{k=1}^{m} t_k b_k\right| \le \varepsilon \ \forall m, p \ge N$. Hence, $|x_p - x|' \le \varepsilon$ for $p \ge N$.

Corollary 13. (i) $|\ |$ and $|\ |'$ in Theorem 12 are equivalent. (ii) The coordinate functionals $\{f_k\}$ with respect to $\{b_k\}$ are continuous.

Proof. $|\ |'$ is stronger than $|\ |$ so (i) follows from the completeness in Theorem 12 and 10.5.

(ii) follows from part (i), Lemma 11 and Theorem 10.

From Corollary 13, it follows that $L^0[0, 1]$ is a complete, separable quasi-normed space which does not possess a Schauder basis (Example 5.18). For the same reason $L^p[0, 1]$, $0 < p < 1$, does not have a Schauder basis (see the remarks following 5.18).

Corollary 14. For each k define $P_k : X \to X$ by $P_k x = \sum_{j=1}^{k} \langle f_j, x \rangle b_j$, where $\{b_k\}$ is a Schauder basis for the complete quasi-normed space X. Then $\{P_k\}$ is equicontinuous.

Proof. By (2), $|P_k x| \le 2|x|' \; \forall x \in X, k \in \mathbb{N}$.

If X and Y are TVS with Schauder bases $\{x_k\}$ and $\{y_k\}$, respectively, and if $T \in L(X, Y)$, then T has a matrix representation in the following sense. If $x = \sum_k \beta_k x_k \in X$, then $y = Tx = \sum_k \beta_k T x_k$. Each Tx_k has a representation $Tx_k = \sum_j t_{jk} y_j$ so $y = Tx = \sum_k \beta_k \sum_j t_{jk} y_j = \sum_j \langle f_j, y \rangle y_j$, where $\{f_j\}$ are the coordinate functionals with respect to $\{y_j\}$. If the coordinate functionals are continuous,

$$\langle f_i, y \rangle = \langle f_i, Tx \rangle = \sum_k \beta_k \sum_j t_{jk} \langle f_i, y_j \rangle = \sum_k \beta_k t_{ik} .$$

Thus, T can be represented by the infinite matrix $[t_{jk}]$. If $x = \sum_k \beta_k x_k$ is represented by the column vector $\{\beta_k\}$, then the representation of Tx as a column vector can be computed by evaluation the formal matrix product $[t_{jk}][\beta_k] = [\langle f_j, Tx\rangle]$ $(Tx = y)$.

For an interesting expository article on Schauder bases see [J3].

We have a nice criteria for compactness in QNLS with a Schauder basis.

Theorem 15. Let $(X, |\ |)$ be a complete QNLS with Schauder basis $\{b_k\}$ and associated coordinate functionals $\{f_k\}$. A subset $K \subseteq X$ is relatively compact if and only if K is bounded and $\lim_N \sum_{k=N}^{\infty} \langle f_k, x\rangle b_k = 0$ uniformly for $x \in K$.

Proof. Suppose K is relatively compact. By Corollay 13, we may assume that $|\ |'$ is the quasi-norm of X. For each n, $x \in X$ put

$$R_n(x) = \sum_{k=n+1}^{\infty} \langle f_k, x\rangle b_k,\ s_n(x) = \sum_{k=1}^{n} \langle f_k, x\rangle b_k .$$

Let $\varepsilon > 0$. $\exists$ a finite ε-net $\{x_1, ..., x_p\} \subseteq K$. Since $\lim_n R_n(x_j) = 0$ for $j = 1, ..., p$, $\exists\, N$ such that $n \geq N$ implies $|R_n(x_j)|' < \varepsilon$. If $x \in K$, $\exists\, j$ such that $|x - x_j|' < \varepsilon$ so if $n \geq N$,

$$|R_n(x)|' = |x - s_n(x)|' \leq |x - x_j|' + |s_n(x_j) - s_n(x)|' + |R(x_j)|'$$
$$< \varepsilon + |x_j - x|' + \varepsilon < 3\varepsilon .$$

Conversely, $\exists\, N$ such that $|R_n(x)|' < \varepsilon$ for $n \geq N$, $x \in K$. Let $K_N = \{s_N(x) : x \in K\}$. Then $K_N \subseteq \text{span}\{b_1, ..., b_N\}$ and K_N is bounded so K_N is relatively compact, and, therefore, has a finite ε-net, $\{x_1, ..., x_p\}$. Then $\{x_1, ..., x_p\}$ is a finite 2ε-net for K since if $x \in K$, $\exists\, j$ such that $|s_N(x) - x_j|' < \varepsilon$ and $|x - x_j|' \le |s_N(x) - x_j|' + |R_N(x)|' < 2\varepsilon$.

Exercise 1. Show that a TVS with a Schauder basis is separable.

Exercise 2. Let $\{t_k\} \in \ell^\infty$ and define $T : \ell^1 \to \ell^1$ by $T\{s_k\} = \{s_k t_k\}$. Give the matrix representation of T with respect to $\{e_k\}$.

Exercise 3. Give the matrix representation of the operators R^k, L^k of Exer. 5.2 with respect to $\{e_k\}$.

Exercise 4. Show that the matrix $T = [1/(i + j)]$ induces a continuous linear operator on ℓ^2.

Part III
Locally Convex TVS (LCS)

In part III we consider an important class of TVS, the locally convex spaces. A TVS is locally convex if its topology has a neighborhood base which consists of convex sets. We use the Hahn-Banach Theorem to show that these spaces have large dual spaces and share many of the important properties of NLS. Most of the important spaces in analysis are locally convex spaces.

We begin by developing some elementary properties of convex sets.

11
Convex Sets

Let X be a vector space.

Definition 1. A subset $K \subseteq X$ is convex if $x, y \in K$, $0 \le t \le 1$, implies $tx + (1 - t)y \in K$.

For example, any sphere $S(x, r) = \{y \in X : \|x - y\| < r\}$ or ball $B(x, r) = \{y \in X : \|x - y\| \le r\}$ in a semi-NLS is convex [the analogous statement in a quasi-norm space is false; see Example 13.11].

Proposition 2. Let K be convex. If $x_1, \ldots, x_n \in K$ and $t_i \ge 0$, $\sum_{i=1}^{n} t_i = 1$, then $\sum_{i=1}^{n} t_i x_i \in K$. [The element $\sum_{i=1}^{n} t_i x_i$ is called a convex combination of the $\{x_i\}$.]

Proof: The result holds for $n = 2$ by definition. Assume that it holds for $n \le m$. We show that it holds for $n = m + 1$. Set $b = \sum_{i=2}^{m+1} t_i$. If $b = 0$, then $t_1 = 1$ and $\sum_{i=1}^{m+1} t_i x_i = x_1 \in K$. Suppose $b > 0$ and set $y = \sum_{i=2}^{m+1} (t_i/b)x_i$. By the induction hypothesis, $y \in K$. Since $t_1 + b = 1$, $t_1x_1 + by = \sum_{i=1}^{m+1} t_i x_i \in K$.

If $A \subseteq X$, the smallest convex set containing A is called the <u>convex hull</u> of A and is denoted by $\text{co}A$ (see Exer. 1).

Proposition 3. For $A \subseteq X$, $\text{co}A$ is the set of all convex combinations $\sum_{i=1}^{n} t_i x_i$ where $x_i \in A$, $t_i \ge 0$, $\sum_{i=1}^{n} t_i = 1$.

Proof: The set of all such convex combinations is convex and must be $\text{co}A$ by Proposition 2.

Definition 4. A subset $A \subseteq X$ is <u>absolutely convex</u> if it is balanced and convex.

Proposition 5. $A \subseteq X$ is absolutely convex $\Leftrightarrow$ $x, y \in A$ and $|t| + |s| \le 1$ implies $tx + sy \in A$.

Proof: $\Leftarrow$: clear.

$\Rightarrow$: If either $s = 0$ or $t = 0$, then $tx + sy \in A$; if $s \neq 0, t \neq 0$, then $(t/|t|)x \in A$, $(s/|s|)y \in A$ and $\frac{|t|}{|t|+|s|} + \frac{|s|}{|t|+|s|} = 1$ so

$$tx + sy = (|t| + |s|)\left(\frac{|t|}{|t|+|s|}\frac{tx}{|t|} + \frac{|s|}{|t|+|s|}\frac{sy}{|s|}\right) \in A .$$

Proposition 6. Let $A \subseteq X$ $(A \neq \emptyset)$ be absolutely convex. Then

(i) $0 \in A$,

(ii) $tA \subseteq sA$ when $|t| \leq |s|$,

(iii) $\sum_{i=1}^{n} t_i A \subseteq (\sum_{i=1}^{n} |t_i|)A$ when $t_i \in \mathbb{F}$.

Proof: (i) is clear. (ii): If $s = 0$, trivial; if $s \neq 0$ and $x \in A$, then $(t/s)x \in A$ so $tx \in sA$. (iii): for $n = 2$, (iii) is

$$sA + tA \subseteq (|s| + |t|)A .$$

If $s = t = 0$, this is trivial; otherwise,

$$\frac{s}{|s|+|t|}A + \frac{t}{|s|+|t|}A \subseteq A \Rightarrow sA + tA \subseteq (|s| + |t|)A .$$

Induction gives the general result.

The <u>absolutely convex hull</u> of a subset A of X is the smallest absolutely convex set containing A and is denoted by abco$A$ (see Exer. 1). As in Proposition 3,

Proposition 7. For $A \subseteq X$ the absolutely convex hull of A is the set of all $\sum_{i=1}^{n} t_i x_i$, where $x_i \in A$, $\sum_{i=1}^{n} |t_i| \leq 1$.

Proposition 8. Let X be a TVS.

(i) If $K \subseteq X$ is convex, $\bar{K}$ is convex.

(ii) If $K \subseteq X$ is absolutely convex, $\bar{K}$ is absolutely convex.

(iii) If $K \subseteq X$ is open, then $\operatorname{co}K$ is open.

Proof: (i): Let $I = [0, 1]$. Define $\Psi : X \times X \times I \to X$ by $\Psi(x, y, t) = tx + (1 - t)y$. Then K is convex if and only if $\Psi(K, K, I) \subseteq K$. Now Ψ is continuous and $\overline{K \times K \times I} = \bar{K} \times \bar{K} \times I$ so

$$\Psi(\bar{K}, \bar{K}, I) \subseteq \overline{\Psi(K, K, I)} \subseteq \bar{K}.$$

(ii) follows from (i) and 1.8.

(iii): Let $x \in \operatorname{co}K$. Then $x = \sum_{i=1}^{n} t_i x_i$, with $x_i \in K$, $t_i \geq 0$, $\sum_{i=1}^{n} t_i = 1$. Since K is open, for each $k = 1, \ldots, n$ $\exists$ an open neighborhood V_k of x_k such that $V_k \subseteq K$. Then $U = \sum_{k=1}^{n} t_k V_k$ is open and $x \in U \subseteq \operatorname{co}K$.

If X is a TVS and $A \subseteq X$, the <u>closed convex hull</u> of A, denoted by $\overline{\operatorname{co}}A$, is the smallest closed convex set containing A.

Proposition 9. $\overline{\operatorname{co}}A = \overline{(\operatorname{co}A)}$.

Proof: Since $\overline{coA}$ is closed, convex and contains A, $\overline{coA} \supseteq \overline{(coA)}$. By Proposition 8, $\overline{(coA)}$ is convex and contains A so $\overline{(coA)} \supseteq \overline{coA}$.

Definition 10. Let $K \subseteq X$ be a convex absorbing set which contains 0. The function $p_K = p : X \to \mathbb{R}$ defined by $p(x) = \inf\{t > 0 : x \in tK\}$ is called the <u>Minkowski functional of K</u> (<u>gauge of K</u>).

For example, if X is a semi-NLS and $K = \{x : \|x\| \le r\}$, then $p_K(x) = \|x\|/r$. In particular, if $r = 1$, $p_K = \| \ \|$.

p_K has the following properties.

Proposition 11. (i) $p(x + y) \le p(x) + p(y)$ $\forall x, y \in X$ (sublinear).

(ii) $p(tx) = tp(x)$ $t \ge 0$, $x \in X$ (positive homogeneous).

(iii) If K is balanced (i.e., if K is absolutely convex and absorbing), then p is a semi-norm.

Proof: Note that since K is absorbing, $p(x) < \infty$ $\forall x \in X$. Clearly $p \ge 0$.

(i): If $x, y \in X$ and $x \in tK$, $y \in sK$ $(s, t > 0)$, then

$$x + y \in tK + sK = (t + s)K$$

since K is convex. Hence, $p(x + y) \le s + t$ and $p(x + y) \le p(x) + p(y)$.

(ii): If $t > 0$,

$$p(tx) = \inf\{a : tx \in aK\} = \inf\{t \tfrac{a}{t} : x \in \tfrac{a}{t} K\} = tp(x).$$

(iii): If K is balanced, $tx \in aK \Leftrightarrow |t|x \in aK$ since aK is also balanced. Therefore, $p(tx) = p(|t|x) = |t|p(x)$ by (ii).

For the relationship between K and p_K, we have

Proposition 12. If K is absolutely convex and absorbing, then

$$\{x : p_K(x) < 1\} \subseteq K \subseteq \{x : p_K(x) \le 1\} .$$

Proof: If $p(x) < 1$, $\exists\, t$, $0 < t < 1$, such that $x \in tK$ and since K is balanced, $x \in K$. If $x \in K$, then clearly $p(x) \le 1$.

If K is a subset of a TVS, we have the following relationship between K and p_K.

Theorem 13. Let X be a TVS and $K \subseteq X$ absolutely convex and absorbing. Then p_K is continuous if and only if K is a neighborhood of 0. In this case, $\{x : P_K(x) < 1\}$ is the interior of K and $\{x : p_K(x) \le 1\}$ is the closure of K.

Proof: Suppose K is a neighborhood of 0. Let $\varepsilon > 0$. Then $p(x) \le \varepsilon$ for $x \in \varepsilon K$ so p is continuous at 0 and, therefore, continuous (Exer. 2.14).

If p is continuous, then $V = \{x : p(x) < 1\}$ is open and $V \subseteq K$ by Proposition 12.

Suppose $x \in \text{int}K$. Then $tx \in K$ for t sufficiently close to 1 so $\exists\; t > 1$ such that $tx \in K$ or $x \in \frac{1}{t}K$ and $p(x) \le \frac{1}{t} < 1$. Hence, $\text{int}K \subseteq V \subseteq K$ and $V = \text{int}K$.

If $p(x) \le 1$, then $tx \in V$ if $0 < t < 1$. Letting $t \to 1$, we have

$tx \to x$ so $x \in \overline{V}$. That is, $\overline{K} \subseteq \{x : p(x) \le 1\} \subseteq \overline{V} \subseteq \overline{K}$.

Exercise 1. Show the intersection of (absolutely) convex sets is (absolutely) convex.

Exercise 2. If K_1 and K_2 are convex, show that $K_1 \pm K_2$, tK_1 are convex for $t \in \mathbb{F}$.

Exercise 3. Show a linear map preserves convex sets.

Exercise 4. Show $\text{abco}A = \text{co}(\text{bal}A)$. What about $\text{bal}(\text{co}A)$?

Exercise 5. If X is a TVS and $K \subseteq X$ is convex, show $\text{int}K$ is convex.

Exercise 6. Show $\text{co}(tA) = t\,\text{co}A$ and $\overline{\text{co}}(tA) = t\,\overline{\text{co}}A$.

12

Separation of Convex Sets

In this section we use the Minkowski functional and the Hahn-Banach Theorem to establish several results on the separation of convex sets by linear functionals and hyperplanes.

Lemma 1. If X is a TVS and $f : X \to \mathbb{F}$ a non-zero linear functional, then f is open.

Proof: Let $A \subseteq X$ be open and $x \in A$. Then $A - x$ is an open neighborhood of 0 and, hence, absorbing. Since $f \neq 0$, $\exists\ a \in X$ such that $\langle f, a\rangle = 1$. $\exists\ \varepsilon > 0$ such that $ta \in A - x$ for $|t| \leq \varepsilon$. Hence,

$$\langle f, x\rangle + t \in f(A) \text{ for } |t| \leq \varepsilon .$$

Theorem 2. Let A and B be non-void, disjoint, convex subsets of a TVS X. If A is open, then $\exists\ x' \in X'$ and $c \in \mathbb{R}$ such that

$$(1) \qquad \mathcal{R}\langle x', x\rangle < c \leq \mathcal{R}\langle x', y\rangle \ \ \forall x \in A, y \in B .$$

Proof: First, assume $\mathbb{F} = \mathbb{R}$. Fix $a_0 \in A$, $b_0 \in B$. Put $x_0 = b_0 - a_0$, $C = A - B + x_0$. Then C is an open convex neighborhood of 0 such that $x_0 \notin C$ since A and B are disjoint. Let p be the Minkowski functional of C. Note $p(x_0) \geq 1$ by 11.13.

Let $M = \text{span}\{x_0\}$ and define f on M by $\langle f, tx_0\rangle = t$. If $t \geq 0$, then $\langle f, tx_0\rangle = t \leq tp(x_0) = p(tx_0)$; if $t < 0$, then $\langle f, tx_0\rangle = t \leq p(tx_0)$. Thus, $f \leq p$ on M. By the Hahn-Banach Theorem (8.2), f has a linear extension, x', to X which satisfies $x' \leq p$ on X. In particular, $x' \leq 1$ on C so $x' \geq -1$ on $-C$. Hence, $|x'| \leq 1$ on the neighborhood $C \cap (-C)$ of 0 and x' is continuous (Exer. 5.12).

If $a \in A$ and $b \in B$, then

$$\langle x', a\rangle - \langle x', b\rangle + 1 = \langle x', a - b + x_0\rangle \leq p(a - b + x_0) < 1$$

by 11.13. Hence, $\langle x', a\rangle < \langle x', b\rangle$. Thus, $x'(A)$ and $x'(B)$ are disjoint, convex subsets of $\mathbb{R}$ and $x'(A)$ lies to the left of $x'(B)$. $x'(A)$ is open by Lemma 1 and is, therefore, an open interval in $\mathbb{R}$. Pick c to be the right hand endpoint of $x'(A)$.

If $\mathbb{F} = \mathbb{C}$, by the first part there is a real-linear continuous functional f on X satisfying (1). Put $\langle x', x\rangle = \langle f, x\rangle - i\langle f, ix\rangle$ (8.3).

If f is a non-zero linear functional on a vector space X, then any level set of f, $\{x : f(x) = c\}$, is called a <u>hyperplane</u>. In Theorem 2, the set A lies in the open half-space $\{x : \langle x', x\rangle < c\}$ and B lies in the closed half-space $\{x : \langle x', x\rangle \geq c\}$ and the closed hyperplane

$$H = \{x : \langle x', x\rangle = c\}$$

is said to <u>separate</u> A and B.

Without some additional assumption on the convex sets it may not be possible to separate the convex sets by a hyperplane.

Example 3. In c_{00}, let $K = \{\{t_i\} :$ last non-zero t_i is positive$\}$. Let f be any non-zero linear functional on c_{00}. Then $<f, e_k> \neq 0$ for some k. Then $te_k + e_{k+1} \in K$ for any $t \in \mathbb{R}$ and

$$<f, te_k + e_{k+1}> = t<f, e_k> + <f, e_{k+1}>$$

may take on any value by choosing t appropriately. Note $0 \notin K$ so $\{0\}$ and K cannot be separated by any non-zero linear functional.

It is of interest, particularly in optimization theory, to know that arbitrary disjoint convex subsets of $\mathbb{R}^n$ can always be separated by hyperplanes ([Fr] 6.3).

Exercise 1. If B in Theorem 2 is also open, show that A and B can be strictly separated in the sense that $\exists\ \delta > 0$ such that

$$\mathcal{R}<x', x> \leq c - \delta < c < c + \delta \leq \mathcal{R}<x', y> \ \forall x \in A, y \in B .$$

13
Locally Convex TVS

Definition 1. A TVS (X, τ) is a <u>locally convex TVS</u> (LCS) or τ is a <u>locally convex topology</u> if and only if τ has a neighborhood base at 0 consisting of convex sets.

Example 2. Any semi-NLS is a LCS.

Example 3. Let $\mathcal{P}$ be a family of semi-norms on a vector space X. Let $\sigma(X, \mathcal{P})$ be the weakest topology on X such that each semi-norm $p \in \mathcal{P}$ is continuous. Thus, a net $\{x_\delta\}$ in X converges to 0 in $\sigma(X, \mathcal{P})$ if and only if $p(x_\delta) \to 0$ $\forall p \in \mathcal{P}$. The subsets, U, of the form

$$U = \{x : p(x) < \varepsilon, p \in \mathcal{P}_0\},$$

where $\varepsilon > 0$ and $\mathcal{P}_0$ is a finite subset of $\mathcal{P}$, form an open neighborhood base at 0 for $\sigma(X, \mathcal{P})$. Thus, $\sigma(X, \mathcal{P})$ is a locally convex topology for X. Note $\sigma(X, \mathcal{P})$ is Hausdorff if and only if $\forall x \neq 0 \ \exists p \in \mathcal{P}$ such that

$p(x) \neq 0$, i.e., if and only if $\mathcal{P}$ separates the points of X. A net $\{x_\delta\}$ in X is a Cauchy net in $\sigma(X, \mathcal{P})$ if and only if for every $\varepsilon > 0$ and every $p \in \mathcal{P} \ \exists \ \delta$ such that $\alpha, \beta \geq \delta$ implies that $p(x_\alpha - x_\beta) < \varepsilon$.

We show below in Theorem 6 that Example 3 is a canonical example in the sense that every locally convex topology is $\sigma(X, \mathcal{P})$ for some family $\mathcal{P}$.

Lemma 4. If U is an open, convex neighborhood of 0 in the TVS X, then U contains an absolutely convex open neighborhood of 0.

Proof: U contains an open, balanced neighborhood, V, of 0. Set W = coV. Then W is an open neighborhood of 0 (11.8), and since V is balanced, W is absolutely convex (Exer. 11.4) with $W \subseteq U$.

Theorem 5. If X is a LCS, then X has an open (closed) neighborhood basis at 0 consisting of absolutely convex sets.

Proof: The statement about open sets follows from Lemma 4 and Exer. 11.5.

Let U be a closed neighborhood of 0 in X. By Exer. 11.5 and Lemma 4 $\exists$ an absolutely convex neighborhood, V, of 0 such that $V \subseteq U$. Then $\bar{V} \subseteq U$ and $\bar{V}$ is absolutely convex (11.8).

Theorem 6. If (X, τ) is a LCS, then $\tau = \sigma(X, \mathscr{P})$ for some family of semi-norms $\mathscr{P}$.

Proof: Let $\mathscr{U}$ be a neighborhood base at 0 of open, absolutely convex sets. For $U \in \mathscr{U}$ let p_U be the Minkowski functional of U and set $\mathscr{P} = \{p_U : U \in \mathscr{U}\}$. If $U \in \mathscr{U}$, then $U = \{x : p_U(x) < 1\}$ (11.13) so $\tau \subseteq \sigma(X, \mathscr{P})$. But, each p_U is τ-continuous (11.13) so $\sigma(X, \mathscr{P}) \subseteq \tau$.

Corollary 7. If (X, τ) is a LCS, then $\tau = \sigma(X, \mathscr{P})$, where $\mathscr{P}$ is the family of all continuous semi-norms on X.

We have the following criteria for continuity and equicontinuity in LCS.

Theorem 8. Let X and Y be LCS and assume that $\sigma(Y, Q)$ is the topology of Y. Let $T : Y \to Y$ be linear [$\mathscr{T}$ a family of linear maps from X into Y]. Then T is continuous [$\mathscr{T}$ is equicontinuous] if and only if $\forall q \in Q$ $\exists$ a continuous semi-norm p on X such that $q(Tx) \le p(x)$ $\forall x$ [$\forall T \in \mathscr{T}$].

Proof: $\Leftarrow$: Clear.

$\Rightarrow$: $V = \{y \in Y : q(y) < 1\}$ is an open neighborhood of 0 in Y. Therefore, $\exists$ $\varepsilon > 0$ and continuous semi-norms $p_1, ..., p_n$ on X such that $TU \subseteq V$ [$TU \subset V$ $\forall T \in \mathscr{T}$], where $U = \{x : p_i(x) < \varepsilon, i = 1, ..., n\}$. That is, $q(Tx) < 1 \ \forall x \in U \ [\forall T \in \mathscr{S}]$. Put $p(x) = (1/\varepsilon) \max_{1 \le i \le n} p_i(x)$. Then p is a continuous semi-norm on X. If $x \in X$ and t is chosen such that $p(x) < t$, then $x/t \in U$ so $q(T(x/t)) = \frac{1}{t} q(Tx) < 1$ or $q(Tx) < t \ [\forall T \in \mathscr{S}]$. Letting $t \to p(x)$ gives $q(Tx) \le p(x) \ [\forall T \in \mathscr{S}]$.

This result should be compared with 5.5 and 5.21 for NLS.

We next show that a LCS has a large dual space.

Theorem 9. Let X be a LCS and X_0 a closed linear subspace with $x_0 \notin X_0$. Then $\exists$ a continuous linear functional $x' \in X'$ such that $\langle x', x_0 \rangle = 1$ and $\langle x', X_0 \rangle = 0$.

Proof: Let $M = \operatorname{span}\{x_0, X_0\}$. $\exists$ an open, absolutely convex neighborhood, V, of 0 such that $(x_0 + V) \cap X_0 = \emptyset$. Let p be the Minkowski functional of V. Thus, p is a continuous semi-norm and $V = \{x : p(x) < 1\}$ (11.13). Define m' on M by $\langle m', tx_0 + x \rangle = t$, where $t \in \mathbb{F}$, $x \in X_0$. We claim that $|\langle m', m \rangle| \le p(m)$ $\forall m = tx_0 + x \in M$. If $t = 0$, this is clear; if $t \ne 0$, then

$$p(m) = p(tx_0 + x) = |t| p(x_0 + x/t) \ge |t| = |\langle m', m \rangle|$$

since $x_0 + x/t \notin V$. By the Hahn-Banach Theorem m' has a linear extension x' to X satisfying $|\langle x', x \rangle| \le p(x) \ \forall x \in X$. Then $x' \in X'$ by Theorem 8 and $\langle x', x_0 \rangle = 1$ and $\langle x', X_0 \rangle = 0$.

Corollary 10. Let X be a Hausdorff LCS. If $x_0 \ne 0$, $x_0 \in X$, then $\exists$

$x' \in X'$ such that $<x', x_0> \neq 0$. That is, the dual space X' separates the points of X.

Proof: Put $X_0 = \{0\}$ in Theorem 9.

Example 11. It follows that the space $L^0[0, 1]$ $(L^p[0, 1], 0 < p < 1)$ is not locally convex. In particular, the spheres $\{x : |x| < r\}$ in $L^0[0, 1]$ are not convex. (Recall Example 5.18.)

Theorem 9 and Corollary 10 should be compared with 8.1.2 and 8.1.4. Results analogous to 8.1.4 and 8.1.5 are given by the following two propositions.

Proposition 12. Let X be a LCS and p a continuous semi-norm on X. If $x_0 \in X$, $\exists\, x_0' \in X'$ such that $<x_0', x_0> = p(x_0)$ and $|<x_0', x>| \leq p(x)$ $\forall\, x \in X$.

Proof: Let $E = \text{span}\{x_0\}$ and define $x' : E \to \mathbb{F}$ by

$$<x', tx_0> = tp(x_0).$$

Then $|<x', x>| \leq p(x)$ $\forall\, x \in E$. By the Hahn-Banach Theorem x' can be extended to a linear functional x_0' on X satisfying $|<x_0', x>| \leq p(x)$ $\forall\, x \in X$. $x_0' \in X'$ by Theorem 8.

Proposition 13. Let K be a closed convex subset of a Hausdorff LCS X and $x \notin K$. Then $\exists\, x' \in X'$ such that

$$\mathcal{R}<x', x> > \sup\{\mathcal{R}<x', y> : y \in K\}.$$

If K is absolutely convex, x' can be chosen such that

(1) $$<x', x> > 1 \geq \sup\{|<x', y>| : y \in K\} .$$

Proof: $\exists$ an open absolutely convex neighborhood of 0, U, such that $(x + U) \cap K = \emptyset$. By 12.2 $\exists\, x' \in X'$ and $c \in \mathbb{R}$ such that

$$\mathcal{R}<x', x> > c \geq \mathcal{R}<x', y> \ \forall\, y \in K .$$

This gives the first part. If K is balanced, then $<x', K>$ is a balanced set of scalars so

$$|<x', x>| \geq \mathcal{R}<x', x> > c \geq \sup\{ \mathcal{R}<x', y> : y \in K\}$$
$$= \sup\{|<x', y>| : y \in K\} .$$

Multiplication of x' by a suitable scalar gives (1).

Proposition 13 has an interesting corollary pertaining to the closures of convex sets.

Corollary 14. Let X be a vector space with two locally convex topologies, τ_1 and τ_2, which have the same continuous linear functionals. A convex set $K \subseteq X$ is τ_1-closed if and only if it is τ_2-closed.

Proof: Suppose K is τ_1-closed and $x \notin K_1$. By Proposition 13 $\exists$ a τ_1-continuous linear functional x' and $\delta > 0$ such that

$$|<x', x>| > |<x', y>| + \delta \ \forall\, y \in K .$$

Then $|<x', x> - <x', y>| > \delta \ \ \forall\, y \in K$. Hence, $U = \{y : |<x', y>| < \delta\}$ is a τ_2-neighborhood of 0 such that $(x + U) \cap K = \emptyset$ and K is τ_2-closed.

Bounded Sets in LCS:

Theorem 15. A subset B of a LCS X with topology $\sigma(X, \mathscr{P})$ is bounded if and only if $p(B)$ is bounded $\forall\, p \in \mathscr{P}$.

Proof: Suppose B is bounded and $p \in \mathscr{P}$. Let p also denote the semi-norm topology induced by p. The identity $(X, \sigma(X, \mathscr{P})) \to (X, p)$ is continuous so the implication follows from 5.3.

Let U be a neighborhood of 0 in X. Then $\exists\ \varepsilon > 0$, $p_1, \ldots, p_n \in \mathscr{P}$ such that $\{x : p_i(x) < \varepsilon, i = 1, \ldots, n\} \subseteq U$. $\exists\ M > 0$ such that $p_i(x) < M$ $\forall\, x \in B$, $i = 1, \ldots, n$. Hence $(\varepsilon/M)B \subseteq U$ and B is $\sigma(X, \mathscr{P})$ bounded.

Analogous to 9.12, we have

Theorem 16 (Mackey). Let (X, τ) be a LCS. A subset $B \subseteq X$ is bounded if and only if $\langle x', B\rangle$ is bounded $\forall\, x' \in X'$.

Proof: $\Rightarrow$: 5.3.

$\Leftarrow$: By Theorem 15 it suffices to show that $p(B)$ is bounded for each continuous semi-norm p on X, i.e., B is bounded in the semi-NLS (X, p). By 9.12, B is bounded in (X, p) if and only if $\langle x', B\rangle$ is bounded $\forall\, x' \in (X, p)'$. But $(X, p)' \subseteq (X, \tau)'$ so $\langle x', B\rangle$ is bounded $\forall\, x' \in (X, p)'$ by hypothesis.

Recall that a sequence $\{x_k\}$ in a TVS X is said to be a <u>Cauchy</u>

<u>sequence</u> if $\forall$ neighborhood of 0, U, in X $\exists$ N such that $j, k \geq N$ implies $x_j - x_k \in U$. If X is a LCS, this is equivalent to the requirement that $\forall$ continuous semi-norm p on X and for each $\varepsilon > 0$ $\exists$ N such that $j, k \geq N$ implies $p(x_j - x_k) < \varepsilon$. (See Example 3.) We say that a subset A of a TVS X is <u>sequentially complete</u> if every Cauchy sequence in A converges to a point of A. For sequentially complete LCS, we have the following criterion for $\mathscr{K}$ boundedness.

Proposition 17. Let X be a LCS. If $A \subseteq X$ is absolutely convex, bounded and sequentially complete, then A is $\mathscr{K}$ bounded.

Proof: Let $\{x_j\} \subseteq A$ and $t_j \to 0$. Given any subsequence of $\{t_j\}$, pick a further subsequence $\{t_{n_j}\}$ satisfying $\sum_{j=1}^{\infty} |t_{n_j}| \leq 1$. Let p be a continuous semi-norm on X. The partial sums, $s_k = \sum_{j=1}^{k} t_{n_j} x_{n_j}$, form a Cauchy sequence in A since each $s_k \in A$ and

$$p(x_k - x_\ell) \leq \sum_{j=k+1}^{\ell} |t_{n_j}| p(x_{n_j}) \leq \sup\{p(x_i) : i \in \mathbb{N}\} \sum_{j=k+1}^{\ell} |t_{n_j}|$$

for $k < \ell$ (Theorem 15). Thus, the series $\sum_{j=1}^{\infty} t_{n_j} x_{n_j}$ converges to a point in A, and A is $\mathscr{K}$ bounded.

The absolute convexity is important; see Exercise 20.3.

Corollary 18. If X is a sequentially complete LCS, then X is an

$\mathcal{A}$-space (§4).

This result now gives us a large number of examples of $\mathcal{A}$-spaces (§4) and, thus, spaces for which the UBP holds (§9). Klis's example of a non-complete normed $\mathcal{K}$-space shows that the converse of Corollary 18 is false ([Kℓ]).

Convex Hulls of Compact Sets:

Recall that a subset E of a metric space X is totally bounded or precompact if $\forall\ \varepsilon > 0\ \exists$ a finite set $F \subseteq X$ such that

$$E \subseteq \cup\{S(x, \varepsilon) : x \in F\},$$

where $S(x, \varepsilon)$ is the ε-sphere centered at x ([DS]). This idea is readily generalized to TVS.

Definition 19. A subset E of a TVS X is <u>precompact</u> or <u>totally bounded</u> if $\forall$ neighborhood of 0, U, in X $\exists$ a finite set F such that $E \subseteq F + U$.

Theorem 20. If X is a LCS and $E \subseteq X$ is precompact, then coE is precompact.

Proof: Let U be a neighborhood of 0 in X. Then $\exists$ a convex neighborhood of 0, V, in X such that $V + V \subseteq U$, and $\exists$ a finite set F such that $E \subseteq F + V$. Let $E_1 = coE$, $F_1 = coF$.

If $\{y_1, ..., y_n\}$ are the points of F, then F_1 is the image of the compact set $\{(t_1, ..., t_n) : t_i \geq 0, \sum_{i=1}^{n} t_i = 1\} \subseteq \mathbb{R}^n$ under the mapping

$(t_1, \ldots, t_n) \to \sum_{i=1}^{n} t_i y_i$ and is, therefore, compact.

If $x \in E_1$, then $x = \sum_{i=1}^{k} s_i x_i$, where $s_i \geq 0$, $\sum_{i=1}^{k} s_i = 1$ and $x_i \in E$. For each i, $\exists$ $z_i \in F$ such that $x_i - z_i \in V$. Decompose x into a sum $x = x^1 + x^2$, where $x^1 = \sum_{i=1}^{k} s_i z_i$, $x^2 = \sum_{i=1}^{k} s_i(x_i - z_i)$. Now $x^1 \in F_1$ and $x^2 \in V$ since V is convex. Thus, $E_1 \subseteq F_1 + V$. But, F_1 is compact so $\exists$ a finite set F_2 such that $F_1 \subseteq F_2 + V$ (Exer. 5). Hence,

$$E_1 \subseteq F_1 + V + V \subseteq F_2 + U$$

and E_1 is totally bounded.

From Theorem 20, we obtain a result of Mazur.

Corollary 21 (Mazur). Let X be a complete metrizable LCS. If $E \subseteq X$ is compact, then $\overline{\text{co}}E$ is compact.

Proof: Recall that a subset of a complete metric space is compact if and only if it is closed and precompact ([DS]); apply Exer. 5.

It is interesting to note that the convex hull of a compact subset of a finite dimensional space is compact ([Ru] 3.25). The completeness in Corollary 21 is important (see Exer. 6).

Completion of a LCS:

Let E be a Hausdorff LCS whose topology is generated by the family of semi-norms $\mathscr{P}$. For $p \in \mathscr{P}$ let $N_p = \{x : p(x) = 0\}$. Form the quotient space $E_p = E/N_p$ and denote the coset $x + N_p = \bar{x}_p$. Equip E_p with the quotient norm $\bar{p}$ from p, $\bar{p}(\bar{x}_p) = p(x)$ (6.4), and let F_p be the completion of E_p under $\bar{p}$ (8.1.14). Form the product space $F = \Pi F_p$ for $p \in \mathscr{P}$; F is then a complete LCS (Exer. 1.9 and Exer. 13).

With any $x \in E$ associate the element $\bar{x} = (\bar{x}_p)$ of F. The map $x \to \bar{x}$ is 1-1, linear and is a homeomorphism from E onto a linear subspace E^{-} of F. This construction shows that any Hausdorff LCS can be imbedded in a product of B-spaces. The closure of E^{-} in F then furnishes a completion of E. Furthermore, it is easy to see that E is complete if and only if E^{-} is closed in F.

Exercise 1. If $\mathscr{P} \supseteq \mathscr{Q}$, show $\sigma(X, \mathscr{P}) \supseteq \sigma(X, \mathscr{Q})$. Show that it is possible to have $\sigma(X, \mathscr{P}) = \sigma(X, \mathscr{Q})$ with $\mathscr{P} \neq \mathscr{Q}$.

Exercise 2. If X is a LCS whose topology is induced by a sequence of semi- norms $\{p_k\}$, show $\exists$ a sequence of semi-norms $\{p_k'\}$ such that $p_k' \le p_{k+1}'$ and $\{p_k'\}$ induces the same locally convex topology as the $\{p_k\}$.

Exercise 3. Let X be a TVS and $(Y, \sigma(Y, \mathscr{P}))$ a LCS. If $T : X \to Y$ is linear, show T is continuous if and only if pT is continuous $\forall\, p \in \mathscr{P}$.

Exercise 4. Let $(X, \sigma(X, \mathscr{P}))$ be a LCS and Z a linear subspace. Show

Z is a LCS with respect to the induced topology and the semi-norms $\{p|_Z : p \in \mathscr{P}\}$ generate the induced topology.

Exercise 5. Show a precompact set is bounded. Show a compact set is precompact. Show the closure of a precompact set is precompact.

Exercise 6. Show Corollary 21 is false if completeness is dropped. [Hint: Consider c_{00} with the ℓ^2-norm. Let $E = \{e_k/k : k \in \mathbb{N}\} \cup \{0\}$. Pick $t_k > 0$, $\sum_{k=1}^{\infty} t_k = 1$, and consider the vectors $(\sum_{k=1}^{n} t_k)^{-1} \sum_{k=1}^{n} t_k e_k/k$.]

Exercise 7. Let X, Y be LCS and $B : X \times Y \to \mathbb{F}$ bilinear. Show B is continuous if and only if $\exists$ a continuous semi-norm p(q) on X(Y) such that $|B(x, y)| \le p(x)q(y) \ \forall \ x \in X, y \in Y$.

Exercise 8. Let $\{t_k\} \in \ell^1$. Let X, Y be LCS with $\{x_k'\} \in X'$ and $\{y_k'\} \subseteq Y'$ equicontinuous. Show $B : X \times Y \to \mathbb{R}$,

$$B(x, y) = \sum_{k=1}^{\infty} t_k \langle x_k', x\rangle \langle y_k', y\rangle$$

is continuous. Such bilinear forms are called <u>nuclear</u>.

Exercise 11. Let X be a Hausdorff LCS. Show that X is metrizable if and only if its topology is induced by a countable number of semi-norms.

Exercise 10. If X is a LCS and M a linear subspace, show that M′ is algebraically isomorphic to $X'/M^{\perp}$, where

$$M^{\perp} = \{x' \in X' : \langle x', m\rangle = 0 \ \forall m \in M\} .$$

Exercise 11. Let X be a LCS and M a linear subspace. Show that X/M is a LCS under the quotient topology. If X is metrizable and M is closed, show X/M is metrizable.

Exercise 12. Let X, Y be Hausdorff LCS and $T : X \to Y$ a linear, continuous, open mapping onto Y. If X is a complete metrizable space, show that Y is also.

Exercise 13. Show that the product of LCS is a LCS. Describe the semi-norms which generate the product topology.

Exercise 14. If B is a bounded subset of a LCS, show $\overline{\text{co}}B$ is bounded.

Exercise 15. Let E be a vector space and let $\mathcal{U}$ be the family of all absolutely convex, absorbing subsets of E. Show $\mathcal{U}$ is a base for a locally convex topology τ on E; τ is called the strongest locally covex topology of E. Show:

(i) τ is Hausdorff. [Hint: Use a Hamel basis.]

(ii) τ is generated by the family of all semi-norms on E.

(iii) If F is a LCTVS and $T : E \to F$ is linear, show T is continuous with respect to τ.

(iv) Every τ bounded set is finite dimensional.

(v) Every linear subspace is τ closed.

(vi) If E is infinite dimensional, τ is not metrizable.

13.1 Normability

We now have the machinery available to characterize normable TVS. It turns out that the simplest necessary condition is also sufficient.

Theorem 1 (Kolmogorov [K]). Let X be a Hausdorff TVS. The topology of X is induced by a norm (i.e., is normable) if and only if X contains a bounded, convex open neighborhood of 0.

Proof: $\Rightarrow$: Clear.

$\Leftarrow$: Let U be a bounded, open convex neighborhood of 0. Then $\exists$ an absolutely convex open neighborhood of 0, V, such that $V \subseteq U$ (13.4). V is clearly bounded. Let p be the Minkowski functional of V.

Now p is a norm; for if $x \neq 0$, then since X is Hausdorff $\exists$ a balanced neighborhood W of 0 such that $x \notin W$. Since V is bounded, $\exists\ t > 0$ such that $V \subseteq tW$. There exists $s > p(x)$ so that $x \in sV$. Then $x/s \in tW$ or $x \in stW$. Since W is balanced and $x \notin W$, $st > 1$ or $s > 1/t$ and $p(x) \geq 1/t > 0$.

Next we claim that the topology induced by p (and denoted by p) coincides with the original topology, τ, of X. Since V is bounded, if W is a τ-neighborhood of $0\ \exists\ \varepsilon > 0$ such that $\varepsilon V \subseteq W$. Hence, $\{\varepsilon V : \varepsilon > 0\}$ is a base at 0 for τ. Since $V = \{x : p(x) < 1\}$, $\{\varepsilon V : \varepsilon > 0\}$ is also a base at 0 for p and $p = \tau$.

Corollary 2. A bounded set in a non-normable Hausdorff LCS X is nowhere dense.

Proof: Suppose B is bounded and $x_0 \in \text{int}(\bar{B})$. $\exists$ an open convex neighborhood of 0, U, such that $x_0 + U \subseteq \bar{B}$. Then $U \subseteq -x_0 + \bar{B}$ implies that U is bounded and X is normable by Theorem 1.

Corollary 3. A complete, quasi-normed Hausdorff LCS X is normable if and only if X is a countable union of bounded sets.

Proof: $\Rightarrow$: Clear.

$\Leftarrow$: If $X = \bigcup_{n=1}^{\infty} B_n$ with B_n bounded, some B_n is not nowhere dense by the Baire Category Theorem. X is normable by Corollary 2.

13.2 Krein-Milman Theorem

Let X be a vector space and $K \subseteq X$. A subset $E \subseteq K$ is an <u>extremal subset</u> of K if $E \neq \emptyset$ and whenever $x, y \in K$ and $0 < t < 1$ implies $tx + (1 - t)y \in E$, then $x, y \in E$. An <u>extremal point</u> (or <u>extreme point</u>) of K is an extremal subset consisting of the point. Thus, a point $x \in K$ is an extreme point of K if $x = tx_1 + (1 - t)x_2$ with $x_i \in K$, $0 \leq t \leq 1$, implies $x_1 = x_2 = x$. It is easy to see that if E_1 is an extremal subset of an extremal subset E_2 of K, then E_1 is an extremal subset of K.

Example 1. In $\mathbb{R}^2$, the extreme points of $\{(x, y) : \|(x, y)\|_1 \leq 1\}$ are $(\pm 1, 0)$ and $(0, \pm 1)$; the extreme points of $\{(x, y) : \|x, y)\|_2 \leq 1\}$ are $\{(x, y) : \|(x, y)\|_2 = 1\}$.

In this example the sets are the closed unit balls of NLS. All such balls don't have extreme points.

Example 2. Let $K = \{\{t_k\} \in c_0 : \|\{t_k\}\| \leq 1\}$. Then K has no extreme points. It is clear that an interior point of any convex subset of a TVS cannot be an extreme point so let $\{t_k\} \in K$ with $\|\{t_k\}\| = 1$.

Select any t_j with $|t_j| < 1/2$. Define $x = \{x_k\}, y = \{y_k\} \in K$ by $y_k = x_k = t_k$ if $k \neq j$ and $x_j = t_j + 1/2$, $y_j = t_j - 1/2$. Then $\frac{1}{2}x + \frac{1}{2}y = \{t_j\}$ and $x \neq y$.

Further examples of extreme points in function spaces are given in the exercises.

Note in Example 1 that the compact convex sets considered are the convex hulls of their extreme points. The Krein-Milman Theorem gives a general relation between a compact convex set and the convex hull of its extreme points.

Lemma 3. Let $K \neq \emptyset$ be a compact subset of a Hausdorff TVS X. For $x' \in X'$ let $a = \sup\{ \mathcal{R}\langle x', x\rangle : x \in K\}$. Then

$$E = \{x \in K : \mathcal{R}\langle x', x\rangle = a\}$$

is an non-empty, compact, extremal subset of K.

Proof: Since K is compact, E is non-empty and compact. Suppose $x, y \in K$, $0 < t < 1$ and $tx + (1 - t)y \in E$. Then

$$a = \mathcal{R}\langle x', tx + (1 - t)y\rangle = t\,\mathcal{R}\langle x', x\rangle + (1 - t)\,\mathcal{R}\langle x', y\rangle$$

and since $\mathcal{R}\langle x', x\rangle \leq a$, $\mathcal{R}\langle x', y\rangle \leq a$, we must have $\mathcal{R}\langle x', x\rangle = \mathcal{R}\langle x', y\rangle = a$ and $x, y \in E$.

Theorem 4. Let K be a non-empty compact subset of a Hausdorff LCS X. Then K has at least one extreme point.

Proof: Let $\mathcal{E}$ be the family of all non-empty, compact extremal subsets of K. Then $\mathcal{E} \neq \emptyset$ since $K \in \mathcal{E}$. Partially order $\mathcal{E}$ by reverse set inclusion; $E_1 \leq E_2$ if and only if $E_1 \supseteq E_2$. If $\mathcal{C}$ is any chain in $\mathcal{E}$, then $\cap\,\mathcal{C} = E$ is non-empty since K is compact and E is an upper bound for $\mathcal{C}$. By Zorn's Lemma $\mathcal{E}$ has a maximal element E_0. We claim that E_0 must be a singleton, establishing the result. Since E_0 cannot contain a proper non-empty compact extremal subset, it follows from Lemma 3 that

$\mathcal{R}x'$ must be constant on E_0 $\forall x' \in X'$. By 13.10, E_0 must be a singleton.

Theorem 5 (Krein-Milman). Let K be a non-empty compact subset of a Hausdorff LCS X. Then K is contained in the closed convex hull of its extreme points.

Proof: Let H be the closed convex hull of the extreme points of K. Suppose $\exists\, x \in K \backslash H$. By 13.13 $\exists\, x' \in X'$ such that

$$\sup\{\, \mathcal{R}\langle x', y\rangle : y \in H\} < \mathcal{R}\langle x', x\rangle .$$

If $a = \sup\{\, \mathcal{R}\langle x', y\rangle : y \in K\}$ and $E = \{y \in K : \mathcal{R}\langle x', y\rangle = a\}$, $H \cap E = \emptyset$. Now E is non-empty and compact so E has an extreme point x_0 by Theorem 4. E is an extremal subset of K by Lemma 3 so x_0 is an extreme point of K. Since $H \cap E = \emptyset$, this is impossible. Hence, $K \subseteq H$.

Corollary 6. If K is a non-empty, convex compact subset of a Hausdorff LCS X, then K is the closed convex hull of its extreme points.

Despite its somewhat esoteric appearance, the Krein-Milman Theorem has a surprising number of applications. We give such applications to the existence of preduals for NLS, Liapounoff's theorem on the convexity of the range of a vector measure and the Stone-Weierstrass Theorem in §15.

Another interesting application of the Krein-Milman Theorem which we will not give is to a result of Banach and Stone which we now describe. If S and S' are compact Hausdorff spaces such that $C(S)$ and $C(S')$ are linearly isometric, then the Banach-Stone Theorem asserts that S and S' are homeomorphic; for a proof use [DS] V.8.8.

Exercise 1. Show the set of extreme points of a compact convex set needn't be compact. [Hint: In $\mathbb{R}^3$, consider the convex hull of $(1, 0, \pm 1)$ and $(\cos t, \sin t, 0)$ $0 \le t \le 2\pi$.]

Exercise 2. Show the closed unit ball of $L^1[0, 1]$ has no extreme points. [Hint: If $\|f\|_1 = 1$, pick c such that $\int_0^c |f| = 1/2$ and consider $[0, c)$, $[c, 1]$.]

Exercise 3. Show $f \in L^\infty[0, 1]$ is an extreme point of the closed unit ball if and only if $f(t) = 1$ a.e.

Exercise 4. Show compact cannot be replaced by closed and bounded in Theorem 4.

Exercise 5. Show $\{te_k : |t| = 1, k \in \mathbb{N}\}$ are the extreme points of the closed unit ball of ℓ^1.

Exercise 6. If S is a compact Hausdorff space, show the extreme points of the closed unit ball of $C(S)$ consists of those functions with $|f(t)| = 1$

$\forall\, t \in S$.

Exercise 7. Show the extreme points of the closed unit ball of $L^p[0, 1]$ consists of those functions with $\|f\|_p = 1$ provided $1 < p < \infty$.

Exercise 8. Show the closed unit ball of c has two extreme points, $x = (1, 1, ...)$ and $-x$.

14
Duality and Weak Topologies

We next consider an abstract construction that can be used to define important locally convex topologies, called weak topologies, on vector spaces.

If X is a vector space, its algebraic dual, the set of all linear functionals on X, will be denoted by $X^{\#}$.

Definition 1. Let X, Y be vector spaces (over $\mathbb{F}$) and $b : X \times Y \to \mathbb{F}$ bilinear. Then X and Y are said to be in <u>duality</u> (with respect to b) if

(i) $\{b(\cdot, y) : y \in Y, y \neq 0\}$ separates the points of X,

(ii) $\{b(x, \cdot) : x \in X, x \neq 0\}$ separates the points of Y.

Example 2. Let X be a vector space. Then X and $X^{\#}$ are in duality with respect to the natural bilinear mapping, $b(x, y) = y(x) = \langle y, x \rangle$, between X and $X^{\#}$. Actually if Y is any linear subspace of $X^{\#}$ which separates the points of X, then X and Y are in duality with respect to $< , >$.

Example 3. If X is a Hausdorff LCS, then X and X' are in duality as in Example 2 (13.10).

Suppose that X and Y are in duality. For each $y \in Y$, $y^{\#} = b(\cdot, y)$ defines an element of $X^{\#}$ and from (ii) the map $y \to y^{\#}$ defines an isomorphism of Y into $X^{\#}$. Therefore, we may consider Y to be a linear subspace of $X^{\#}$ which separates the points of X with the bilinear mapping between X and Y to be the natural bilinear pairing given in Example 2. Henceforth, we adopt this notational convenience.

Let X and Y be in duality. To each $y \in Y$, there is associated a semi-norm p_y on X defined by $p_y(x) = |<y, x>|$. The locally convex topology $\sigma(X, \{p_y : y \in Y\})$ is called the <u>weak topology</u> on X (from the duality between X and Y) and is denoted by <u>$\sigma(X, Y)$</u>. We have the following properties of the weak topology (13.3).

Proposition 4. (i) $\sigma(X, Y)$ is the weakest topology on X such that all the elements of Y are continuous.

(ii) A base at 0 consists of $\{x : |<y_i, x>| < \varepsilon, i = 1, ..., k\}$, where $y_1, ..., y_k \in Y$ and $\varepsilon > 0$.

(iii) A net $\{x_\delta\}$ in X converges to 0 in $\sigma(X, Y)$ if and only if $<y, x_\delta> \to 0 \ \forall y \in Y$.

(iv) $\sigma(X, Y)$ is Hausdorff (Definition 1.(i)).

(v) A sequence $\{x_k\}$ in X is $\sigma(X, Y)$ Cauchy if and only if

$\forall\, y \in Y\ \{<y, x_k>\}$ is Cauchy.

It follows from Proposition 4 (i) that the dual of $(X, \sigma(X, Y))$ contains Y; we now show that the dual is exactly Y.

Lemma 5. If $g \neq 0$ and $f_1, \ldots, f_k$ are linear functionals on the vector space X, then either g is a linear combination of $\{f_1, \ldots, f_k\}$ or else $\exists\, x_0 \in X$ such that $<g, x_0> = 1$ and $<f_i, x_0> = 0$ for $i = 1, \ldots, k$.

Proof: We may assume that $\{f_1, \ldots, f_k\}$ is linearly independent. For $k = 0$ the result is trivial; assume that it holds for $k \leq n - 1$. Then for each i, $1 \leq i \leq n$, f_i is not a linear combination of $\{f_j : 1 \leq j \leq n, j \neq i\}$ so by the induction hypothesis $\exists\, a_i \in X$ such that $<f_j, a_i> = \delta_{ij}$, $i = 1, \ldots, n$. For each $x \in X$, $x - \sum_{i=1}^{n} <f_i, x>a_i \in K = \bigcap_{i=1}^{n} \ker f_i$. Then either $\exists\, a \in K$ such that $<g, a> = 1$ (and $<f_i, a> = 0$ $i = 1, \ldots, n$) or $<g, a> = 0\ \forall\, a \in K$. In the latter case, $\forall\, x \in X$ $<g, x> = \sum_{i=1}^{n} <f_i, x><g, a_i>$ and $g = \sum_{i=1}^{n} <g, a_i>f_i$.

Theorem 6. Let X and Y be in duality. Then the dual of $(X, \sigma(X, Y))$ is Y.

Proof: Let f be a linear functional on X which is continuous with respect to $\sigma(X, Y)$. Then $|<f, x>| \leq t < 1$ on some neighborhood, U, of 0 of the form $U = \{x : |<y_i, x>| \leq \varepsilon, i = 1, \ldots, n\}$, $y_i \in Y$, $\varepsilon > 0$. By

Lemma 5, either f is a linear combination of $y_1, \ldots, y_n$ or $\exists\, a \in X$ such that $\langle f, a \rangle = 1$ and $|\langle y_i, a \rangle| = 0$, $i = 1, \ldots, n$. But in the latter case, $a \in U$ and $\langle f, a \rangle = 1 > t$. Hence, $f \in Y$.

If X and Y are in duality, then certainly Y and X are in duality so $(Y, \sigma(Y, X))' = X$ by Theorem 6.

Example 7. If X is a Hausdorff LCS, then $(X, \sigma(X, X'))' = X'$ and $(X', \sigma(X', X))' = X$. In this situation, $\sigma(X, X')$ is called the *weak topology* of X and $\sigma(X', X)$ is called the *weak* topology* of X'.

Definition 8. Let X and Y be in duality. A locally convex Hausdorff topology τ on X is said to be *compatible* with the duality between X and Y if $(X, \tau)' = Y$.

By Theorem 6, $\sigma(X, Y)$ is such that topology, and it is the weakest such topology.

Example 9. Let X be a non-reflexive B-space. Then X' and X are in duality. On X the topology $\sigma(X, X')$ and the norm topology are compatible with respect to this duality. On X' the weak* topology $\sigma(X', X)$ is compatible with the duality between X' and X, but the norm topology and the weak topology $\sigma(X', X'')$ on X' are not compatible with this duality since their duals are $X'' \neq X$.

From 13.14 we have the following important property of the closures of convex sets with respect to the weak topology.

Theorem 10. Let X and Y be in duality. If $K \subseteq X$ is convex, then the closure of K is the same for any locally convex topology which is compatible for the duality between X and Y.

In particular, if (X, τ) is a Hausdorff LCS, then any convex subset of X is τ-closed if and only if it is $\sigma(X, X')$ closed.

Continuity of Linear Operators with respect to Weak Topologies:

Theorem 11. Let X and Y be Hausdorff LCS and $T : X \to Y$ linear. If T is continuous with respect to the original topologies of X and Y, then T is continuous with respect to $\sigma(X, X')$ and $\sigma(Y, Y')$.

Proof: Let $\{x_\delta\}$ be a net in X which is $\sigma(X, X')$ convergent to 0. If $y' \in Y'$, then $y'T \in X'$ so $\langle y', Tx_\delta \rangle \to 0$ or $\{Tx_\delta\}$ is $\sigma(Y, Y')$ convergent to 0.

The converse of Theorem 11 is false. To show this we need a preliminary observation.

Lemma 12. Let X and Y be in duality. If X is infinite dimensional, every $\sigma(X, Y)$ neighborhood of 0 contains an infinite dimensional subspace of X.

Proof: Let $U = \{x : |\langle y_i, x\rangle| \le \varepsilon, i = 1, ..., n\}$, $y_i \in Y$, $\varepsilon > 0$, be a basic $\sigma(X, Y)$ neighborhood of 0. Then $L = \bigcap_{i=1}^{n} \ker y_i$ is a linear subspace of co-dimension at most n which is contained in U.

Corollary 13. If X is an infinite dimensional NLS, then the weak topology of X, $\sigma(X, X')$, is strictly weaker than the norm topology of X and is not normable.

We can now show that the converse of Theorem 11 is false.

Example 14. Let Y be an infinite dimensional NLS. Take X to be Y equipped with the weak topology $\sigma(Y, Y')$. Then the identity map from X onto Y is $\sigma(X, X')$ - $\sigma(Y, Y')$ continuous but is not continuous with respect to the original topologies of X and Y by Corollary 13.

We can establish a partial converse to Theorem 11. For this we need the following important observation concerning bounded sets.

Theorem 15 (Mackey). Let (X, τ) be a Hausdorff LCS. A subset $B \subseteq X$ is τ bounded if and only if it is $\sigma(X, X')$ bounded.

Proof: 13.16.

Thus, any two locally convex topologies which are compatible with the duality between X and X′ have the same bounded sets. However, two locally convex topologies can have the same bounded sets and not be compatible (see 16.8).

Proposition 16. Let X, Y be Hausdorff LCS and $T : X \to Y$ linear. If T is $\sigma(X, X') - \sigma(Y, Y')$ continuous, then T is bounded with respect to the original topologies of X and Y.

Proof: T is bounded with respect to the weak topologies of X and Y (5.3) so the result follows from Theorem 15.

Corollary 17. Let X, Y be Hausdorff LCS and $T : X \to Y$ linear. If X is quasi-normed and T is $\sigma(X, X') - \sigma(Y, Y')$ continuous, then T is continuous with respect to the original topologies of X and Y.

Proof: Proposition 16 and 5.4.

We will give a generalization of this result to more general domain spaces in §21.

We can now give an example of a bounded linear operator which is not continuous (see 5.3).

Example 18. Let X be an infinite dimensional NLS. The identity operator from $(X, \sigma(X, X'))$ onto $(X, \| \|)$ is bounded (Theorem 15) but not continuous (Corollary 13).

$\mathcal{A}$-spaces which are not $\mathcal{K}$-spaces:

The results above can be used to give examples of $\mathcal{A}$-spaces which are not $\mathcal{K}$-spaces (§4). For this we require a preliminary observation.

Lemma 19. Let X and Y be in duality, and let σ, τ be two locally convex topologies on X which are compatible with the duality between X and Y and assume $\sigma \subseteq \tau$. If (X, τ) is an $\mathcal{A}$-space, then (X, σ) is an $\mathcal{A}$-space.

Proof: The identity map from (X, τ) onto (X, σ) is continuous so the result follows from Theorem 15 and Exer. 5.13.

Example 20. If X is a B-space, then $(X, \sigma(X, X'))$ is an $\mathcal{A}$-space by Lemma 19. In particular, if $X = \ell^p$, $1 < p < \infty$, then $(X, \sigma(X, X'))$ is an $\mathcal{A}$-space but is not a $\mathcal{K}$-space [consider the unit vectors e_k].

Exercise 1. If τ is a locally convex topology lying between two compatible topologies, show τ is compatible.

Exercise 2. If X is finite dimensional and X and Y are in duality, show Y is finite dimensional and $\dim X = \dim Y$.

Exercise 3. Let X be a Hausdorff LCS and L a linear subspace of X'. Show L is $\sigma(X', X)$ dense in X' if and only if L separates the points of X.

Exercise 4. Let X be a NLS. Show $\sigma(X, X')$ is metrizable if and only if $\sigma(X, X')$ is normable if and only if X is finite dimensional.

Exercise 5. Let X and Y be in duality and M a linear subspace of Y which separates the points of X with $M \neq Y$. Show $\sigma(X, Y)$ is strictly stronger than $\sigma(X, M)$.

Exercise 6. Let X be a Hausdorff LCS and $M \subseteq X$ a linear subspace. Show $\sigma(M, M')$ and the relative $\sigma(X, X')$ topology on M coincide.

Exercise 7. Let X be a B-space. When is $\sigma(X', X) = \sigma(X', X'')$?

Exercise 8. Compare the metric and weak topologies of s.

Exercise 9. Does the analogue of Corollary 13 hold for quasi-normed spaces?

Exercise 10. In $(\ell^2, \sigma(\ell^2, \ell^2))$ show $\{e_k\}$ converges to 0 but is not Mackey convergent to 0.

Exercise 11. Let X be a Hausdorff LCS and M a linear subspace. Show that $\sigma(X/M, (X/M)')$ is $\sigma(X, X')|_M$ (see Exer. 6.4).

Exercise 12. Let E be a vector space and $E^{\#}$ its algebraic dual. Show that $E^{\#}$ is complete under $\sigma(E^{\#}, E)$.

15

The Bipolar and Banach-Alaoglu Theorems

In this section we establish two very important results for weak topologies, the Bipolar Theorem and the Banach-Alaoglu Theorem. We use these results along with the Krein-Milman Theorem to prove the Liapounoff Theorem on the range of a vector-valued measure and a generalization of the Stone- Weierstrass Theorem due to de Branges.

Let X and Y be in duality.

Definition 1. Let $A \subseteq X$. The polar of A (with respect to the duality between X and Y), A^0, is defined by $\{y \in Y : |\langle y, x\rangle| \le 1 \ \forall x \in A\}$. Let $B \subseteq Y$. The polar of B, B_0, is defined by

$$\{x \in X : |\langle y, x\rangle| \le 1 \quad \forall y \in B\}.$$

The notation is designed to help indicate the space in which the polars are being computed. If X and Y are in duality, we think of X as being the "ground space" and Y the "dual" or "upstairs space" as in the diagram below

$$\begin{array}{ccc} Y & A^0 & B \\ X & A & B_0 \end{array}$$

If A is a subset of the ground space, A^0 is computed in the upstairs space; if B is a subset of the upstairs space, B_0 is computed in the ground space.

Example 2. Let X be a NLS. Consider the duality between X and X'. If $S = \{x : \|x\| \le 1\}$, then $S^0 = \{x' \in X' : \|x'\| \le 1\}$; if

$$S' = \{x' \in X' : \|x'| \le 1\} ,$$

then $(S')_0 = S$. Note in this example that $(S^0)_0 = S$ and $((S')_0)^0 = S'$. The Bipolar Theorem gives a general form of this phenomena.

Polars satisfy the following elementary properties.

Proposition 3. Let $A, B \subseteq X$.

(i) If $A \subseteq B$, then $B^0 \supseteq A^0$.

(ii) $A \subseteq (A^0)_0$ and $A^0 = ((A^0)_0)^0$.

(iii) A^0 is absolutely convex and $\sigma(Y, X)$ closed.

(iv) If $t \in \mathbb{F}$, $t \ne 0$, then $(tA)^0 = (1/t)A^0$.

(v) If H is the $\sigma(X, Y)$ closed, absolutely convex hull of A, then $H^0 = A^0$.

Proof: (i) is clear.

(ii): If $a \in A$, then for $a' \in A^0$, $|\langle a', a\rangle| \le 1$ so $a \in (A^0)_0$ and

$A \subseteq (A^0)_0$. By the first part $((A^0)_0)^0 \supseteq A^0$ and by (i) $A^0 \supseteq ((A^0)_0)^0$.

(iii): The absolute convexity is clear and

$$A^0 = \bigcap_{x \in A} \{x' \in Y : |\langle x', x\rangle| \le 1\}$$

implies A^0 is $\sigma(Y, X)$ closed since each $x \in A$ is $\sigma(Y, X)$ continuous.

(iv): Let $x' \in (1/t)A^0$. For $x \in tA$, $x = ta$ for some $a \in A$. Also $x' = (1/t)y'$ with $y' \in A^0$. Hence, $|\langle x', x\rangle| = |\langle y', a\rangle| \le 1$ and $x' \in (tA)^0$. Hence, $(1/t)A^0 \subseteq (tA)^0$. Now $(tA)^0 = (1/t)t(tA)^0 \subseteq 1/t(A^0)$.

(v): From (iii) and (ii), $A \subseteq H \subseteq (A^0)_0$; by (i) and (ii), $A^0 \supseteq H^0 \supseteq ((A^0)_0)^0 = A^0$.

Proposition 4. $A \subseteq X$ is $\sigma(X, Y)$ bounded if and only if A^0 is absorbing in Y.

Proof: Suppose A is $\sigma(X, Y)$ bounded. Let $x' \in Y$. Then $t = \sup\{|\langle x', x\rangle| : x \in A\} < \infty$. If $t = 0$, $x' \in A^0$; if $t > 0$, then for $|s| \le 1/t$, $sx' \in A^0$ so A^0 is absorbing.

Suppose A^0 is absorbing. Let $x' \in Y$. Then $\exists\, t > 0$ such that $x' \in tA^0$. Then $|\langle x', x\rangle| \le t\ \forall x \in A$ and A is $\sigma(X, Y)$ bounded.

Theorem 5 (Bipolar Theorem). Let X and Y be in duality. For $A \subseteq X$, $(A^0)_0$ is the $\sigma(X, Y)$ closure of the absolutely convex hull of A. In particular, if A is $\sigma(X, Y)$ closed and absolutely convex, then $A = (A^0)_0$.

Proof: Let H be the $\sigma(X, Y)$ closure of abco$A$. By Proposition 3 $(A^0)_0 \supseteq H$. Suppose $\exists\, x \in (A^0)_0 \backslash H$. By Lemma 13.13, $\exists\, x' \in Y$ such that $\langle x', x\rangle > 1$ and $|\langle x', a\rangle| \le 1 \;\forall\, a \in H$. Since $H \supseteq A$, $x' \in A^0$ so $x \notin (A^0)_0$.

We now prove an important compactness result for weak topologies, the Banach-Alaoglu Theorem.

Lemma 6. Let X be a vector space and Y a linear subspace of $X^{\#}$ which separates the points of X. Then $Y \subseteq \prod_X \mathbb{F} = \mathbb{F}^X$ and the topology $\sigma(Y, X)$ and the induced product topology on Y from $\prod_X \mathbb{F}$ coincide.

Proof: A net $\{y_\delta\}$ in Y converges to $y \in Y$ in either topology if and only if $\langle y_\delta, x\rangle \to \langle y, x\rangle \;\forall\, x \in X$.

Theorem 7 (Banach-Alaoglu). Let X and Y be in duality. Suppose τ is a compatible topology on X. If U is a τ neighborhood of 0 in X, then $U^0 \subseteq Y$ is $\sigma(Y, X)$ compact.

Proof: By Proposition 3 we may assume that U is $\sigma(X, Y)$ closed and absolutely convex. By 14.10, U is also τ-closed.

Let p be the Minkowski functional of U. We claim that $U^0 = \{x' \in Y : |\langle x', x\rangle| \le p(x) \;\forall\, x \in X\}$. For if $x' \in Y$ is such that $|\langle x', x\rangle| \le p(x) \;\forall\, x \in X$, then for $x \in U$ $|\langle x', x\rangle| \le p(x) \le 1$ (11.13), so $x' \in U^0$. On the other hand, if $x' \in U^0$ and $x \in X$ with $p(x) > 0$, then $x/p(x) \in U$ (11.13) so $|\langle x', x\rangle| \le p(x)$, and if $p(x) = 0$, then $nx \in U\ \forall\, n \ge 1$ so $|\langle x', x\rangle| \le 1/n$ and $|\langle x', x\rangle| = 0 = p(x)$.

For $x \in X$, set $D(x) = \{t \in \mathbb{F} : |t| \le p(x)\}$. From above $U^0 \subseteq \prod_{x\in X} D(x)$ and from Lemma 6 the $\sigma(Y, X)$ topology on U^0 is the relative product topology on U^0 from the product $\prod_X \mathbb{F}$. From Tychonoff's Theorem U^0 is relatively compact in the product topology so it suffices to show that U^0 is closed in the product topology. Let $\{x'_\delta\}$ be a net in U^0 such that $x'_\delta \to f \in \prod \mathbb{F}$ in the product topology. f is clearly linear and

$$|\langle f, x\rangle| \le p(x) \quad \forall\, x \in X$$

so f is also τ continuous since p is τ continuous (11.13). Hence, $f \in U^0$ and U^0 is closed in the product topology.

Corollary 8. Let X be a NLS. The closed unit ball,

$$S' = \{x' \in X' : \|x'\| \le 1\},$$

of X' is weak* ($\sigma(X', X)$) compact.

Proof: Theorem 7 and Example 2.

From 8.1.6, we obtain

Corollary 9. Let X be a NLS. Then X is linearly isometric to a subspace of $C(S')$, where $S' = \{x' : \|x'\| \le 1\}$ has the weak* topology.

Corollary 8 is false for the unit ball of a NLS (or even a B-space) and the weak topology. In fact, the unit ball of c_0 is not compact for <u>any</u> locally convex topology on c_0. For if this were the case, the closed unit ball of c_0 would have extreme points by the Krein-Milman Theorem. In Example 13.2.2 it was shown that the closed unit ball of c_0 has no extreme points.

In 5.8 it was shown that the dual of a NLS is a B-space under the dual norm. It is reasonable to ask if every B-space is the dual of some NLS, i.e., does every B-space have a <u>predual</u>. From the discussion above it follows that c_0 is not the dual of any NLS for if this were the case, the unit ball of c_0 would be weak* compact and have extreme points.

We give another interesting and important application of the Banach-Alaoglu and Krein-Milman Theorems to vector-valued measures. Let Σ be a σ-algebra of subsets of a set S. A real-valued measure λ on Σ is <u>non-atomic</u> if every $E \in \Sigma$ with $0 < |\lambda|(E) = \text{var}(\lambda)(E)$ contains a subset $A \in \Sigma$ with $0 < |\lambda|(A) < |\lambda|(E)$. For example, Lebesgue measure is non-atomic. Concerning the range of a non-atomic measure with values in a finite dimensional space, we have the following result of Liapounoff.

Theorem 10 (Liapounoff). Let $\mu_1, \ldots, \mu_n : \Sigma \to \mathbb{R}$ be countably additive and non-atomic. If $\mu : \Sigma \to \mathbb{R}^n$ is given by $\mu(E) = (\mu_1(E), \ldots, \mu_n(E))$, then $\mu(\Sigma)$ is a convex, compact subset of $\mathbb{R}^n$.

Proof: Let $\lambda = \sum_{i=1}^{n} |\mu_i|$ and note that λ is non-atomic. Define

$T : L^\infty(\lambda) \to \mathbb{R}^n$ by $Tg = (\int_S g d\mu_1, \ldots, \int_S g d\mu_n)$ (note T is well-defined since each μ_i is absolutely continuous with respect to λ). By the Radon-Nikodym Theorem, T is continuous with respect to the weak* topology of $L^\infty(\lambda)$ and the norm topology of $\mathbb{R}^n$.

Put $K = \{g \in L^\infty(\lambda) : 0 \le g \le 1\}$. Now K is convex and is weak* closed since $g \in K$ if and only if $0 \le \int_S fg d\lambda \le \int_S f d\lambda \ \forall\ f \ge 0,\ f \in L^1(\lambda)$. Since K is contained in the closed unit ball of $L^\infty(\lambda)$, K is weak* compact by the Banach-Alaoglu Theorem. Thus, TK is compact in $\mathbb{R}^n$.

We claim that $\mu(\Sigma) = TK$ establishing the result. If $E \in \Sigma$, then $C_E \in K$ so $T(C_E) = \mu(E) \in TK$, and $\mu(\Sigma) \subseteq TK$. Suppose $y \in TK$ and set $K_y = T^{-1}(y)$. We must show that K_y contains the characteristic function of a set in Σ. But K_y is weak* closed so weak* compact and is convex. Therefore, by the Krein-Milman Theorem K_y has an extreme point. We claim that such an extreme point must be a characteristic function. For suppose $g \in K_y$ is not a characteristic function. Then $\exists\ E \in \Sigma$ and $\varepsilon > 0$ such that $\lambda(E) > 0$ and $\varepsilon \le g \le 1 - \varepsilon$ on E. If $L = C_E L^\infty(\lambda)$, then since $\lambda(E) > 0$ and λ is non-atomic, L must be infinite dimensional (Exer. 4). Therefore, $\exists\ f \in L$, $f \ne 0$, such that $Tf = 0$ with $-\varepsilon < f < \varepsilon$. Then $g \pm f \in K_y$, $\pm f \ne g$ and $g = \frac{1}{2}(g + f) + \frac{1}{2}(g - f)$ implies that g is not an extreme point of K_y.

Liapounoff's Theorem does not hold in general for infinite dimensional B-spaces. Let Σ be the Lebesgue measurable subsets of $[0, 1]$ and let m be Lebesgue measure on $[0, 1]$. Define $\mu : \Sigma \to L^1(m)$

by $\mu(E) = C_E$. Then μ is countably additive and by the Dominated Convergence Theorem, $\mu(E)$ is closed. However, $\mu(\Sigma)$ is not convex [$1/2 \in co\mu(\Sigma)$, but if $E \in \Sigma$, then

$$\|1/2 - \mu(E)\|_1 = (m([0, 1]\backslash E) + m(E))/2 = 1/2],$$

and also $\mu(\Sigma)$ is not compact since if $E_n = \{t : \sin 2^n \pi t > 0\}$, then $\|\mu(E_j) - \mu(E_k)\|_1 = 1/4$ for $j \neq k$. Note that μ is non-atomic in the (weak) sense that $x'\mu$ is non-atomic $\forall\, x' \in L^\infty(m)$.

Liapounoff's Theorem has important applications in control theory so there is a great interest in obtaining infinite dimensional versions of the theorem. For such generalizations, see [DU]; for applications of the theorem to control theory, see [HL].

We give a further application of the Krein-Milman and Banach-Alaoglu Theorems to the Stone-Weierstrass Theorem. Let S be a compact Hausdorff space and let $C(S)$ be the space of all complex-valued continuous functions defined on S equipped with the sup-norm. The dual of $C(S)$ is the space, $rca(S)$, of all regular, complex Borel measures on S with the total variation norm, $\|\mu\| = var(\mu)(S) = |\mu|(S)$ (Example 5.16). The <u>support</u> of a measure $\mu \in rca(S)$ is the set of all $t \in S$ such that $|\mu|(U) > 0$ $\forall$ open neighborhood U of t. The support is a closed set. If the support of μ is empty, then $\forall\, t \in S$ $\exists$ an open neighborhood $U(= U_t)$ of t such that $|\mu|(U) = 0$ and since S is compact, S is a finite union of such neighborhoods; it follows that $\mu = 0$ if and only if the support of μ is empty.

Notice that if A is linear subspace of a LCS E, then its polar is $A^0 = \{x' \in E' : \langle x', x\rangle = 0\ \forall\, x \in A\}$.

Theorem 11 (de Branges). Let $\mathscr{A}$ be a subalgebra of C(S) and let $K = \{\mu \in \mathscr{A}^0 : \|\mu\| \le 1\}$. Let μ be an extreme point of K and let f be a real-valued function in $\mathscr{A}$ such that $0 < f < 1$. Then f is constant on the support of μ.

Proof: If $\mu = 0$, trivial. If $\mu \ne 0$, then $\|\mu\| = 1$. Define λ, $\nu \in rca(S)$ by $\nu(E) = \int_E f d\mu$, $\lambda(E) = \int_E (1 - f) d\mu$. Since $f \in \mathscr{A}$ and $\mathscr{A}$ is an algebra, $\nu, \lambda \in \mathscr{A}^0$, and since $0 < f < 1$, λ and ν are non-zero. Now $\mu = \|\nu\|(\nu/\|\nu\|) + \|\lambda\|(\lambda/\|\lambda\|)$ and

$$\|\nu\| + \|\lambda\| = \int_S f d|\mu| + \int_S (1 - f) d|\mu| = |\mu|(S) = 1$$

so μ is a convex combination of $\nu/\|\nu\|$ and $\lambda/\|\lambda\|$. Since μ is an extreme point of K, $\mu = \nu/\|\nu\| = \lambda/\|\lambda\|$ or $\nu = \|\nu\|\mu$. Hence, $\nu(E) = \int_E f d\mu = \int_E \|\nu\| d\mu$ for each Borel set E, and since f is continuous, $f = \|\nu\|$ on the support of μ.

Theorem 12 (Stone-Weierstrass). Let $\mathscr{A}$ be a closed subalgebra of C(S) satisfying

(i) $1 \in \mathscr{A}$,

(ii) $\mathscr{A}$ separates the points of S and

(iii) $f \in \mathscr{A} \Rightarrow \bar{f} \in \mathscr{A}$. Then $\mathscr{A} = C(S)$.

Proof: Let $K = \{\mu \in \mathscr{A}^0 : \|\mu\| \le 1\}$. It suffices to show that $K = \{0\}$. Now K is non-void, convex and compact in the weak* topology by the Banach- Alaoglu Theorem. Hence, K contains an extreme point

μ(13.2.4). We claim that the support of μ is a singleton. Suppose s, t belong to the support of μ. By (ii) and (iii), $\exists$ f $\in \mathcal{A}$ such that $0 < f < 1$ and $f(s) \neq f(t)$. This is impossible by Theorem 11.

Thus, $\mu = t\delta_s$, where $t \in \mathbb{C}$ and δ_s is the Dirac measure or point mass at $s \in S$. However, $\mu \in \mathcal{A}^0$ so $\langle\mu, f\rangle = tf(s) = 0 \ \forall f \in \mathcal{A}$. By (i), $t = 0$ or $\mu = 0$. By the Krein-Milman Theorem, $K = \{0\}$.

Exercise 1. Show $(\cup A_a)^0 = \cap A_a^0$.

Exercise 2. If $s, t > 0$, show $(1/s\ A)^0 + (1/t\ A)^0 = (1/(s + t)A)^0$.

Exercise 3. Does $L^1[0, 1]$ have a predual? (Hint: Exer. 13.2.2.)

Exercise 4. If λ is a non-atomic, positive measure, show $L^\infty(\lambda)$ is infinite dimensional. Can the non-atomic assumption be dropped?

Exercise 5. Show the non-atomic assumption in Theorem 10 cannot be dropped.

Exercise 6. Does c have a predual? (Hint: Exer. 13.2.8.)

16
Duality in NLS

In this section we examine properties of the weak and weak* topologies on NLS and their duals. These topologies have many special properties in NLS which are important in applications.

Let X be a NLS. Then X carries two natural topologies, the norm topology and the weak topology $\sigma(X, X')$ from the duality between X and X'. If X is infinite dimensional, then $\sigma(X, X')$ is always strictly weaker than the norm topology (14.13) and is never metrizable (Exer. 14.4).

The dual space X' carries three natural topologies, the dual norm topology, the weak* topology $\sigma(X', X)$ from the duality between X' and X and the weak topology $\sigma(X', X'')$ from the duality between X' and X''. We have $\sigma(X', X) \subseteq \sigma(X', X'') \subseteq \| \;\|$. As above, if X is infinite dimensional, $\sigma(X', X'')$ is strictly weaker than the norm topology. From 14.6, X is reflexive if and only if $\sigma(X', X) = \sigma(X', X'')$. Thus, if X is non- reflexive, $\sigma(X', X)$ is strictly weaker than $\sigma(X', X'')$.

The Closed Unit Ball in NLS:

Compactness of the closed unit ball in a NLS and its dual carries a great deal of information about the NLS. For example, we know that the closed unit ball of a NLS X is norm compact if and only if X is finite dimensional (7.9). On the other hand, we know from the Banach-Alaoglu Theorem that the closed unit ball in the dual of any NLS is always weak* compact (15.8). We now consider the compactness of the closed unit ball of a NLS in the weak topology. Throughout the discussion which follows let X be a NLS and let $J : X \to X''$ be the canonical imbedding of X into its second dual (§8.1), and let S, S', S'' be the closed unit balls of X, X' and X'', respectively.

Lemma 1. $S^0 = S' = (S'')_0 = (JS)_0$.

Proof: $S^0 = S' = (S'')_0$ by Example 15.2; $S^0 = (JS)_0$ by definition.

Theorem 2 (Goldstine). JS is weak* dense in S''. Indeed, S'' is the weak* closure of JS.

Proof: Since JS is absolutely convex, the Bipolar Theorem implies that $((JS)_0)^0$ is the $\sigma(X'', X')$ closure of JS. By Lemma 1, $S'' = (S')^0 = ((JS)_0)^0$.

Corollary 3. JX is $\sigma(X'', X')$ dense in X''.

Proof: The $\sigma(X'', X')$ closure of JX contains the unit ball of X''.

Lemma 4. The canonical imbedding $J : X \to X''$ is a linear homeomorphism with respect to the topologies $\sigma(X, X')$ and $\sigma(X'', X')$.

Proof: 14.4.(iii).

Concerning compactness of the closed unit ball of a NLS in the weak topology, we have

Theorem 5. A B-space X is reflexive if and only if S is $\sigma(X, X')$ compact.

Proof: If $JX = X''$, then J is a homeomorphism from S onto S'' when S has the relative $\sigma(X, X')$ topology and S'' has the relative $\sigma(X'', X')$ topology (Lemma 4). But S'' is $\sigma(X'', X')$ compact (15.8) so S is $\sigma(X, X')$ compact.

If S is $\sigma(X, X')$ compact, then JS is $\sigma(X'', X')$ compact by Lemma 4. In particular, JS is $\sigma(X'', X')$ closed so Goldstine's Theorem 2 implies that $JS = S''$. Hence, $JX = X''$ and X is reflexive.

Corollary 6. A B-space X is reflexive if and only if norm bounded sets of X are relatively $\sigma(X, X')$ compact.

Bounded Sets:

We now consider bounded sets in the norm, weak and weak* topologies. First, concerning the norm and weak topologies we have

Theorem 7. A subset of X is norm bounded if and only if it is $\sigma(X, X')$ bounded.

Proof: Mackey's Theorem (13.16).

For the duals of B-spaces, we have

Theorem 8. Let X be a $\mathscr{K}$-space. In X' the norm, weak and weak* bounded sets are the same.

Proof: Let $B \subseteq X'$ be weak* bounded. Then

$$\sup\{|\langle x', x\rangle| : x' \in B\} < \infty \quad \forall x \in X.$$

By the UBP 9.7, $\sup\{\|x'\| : x' \in B\} < \infty$.

Note, in particular, that Theorem 8 applies to B-spaces.

Without some completeness type assumption, Theorem 8 is false.

Example 9. Let $X = c_{00}$. Consider $B = \{ke_k : k \in \mathbb{N}\} \subseteq X' = \ell^1$. B is weak* bounded but is not norm bounded.

Example 9 shows that weak* bounded subsets of the dual of a NLS needn't be norm bounded. However, using the stronger notion of $\mathcal{K}$ boundedness, we show that weak*-$\mathcal{K}$ bounded sets are norm bounded.

Proposition 10. If $B \subseteq X'$ is $\sigma(X', X)$-$\mathcal{K}$ bounded, then B is norm bounded and, hence, norm-$\mathcal{K}$ bounded since X' is complete.

Proof: Let $\{x_j'\} \subseteq B$ and $\{t_j\}$ be a sequence of positive scalars which converges to 0. For each j pick $x_j \in X$, $\|x_j\| = 1$ such that

$$|<x_j', x_j>| \geq \|x_j'\| - 1/j \,.$$

Consider the matrix $M = [z_{ij}] = [<\sqrt{t_j}\, x_j', \sqrt{t_i}\, x_i>]$. Since $\|\sqrt{t_i}\, x_i\| \to 0$ and $\{\sqrt{t_j}\, x_j'\}$ is weak*-$\mathcal{K}$ convergent to 0, M is a $\mathcal{K}$-matrix. By the Basic Matrix Theorem, $\lim z_{jj} = \lim t_j <x_j', x_j> = 0$ so $\lim t_j \|x_j'\| = 0$, and B is norm bounded.

Metrizability of the Closed Unit Balls:

Although the weak topology of an infinite dimensional NLS is not metrizable the closed unit ball can be metrizable in the weak topology as we show in Theorem 13 below. We begin by considering the metrizability of the closed unit ball in a dual space with the weak* topology.

Theorem 11. The closed unit ball S' of X' is metrizable in the weak* topology if and only if X is separable.

Proof: Assume that X is separable and let $\{x_k\}$ be a countable dense subset of X. For $x', y' \in S'$ set

$$d(x', y') = \sum_{k=1}^{\infty} \frac{|<x'-y', x_k>|}{2^k(1+|<x'-y', x_k>|)}.$$

Then d is a metric on S' which induces a topology on S' which is weaker than the weak* topology (2.3). But, S' is a compact Hausdorff space under the weak* topology (15.8) so the metric topology is exactly the weak* topology on S'.

Conversely, assume that the weak* topology of S' is metrizable. Then $\exists$ a sequence $\{U_n\}$ of weak* neighborhoods of 0 in X' such that $\bigcap_{n=1}^{\infty} U_n = \{0\}$. We may assume that

$$U_n = \{x' \in X' : |<x', x>| < \varepsilon_n \ \forall \, x \in A_n\},$$

where $\varepsilon_n > 0$ and A_n is a finite subset of X. Put $A = \bigcup_{n=1}^{\infty} A_n$, and $X_1 = \overline{\text{span}} A$. It suffices to show $X_1 = X$. Suppose $x' \in X'$ is such that $<x', X_1> = 0$. Then $x' \in U_n$ $\forall n$ so $x' = 0$ and $X_1 = X$ by 8.1.3.

We have the following important corollary of Theorem 11 and the Banach-Alaoglu Theorem.

Corollary 12. Let X be separable. If $\{x_k'\}$ is a norm bounded sequence in X', then $\{x_k'\}$ has a weak* convergent subsequence.

Concerning the metrizability of the closed unit ball of X, we have

Theorem 13. The closed unit ball of S is metrizable in the weak topology of X if and only if X' is separable.

Proof: Assume X' is separable. By Theorem 11 the closed unit ball S'' of X'' is metrizable in the weak* topology. By Lemma 4 S is metrizable in the weak topology of X.

Conversely, assume the weak topology of S is metrizable. Then $\exists$ a sequence of weak neighborhoods of 0 in S,

$$U_n = \{x \in S : |\langle x', x\rangle| < \varepsilon_n, x' \in A_n\}$$

where $\varepsilon_n > 0$ and A_n is a finite subset of X', such that every weak neighborhood of 0 in S is contained in some U_n. Set $A = \bigcup_{n=1}^{\infty} A_n$ and $Y = \overline{\text{span}A}$. It suffices to show that $Y = X'$. Suppose $\exists\, x_0' \in X' \backslash Y$. Set $d = \inf\{\|x_0' - x'\| : x' \in Y\}$. By 8.1.2 $\exists\, x'' \in X''$ with $\|x''\| = 1/d$, $\langle x'', x_0'\rangle = 1$ and $\langle x'', Y\rangle = 0$. $V = \{x \in S : |\langle x_0', x\rangle| < d/2\}$ is a weak neighborhood of 0 in S so $V \not\supseteq U_n$ for some n. Since $dx'' \in S''$, by Goldstine's Theorem 2, $\exists\, x_1 \in S$ such that

$$|\langle dx'', x'\rangle - \langle x', x_1\rangle| = |\langle x', x_1\rangle| < \varepsilon_n \text{ for } x' \in A_n,$$

and

$$|\langle dx'', x_0'\rangle - \langle x_0', x_1\rangle| = |d - \langle x_0', x_1\rangle| < d/2 .$$

Hence, $|\langle x_0', x_1\rangle| > d/2$ and $|\langle x', x_1\rangle| < \varepsilon_n$ for $x' \in A_n$ so $x_1 \in U_n$ and $x_1 \notin V$ which is impossible.

Sequential Convergence:

As noted earlier the norm topology of an infinite dimensional NLS is always strictly stronger than the weak topology. Despite this fact, it follows from the summability theorems of Schur and Hahn (9.5.3) that a sequence in ℓ^1 is weakly convergent to 0 if and only if it is norm convergent to 0. More generally, we can rephrase Theorem 9.5.3 in terms of weak topologies in the form

Theorem 14 (Schur). Let $x_i \in \ell^1$ for $i \in \mathbb{N}$. The following are equivalent:

(i) $\{x_i\}$ is norm convergent in ℓ^1

(ii) $\{x_i\}$ is weakly $(\sigma(\ell^1, \ell^\infty))$ convergent in ℓ^1

(iii) $\{x_i\}$ is $\sigma(\ell^1, m_0)$ convergent in ℓ^1

(iv) $\{x_i\}$ is a Cauchy sequence with respect to $\sigma(\ell^1, m_0)$.

Proof: If $x_i = \{a_{ij}\}_{j=1}^\infty$ and $A = [a_{ij}]$, then condition (iv) is equivalent to the matrix A belonging to the class (m_0, c). The equivalence of (i)-(iv) now follows readily from 9.5.3.

The equivalence of (i) and (ii) is often referred to as Schur's Theorem; a NLS in which weakly convergent sequences are norm convergent is called a Schur space. This example is often used to show that sequences are inadequate to characterize a topology ([C2]).

We next use Schur's Theorem to obtain another remarkable result concerning sequential convergence in ba due to Phillips ([Ph]). This result is not used until §27 and can be skipped at this point if desired.

Theorem 15 (Phillips). Let $\mu_i \in ba$. If $\lim \mu_i(E) = 0$ $\forall E \subseteq \mathbb{N}$, then

$$\lim_i \sum_{j=1}^\infty |\mu_i(\{j\})| = 0.$$

Proof: By Schur's Theorem it suffices to show $\lim_i \sum_{j \in E} \mu_i(\{j\}) = 0$ $\forall E \subseteq \mathbb{N}$. If this fails, we may assume by passing to a subsequence if necessary, that $\exists\ \varepsilon > 0$ and an infinite set $E \subseteq \mathbb{N}$ such that $|\sum_{j \in E} \mu_i(\{j\})| > \varepsilon$. Arrange the elements of E into a sequence $\{p_j\}$ and set $\mu_i(p_j) = \mu_i(\{p_j\})$.

$\exists\ n_1$ such that $|\sum_{j=1}^{n_1} \mu_1(p_j)| > \varepsilon$. $\exists\ m_1$ such that

$$\sum_{j=1}^{n_1} |\mu_i(p_j)| < \varepsilon/2$$

for $i \geq m_1$. $\exists\ n_2 > n_1$ such that $|\sum_{j=1}^{n_2} \mu_{m_1}(p_j)| > \varepsilon$. Hence,

$$|\sum_{j=n_1+1}^{n_2} \mu_{m_1}(p_j)| \geq |\sum_{j=1}^{n_2} \mu_{m_1}(p_j)| - \sum_{j=1}^{n_1} |\mu_{m_1}(p_j)| > \varepsilon/2 .$$

Continuing produces increasing sequences $\{m_i\}$, $\{n_i\}$ satisfying

$$|\sum_{j=n_i+1}^{n_{i+1}} \mu_{m_i}(p_j)| > \varepsilon/2 .$$

Put $E_j = \{p_k : n_j + 1 \leq k \leq n_{j+1}\}$. Then $\{E_j\}$ is a pairwise disjoint sequence of subsets of $\mathbb{N}$ with $|\mu_{m_i}(E_i)| > \varepsilon/2$.

Consider the matrix $M = [\mu_{m_i}(E_j)]$. The rows and columns of M converge to 0 and if $\{r_j\}$ is any increasing sequence of positive integers, $\exists$ a subsequence $\{s_j\}$ of $\{r_j\}$ such that each μ_i is countably additive on the σ-algebra generated by $\{E_{s_j}\}$ (Drewnowski's Lemma 9.2.2). Thus,

$$\sum_{j=1}^{\infty} \mu_{m_i}(E_{s_j}) = \mu_{m_i}(\bigcup_{j=1}^{\infty} E_{s_j}) \to 0 .$$

Hence, M is a $\mathscr{K}$-matrix and by the Basic Matrix Theorem 9.2, $\mu_{m_i}(E_i) \to 0$. This contradicts the construction above.

We will give an application of Phillip's Theorem later to show that there is no continuous projection from ℓ^{∞} onto c_0 (27.5). This was the original application due to Phillips.

In general, a weakly convergent sequence in a NLS is not norm convergent. For example, the sequence $\{e_k\}$ in ℓ^p, $1 < p < \infty$, is weakly convergent to 0 but is not norm convergent. For $\mathscr{K}$ convergent sequences, we have

Proposition 16. If $\{x_j\} \subseteq X$ is $\sigma(X, X')$-$\mathscr{K}$ convergent to 0, then $\|x_j\| \to 0$.

Proof: By replacing X by the closed linear subspace spanned by $\{x_k\}$, we may assume that X is separable (Exer. 14.6). For each j pick $x_j' \in X'$, $\|x_j'\| = 1$, such that $\langle x_j', x_j \rangle = \|x_j\|$ (8.1.4). By Corollary 12, $\{x_j'\}$ has a subsequence $\{x_{k_j}'\}$ which is weak* convergent to some $x' \in X'$. Consider the matrix $M = [\langle x_{k_i}', x_{k_j} \rangle]$. By the weak* convergence of $\{x_{k_j}'\}$ and the weak-$\mathscr{K}$ convergence of $\{x_j\}$, M is a

$\mathcal{K}$-matrix. It follows from the Basic Matrix Theorem 9.2 that $\lim\langle x'_{k_j}, x_{k_j}\rangle = \lim\|x_{k_j}\| = 0$. Since the same argument can be applied to any subsequence of $\{x_j\}$, it follows that $\|x_j\| \to 0$.

The analogue of Proposition 16 is false for weak* topologies. For example, the sequence $\{e_k\}$ in ℓ^∞ is weak*-$\mathcal{K}$ convergent but is not norm convergent.

We next show that the conclusion of Proposition 16 can be improved to assert that a weakly-$\mathcal{K}$ convergent sequence in a NLS is actually norm-$\mathcal{K}$ convergent. For this, we need the following useful concept.

Definition 17. Let σ and τ be two vector topologies on the vector space E. We say that σ is <u>linked</u> to τ if σ has a neighborhood base at 0 which consists of τ closed sets.

For example, the topology τ of any Hausdorff LCS E is linked to the weak topology $\sigma(E, E')$.

Lemma 18. Let σ, τ be two vector topologies on the vector space E and assume that σ is linked to τ. If $\{x_k\}$ is a sequence in E which is σ-Cauchy and if τ - $\lim x_k = x$, then σ - $\lim x_k = x$.

Proof: Let U be a basic σ neighborhood of 0 which is τ-closed. $\exists$ N such that $j, k \geq N$ implies $x_j - x_k \in U$. Since U is τ-closed,

$x_j - x \in U$ for $j \geq N$.

Theorem 19. If $\{x_k\} \subseteq X$ is $\sigma(X, X')$-$\mathscr{K}$ convergent to 0, then $\{x_k\}$ is $\| \;\|$-$\mathscr{K}$ convergent.

Proof: By Proposition 16 $\{x_k\}$ is norm convergent to 0. Let $\{y_k\}$ be a subsequence of $\{x_k\}$ such that $\sum_{k=1}^{\infty} \|y_k\| < \infty$. Next, let $\{z_k\}$ be a subsequence of $\{y_k\}$ such that $\sum_{k=1}^{n} z_k$ is weakly convergent to some $z \in X$. The partial sums $s_n = \sum_{k=1}^{n} z_k$ form a norm Cauchy sequence since the series Σz_k is absolutely convergent. Since the norm topology of X is linked to the weak topology, by Lemma 18 $\{s_n\}$ converges to z in norm. Since the same argument can be applied to any subsequence of $\{x_k\}$, $\{x_k\}$ is norm-$\mathscr{K}$ convergent.

As an application of Proposition 16 we give a proof of the Orlicz-Pettis Theorem. A series $\sum_{k=1}^{\infty} x_k$ in a TVS E is said to be <u>subseries convergent</u> if for each subsequence $\{x_{n_k}\}$ of $\{x_k\}$, the subseries $\sum_{k=1}^{\infty} x_{n_k}$ is convergent in E. The remarkable result of Orlicz-Pettis asserts that a series Σx_k in a NLS which is subseries convergent in the weak topology is subseries convergent in the norm topology. A result which asserts that any series which is subseries convergent in a given vector topology is also subseries convergent in a stronger vector topology is usually referred to as an Orlicz-Pettis Theorem. We first establish a general lemma which is useful in establishing Orlicz-Pettis Theorems.

Lemma 20. Let σ and τ be two vector topologies on the vector space E with σ linked to τ. If every series Σx_k in E which is subseries convergent with respect to τ satisfies $\sigma - \lim x_k = 0$, then every series in X which is τ-subseries convergent is also σ-subseries convergent.

Proof: By Lemma 18 it suffices to show that every τ-subseries convergent series Σx_k is such that its sequence of partial sums $s_n = \sum_{k=1}^{n} x_k$ forms a σ-Cauchy sequence. If Σx_k is τ-subseries convergent, but $\{s_n\}$ fails to be σ-Cauchy, there is a σ-neighborhood of 0, U, and a pairwise disjoint sequence of finite subsets of $\mathbb{N}$, $\{\sigma_n\}$, satisfying $\max \sigma_n < \min \sigma_{n+1}$ and $\sum_{k \in \sigma_n} x_k = z_n \notin U$. The series $\sum_n z_n$ is τ-subseries convergent being a subseries of $\sum_k x_k$ so $\sigma - \lim z_n = 0$ by hypothesis. This contradicts $z_n \notin U$ and establishes the result.

Theorem 21 (Orlicz-Pettis). If Σx_k is weak $(\sigma(X, X'))$ subseries convergent in X, then Σx_k is norm subseries convergent in X.

Proof: It follows from Proposition 16, that $\|x_k\| \to 0$ since the sequence $\{x_k\}$ is clearly $\sigma(X, X')$-$\mathscr{K}$ convergent. Since the norm topology is linked to $\sigma(X, X')$, the theorem follows from Lemma 20.

For the conclusion of Lemma 20 (and the Orlicz-Pettis Theorem) to hold, it is important that the two topologies are linked. See Exer. 15 for an example.

The Orlicz-Pettis Theorem has many important applications, especially in the theory of vector-valued measures, so there have been many generalizations of the theorem. For a history of the theorem and references to its many generalizations, see [DU]. We give a version for locally convex spaces in 18.11.

Weak Compactness:

Although the weak topology of an infinite dimensional NLS is not metrizable, it is a remarkable result due to Smulian and Eberlein that weak compact subsets of a NLS have the same compactness properties as those of a metrizable space. In this section we establish these compactness criteria.

If X is a topological space, recall the following types of compactness.

c_1: $E \subseteq X$ is relatively compact if $\bar{E}$ is compact.

c_2: $E \subseteq X$ is conditionally countably compact if every sequence in E has a cluster point in E [x is a cluster point of $\{x_k\}$ if the sequence is frequently in every neighborhood of x].

c_3: $E \subseteq X$ is <u>conditionally sequentially compact</u> if every sequence in E has a convergent subsequence.

It is always the case that $c_1 \Rightarrow c_2$ and $c_3 \Rightarrow c_2$. If X is first countable, then $c_2 \Leftrightarrow c_3$, and for metric spaces $c_1 \Leftrightarrow c_2 \Leftrightarrow c_3$ ([Ke] p. 138). If X is a B-space, then $\sigma(X, X')$ is first countable if and only if X is finite dimensional, but, nevertheless, we show that $c_1 \Leftrightarrow c_2 \Leftrightarrow c_3$ for the weak topology.

Proposition 22. Let X be a NLS whose dual contains a countable set D which separates the points of X. Then the weak topology on a weakly compact subset K of X is metrizable.

Proof: Let $D = \{x'_n : n \in \mathbb{N}\}$. Define a metric on K by

$$d(x, y) = \sum_{n=1}^{\infty} \frac{1}{2^n} \frac{|\langle x'_n, x-y\rangle|}{1+|\langle x'_n, x-y\rangle|}.$$

The topology on K induced by the metric is weaker than the weak topology which is compact so the weak topology is given by the metric.

Recall that Proposition 22 is applicable to separable spaces (8.1.11).

Theorem 23. If X is a NLS, then $c_1 \Rightarrow c_3$ fo $\sigma(X, X')$.

Proof: Let $K \subseteq X$ be relatively compact for $\sigma(X, X')$. Let $\{x_k\} \subseteq K$ and let E be the norm (weak) closure of span$\{x_k\}$. Then

$K \cap E$ is relatively $\sigma(X, X')$ compact and, therefore, relatively $\sigma(E, E')$ compact (Exer. 14.6). Since E is separable, Proposition 22 implies that the weak topology $\sigma(E, E')$ on $K \cap E$ is metrizable so $K \cap E$ is conditionally sequentially compact with respect to $\sigma(E, E')$. Hence, $\{x_k\}$ has a subsequence, $\{x_{n_k}\}$, which is $\sigma(E, E')$ convergent to some $x \in E$ and, therefore, $\sigma(X, X')$ convergent to x.

We next show that $c_2 \Rightarrow c_1$ for the weak topology.

Lemma 24. Let X be a NLS and H a finite dimensional subspace of X'. $\exists\, z_1, \ldots, z_n \in X$, $\|z_j\| = 1$ for $j = 1, \ldots, n$, such that

$$\max\{|\langle x', z_j\rangle| : i \le j \le n\} \ge \|x'\|/2 \quad \forall\, x' \in H .$$

Proof: $S = \{x' \in H : \|x'\| = 1\}$ is compact in H so $\exists$ a $\frac{1}{4}$ - net, $\{x_1', \ldots, x_n'\}$, for S. For each $j = 1, \ldots, n$ choose z_j of norm 1 such that $|\langle x_j', z_j\rangle| > 3/4$. Then for $0 \ne x' \in H$,

$$\max\{|\langle x'/\|x'\|, z_j\rangle| : 1 \le j \le n\} \ge 1/2$$

since

$$|\langle x'/\|x'\|, z_j\rangle| \ge |\langle x_j', z_j\rangle| - |\langle x'/\|x'\| - x_j', z_j\rangle|$$
$$\ge \frac{3}{4} - \|x'/\|x'\| - x_j'\| \ge \frac{1}{2}$$

if j is chosen such that $\|x'/\|x'\| - x_j\| < 1/4$.

Theorem 25. If X is a NLS, then $c_2 \Rightarrow c_1$ for $\sigma(X, X')$.

Proof: Let $K \subseteq X$ be conditionally countably compact with respect to $\sigma(X, X')$. If $x' \in X'$, then $\langle x', K\rangle$ is a conditionally countably compact set of scalars and is, therefore, bounded so K is $\sigma(X, X')$ and, therefore, norm bounded. Since JK is bounded, the weak* closure of JK in X'', $w^*(JK)$, is weak* compact (Banach-Alaoglu). Therefore, it suffices to show that $w^*(JK) \subseteq JX$ since J is a homeomorphism from $(X, \sigma(X, X'))$ onto $(JX, \sigma(X'', X'))$.

We show $w^*(JK) \subseteq JX$: Let $x'' \in w^*(JK)$. We begin an induction procedure by choosing $x_1' \in X'$, $\|x_1'\| = 1$. $\exists\ a_1 \in K$ such that $|\langle x'' - Ja_1, x_1'\rangle| < 1$ since $x'' \in w^*(JK)$. The space spanned by x'' and $x'' - Ja_1$ is finite dimensional so by Lemma 24 $\exists\ x_2', \ldots, x_{n_2}'$ in X' with norm 1 such that

$$\max\{|\langle y'', x_m'\rangle| : 2 \le m \le n_2\} \ge \|y''\|/2 \quad \forall\ y'' \in \text{span}\{x'', x'' - Ja_1\}.$$

Since $x'' \in w^*(JK)$, $\exists\ a_2 \in K$ such that

$$\max\{|\langle x'' - Ja_2, x_m'\rangle : 1 \le m \le n_2\} < 1/2 .$$

$\exists\ x_{n_2+1}', \ldots, x_{n_3}'$ of norm 1 in X' such that

$$\max\{|\langle y'', x_m'\rangle| : n_2 < m \le n_3\} \ge \|y''\|/2$$

for $y'' \in \text{span}\{x'' - Ja_1, x'' - Ja_2, x''\}$. Continue.

By hypothesis, $\exists\ x \in X$ which is a cluster point of the sequence $\{a_n\}$ with respect to $\sigma(X, X')$. We show that $Jx = x''$. Since $\overline{\text{span}}\{a_n : n \in \mathbb{N}\}$ (all closures are in norm topologies) is weakly closed, $x \in \overline{\text{span}}\{a_n\}$ so $x'' - Jx \in \overline{\text{span}}\{x'', x'' - Ja_1, x'' - Ja_2, \ldots\}$.

By construction if $y'' \in \text{span}\{x'' - Ja_n : n\} + \text{span}\{x''\}$,

$$(1) \qquad \sup\{|\langle y'', x_m'\rangle| : m\} \ge \|y''\|/2$$

so the same holds for any point in the norm closure of this subspace and, in particular, for $x'' - Jx$. Also $|\langle x'' - Ja_n, x'_m\rangle| < 1/p$ for $n > n_p > m$ implies

$$(2) \qquad |\langle x'' - Jx, x'_m\rangle| \le |\langle x'' - Ja_n, x'_m\rangle| + |\langle x'_m, a_n - x\rangle| \le 1/p + |\langle x'_m, a_n - x\rangle|$$

for $n > n_p > m$. Since x is a weak cluster point of $\{a_n\}$, given x'_m and $N > m$ $\exists$ a_n such that $|\langle x'_m, a_n - x\rangle| < 1/N$ where $n > n_N > m$. For such an a_n, we have from (2), $|\langle x'' - Jx, x'_m\rangle| \le 2/N$ so

$$|\langle x'' - Jx, x'_m\rangle| = 0 \ \forall m.$$

But, (1) implies that $\sup\{|\langle x'' - Jx, x'_m\rangle| \ge \|x'' - Jx\|/2$ so $x'' = Jx$ as desired.

Combining Theorems 23 and 25, we have

Theorem 26 (Smulian-Eberlein). Let X be a NLS and $K \subseteq X$. For the weak topology, $\sigma(X, X')$, c_1, c_2 and c_3 are equivalent.

See [DS] p. 466 for a historical sketch of the evolution of Theorem 26. The proof given here is due to Whitley [Wh1]. It should be remarked that an analogous result for the weak* topology is false; see Exer.'s 7 and 11.

Exercise 1. Give an example of a B-space which is not weakly sequentially complete.

Exercise 2. Show that a subset of ℓ^1 is compact if and only if it is weakly compact.

Exercise 3. Show that in Theorem 14(iv), $\sigma(\ell^1, m_0)$ cannot be replaced by $\sigma(\ell^1, c_{00})$.

Exercise 4. Let X be a NLS. Show a sequence $\{x_k\} \subseteq X$ converges weakly to 0 if and only if $\{\|x_k\|\}$ is bounded and $\langle x', x_k\rangle \to 0$ $\forall\, x'$ belonging to a norm dense linear subspace of X'.

Exercise 5. Let X be a B-space. Show a sequence $\{x_k'\} \subseteq X'$ converges weak* to 0 if and only if $\{\|x_k'\|\}$ is bounded and $\langle x_k', x\rangle \to 0$ $\forall\, x$ belonging to a norm-dense linear subspace of X.

Exercise 6. Let $\{f_k\} \subseteq C(S)$. Show $\{f_k\}$ is weakly convergent to 0 if and only if $\{\|f_k\|\}$ is bounded and $f_k(t) \to 0$ $\forall\, t \in S$.

Exercise 7. Let S be a compact Hausdorff space and set

$$S_0 = \{\delta_t : t \in S\} \subseteq rca(S) = C(S)'.$$

Show S and S_0 are homeomorphic if S_0 is equipped with the relative $\sigma(rca(S), C(S))$ topology. Compare Theorem 26.

Exercise 8. A TVS is said to be <u>quasicomplete</u> if every bounded Cauchy net is convergent. Show that the dual of a B-space is quasicomplete under the weak* topology. Show the completeness can't be dropped.

Exercise 9. Let X be a B-space and Y a NLS and $\mathcal{F} \subseteq L(X, Y)$. Suppose $\{\langle y', Tx\rangle : T \in \mathcal{F}\}$ is bounded $\forall\, y' \in Y'$, $x \in X$. Show $\{\|T\| : T \in \mathcal{F}\}$ is bounded.

Exercise 10. If X is a reflexive B-space, show that X is weakly sequentially complete.

Exercise 11. Let S be uncountable and let μ be counting measure on S. Show $\ell^\infty(S) = B(S)$ is the dual of $\ell^1(S) = L^1(\mu)$. Show the closed unit ball of $\ell^\infty(S)$ is weak* compact but not weak* sequentially compact. Compare Theorem 26.

Exercise 12. Show that a subspace of ℓ^1 is reflexive if and only if it is finite dimensional.

Exercise 13. Give an example of a non-reflexive B-space that is weakly sequentially complete.

Exercise 14. Let X be a vector space with two vector topologies σ, τ. If $\sigma \subseteq \tau$, show σ is linked to τ. If $\sigma \subseteq \tau$, τ is linked to σ and σ is complete, show τ is complete. Give an example where $\sigma \subseteq \tau$ and τ is not linked to σ. (Hint: Exer. 10.9.)

Exercise 15. Consider c and its weak topology, $\sigma = \sigma(c, \ell^1)$ (5.12), along with the topology, p, of coordinatewise convergence. Show Σe_k is p-subseries convergent and the partial sums of any subseries are σ-Cauchy, but Σe_k is not σ-subseries convergent.

Exercise 16. Give an example of a sequentially continuous linear operator which is not continuous.

Exercise 17. Let X be a B-space and M a closed subspace. Show X is reflexive if and only if M and X/M are reflexive.

17
Polar Topologies

In this section we consider a very general method for constructing locally convex topologies. In fact, we will show below in Theorem 7 that any Hausdorff locally convex topology can be obtained from the construction which we now describe.

Let X and X' be in duality and let $\mathcal{A}$ be a family of $\sigma(X, X')$ bounded subsets of X. For each $A \in \mathcal{A}$, A^0 is absolutely convex and absorbing (15.4). Let p_{A^0} be the Minkowski functional of A^0. Thus,

$$(1)\quad p_{A^0}(x') = \inf\{t > 0 : x' \in tA^0\}$$

$$= \inf\{t > 0 : |\langle x', x\rangle| \le t \ \forall\, x \in A\}$$

$$= \inf\{t > 0 : \sup_{x \in A} |\langle x', x\rangle| \le t\} = \sup\{|\langle x', x\rangle| : x \in A\}\,.$$

The locally convex topology on X' generated by the semi-norms $\{p_{A^0} : A \in \mathcal{A}\}$ is denoted by $\tau_{\mathcal{A}}$. By (1), a net $\{x'_\delta\}$ in X' converges to 0 in $\tau_{\mathcal{A}}$ if and only if $\langle x'_\delta, x\rangle \to 0$ uniformly for $x \in A$, and for this reason the topology $\tau_{\mathscr{A}}$ is called the <u>topology of uniform convergence on</u> $\mathscr{A}$. We call a topology of the form $\tau_{\mathscr{A}}$ a <u>polar topology</u> generated by the family $\mathscr{A}$.

Proposition 1. If $X = \bigcup_{A \in \mathscr{A}} A$, then $\tau_{\mathscr{A}}$ is Hausdorff.

Proof: Let $x' \in X'$, $x' \neq 0$. Then $\exists\, x \in X$ such that $\langle x', x\rangle \neq 0$. If $x \in A$, then $p_{A^0}(x') = \sup\{|\langle x', y\rangle| : y \in A\} > 0$.

Example 2. Let $\mathscr{A}$ be the family of all finite subsets of X. Then $\tau_{\mathscr{A}}$ is just the weak topology $\sigma(X, X')$.

Example 3. Let $\mathscr{A}$ be the family of all absolutely convex, $\sigma(X, X')$ compact subsets of X. The topology $\tau_{\mathscr{A}}$ in this case is called the <u>Mackey topology</u> of X'. It will be considered in detail in §18.

Example 4. Let $\mathscr{A}$ be the family of all $\sigma(X, X')$-$\mathscr{K}$ bounded subsets of X.

Example 5. Let $\mathscr{A}$ be the family of all $\sigma(X, X')$ bounded subsets of X. The topology $\tau_{\mathscr{A}}$ is called the <u>strong topology</u> of X'. If X is a NLS and X' is its dual, this topology is just the norm topology of X'. We consider this topology in §19.

We now show that every Hausdorff locally convex topology is a topology of uniform convergence on an appropriate family of $\sigma(X, X')$ bounded sets, i.e., is a polar topology.

Theorem 6. Let X be a Hausdorff LCS and $A \subseteq X'$. The following are equivalent.

(i) A is equicontinuous,

(ii) A is contained in the polar, U^0, of some neighborhood of 0, U, in X,

(iii) $\exists$ a continuous semi-norm p on X such that

$$|\langle x', x\rangle| \le p(x) \ \forall\, x' \in A, x \in X.$$

Proof: (i) $\Rightarrow$ (ii): $\exists$ a neighborhood of 0, U, such that $|\langle x', x\rangle| \le 1$ $\forall\, x \in U, x' \in A$. Whence, $A \subseteq U^0$.

(ii) $\Rightarrow$ (iii): Let V be an open, absolutely convex neighborhood of 0 such that $V \subseteq U$. Then $A \subseteq V^0$. Let p be the Minkowski functional of V. Then p is continuous (11.13). If $p(x) > 0$, then $x/p(x) \in \bar{V}$ so $|\langle x', x\rangle| \le p(x)$ $\forall\, x' \in A$ (11.13); while if $p(x) = 0$, then $nx \in V$ $\forall\, n \in \mathbb{N}$ so $|\langle x', nx\rangle| \le 1$ $\forall\, x' \in A$ or $|\langle x', x\rangle| \le 0 = p(x)$ $\forall\, x' \in A$ and (iii) holds.

(iii) $\Rightarrow$ (i) by 13.8.

Theorem 7. Let (X, τ) be a Hausdorff LCS. Then τ is a polar topology, namely $\tau_{\mathscr{A}}$, where $\mathscr{A}$ is the family of all equicontinuous subsets of X'.

Proof: Let $\mathscr{U}$ be a neighborhood base at 0 for τ consisting of τ closed, absolutely convex sets. Put $\mathscr{A} = \{U^0 : U \in \mathscr{U}\}$. Note that each $U^0 \in \mathscr{A}$ is $\sigma(X', X)$ bounded (15.4). By the Bipolar Theorem $(U^0)_0 = U \quad \forall U \in \mathscr{U}$ so $\{x : p_{(U^0)_0}(x) \le 1\} = \{x : p_U(x) \le 1\} = U$ (11.13). Thus, both τ and $\tau_{\mathscr{A}}$ have the same neighborhood base at 0 and by Theorem 6 every equicontinuous subset of X' is contained in some member of $\mathscr{A}$.

For bounded sets in $\tau_{\mathscr{A}}$ we obtain from (1).

Proposition 8. A subset $B \subseteq X'$ is $\tau_{\mathscr{A}}$ bounded if and only if

$$|\langle B, A\rangle| = \sup\{|\langle x', x\rangle| : x' \in B, x \in A\} < \infty \quad \forall A \in \mathscr{A}.$$

Proof: (1) and 13.15.

Finally, if the family $\mathscr{A}$ satisfies the conditions

(i) $tA \in \mathscr{A}$ when $A \in \mathscr{A}, t \in \mathbb{F}$

(ii) if $A, B \in \mathscr{A}$, there exists $C \in \mathscr{A}$ such that $C \supseteq A \cup B$,

we can give a convenient description of the neighborhood base at 0 for $\tau_{\mathscr{A}}$.

Theorem 9. If $\mathscr{A}$ satisfies conditions (i) and (ii), then the set $\mathscr{A}^0 = \{A^0 : A \in \mathscr{A}\}$ forms a neighborhood base at 0 for $\tau_{\mathscr{A}}$.

Proof: It is easily checked that the family $\mathscr{A}^0$ satisfies the conditions of 1.18 and, therefore, is a neighborhood base at 0, for a locally convex topology τ on X'. A net $\{x'_\delta\}$ in X' converges to 0 in τ if and only if $\{x'_\delta\}$ converges to 0 uniformly on elements of $\mathcal{A}$ so the topology τ is exactly the topology $\tau_{\mathcal{A}}$.

Exercise 1. Let $\mathcal{A}$ be a family of $\sigma(X, X')$ bounded subsets of X. Set $\mathcal{B} = \{A \subseteq X : A$ is the $\sigma(X, X')$ closed convex hull of a number of $\mathcal{A}\}$. Show $\tau_{\mathcal{A}} = \tau_{\mathcal{B}}$.

Exercise 2. A class, $\mathcal{A}$, of $\sigma(X, X')$ bounded sets is <u>saturated</u> if it satisfies condition (i) and

(a) if $A \in \mathcal{A}$ and $B \subseteq A$, then $B \in \mathcal{A}$,

(b) if **A, B** $\in \mathcal{A}$, then the weakly closed absolutely convex hull of $A \cup B$ belongs to $\mathcal{A}$

Suppose that $\mathcal{A}$ and $\mathcal{B}$ are saturated. Show $\mathcal{A} = \mathcal{B}$ if and only if $\tau_{\mathcal{A}} = \tau_{\mathcal{B}}$.

18
The Mackey-Arens Theorem

Let X and X' be in duality. Recall that a locally convex Hausdorff topology τ on X is said to be compatible with the duality between X and X' if $(X, \tau)' = X'$. By 14.6 $\sigma(X, X')$ is such a topology and by definition is the weakest such topology. In this section we show that there is a strongest locally convex topology, called the Mackey topology, which is compatible with the duality.

Definition 1. Let $\mathscr{M}$ be the family of absolutely convex, $\sigma(X', X)$ compact subsets of X'. The polar topology, $\tau_{\mathscr{M}}$, on X is called the Mackey topology of X and is denoted by $\tau(X, X')$. Thus, the Mackey topology is the topology of uniform convergence on absolutely convex, $\sigma(X', X)$ compact subsets of X'.

Proposition 2. (i) $\tau(X, X') \supseteq \sigma(X, X')$.

(ii) $\tau(X, X')$ is Hausdorff.

Proof: (i) is clear and (ii) follows from (i).

We next show that $\tau(X, X')$ is a compatible topology for the duality between X and X'. For this we require two lemmas.

Lemma 3. Let N be a $\tau(X, X')$ neighborhood of 0. Then N contains a set G which is absolutely convex, $\sigma(X, X')$ closed and such that G^0 is $\sigma(X', X)$ compact.

Proof: By 17.9, $N \supseteq S_0$, where $S \subseteq X'$ is absolutely convex and $\sigma(X', X)$ compact. Set $G = S_0$ and observe $G^0 = (S_0)^0 = S$ by the Bipolar Theorem.

Lemma 4. Let Y be a Hausdorff LCS. Let G be an absolutely convex closed set in Y such that G^0 is $\sigma(Y', Y)$ compact. If f is a linear functional on Y which is bounded on G, then f is continuous.

Proof: Consider Y and $Y^\#$ in duality. If $A \subseteq Y$, write A^p for its polar with respect to this duality. Note for $S \subseteq Y'$, S_0 is unambiguous.

We claim that $((G^0)_0)^p = G^0$. G^0 is $\sigma(Y^\#, Y)$ compact since it is $\sigma(Y', Y)$ compact and, therefore, G^0 is $\sigma(Y^\#, Y)$ closed. Our claim now follows from the Bipolar Theorem since G^0 is absolutely convex.

However, $(G^0)_0 = G$ by the Bipolar Theorem. Thus, $G^p = G^0$ so $G^p \subseteq Y'$ and f must be continuous since some multiple of f belongs to G^p.

Theorem 5. $\tau(X, X')$ is compatible with the duality between X and X'.

Proof: For each $x' \in X'$, $\langle x', \cdot \rangle$ is $\sigma(X, X')$ continuous and, hence, $\tau(X, X')$ continuous by Proposition 2. Hence,

$$(X, \tau(X, X'))' \supseteq X' = (X, \sigma(X, X'))' .$$

Let $f \in (X, \tau(X, X'))'$. Let $N = \{x : |\langle f, x \rangle| < 1\}$. Then N is a $\tau(X, X')$ neighborhood of 0 so choose G as in Lemma 3. If X is equipped with $\sigma(X, X')$, the conditions of Lemma 4 are satisfied since f is bounded on $G \subseteq N$. Hence, f is $\sigma(X, X')$ continuous and $f \in X'$. Hence, $(X, \tau(X, X'))' \subseteq X'$, and equality holds.

We next show that the Mackey topology is the strongest locally convex topology which is compatible with the duality between X and X'.

Lemma 6. If τ is compatible with the duality between X and X' and U is an absolutely convex τ closed neighborhood of 0 in X, $\exists$ an absolutely convex $\sigma(X', X)$ compact subset $S \subseteq X'$ such that $S_0 = U$.

Proof: Put $S = U^0$ and apply the Bipolar Theorem, the Banach-Alaoglu Theorem and 13.14.

Theorem 7 (Mackey-Arens). Let τ be a Hausdorff locally convex topology on X. Then τ is compatible with the duality between X and X' if and only if $\sigma(X, X') \subseteq \tau \subseteq \tau(X, X')$.

Proof: $\Leftarrow$ follows from Exer. 14.1 and Theorem 5.

$\Rightarrow$: If τ is compatible, $\sigma(X, X') \subseteq \tau$ by definition, and Lemma 6 and 17.9 show that $\tau \subseteq \tau(X, X')$.

A LCS (X, τ) is said to be a Mackey space if $\tau = \tau(X, X')$. For example,

Proposition 8. If (X, τ) is a metrizable LCS, then X is a Mackey space.

Proof: The identity from (X, τ) into $(X, \tau(X, X'))$ is bounded (14.15) and, hence, continuous (5.4). Therefore, $\tau(X, X') \subseteq \tau$. But, always $\tau \subseteq \tau(X, X')$.

We next establish a version of the Orlicz-Pettis Theorem for LCS. We establish a version of this result due to Bennett and Kalton which is stronger than the usual versions given ([BK]). We first require a preliminary result.

Theorem 9. Let E be a separable Hausdorff LCS and $K \subseteq E'$ a weak* compact subset. The weak* topology on K is metrizable.

Proof: Let $\{x_k\}$ be dense in E. For each k define a semi-norm on K by $|x'|_k = |\langle x', x_k\rangle|$, and let $|\ |$ be the Frechet quasi-norm induced by the sequence $\{|\ |_k\}$ (2.3). Then $d(x', y') = |x' - y'|$ defines a metric on K and the metric topology is weaker than the weak*

topology on K. Since K is weak* compact, the metric topology is exactly the weak* topology.

We next establish an analogue of 16.16 for LCS. Let E be a Hausdorff LCS and let $\lambda(E, E')$ be the topology of uniform convergence on $\sigma(E', E)$ compact subsets of E'. Thus, $\lambda(E, E')$ is stronger than the Mackey topology, $\tau(E, E')$, and can be strictly stronger than the Mackey topology ([K1] 21.4).

Theorem 10. If $\{x_k\}$ is $\sigma(E, E')$-$\mathscr{K}$ convergent to 0, then $\{x_k\}$ is $\lambda(E, E')$ convergent to 0.

Proof: It suffices to show that $<x_k', x_k> \to 0$ for an arbitrary sequence $\{x_k'\}$ belonging to any $\sigma(E', E)$ compact subset K of E'. We may assume, by replacing E by the linear subspace spanned by $\{x_k\}$ if necessary, that E is separable. Since the weak* topology of K is metrizable (Theorem 9), $\{x_k'\}$ has a subsequence $\{x_{n_k}'\}$ which is weak* convergent to an element $x' \in K$. Consider the matrix $M = [<x_{n_i}', x_{n_j}>]$. It is easily checked that M is a $\mathscr{K}$-matrix so by the Basic Matrix Theorem $<x_{n_k}', x_{n_k}> \to 0$. Since the same argument can be applied to any subsequence of $\{<x_k', x_k>\}$, it follows that $<x_k', x_k> \to 0$.

From 16.20 we obtain our locally convex version of the Orlicz-Pettis Theorem.

Theorem 11 (Orlicz-Pettis). If Σx_k is $\sigma(E, E')$ subseries convergent, then Σx_k is $\lambda(E, E')$ subseries convergent.

Proof: $\{x_k\}$ is $\sigma(E, E')$-$\mathcal{K}$ convergent so $x_k \to 0$ in $\lambda(E, E')$ by Theorem 10. 16.20 now gives the result.

Since $\lambda(E, E')$ is stronger than the Mackey topology, if Σx_k is $\sigma(E, E')$ subseries convergent it is also $\tau(E, E')$ and τ subseries convergent (Theorem 7). This is the usual statement of the Orlicz-Pettis Theorem for LCS.

Exercise 1. Give an example of a LCS whose topology is strictly weaker than the Mackey topology.

Exercise 2. Show that all compatible topologies on s coincide.

Exercise 3. If E is a separable Hausdorff LCS and U is a neighborhood of 0 in E, show that the weak* topology on U^0 is metrizable.

Exercise 4. Let E be separable and metrizable. Show that E′ is weak* sequentially separable. [Hint: $\{0\} = \bigcap_{n=1}^{\infty} U_n$.]

Exercise 5. Let K be a compact metric space. If Σf_k is subseries convergent in the topology of pointwise convergence on C(K), show that Σf_k is subseries convergent in the sup-norm topology. Use this result to prove Theorem 11.

19
The Strong Topology and the Bidual

In this section we study an important polar topology, the strong topology, which is a generalization of the norm topology on the dual of a NLS.

Let X and X′ be in duality. If $\mathscr{B}$ is the family of all $\sigma(X, X')$ bounded subsets of X, the polar topology, $\tau_{\mathscr{B}}$, on X′ generated by $\mathscr{B}$ is called the <u>strong topology</u> of X′. The topology $\tau_{\mathscr{B}}$ is denoted by <u>$\beta(X', X)$</u>, and <u>X'_b</u> denotes X′ with the topology $\beta(X', X)$. $\beta(X', X)$ is the topology of uniform convergence on $\sigma(X, X')$ bounded subsets of X and has for a neighborhood base at 0 the family $\{A^0 : A \in \mathscr{B}\}$ (17.9).

Example 1. If X is a NLS, then the strong topology of X′ is just the (dual) norm topology (16.7). However, the strong topology, $\beta(X, X')$, of X may be strictly stronger than the norm topology. For example, if $X = c_{00}$ and $B = \{ke_k : k \in \mathbb{N}\} \subseteq \ell^1 = X'$, then

$$B_0 = \{\{t_k\} \in c_{00} : |t_k| \le 1/k \ \forall\, k\}$$

is a strong neighborhood of 0 which is not a norm neighborhood of 0. If X is a B-space, then the strong topology of X is just the norm topology (16.8).

Example 2. $\beta(X', X)$ is not, in general, compatible with the duality between X and X'. For example, if X is a non-reflexive B-space, then $(X', \beta(X', X))' = (X', \| \|)' = X'' \neq X$.

Since any compact subset of a TVS is bounded (Exer. 4.2), we have

Proposition 3. $\sigma(X', X) \subseteq \tau(X', X) \subseteq \beta(X', X)$. In particular, if (X, τ) is a Hausdorff LCS, then $\sigma(X, X') \subseteq \tau \subseteq \tau(X, X') \subseteq \beta(X, X')$.

Exercise 1 shows the containments above can be proper.

From 17.8 we have the following characterization of strongly bounded sets.

Proposition 4. $B \subseteq X'_b$ is $\beta(X', X)$ bounded if and only if $\forall\ \sigma(X, X')$ bounded subset $A \subseteq X$, $\sup\{|\langle x', x\rangle| : x' \in B, x \in A\} = |\langle B, A\rangle| < \infty$.

Corollary 5. Let (X, τ) be a Hausdorff LCS. Every equicontinuous subset $B \subseteq X'$ is strongly bounded.

Proof: $\exists$ a τ neighborhood of 0, V, such that $B \subseteq V^0$. If A is $\sigma(X, X')$ $(= \tau)$ bounded, $\exists$ $t > 0$ such that $A \subseteq tV$. Thus,

$$|<B, A>| \le |<B, tV>| = t|<B, V>| \le t,$$

and the result follows from Proposition 4.

Example 6. The converse of Corollary 5 is false. Let $X = (c_0, \sigma(c_0, \ell^1))$. Then $X' = \ell^1$. The closed unit ball of ℓ^1 is strongly (= norm) bounded but is not equicontinuous with respect to $\sigma(c_0, \ell^1)$ [$e_k \to 0$ in $\sigma(c_0, \ell^1)$ and $\|e'_k\|_1 = 1$ but $<e_k, e_k> = 1$ (5.20)].

We give one further characterization of strongly bounded sets in LCS.

Definition 7. A subset B of a TVS E is <u>bornivorous</u> or a <u>bornivore</u> if it absorbs bounded sets.

For example, a neighborhood of 0 is bornivorous by definition. For an example of a bornivore which is not a neighborhood of 0, let X be an infinite dimensional NLS and let B be any weak neighborhood of 0. Then B is a bornivore (16.7) but is not a norm neighborhood of 0 (14.12).

Definition 8. A subset B of a LCS E is a <u>barrel</u> if B is absolutely convex, closed and absorbing.

For example, any absolutely convex, closed neighborhood of 0 is a barrel. For an important class of LCS, the barrelled spaces, the converse holds (see §24).

Proposition 9. Let E be a Hausdorff LCS. A subset $B \subseteq E'$ is $\beta(E', E)$ bounded if and only if B is contained in the polar, A^0, of a bornivorous barrel A in E.

Proof: $\Rightarrow$: $B_0 = A$ is absolutely convex, $\sigma(E, E')$ closed and, therefore, closed in E (14.10). Moreover, A is a bornivore since if C is a bounded set in E, $\exists\ t > 0$ such that $B \subseteq tC^0$ since C^0 is a basic strong neighborhood of 0. Thus, $tB_0 = tA \supseteq (C^0)_0 \supseteq C$. Since $B \subseteq (B_0)^0 = A^0$, the result follows.

$\Leftarrow$: If $B \subseteq A^0$, where A is a bornivorous barrel in E, let C be a bounded set in E. Then $\exists\ t > 0$ such that $C \subseteq tA$ so $tC^0 \supseteq A^0 \supseteq B$ and B is $\beta(E', E)$ bounded since C^0 is a basic $\beta(E', E)$ neighborhood of 0.

The Bidual:

If E is a Hausdorff LCS, the <u>bidual</u> E'' of E is the dual of E'_b, $E'' = (E'_b)'$. In the case when E is a NLS this agrees with our previous definition (§8.1).

The strong topology of E' is stronger than $\sigma(E', E)$ so the dual of $(E', \sigma(E', E))$ $(= E)$ can be regarded as a linear subspace of E''. That is, the canonical map $J : E \to E''$, $\langle Je, x'\rangle = \langle \hat{e}, x'\rangle = \langle x', e\rangle$, $x' \in E'$, imbeds E into E''; the imbedding $Je = \hat{e}$ is again referred to as the <u>canonical imbedding</u>.

In the NLS case, J is an isometry and is, therefore, a homeomorphism from E onto JE with the strong (= norm) topology. However, the LCS case is quite different. We say that E is <u>semi-reflexive</u> if $JE = E''$ and E is <u>reflexive</u> if E is semi-reflexive and J is a homeomorphism from E onto E_b''. Thus, a NLS is reflexive if and only if it is semi-reflexive, but as the example below indicates the situation is different for LCS.

Example 10. Let $E = (\ell^\infty, \sigma(\ell^\infty, \ell^1))$. Then $\ell^1 = E'$ and the strong topology on ℓ^1 from this duality is just the norm topology. Hence, ℓ^∞ is the bidual of E and the strong topology of ℓ^∞ is just the norm topology which is strictly stronger than $\sigma(\ell^\infty, \ell^1)$. Thus, E is semi-reflexive but not reflexive.

For semi-reflexive spaces, we have the analogue of 16.6.

Theorem 11. A Hausdorff LCS (E, τ) is semi-reflexive if and only if every $\sigma(E, E')$ closed, bounded subset $B \subseteq E$ is $\sigma(E, E')$ compact.

Proof: $\Leftarrow$: Suppose $x'' \in E''$. Since x'' is continuous on E_b', $\exists$ a closed, bounded $(\sigma(E, E')$ or $\tau)$, absolutely convex subset $B \subseteq E$ such that

(1) $$|\langle x'', u\rangle| \le 1 \ \forall u \in B^0 .$$

Now J is a homeomorphism from $(E, \sigma(E, E'))$ onto JE with the relative $\sigma(E'', E')$ topology so JB is $\sigma(E'', E')$ compact in E''

and, in particular, JB is $\sigma(E'', E')$ closed. Since $(JB)_0 = B^0$ (definition), $((JB)_0)^0 = B^{00}$ and by the Bipolar Theorem $((JB)_0)^0 = JB = B^{00}$. But, (1) implies $x'' \in B^{00}$ so $x'' \in JB \subseteq JE$ and J is onto E''.

$\Rightarrow$: J is a homeomorphism from $(E, \sigma(E, E'))$ onto $(E'', \sigma(E'', E'))$. Now B^0 is a strong neighborhood of 0 in E' so by the Banach-Alaoglu Theorem $B^{00} = ((JB)_0)^0$ is $\sigma(E'', E')$ compact. But, $((JB)_0)^0 \supseteq JB$ so JB is $\sigma(E'', E')$ compact and B is $\sigma(E, E')$ compact.

Let $\varepsilon(E'', E')$ be the locally convex topology on E'' of uniform convergence on equicontinuous subsets of E'. By Theorem 17.7, J is a homeomorphism from (E, τ) onto JE with the relative $\varepsilon(E'', E')$ topology. Thus, if a space E is to be reflexive, the strong topology $\beta(E'', E')$, and $\varepsilon(E'', E')$ must coincide. We have the following criteria for this to occur.

Proposition 12. The following are equivalent.

(i) $\beta(E'', E') = \varepsilon(E'', E')$.

(ii) every $\beta(E', E)$ bounded subset of E' is equicontinuous [note Corollary 5].

(iii) every bornivorous barrel in E is a neighborhood of 0.

Proof: The equivalence of (i) and (ii) follows from the definitions of the topologies $\beta(E'', E')$ and $\varepsilon(E'', E')$ as polar topologies.

(iii) ⇒ (ii): Let $B \subseteq E'$ be $\beta(E', E)$ bounded. By Proposition 9 ∃ a bornivorous barrel $A \subseteq E$ such that $B \subseteq A^0$. But, A is a neighborhood of 0 so B must be equicontinuous (17.6).

(ii) ⇒ (iii): Let B be a bornivorous barrel in E. By Proposition 9, B^0 is $\beta(E', E)$ bounded and, hence, equicontinuous. Thus, $(B^0)_0 = B$ is a neighborhood of 0 (17.7).

Definition 13. A Hausdorff LCS E is quasi-barrelled (infra-barrelled) if it satisfies any of the equivalent conditions of Proposition 12.

From the discussion above, we have a characterization of reflexive spaces.

Theorem 14. A Hausdorff LCS E is reflexive if and only if it is semi-reflexive and quasi-barrelled.

Exercise 1. Let X be a non-reflexive B-space. Show $\sigma(X, X') \subsetneq \tau(X, X')$ and $\tau(X', X) \subsetneq \beta(X', X)$.

Exercise 2. Show that a closed linear subspace of a semi-reflexive space is semi-reflexive. [The corresponding result for reflexive spaces is false ([K1] 23.5.]

Exercise 3. Let E be a Hausdorff LCS and $B \subseteq E'$. If there is a neighborhood of 0, U, in E such that $|\langle B, U\rangle| < \infty$, show B is strongly bounded.

Exercise 4. Let E be a metrizable LCS. If $B \subseteq E'$ is strongly bounded, show there is a neighborhood of 0, U, in E such that $|\langle B, U\rangle| < \infty$. [Hint: Let $U_1 \supseteq U_2 \supseteq \ldots$ be a neighborhood base at 0 in E and suppose the conclusion is false.]

Exercise 5. Show that a subset of a LCS is strongly bounded if and only if it is absorbed by every barrel.

Exercise 6. Show that if $\{x_k\} \subseteq X$ is $\sigma(X, X')$-$\mathscr{K}$ convergent to 0, then $\{x_k\}$ is $\beta(X, X')$ bounded.

Exercise 7. Let E be a Hausdorff LCS. Show that a subset $B \subseteq E'$ is $\sigma(E', E)$ bounded if and only if B is contained in the polar of a barrel in E.

20

Quasi-barrelled Spaces and the Topology β^* (E, E′)

Let E, E′ be in duality. We use Proposition 19.12 to introduce another locally convex topology which is closely related to quasi-barrelled spaces. We define $\underline{\beta^*(E, E')}$ to be the topology of uniform convergence on the $\beta(E', E)$ bounded subsets of E′. Since any $\beta(E', E)$ bounded subset of E′ is $\sigma(E', E)$ bounded, $\beta(E, E')$ is stronger than $\beta^*(E, E')$ and can be strictly stronger (Exer. 1). From 19.12, we obtain

Theorem 1. A Hausdorff LCS (E, τ) is quasi-barrelled if and only if $\tau = \beta^*(E, E')$.

We show in the next section that any metrizable LCS is always quasi- barrelled (21.3) so the class of quasi-barrelled spaces is quite large.

We now compare $\beta^*(E, E')$ with the Mackey topology.

Lemma 2. (i) If $B \subseteq E'$ is absolutely convex and $\sigma(E', E)$ compact, then B is $\sigma(E', E)$-$\mathscr{K}$ bounded.

(ii) If $B \subseteq E'$ is $\sigma(E', E)$-$\mathscr{K}$ bounded, then B is $\beta(E', E)$ bounded [compare with 16.10].

Proof: (i) follows from 13.17 since a compact subset of a TVS is always complete (Exer. 1.7).

(ii): Let $A \subseteq E$ be $\sigma(E, E')$ bounded. It suffices to show that $\{<x_i', x_i>\}$ is bounded whenever $\{x_i'\} \subseteq B$ and $\{x_i\} \subseteq A$ (19.4). Let $t_i \to 0$ with $t_i > 0$. Consider the matrix $M = \{<\sqrt{t_j}\, x_j', \sqrt{t_i}\, x_i>\}$. Since $\{\sqrt{t_i}\, x_i\}$ is $\sigma(E, E')$ convergent to 0 and $\{\sqrt{t_j}\, x_j'\}$ is $\sigma(E', E)$-$\mathscr{K}$ convergent to 0, M is a $\mathscr{K}$-matrix so by the Basic Matrix Theorem 9.2, $t_i<x_i', x_i> \to 0$ and $\{<x_i', x_i>\}$ is bounded.

The absolute convexity in (i) is important; see Exercise 3.

Thus, from Lemma 2, any absolutely convex $\sigma(E', E)$ compact subset of E' is $\beta(E', E)$ (strongly) bounded, and we have

Theorem 3. $\tau(E, E') \subseteq \beta^*(E, E')$.

The containment in Theorem 3 can be proper.

Example 4. Let $E = \ell^\infty$, $E' = \ell^1$. The family of $\beta(\ell^1, \ell^\infty)$ bounded subsets of ℓ^1 is just the family of norm bounded subsets of ℓ^1 so $\beta^*(\ell^\infty, \ell^1)$ is the norm topology of ℓ^∞. Since the dual of ℓ^∞ under the norm topology is ba (5.15), $\beta^*(\ell^\infty, \ell^1)$ is strictly stronger than the Mackey topology $\tau(\ell^\infty, \ell^1)$ whose dual is ℓ^1.

There are examples of LCS where the strong topology on the dual space and $\beta^*(E', E)$ coincide. Using Lemma 2, we give such an example.

Theorem 5. If E is an $\mathscr{A}$-space for any locally convex topology which is compatible with the duality between E and E′, then every weakly bounded subset of E is strongly bounded and $\beta(E', E) = \beta^*(E', E)$.

Proof: $(E, \sigma(E, E'))$ is an $\mathscr{A}$-space (14.19) so from Lemma 2 (ii), it follows that any $\sigma(E, E')$ bounded set is $\beta(E, E')$ bounded and $\beta(E', E) = \beta^*(E', E)$.

A locally convex space with the property that bounded sets are strongly bounded is called a <u>Banach-Mackey space</u> ([W], 10.4). From Theorem 5 any $\mathscr{A}$-space is a Banach-Mackey space. In particular, sequentially complete Hausdorff LCS are Banach-Mackey spaces (Theorem 5 and 13.18); this result is ofter referred to as the <u>Banach-Mackey Theorem</u> ([K1], 20.11(8); [W], 10.4.8).

Bounded Sets:

The topology $\beta^*(E, E')$ has another interesting property which is quite useful. Even though $\beta^*(E, E')$ can be strictly stronger than the Mackey topology (Example 4), these two topologies have exactly the same bounded sets.

Theorem 6. Let (E, τ) be a Hausdorff LCS. A subset $B \subseteq E$ is τ-bounded if and only if B is $\beta^*(E, E')$ bounded.

Proof: Let $B \subseteq E$ be τ-bounded. Let $A \subseteq E'$ be $\beta(E', E)$ bounded (equivalently, $\sigma(E', E'')$ bounded). Then A_0 is a basic $\beta^*(E, E')$ neighborhood of 0, and we show that B is absorbed by A_0. Now B^0 is a basic $\beta(E', E)$ neighborhood of 0 in E' so there is $t > 0$ such that $A \subseteq tB^0$. Then $A_0 \supseteq (1/t)(B^0)_0 \supseteq (1/t)B$ so B is absorbed by A_0.

Since $\beta^*(E, E')$ is stronger than τ (Theorem 3), the converse is clear.

Theorem 6 can be used to obtain a UBP which holds without any completeness assumptions on the domain space much in the spirit of the UBP of §9.

Theorem 7 (UBP). Let E and F be Hausdorff LCS and $\mathcal{F} \subseteq L(E, F)$. If $\mathcal{F}$ is pointwise bounded on E, then $\mathcal{F}$ is uniformly bounded on $\beta(E, E')$ (strongly) bounded subsets of E.

Proof: Let $B \subseteq E$ be $\beta(E, E')$ bounded. We show

$$\mathcal{F}B = \{Tx : T \in \mathcal{F}, x \in B\}$$

is $\sigma(F, F')$ bounded in F. For this let $y' \in F'$. Since $\mathcal{F}$ is pointwise bounded on E, the subset $\{y'T : T \in \mathcal{F}\} \subseteq E'$ is $\sigma(E', E)$ bounded and is, therefore, $\beta^*(E', E)$ bounded by Theorem 6. Thus,

$$\{\langle y'T, x\rangle : T \in \mathcal{F}, x \in B\}$$

is bounded and the result is established.

Exercise 1. Give an example where $\beta(E, E')$ is strictly stronger than $\beta^*(E, E')$ [Hint: 19.1].

Exercise 2. A LCS E is called a <u>Banach-Mackey space</u> if weakly bounded sets are strongly bounded ([W]). Show E is a Banach-Mackey space if and only if every barrel is a bornivore.

Exercise 3. Let $E = c_{00}$ and $E' = \ell^1$. Let $B = \{ke_k : k \in \mathbb{N}\} \cup \{0\}$. Show B is $\sigma(\ell^1, c_{00})$ compact but not $\beta(\ell^1, c_{00})$ bounded. Show $(B_0)^0$ is not $\sigma(\ell^1, c_{00})$ compact.

20.1 Perfect Sequence Spaces

A classic theorem due to Hellinger and Toeplitz asserts that if an infinite (scalar) matrix $A = [a_{ij}]$ is such that the formal matrix product $Ax = \{\sum_{j=1}^{\infty} a_{ij}x_j\}$ belongs to ℓ^2 for every $x = \{x_j\} \in \ell^2$, then $A : \ell^2 \to \ell^2$ is bounded ([HT]). In this section we use the general UBP 9.6 to give a generalization of the Hellinger-Toeplitz Theorem to more general sequence spaces than ℓ^2.

If X is a subspace of s, the α-dual of X is

$X^x = \{y = \{y_i\} \in s : \sum_{i=1}^{\infty} y_i x_i$ is absolutely convergent for every $x = \{x_i\} \in X\}$.

If $y \in X^x$, $x \in X$, we write $\langle y, x\rangle = \sum_{i=1}^{\infty} y_i x_i$; if X contains c_{00}, then X and its α-dual, X^x, are in duality with respect to $\langle\ ,\ \rangle$.

Example 1. $s^x = c_{00}$, $c_{00}^x = s$, $(c_0)^x = \ell^1$, $(\ell^1)^x = \ell^\infty$, $(\ell^\infty)^x = \ell^1$, $(\ell^p)^x = \ell^q$ when $1 < p < \infty$ and $1/p + 1/q = 1$.

X is said to be perfect if $X^{xx} = (X^x)^x = X$; c_{00}, s, $\ell^p (1 \le p \le \infty)$ are perfect while c_0 is not. X is said to be normal if $x = \{x_i\} \in X$ and $y = \{y_i\}$ satisfies $|y_i| \le |x_i|$ implies $y \in X$. We have the following elementary properties of α-duals.

Proposition 2. (i) If $X \subseteq Y \subseteq s$, then $X^x \supseteq Y^x$.

(ii) $X^{xx} \supseteq X$

(iii) X^x is perfect

(iv) X^x is normal and $X^x \supseteq c_{00}$.

(v) If X is perfect, then X is normal and $X \supseteq c_{00}$.

Proof: (i) and (ii) are clear. By (i) and (ii), $X^{xx} \supseteq X$ and $X^{xxx} \subseteq X^x$. By (ii) $X^{xxx} \supseteq X^x$ so (iii) holds. (iv) is clear and (iv) implies (v).

Henceforth, we assume that $X \supseteq c_{00}$. Then X and X^x are in duality. In terms of the weak topology, $\sigma(X, X^x)$, we have the following interesting result.

Theorem 3. X is perfect if and only if X is $\sigma(X, X^x)$ sequentially complete.

Proof: $\Rightarrow$: Let $\{x^i\} \subseteq X$ be $\sigma(X, X^x)$ Cauchy with $x^i = \{x_{ij}\}_{j=1}^{\infty}$. For each e_j, $\lim_i \langle e_j, x^i \rangle = \lim_i x_{ij} = x_j$ exists. Set $x = \{x_j\}$. For any $y = \{y_j\} \in X^x$ and any $\sigma \subseteq \mathbb{N}$, since X^x is normal (Proposition 2),

$$\lim_i \sum_{j \in \sigma} y_j x_{ij}$$

exists. By the Schur Theorem (16.14) $\{y_j x_j\} \in \ell^1$ and

$$\lim_i \sum_{j=1}^{\infty} y_j x_{ij} = \sum_{j=1}^{\infty} y_j x_j$$

so $x \in X^{xx} = X$ and $\{x^i\}$ is $\sigma(X, X^x)$ convergent to x.

$\Leftarrow$: Let $x = \{x_j\} \in X^{xx}$. For each $y = \{y_i\} \in X^x$,

$$\lim_n \sum_{i=1}^n y_i x_i = \lim_n \langle y, \sum_{i=1}^n x_i e_i \rangle$$

exists so $\{\sum_{i=1}^n x_i e_i\}$ is $\sigma(X, X^x)$ Cauchy in X and, therefore, converges to some sequence in X which must be x.

Corollary 4. If X is perfect, $(X, \sigma(X, X^x))$ is an $\mathscr{A}$-space, every $\sigma(X, X^x)$ bounded set is $\beta(X, X^x)$ bounded and $\beta(X^x, X) = \beta^*(X^x, X)$.

Proof: Theorem 3, 13.18 and 20.5.

We now give a generalization of the Hellinger-Toeplitz Theorem to certain sequence spaces. Let Y be a subspace of s containing c_{00}. Let $A = [a_{ij}]$ be a matrix belonging to the class (X, Y)

$$\text{(i.e., } Ax = \{\sum_{j=1}^{\infty} a_{ij} x_j\} \in Y \ \forall x \in X, \ \S 9.5).$$

Theorem 5. Let X be normal and such that $(X, \sigma(X, X^x))$ is an $\mathscr{A}$-space. If $B \subseteq X$ is $\sigma(X, X^x)$ bounded, then $AB \subseteq Y$ is $\sigma(Y, Y^x)$ bounded, i.e., A is $\sigma(X, X^x)$ - $\sigma(Y, Y^x)$ bounded.

Proof: Let $a_i = \{a_{ij}\}_{j=1}^{\infty}$ be the ith row of A. Define $A_n : X \to Y$ by $A_n x = \sum_{i=1}^n \langle a_i, x \rangle e_i$. Since X is normal, each $a_i \in X^x$ (Exer. 5) so

A_n is $\sigma(X, X^x)$ - $\sigma(Y, Y^x)$ continuous. For each $x \in X$, $\{A_n x\}$ is $\sigma(Y, Y^x)$ convergent to Ax (Exer. 3) so $\{A_n\}$ is pointwise bounded on X with respect to $\sigma(Y, Y^x)$. Since X is an $\mathcal{A}$-space, $\{A_n\}$ is uniformly bounded on $\sigma(X, X^x)$ bounded subsets of X (9.6), and the result follows.

By Corollary 4, Theorem 5 is applicable to any perfect space; however, the theorem is also applicable to the non-perfect space c_0 since $(c_0, \sigma(c_0, \ell^1))$ is an $\mathcal{A}$-space (14.20). Since the α-dual of ℓ^2 is ℓ^2, Theorem 5 gives the Hellinger-Toeplitz Theorem for ℓ^2, i.e., any $A = [a_{ij}] : \ell^2 \to \ell^2$ is bounded, and therefore, continuous. A more general version of the Hellinger-Toeplitz Theorem than given in Theorem 5 is established in [KT]. See also [K1, K2] for further information on perfect spaces and other versions of the Hellinger-Toeplitz Theorem.

Exercise 1. Show that ℓ^2 is the only self α-dual sequence space, i.e., the only sequence space satisfying $X^x = X$.

Exercise 2. Show the coordinate functionals $x = \{x_i\} \to x_i$ are $\sigma(X, X^x)$ continuous.

Exercise 3. If $X \supseteq c_{00}$, show $\{e_i : i \in \mathbb{N}\}$ is a Schauder basis in X with respect to $\sigma(X, X^x)$.

Exercise 4. Show $(m_0)^x = \ell^1$ so m_0 is not perfect or normal.

Exercise 5. If X is normal, show that $\{y_i\} \in X^x$ if and only if $\Sigma x_i y_i$

converges for every $\{x_i\} \in X$.

Exercise 6. If X and Y are perfect and $A \in (X, Y)$, show that

$$\{\langle y', Ax\rangle : x \in B_1, y' \in B_2\}$$

is bounded when $B_1 \subseteq X$ $(B_2 \subseteq Y^{x})$ is $\sigma(X, X^{x})$ $(\sigma(Y^{x}, Y))$ bounded (see [KT]).

Exercise 7. Show that X is $\sigma(X^{xx}, X^{x})$ sequentially dense in X^{xx}.

Exercise 8. If the conditions of Theorem 5 are satisfied, show that $\langle C, AB\rangle$ is bounded for every strongly bounded subset, C, of Y^{x}. Compare this statement with the original Hellinger-Toeplitz Theorem.

Exercise 9. If X is normal, show $(X^{x}, \sigma(X^{x}, X))$ is sequentially complete. [Hint: Use 16.14.]

21
Bornological Spaces

Any sequentially continuous linear map between TVS is bounded, but, in general, the converse does not hold. The locally convex spaces for which the converse does hold are called bornological spaces, and we now study some of their properties.

Recall that a subset B of a TVS E is a bornivore if it is absolutely convex and absorbs bounded sets (19.7). Any absolutely convex neighborhood of 0 in a TVS is a bornivore, and we say that a LCS E is bornological if every bornivore is a neighborhood of 0. We have the following characterization of bornological spaces.

Theorem 1. Let E be a LCS. The following are equivalent.

(i) E is bornological.

(ii) Every bounded semi-norm q on E is continuous [a semi-norm is bounded if it carries bounded sets to bounded sets].

(iii) Every bounded linear map T from E into an arbitrary LCS F is continuous.

Proof: (i) $\Rightarrow$ (ii): Let $\varepsilon > 0$. It suffices to show that q is continuous at 0 (2.4). If $A = \{x : q(x) \leq \varepsilon\}$, then A is absolutely convex and if $B \subseteq E$ is bounded, $\exists\ t > 0$ such that $q(x) \leq t$ $\forall\, x \in B$ so $B \subseteq (t/\varepsilon)A$. Thus, A is a bornivore and is, therefore, a neighborhood of 0.

(ii) $\Rightarrow$ (iii): Suppose $T : E \to F$ is bounded and linear. Let q be a continuous semi-norm on F. Then qT is a bounded semi-norm on E and is, therefore, continuous. Hence, T is continuous (Exer. 13.3).

(iii) $\Rightarrow$ (i): Suppose V is a bornivore in E. Then V is absolutely convex and absorbing so the Minkowski functional, $p_V = p$, of V is a semi-norm on E. Consider the identity map $I : E \to (E, p)$. If $B \subseteq E$ is bounded, $\exists\ t > 0$ such that $B \subseteq tV$. Since $V \subseteq \{x : p(x) \leq 1\}$, $p(x) \leq t$ $\forall\, x \in B$ so I is a bounded and, hence, continuous linear operator. Since $\{x : p(x) < 1\} \subseteq V$, V is a neighborhood of 0, and E is bornological.

It follows from Theorem 1 and 5.4 that any quasi-normed LCS is bornological. Later we will give an example of a non-metrizable bornological space (25.4.1).

For bornological spaces, we have the following criterion for continuity.

Corollary 2. Let E be a bornological space, F a LCS and $T : E \to F$ linear. The following are equivalent.

(i) T is continuous.

(ii) T is bounded.

(iii) T is sequentially continuous.

Proof: Theorem 1 and 5.3.

For the topology of a bornological space, we have

Proposition 3. If (E, τ) is a Hausdorff bornological space, then E has the topology $\beta^*(E, E')$. Therefore, $\tau = \tau(E, E') = \beta^*(E, E')$.

Proof: The identity map $(E, \tau) \to (E, \beta^*(X, X'))$ is bounded (20.6) and, therefore, continuous. Apply 20.3.

In particular, any metrizable LCS E always carries the topology $\beta^*(E, E')$ and, therefore, the Mackey topology, and is quasi-barrelled (20.1).

The strong dual of a bornological space is always complete.

Proposition 4. Let E be a Hausdorff bornological space. Then E'_b is complete.

Proof: Let $\{x'_\delta\}$ be a Cauchy net in E'_b. Let $\varepsilon > 0$ and B a bounded subset of E. Then $\forall x \in E$ $\{\langle x'_\delta, x\rangle\}$ is a Cauchy net so let $\langle x', x\rangle = \lim \langle x'_\delta, x\rangle$. Then x' is obviously linear; we show $x' \in E'$

and $x'_\delta \to x'$ in $\beta(E', E)$. $\exists\ \delta$ such that α, $\beta \geq \delta$ implies $\sup\{|<x'_\alpha - x'_\beta, x>| : x \in B\} < \varepsilon$. Thus, for $x \in B$, $\alpha \geq \delta$,

(1) $$|<x'_\alpha - x', x>| \leq \varepsilon .$$

$\forall\ x \in B$, and $|<x', x>| \leq \varepsilon + |<x'_\alpha, x>|$ so x' is bounded and, therefore, continuous. (1) implies that $x'_\delta \to x'$ in $\beta(E', E)$.

Compare this result with 5.8.

The bornological topology:

Let (E, τ) be a LCS. The family of all absolutely convex bornivores in E forms a base at 0 for a locally convex topology which we denote by τ^b (1.18). We show below that (E, τ^b) is bornological and $\tau = \tau^b$ if and only if (E, τ) is bornological. For this reason, τ^b is called the bornological topology associated with (E, τ) and (E, τ^b) is called the bornological space associated with (E, τ).

Proposition 5. (i) τ^b is the strongest locally convex topology which has the same bounded sets as (E, τ).

(ii) (E, τ^b) is bornological.

(iii) (E, τ) is bornological if and only if $\tau = \tau^b$.

Proof: (i): Any locally convex topology which has the same bounded sets as (E, τ) must have a neighborhood base at 0 which consists of absolutely convex bornivores.

(ii): Let F be a LCS and $T : E \to F$ a bounded linear map. If V

is an absolutely convex neighborhood of 0 in F, then $U = T^{-1}V$ is absolutely convex, and we claim that U is a bornivore so it must be a τ^b-neighborhood of 0. Let $B \subseteq E$ be bounded. Then TB is bounded so $\exists\ t > 0$ such that $TB \subseteq tV$. Thus, $B \subseteq tU$ and U is a bornivore. Since T is continuous with respect to τ^b, (E, τ^b) is bornological by Theorem 1.

(iii): If $\tau = \tau^b$, (E, τ) is bornological by (ii). If (E, τ) is bornological, the identity $(E, \tau) \to (E, \tau^b)$ is bounded so continuous and $\tau^b \subseteq \tau$. But, always $\tau^b \supseteq \tau$.

Note that we have $\beta^*(E, E') \subseteq \tau^b$ and $\tau^b = \beta^*(E, E')$ when E is bornological.

Exercise 1. Show "LCS F" in Theorem 1 (iii) can be replaced by "semi-NLS F."

Exercise 2. Let E, F be Hausdorff LCS with E bornological. If the linear map $T : E \to F$ is continuous with respect to $\sigma(E, E')$, $\sigma(F, F')$, show T is continuous with respect to the original topologies.

Exercise 3. Give an example of a non-bornological space.

Exercise 4. Show any bornological space is quasi-barrelled.

Exercise 5. Give an example of a sequentially continuous linear map which isn't continuous.

22

Inductive Limits

In this section we present a method of constructing a locally convex topology from a family of LCS and linear maps. Weak topologies (1.18) are another example of such a construction procedure.

Let E be a vector space and (E_α, τ_α) a LCS $\forall\, \alpha \in A$. Suppose $\forall\, \alpha \in A\ \exists$ a linear map $A_\alpha : E_\alpha \to E$ such that $E = \text{span} \bigcup_{\alpha \in A} A_\alpha E_\alpha$.

Proposition 1. $\exists$ a strongest locally convex topology τ on E such that all the linear maps A_α are continuous. A neighborhood base at 0 for τ is $\mathcal{U} = \{U \subseteq E : U$ is absolutely convex and $A_\alpha^{-1}U$ is a τ_α neighborhood of 0 in $E_\alpha\ \forall\, \alpha \in A\}$; the topology τ is generated by the family of semi-norms $\mathcal{P} = \{p : pA_\alpha = p_\alpha$ is a continuous semi-norm on $E_\alpha\ \forall\, \alpha \in A\}$.

Proof: Let τ be the locally convex topology generated by the semi- norms in $\mathcal{P}$, $\tau = \sigma(E, \mathcal{P})$. Then each A_α is continuous with respect to τ (Exer. 13.3). Let τ' be a locally convex topology on E such that each A_α is continuous with respect to τ' and suppose that $\tau' = \sigma(E, \mathcal{P}')$. If $p' \in \mathcal{P}'$, then $p'A_\alpha$ is continuous with respect to τ_α $\forall\, \alpha \in A$ (Exer. 13.3) so $p' \in \mathcal{P}$. Hence, $\mathcal{P}' \subseteq \mathcal{P}$ and $\sigma(E, \mathcal{P}) = \tau \supseteq \sigma(E, \mathcal{P}') = \tau'$.

The statement about the neighborhood base at 0 follows from the definition of $\sigma(E, \mathcal{P})$.

The topology τ of Proposition 1 is called the <u>inductive limit topology</u> (on E) and E is called the inductive limit of the spaces E_α and the maps A_α. We write $E = (E, \tau) = \underline{\text{ind}(E_\alpha, A_\alpha)}$.

For example, the quotient topology, E/M, is an inductive limit topology (§6 and Exer. 13.9).

Corollary 2. Let F be a LCS. A linear map $T : E \to F$ is continuous with respect to τ if and only if $TA_\alpha : E_\alpha \to F$ is continuous $\forall\, \alpha \in A$.

In Exercise 1 the reader is asked to show that the inductive limit of bornological spaces is bornological. In particular, the inductive limit of NLS is bornological; we show, conversely, that any bornological space is the inductive limit of NLS. Let E be a Hausdorff LCS and let $B \subseteq E$ be bounded and absolutely convex. Set $[B] = \text{span}\, B = \bigcup_{n=1}^{\infty} nB$ and note that B is an absolutely convex, absorbing subset of [B] (not in general absorbing in E). Therefore, the Minkowski functional p_B of B (in [B]) is a semi-norm on [B]. Moreover, p_B is actually a norm on [B]; for if U is an arbitrary neighborhood of 0 in E, there exists $t > 0$ such that $B \subseteq tU$ or $(1/t)B \subseteq U \cap [B]$. Therefore, the semi-norm topology on [B] induced by p_B is stronger than the induced topology from E and is, therefore, Hausdorff. This observation also shows that the injection $[B] \to E$ is continuous with respect to p_B. We now have the machinery in place to establish

Theorem 3. If (E, τ) is a Hausdorff, bornological LCS, then E is an inductive limit of NLS.

Proof: Let $\mathscr{B}$ be the family of all absolutely convex τ-bounded subsets of E. For $B \in \mathscr{B}$, let T_B be the continuous injection of $[B] \to E$. Note $E = \cup \mathscr{B}$ and set $(E, i) = \text{ind}([B], T_B)$. By Theorem 1 and the observation above $\tau \subseteq i$. Consider the identity map $I : (E, \tau) \to (E, i)$. If $A \subseteq E$ is τ-bounded, there exists $B \in \mathscr{B}$ such that $A \subseteq B$ so A is bounded in [B] and, hence, bounded in i. That is, I is a bounded map and is, therefore, continuous since E is bornological. Thus, $i = \tau$.

Definition 4. Let E be a vector space and $\{E_n\}$ an increasing sequence of linear subspaces of E with $E = \bigcup_{n=1}^{\infty} E_n$. Suppose each E_n is equipped with a locally convex topology τ_n such that the topology induced on E_n by τ_{n+1} is exactly τ_n. Then $E = \text{ind}\, E_n$ (the mapping $E_n \to E$ is the identity and is suppressed in the notation) is called the <u>strict inductive limit</u>

of the $\{E_n\}$.

Thus, τ is the strongest locally convex topology on E inducing on each E_n a locally convex topology weaker than τ_n.

Example 5. Let $E = c_{00}$ and $E_n = \{\{t_i\} : t_i = 0, i > n\}$. Give E_n the topology induced by the norm $\|\{t_i\}\| = \sum_{i=1}^{\infty} |t_i|$ (or any other convenient norm (7.1)). Then $\text{ind } E_n$ is a strict inductive limit.

Example 6. Let $\Omega \in \mathbb{R}^n$ be open. Let $\mathscr{D}(\Omega)$ be the vector space of all infinitely differentiable functions $\varphi : \Omega \to \mathbb{R}$ with compact support. Then $\mathscr{D}(\Omega) = \cup \mathscr{D}_K$, where the union is over all compact subsets K of Ω (Example 2.27). We give $\mathscr{D}(\Omega)$ the inductive limit topology, $\text{ind } \mathscr{D}_K$, where the map $\mathscr{D}_K \to \mathscr{D}(\Omega)$ is the identity. If $\{K_m\}$ is an increasing sequence of compact subsets of Ω with non-void interiors and $\Omega = \bigcup_{m=1}^{\infty} K_m$, then $\text{ind } \mathscr{D}_K = \text{ind } \mathscr{D}_{K_m}$ independent of the choice of $\{K_m\}$, and $\text{ind } \mathscr{D}_{K_m}$ is a strict inductive limit.

The space $\mathscr{D}(\Omega)$ is the space of test functions of L. Schwartz. It was the study of properties of this space which motivated the introduction of strict inductive limits. We will study some of its properties and those of its dual space in section 26.4.

A strict inductive limit of a sequence of complete metrizable locally convex spaces (B-spaces) is called an <u>LF-space</u> (<u>LB-space</u>). LF spaces are not in general metrizable but have many of the properties of metrizable spaces. The space in Example 5 is an LB-space; $\mathscr{D}(\Omega)$ in Example 6 is an LF-space.

We now establish some of the basic properties of strict inductive limits.

Lemma 7. Let H be a linear subspace of a LCS E and U an absolutely convex neighborhood of 0 in H with respect to the induced topology from E. Then (i) $\exists$ an absolutely convex neighborhood of 0, V, in E such that $V \cap H = U$ and (ii) if $y \notin \overline{H}$, then V in (i) can be chosen such that $y \notin V$.

Proof: $\exists$ an absolutely convex neighborhood of 0, W, in E such that $W \cap H \subseteq U$. Put $V = \text{abco}W \cup U$. Then $V \cap H \supseteq U$. On the other hand, if $x \in V \cap H$, then $x = sw + tu$ with $w \in W$, $u \in U$, $|s| + |t| \le 1$. If $s = 0$, $x = tu \in U$; if $s \ne 0$, $w = (x - tu)/s \in H$ so $w \in U$ and in this case $x \in U$. Therefore, $V \cap H = U$, and (i) holds.

For (ii), choose W as above such that $(y + W) \cap H = \emptyset$. Then if $y \in V$, $y = sw + tu$ with $w \in W$, $u \in U$, $|s| + |t| \le 1$. This implies

$$y - sw = tu \in (y + W) \cap U \subseteq (y + W) \cap H.$$

Hence, $y \notin V$.

Theorem 8. Let (E, τ) be the strict inductive limit of (E_n, τ_n). Then τ induces τ_n on each E_n.

Proof: The topology which τ induces on E_n is weaker than τ_n by Theorem 1. To show the induced topology on E_n is stronger than τ_n, it suffices to show that given an absolutely convex τ_n neighborhood of 0, U_n, $\exists$ a τ-neighborhood of 0, U, in E such that $U_n = U \cap E_n$.

Since τ_{n+1} induces τ_n on E_n by Lemma 7 $\exists$ an absolutely convex τ_{n+1} neighborhood of 0, U_{n+1}, in E_{n+1} with $U_n = U_{n+1} \cap E_n$. Continuing, $\forall\, r > 0$ $\exists$ an absolutely convex τ_{n+r} neighborhood of 0, U_{n+r}, in E_{n+r} with $U_{n+r} \cap E_{n+s} = U_{n+s}$ for $0 \le s \le r$. Set $U = \bigcup_{r=0}^{\infty} U_{n+r}$. Then $U \cap E_{n+r} = U_{n+r}$ for $r \ge 0$ and also for $m < n$, $U \cap E_m = U_n \cap E_m$ so that U is a τ neighborhood of 0 in E with the required property.

Theorem 9. The strict inductive limit, $E = \text{ind}\, E_n$, of a sequence of Hausdorff LCS is Haudsorff.

Proof: Let $x \in E$, $x \neq 0$. $\exists\, n$ such that $x \in E_n$. Since E_n is Hausdorff, $\exists$ a neighborhood of 0, U_n, in E_n with $x \notin U_n$. By Theorem 8, $\exists$ a neighborhood of 0, U, in E such that $U \cap E \subseteq U_n$. Thus, $x \notin U$, and E is Hausdorff.

Note Exercise 2.

Definition 10. The inductive limit $E = \text{ind}(E_\alpha, A_\alpha)$ is <u>regular</u> if whenever $B \subseteq E$ is bounded, then $\exists\, \alpha$ and $B_\alpha \subseteq E_\alpha$ bounded such that $A_\alpha B_\alpha \supseteq B$ [the converse always holds].

Not every inductive limit is regular ([Kom]), but we show below that certain strict inductive limits are regular. There are other conditions which insure that inductive limits (and inductive systems) are regular (see [FW]).

Lemma 11. Let $E = \text{ind } E_n$ be a strict inductive limit. Let $y_n \notin \bar{E}_n \ \forall n$. Then $y_n \not\to 0$.

Proof: We may assume that $y_n \in E_{n+1} \ \forall n$ (Exer. 3). By Theorem 8 y_n does not belong to the closure of E_n computed in E_{n+1}. By Lemma 7 $\exists$ absolutely convex neighborhoods of 0, V_n, in E_n such that $V_{n+1} \cap E_n = V_n$ and $y_n \notin V_{n+1}$ (take $V_1 = E_1$). Put $V = \bigcup_{n=1}^{\infty} V_n$. Then V is a neighborhood of 0 in E and $y_n \notin V \ \forall n$ so $y_n \not\to 0$.

Theorem 12. Let E be the strict inductive limit of $\{E_n\}$, where each E_n is closed in E_{n+1} (this is the case if each E_n is complete). Then $E = \text{ind } E_n$ is regular. If further $E_n \subsetneq E_{n+1} \ \forall n$, then E is not metrizable.

Proof: The first part follows from Lemma 11.

For the second part choose $y_n \in E_{n+1} \backslash E_n$, and suppose $| \ |$ is a quasi-norm which generates the topology of E. Pick t_n such that $|t_n y_n| < 1/n$. Then $t_n y_n \to 0$ contradicting Lemma 11.

A TVS is said to be quasicomplete if every bounded Cauchy net is convergent (Exercise 16.8).

Corollary 13. If E is the strict inductive limit of a sequence of quasicomplete spaces, then E is quasicomplete.

Proof: Theorems 8 and 12.

It is actually the case that the strict inductive limit of complete spaces is complete ([W]).

Example 14. $\mathscr{D}(\Omega)$ is an LF-space so the results above are applicable. In particular, $\mathscr{D}(\Omega)$ is a quasicomplete non-metrizable bornological space (Exer. 1). [In order to show that $\mathscr{D}(\Omega)$ is non-metrizable it is necessary to show that $\mathscr{D}_{K_m} \subsetneq \mathscr{D}_{K_{m+1}}$ in Example 6 so that Theorem 12 is applicable; this is shown to be the case in Lemma 26.4.8.] A sequence $\{\varphi_k\}$ in $\mathscr{D}(\Omega)$ converges to 0 if and only if $\exists$ a compact subset $K \subseteq \Omega$ such that the support of each φ_k is contained in K and $D^\alpha \varphi_k \to 0$ uniformly on K $\forall$ multi-index α. If α is a multi-index, the differential operator $D^\alpha : \mathscr{D}(\Omega) \to \mathscr{D}(\Omega)$ is thus a continuous operator.

Exercise 1. Show an inductive limit of bornological spaces is bornological and the quotient of a bornological space is bornological.

Exercise 2. Show the inductive limit of Hausdorff spaces needn't be Hausdorff.

Exercise 3. If $E = \text{ind}(E_n, A_n)$ and $\{n_k\}$ is a subsequence, show $E = \text{ind}(E_{n_k}, A_{n_k})$.

Exercise 4. Let $E = \text{ind}(E_\alpha, A_\alpha)$ and F be a LCS. Let $\mathcal{F} \subseteq L(E, F)$. Show that $\mathcal{F}$ is equicontinuous if and only if $\forall\, \alpha \in A$ the family $\mathcal{F}_\alpha = \{TA_\alpha : T \in \mathcal{F}\}$ is an equicontinuous subset of $L(E_\alpha, F)$.

Exercise 5. Let $\varphi \in \mathcal{D}(\mathbb{R}^n)$ be defined by

$$\psi(t) = \begin{cases} \exp(-(1/(1 - t^2)) & -1 < t < 1 \\ 0 & |t| \geq 1 \end{cases}$$

and $\varphi(x) = \varphi(x_1, \ldots, x_n) = \psi(x_1)\ldots\psi(x_n)$. Set $\varphi_k(x) = \varphi(x)/k$ and $\psi_k(x) = \varphi(x/k)/k$. Show $\varphi_k \to 0$ in $\mathcal{D}(\mathbb{R}^n)$, $D^\alpha \psi_k \to 0$ uniformly on $\mathbb{R}^n$ $\forall$ multi-index α, but $\psi_k \not\to 0$ in $\mathcal{D}(\mathbb{R}^n)$.

Exercise 6. Let $E = \text{ind}\, E_n$ be an LF space with $E_n \subsetneq E_{n+1}$. Show E is first catégory. Use this to show that E is not metrizable.

Exercise 7. If E is sequentially complete and $B \subseteq E$ is absolutely convex, closed and bounded, show $[B]$ is a B-space.

Exercise 8. Show any Hausdorff, sequentially complete bornological space is the inductive limit of B-spaces.

Exercise 9. For each $\alpha \in A$ let (E_α, τ_α) be a LCS whose topology is generated by the family of semi-norms $\mathcal{P}_\alpha$. Let E be a vector space and suppose that for each $\alpha \in A$ there is a linear map $A_\alpha : E \to E_\alpha$. Let τ be the locally convex topology τ on E generated by the semi-norms $\mathcal{P} = \{p_\alpha A_\alpha : p_\alpha \in \mathcal{P}_\alpha, \alpha \in A\}$; (E, τ) is called the projective limit of $\{E_\alpha, A_\alpha\}$ and is denoted by $E = \text{proj}(E_\alpha, A_\alpha)$. Show that τ is the weakest locally convex topology on E such that each A_α is continuous. Show that a linear map T from a LCS $F \to E$ is continuous if and only if each $A_\alpha T$ is continuous. If $\{A_\alpha : \alpha \in A\}$ separates the points of E and if each E_α is Hausdorff, show τ is Hausdorff.

Exercise 10. Show that any Hausdorff LCS is the projective limit of NLS. [Hint: Let $\mathcal{P}$ generate the topology of E. For $p \in \mathcal{P}$, let

$$N = N_p = \{x : p(x) = 0\}$$

and let $E_p = E/N$ with the quotient norm topology induced by p. Let $A_p : E \to E_p$ be the quotient map and consider $\text{proj}(E_p, A_p)$.]

Exercise 11. Suppose $E = \text{ind}\, E_k$ is strict. If $K \subseteq E$ is compact, show $\exists\, k$ such that $K \subseteq E_k$ and K is compact in E_k.

Part IV
Linear Operators

In part IV we study properties of linear operators between TVS. We begin by presenting a method of constructing locally convex topologies on the space of continuous linear operators between LCS which is very analogous to the construction of polar topologies.

23
Topologies on Spaces of Linear Operators

Let X, Y be LCS. Let $\mathscr{A}$ be a family of bounded (equivalently, $\sigma(X, X')$ bounded) subsets of X and let $\mathscr{Q}$ be the family of all continuous semi-norms on Y. Then the pair $(\mathscr{A}, \mathscr{Q})$ induces a locally convex topology on L(X, Y) via the family of semi-norms:

(1) $$p_{q,A}(T) = \sup\{q(Tx) : x \in A\}, \; q \in \mathscr{Q}, A \in \mathscr{A}.$$

We denote by $L_{\mathscr{A}}(X, Y)$, L(X, Y) with the locally convex topology generated by the semi-norms in (1); this suppresses the fact that the topology also depends upon the topology of Y. A net $\{T_\delta\}$ in L(X, Y) converges to 0 in $L_{\mathscr{A}}(X, Y)$ if and only if $\forall A \in \mathscr{A}$ $\lim T_\delta x = ($ uniformly for $x \in A$; for this reason $L_{\mathscr{A}}(X, Y)$ is called the topology o uniform $\mathscr{A}$-convergence (compare with §17). If $\mathscr{V}$ is a neighborhooc base at 0 in Y, then a base for $L_{\mathscr{A}}(X, Y)$ is given by $W_{A,V} = \{T \in L(X, Y) : TA \subseteq V\}$ for $A \in \mathscr{A}$, $V \in \mathscr{V}$. Also, if $X = \cup\{A : A \in \mathscr{A}\}$ and Y is Hausdorff, then $L_{\mathscr{A}}(X, Y)$ is Hausdorff [if $T \neq 0, \exists\, x \in X$ such that $Tx \neq 0$ and $\exists\, q \in \mathscr{Q}$ such that $q(Tx) \neq 0$].

Example 1. Let $\mathscr{A}$ be the family of all finite subsets of X. In this case, $L_{\mathscr{A}}(X, Y)$ is called the topology of pointwise convergence on X and is denoted by $L_s(X, Y)$.

Example 2. Let $\mathscr{A}$ be the family of all precompact subsets of X (13.19). In this case $L_{\mathscr{A}}(X, Y)$ is denoted by $L_{pc}(X, Y)$.

Example 3. Let $\mathscr{A}$ be the family of all compact subsets of X. In this case $L_{\mathscr{A}}(X, Y)$ is denoted by $L_c(X, Y)$.

Example 4. Let $\mathscr{A}$ be the family of all bounded subsets of X. In this case $L_{\mathscr{A}}(X, Y)$ is denoted by $L_b(X, Y)$ and is called the topology of bounded convergence (compare with §19).

Remark 5. Let X, Y be NLS. Then $L_b(X, Y)$ is just the norm topology induced by the operator norm (§5) and is often called the uniform operator topology. The topology $L_s(X, Y)$ when Y has the norm topology is called the <u>strong operator topology</u>; the topology $L_s(X, Y)$ when Y has the weak topology is called the <u>weak operator topology</u>. Thus, a net $\{T_\delta\} \subseteq L(X, Y)$ converges to 0 in the uniform operator topology (strong operator topology; weak operator topology, respectively) if and only if

$$\|T_\delta\| \to 0 \; (\|T_\delta x\| \to 0 \;\; \forall\, x \in X \,;\, \langle y', T_\delta x\rangle \to 0 \;\; \forall\, y' \in Y', x \in X).$$

For equicontinuous sets of continuous linear operators, we have

Theorem 6. Let $\mathcal{F} \subseteq L(X, Y)$ be equicontinuous. The topologies $L_s(X, Y)$ and $L_{pc}(X, Y)$ coincide on $\mathcal{F}$.

Proof: It suffices to show that the identity map from $\mathcal{F}$ with $L_s(X, Y)$ onto $\mathcal{F}$ with $L_{pc}(X, Y)$ is continuous. Let $K \subseteq X$ be precompact and q a continuous semi-norm on Y. Since $\mathcal{F}$ is equicontinuous, $\exists$ a continuous semi-norm p on X such that $q(Tx) \le p(x) \;\; \forall\, T \in \mathcal{F}, x \in X$ (13.8). Let $\{T_\delta\}$ be a net in $\mathcal{F}$ which converges to $T \in \mathcal{F}$ with respect to $L_s(X, Y)$. Let $\varepsilon > 0$. If $U = \{x : p(x) \le \varepsilon/3\}$, then $\exists$ a finite set $F \subseteq X$ such that $K \subseteq F + U$. There exists δ such that $\alpha \ge \delta$ implies $q((T_\alpha - T)z) < \varepsilon/3$ for $z \in F$. If $x \in K$, $x = z + u$ for some $z \in F$, $u \in U$ so if $\alpha \ge \delta$,

$$q((T_\alpha - T)x) \le q((T_\alpha - T)z) + q(T_\alpha u) + q(Tu) < \varepsilon\,.$$

Thus, $\{T_\delta\}$ converges to T uniformly on precompact subsets of X.

Proposition 7. If $\mathcal{F} \subseteq L(X, Y)$ is equicontinuous, then $\mathcal{F}$ is bounded in

$L_b(X, Y)$.

Proof: If q is a continuous semi-norm on Y, $\exists$ a continuous semi-norm p on X such that $q(Tx) \le p(x)$ $\forall\, T \in \mathcal{T}$, $x \in X$. If $B \subseteq X$ is bounded, $\sup\{q(Tx) : x \in B\} = p_{q,B}(T) \le \sup\{p(x) : x \in B\} < \infty$ $\forall T \in \mathcal{T}$.

Compare this with 19.5; Example 19.6 shows the converse of Proposition 7 is false.

Invertible Operators:

Let X be a B-space and set $L(X) = L(X, X)$. Then $L(X)$ is an algebra with multiplication of operators being composition. Moreover, we have the operator inequality, $\|ST\| \le \|S\|\|T\|$ $\forall$ $S, T \in L(X)$ (5.7). We say that $T \in L(X)$ is <u>invertible</u> if T^{-1} exists and belongs to $L(X)$. If T is 1-1 and onto X, T is invertible by the OMT. We give a sufficient condition for invertibility.

Theorem 8. Suppose $T \in L(X)$ has the property that the series $\sum_{n=0}^{\infty} T^n$ converges in norm in $L(X)$ [here $T^0 = I$, the identity on X]. Then $I - T$ is invertible with $(I - T)^{-1} = \sum_{n=0}^{\infty} T^n$.

Proof: If $S = \sum_{n=0}^{\infty} T^n$, then $ST = TS = \sum_{n=0}^{\infty} T^{n+1}$ so

$(I - T)S = S(I - T) = I.$

The series $\sum_{n=0}^{\infty} T^n$ is called the Neumann series for $(I - T)^{-1}$. A sufficient condition for its convergence is $\|T\| < 1$ since in this case $\|T^n\| \le \|T\|^n$ and the series ΣT^n is absolutely convergent (5.7 and 2.9). Thus, every element in the open sphere of radius 1 and center I is invertible. Note in this case the analogy with the complex series

$$1/(1 - z) = \sum_{n=0}^{\infty} z^n, \ |z| < 1.$$

We can improve the obvious sufficient condition, $\|T\| < 1$, for the convergence of the Neumann series. We show $\|T\| < 1$ can be replaced by the condition, $\lim \sqrt[n]{\|T^n\|} < 1$.

Lemma 9. Let $\{a_n\} \subseteq \mathbb{R}$ satisfy $0 \le a_{n+m} \le a_n a_m \ \forall \ m, n \in \mathbb{N}$. Then $\{\sqrt[n]{a_n}\}$ converges to $a = \inf\{\sqrt[n]{a_n} : n\}$.

Proof: Let $\varepsilon > 0$. Choose m such that $\sqrt[m]{a_m} < a + \varepsilon$. For $n > m$, write $n = qm + r$, where $0 \le r \le m - 1$. Set $M = \max\{a_r : 0 \le r \le m - 1\}$. Then $a_n \le a_m \cdots a_m a_r = (a_m)^q a_r$ so

$$a \le \sqrt[n]{a_n} \le \sqrt[n]{a_r}(a_m^{1/m})^{\frac{qm}{qm+r}} \le \sqrt[n]{M}(a + \varepsilon)^{\frac{qm}{qm+r}} \longrightarrow a + \varepsilon \text{ as } n \to \infty$$

and the result follows.

Corollary 10. Let $T \in L(X)$. Then $\lim \sqrt[n]{\|T^n\|}$ exists and $\lim \sqrt[n]{\|T^n\|} \le \|T\|$.

Proof: Set $a_n = \|T^n\|$ and apply Lemma 9.

The operator in Exer. 5 shows that strict inequality can occur in Corollary 10.

Exactly as in the classical scalar case one has the root test to determine absolute convergence of a series of operators (Exer. 7).

Theorem 11 (Root Test). Let $T_n \in L(X, Y)$ and set $r = \overline{\lim} \sqrt[n]{\|T_n\|}$. If $r < 1$, the series ΣT_n is absolutely convergent and if $r > 1$, the series ΣT_n does not converge (in fact, $\|T_n\| \not\to 0$).

From Corollary 10 and Theorem 11, a sufficient condition for the convergence of the Neumann series is that $\lim \sqrt[n]{\|T^n\|} < 1$. The operator in Exer. 5 shows that this condition improves the sufficient condition, $\|T\| < 1$.

Corollary 12. If $\lim \sqrt[n]{\|T^n\|} < 1$, the Neumann series $\sum_{n=0}^{\infty} T^n$ converges in norm to $(I - T)^{-1}$.

Theorem 13. Let $\mathcal{G}$ be the set of invertible elements in $L(X)$. If $T \in \mathcal{G}$, $S \in L(X)$ and $\|T - S\| < 1/\|T^{-1}\|$, then $S \in \mathcal{G}$ with

(2) $$\|S^{-1}\| \le \|T^{-1}\|/(1 - \|T^{-1}\|\|T - S\|) ,$$

(3) $$\|S^{-1} - T^{-1}\| \le \|T^{-1}\|^2\|T - S\|/(1 - \|T^{-1}\|\|T - S\|) .$$

In particular, $\mathcal{G}$ is open in L(X) and the map $T \to T^{-1}$ is continuous on $\mathcal{G}$.

Proof: $S = T - (T - S) = T(I - T^{-1}(T - S))$ and since $\|T^{-1}(T - S)\| < 1$, Theorem 8 implies that $I - T^{-1}(T - S) \in \mathcal{G}$ with

$$(I - T^{-1}(T - S))^{-1} = \sum_{n=0}^{\infty} [T^{-1}(T - S)]^n .$$

Thus, $S \in \mathcal{G}$ with $S^{-1} = \sum_{n=0}^{\infty} [T^{-1}(T - S)]^n T^{-1}$ so (2) follows. For (3),

$$\|S^{-1} - T^{-1}\| = \|(S^{-1}T - I)T^{-1}\| \le \|T^{-1}\|\|S^{-1}T - I\|$$

$$= \|T^{-1}\|\|I - \sum_{n=0}^{\infty} [T^{-1}(T - S)]^n\| = \|T^{-1}\|\|\sum_{n=1}^{\infty} [T^{-1}(T - S)]^n\|$$

$$\le \|T^{-1}\|^2\|T - S\|/(1 - \|T^{-1}\|\|T - S\|) .$$

The first statement shows that $\mathcal{G}$ is open and the continuity of $T \to T^{-1}$ on $\mathcal{G}$ follows from (3).

Theorem 8 can be used to solve operator equations of the form $(I - T)x = y$ where $T \in L(X)$ and y is given. The solution is, of course, $x = (I - T)^{-1}y$. We give examples of such situations in integral equations.

Let $X = C[a, b]$ equipped with the sup-norm. Let

$$k : [a, b] \times [a, b] \to \mathbb{R}$$

be continuous and for convenience set $M = \|k\|_\infty$. Define an operator

$K : X \to X$ by $Kf(s) = \int_a^b k(s, t)f(t)dt$. An operator of this form is called a <u>Fredholm integral operator</u> and k is called its <u>kernel</u>. Then $\|K\| \le M(b - a)$, and we assume, henceforth, that $M(b - a) < 1$ so by Theorem 8 $(I - K)$ is invertible with $(I - K)^{-1} = \sum_{n=0}^{\infty} K^n$.

We show that each operator K^n is also a Fredholm integral operator, and we compute its kernel. Now

$$K^2f(s) = \int_a^b k(s, t)Kf(t)dt = \int_a^b k(s, t)\int_a^b k(t, u)f(u)du\, dt$$
$$= \int_a^b \left(\int_a^b k(s, t)k(t, u)dt\right)f(u)du = \int_a^b k_2(s, u)f(u)du ,$$

where

$$k_2(s, u) = \int_a^b k(s, t)k(t, u)dt .$$

Thus, k_2 is the kernel of K^2.. Similarly, K^n is determined by the kernel

$$k_n(s, u) = \int_a^b k(s, t)k_{n-1}(t, u)dt .$$

It is easily checked by induction that $|k_n(s, u)| \le M^n(b - a)^{n-1}$ for $n > 1$.

If g is a given function in X, the integral equation

$$f(s) - \int_a^b k(s, t)f(t)dt = g(s)$$

has the solution

$$f = (I - K)^{-1}g = \sum_{n=0}^{\infty} K^n g = g + \sum_{n=1}^{\infty} K^n g .$$

We next show that the operator $\sum_{n=1}^{\infty} K^n$ is a Fredholm integral operator and compute its kernel. Since $|k_n(s, u)| \le M^n(b - a)^{n-1}$ for $n > 1$, the series $\sum_{n=1}^{\infty} k_n(s, u)$ converges uniformly on $[a, b] \times [a, b]$ to a continuous function h. Thus,

$$f(s) = g(s) + \sum_{n=1}^{\infty} \int_a^b k_n(s, u)g(u)du = g(s) + \int_a^b h(s, u)g(u)du$$

or $f = g + Hg$, where H is the integral operator induced by the kernel h. That is, we have $(I - K)^{-1} = I + H$. The kernel h of H has been called the <u>reciprocal kernel</u>.

There is another type of integral operator often encountered, the Volterra operators which we now consider. Let $k : [a, b] \times [a, b] \to \mathbb{R}$ be continuous on the triangular region $\{(s, t) : a \le t \le s \le b\}$ and be 0 for $t > s$. Define $K : X \to X$ by $Kf(s) = \int_a^s k(s, t)f(t)dt$ (note the variable limit of integration in contrast to the Fredholm operator). An operator of this type is called a <u>Volterra integral operator</u> and k is called its <u>kernel</u>. The development above can also be carried out for Volterra operators. There is one major difference; it is not necessary to make any assumptions about bounds on the kernel.

One can verify as above that each K^n is a Volterra operator with its kernel, k_n, given by $k_n(s, t) = \int_t^s k(s, u)k_{n-1}(u, t)du$, $n > 1$, $t \le s$. It is easily established by induction that $|k_n(s, t)| \le M^n(s - t)^{n-1}/(n - 1)!$, where again $M = \|k\|_\infty$. In particular, $|k_n(s, t)| \le M^n(b - a)^{n-1}/(n - 1)!$ so the series $\sum_{n=1}^{\infty} k_n(s, t)$ again converges uniformly on the square $[a, b] \times [a, b]$ to a function h which is continuous for $t \le s$ and 0 for $s < t$. Since $\|K^n\| \le M^n(b - a)^n/n!$, the series $\sum_{n=0}^{\infty} K^n$ converges in $L(X)$ and by Theorem 8, the operator $I - K$ is invertible in $L(X)$. Therefore, the integral equation $f(s) - \int_a^s k(s, t)f(t)dt = g(s)$ has the unique solution,

$$
\begin{aligned}
f(t) &= \sum_{n=0}^{\infty} K^n g(t) = g(t) + \sum_{n=1}^{\infty} \int_a^t k_n(t, s)g(s)ds \\
&= g(t) + \int_a^t h(t, s)g(s)ds = g(t) + Hg(t),
\end{aligned}
$$

where H is the Volterra operator induced by h.

Similar developments can be carried out for L^2-kernels; see [TL].

Existence of a Bounded Inverse:

In order that a linear operator $T : X \to X$, X a NLS, have a continuous inverse, we have the following criterion.

Theorem 14. T has a bounded inverse (on $\mathscr{R}T$) if and only if $\exists\, c > 0$ such that $\|Tx\| \ge c\|x\|$ $\forall\, x \in X$.

Proof: If T has a bounded inverse, then $\|x\| = \|T^{-1}Tx\| \le \|T^{-1}\|\|Tx\|$ $\forall\, x \in X$ so set $c = 1/\|T^{-1}\|$.

Conversely, if the inequality holds, T is certainly 1-1 and $\|T^{-1}Tx\| = \|x\| \le (1/c)\|Tx\|$ implies T^{-1} is continuous on $\mathscr{R}T$.

Exercise 1. Let X be a B-space. Show the three topologies in Remark 5 all have the same bounded sets.

Exercise 2. Describe the dual of L(X, Y) with the weak operator topology. [Hint: This topology is a weak topology.]

Exercise 3. Let X, Y be B-spaces. Show Y is reflexive if and only if the closed unit ball of L(X, Y) is compact in the weak operator topology.

Exercise 4. Let X be bornological and Y a complete Hausdorff LCS. Show $L_b(X, Y)$ is complete (compare 21.4).

Exercise 5. Define $T : C[0, 1] \to C[0, 1]$ by $Tf(t) = \int_0^1 tf(s)ds$. Show $\|T\| = 1$ and $\lim \sqrt[n]{\|T^n\|} = 1/2$.

Exercise 6. Let X, Y be B-spaces and let $T \in L(X, Y)$ be 1-1. Show T has closed range if and only if there exists $c > 0$ such that $c\|x\| \le \|Tx\|$ $\forall x \in X$.

Exercise 7. Prove Theorem 11.

23.1 The Krein-Smulian Theorem

We digress from the study of linear operators to establish the Krein-Smulian Theorem. This result gives a useful criterion for the weak* closedness of convex subsets in the dual of a locally convex F-space. The proof uses the topology of uniform convergence on precompact sets and Theorem 23.6. First we establish an important result of Banach and Dieudonne. We require two preliminary lemmas.

Throughout this section let E be a Hausdorff LCS.

Lemma 1. Let U be a closed, absolutely convex neighborhood of 0 in E and $\mathcal{F}$ the family of finite subsets of U. Then $\cap\{F^0 : F \in \mathcal{F}\} = U^0$.

Proof: Set $A = \cap\{F^0 : F \in \mathcal{F}\}$. Then $U^0 \subseteq A^0$. If $x' \notin U^0$, $\exists$ $x \in U$ such that $|\langle x', x\rangle| > 1$ so $x' \notin A^0$.

Let τ_f be the strongest topology on E' which agrees with $\sigma(E', E)$ on the equicontinuous subsets of $E'[\tau_f = \{A : A \cap M$ is $\sigma(E', E)$ open in M $\forall$ equicontinuous $M \subseteq E'\}$; τ_f is easily checked to be a topology]. Since $\sigma(E', E)$ is translation invariant and equicontinuous subsets of E' are translation invariant, τ_f is translation invariant. In general, however, τ_f is not a vector topology ([Kom]). Note that any $\sigma(E', E)$ closed set is certainly τ_f closed, but the converse does not hold.

Example 2. Let $E = c_{00}$. Set $A = \{ke_k : k \in \mathbb{N}\}$. Then A is not $\sigma(E', E)$ closed since $ke_k \to 0$ weak* but $0 \notin A$. However, A is τ_f

closed since any equicontinuous subset M of $E' = \ell^1$ is $\| \|_1$ bounded and $A \cap M$ is, therefore, finite.

Lemma 3. Let E be metrizable with $\{U_n\}_{n=0}^{\infty}$ a neighborhood base at 0 of closed absolutely convex sets with $U_0 = E$ and $U_n \supseteq U_{n+1}$. Let G be an open τ_f neighborhood of 0. Then $\exists$ a sequence $\{F_n\}_{n=0}^{\infty}$ of finite subsets of E such that

(i) $F_n \subseteq U_n$ $(n \geq 1)$

and

(ii) $(\bigcup_{k=1}^{n-1} F_k)^0 \cap U_n^0 \subseteq G$ $(n \in \mathbb{N})$.

Proof: We use induction on n. Set $\bigcup_{k=1}^{n-1} F_k = H_n$. Suppose $F_k \subseteq U_k$ for $k = 1, \ldots, m - 1$ have been chosen satisfying (ii) (we suppose nothing for $n = 0$). We must show $\exists$ a finite set $F_m \subseteq U_m$ such that

$$(F_m \cup H_{m-1})^0 \cap U_m^0 = H_m^0 \cap U_m^0 \subseteq G .$$

Suppose this fails for every finite set $F \subseteq U_m$. $E' \backslash G$ is τ_f closed in E' and U_m^0 is equicontinuous in E' so $G_1 = (E' \backslash G) \cap U_m^0$ is $\sigma(E', E)$ closed and, therefore, $\sigma(E', E)$ compact (Banach-Alaoglu). If $(F \cup H_{m-1})^0 \cap G_1 \neq \emptyset$ for every finite set $F \subseteq U_m$, then $\exists$ $x' \in \cap\{(F \cup H_{m-1})^0 \cap G_1 : F \subseteq U_m$ finite$\}$. By Lemma 1, $x' \in U_m^0 \cap H_{m-1}^0 \cap G_1$. But $U_m^0 \cap H_{m-1}^0 \subseteq G$ by the induction hypothesis and $G_1 \subseteq E' \backslash G$. This contradiction establishes the result.

Let τ_{pc} (τ_c) be the topology on E' of uniform convergence on the precompact (compact) subsets of E.

Theorem 4 (Banach-Dieudonné). Let E be metrizable and $\mathcal{N}$ be the family of all null sequences in E. Then $\tau_{\mathcal{N}} = \tau_{pc} = \tau_f$.

Proof: The closed, absolutely convex hull of the range of any null sequence is precompact (13.20) so $\tau_{\mathcal{N}} \subseteq \tau_{pc}$ (Exer. 17.1). By 23.6, $\tau_f \supseteq \tau_{pc}$ so it suffices to show $\tau_{\mathcal{N}} \supseteq \tau_f$. Since τ_f is translation invariant, this follows from Lemma 3. For if G is an open τ_f neighborhood of 0, set $S = \bigcup_{n=1}^{\infty} F_n$ with the notation of Lemma 3. Then S is the range of a null sequence in E and $S^0 \cap U_n^0 \subseteq G$ $\forall n$ so $S^0 \subseteq G$ since $\bigcup_{n=1}^{\infty} U_n^0 = E'$. Since $S^0 \in \tau_{\mathcal{N}}$, the result follows.

Corollary 5. Let E be metrizable. Every precompact subset of E is contained in the closed, absolutely convex hull of a null sequence.

Remark 6. For a much sharper description of precompact subsets of a metrizable LCS, see [RR] p. 133 or 28.3.2.

Corollary 7. Suppose that E is complete and metrizable. Then $\tau_f = \tau_{pc} = \tau_c$ and τ_f is compatible with the duality between E' and E.

Proof: Since the closures of precompact sets in a complete metric space are compact, $\tau_c = \tau_{pc}$ and $\tau_f = \tau_{pc}$ by Theorem 4.

If $A \subseteq E$ is precompact, then abcoA is precompact (13.20) so $\overline{\text{abco}}A$ is compact and, in particular, $\sigma(E, E')$ compact. Therefore, $\tau_{pc} \subseteq \tau(E', E)$ (Exer. 17.1).

Theorem 8 (Krein-Smulian). Let E be metrizable and complete. A convex subset $B \subseteq E'$ is $\sigma(E', E)$ closed if and only if $B \cap U^0$ is $\sigma(E', E)$ closed $\forall$ neighborhood of 0, U, in E.

Proof: ⇒: clear.

⇐: Every equicontinuous subset of E' is contained in the polar of some neighborhood of 0 in E so B is τ_f closed in E'. But, τ_f is compatible for the duality between E' and E so B is $\sigma(E', E)$ closed (14.10).

Note that Example 2 shows that completeness cannot be dropped.

Exercise 1. Let X be a B-space.

(a) Show that a convex subset $B \subseteq X'$ is weak* closed if and only if $B \cap S'_r$ is weak* closed $\forall\, r > 0$, where $S'_r = \{x' \in X' : \|x'\| \leq r\}$.

(b) Show that a linear subspace L of X' is weak* closed if and only if $L \cap S'_1$ is weak*-closed.

24
Barrelled Spaces

In this section we study an important class of LCS, the barrelled spaces. The barrelled spaces are the natural class of LCS for which the equicontinuity version of the Uniform Boundedness Principle holds.

Let E be a LCS. Recall that a barrel in E is a closed, absolutely convex, absorbing set (19.8). A closed, absolutely convex neighborhood of 0 in a LCS is always a barrel. We say that a LCS is barrelled if the converse holds.

Definition 1. E is a barrelled LCS if every barrel in E is a neighborhood of 0.

Example 2. A complete, metrizable LCS is barrelled. Let $B \subseteq E$ be a barrel. Then $E = \bigcup_{k=1}^{\infty} kB$. By the Baire Category Theorem some kB is somewhere dense so B contains an open set of the form $x_0 + U$, where

$x_0 \in B$ and U is a balanced open neighborhood of 0. Then, $\text{abco}(x_0 + U) \subseteq B$. If $y \in U$, then $y = \frac{1}{2}(x_0 + y) - \frac{1}{2}(x_0 - y) \in B$ so $U \subseteq B$. Note that we have only used the fact that E is second category; such spaces are called <u>Baire spaces</u>. There are examples of normed Baire spaces which are not complete ([NB] p. 281 or Exer. 13). (See also Exer. 12.)

We give an example of a NLS which is not barrelled showing the importance of the completeness in Example 2. There are, however, non-complete normed spaces which are barrelled; for example, the space m_0 is a non-complete (non-Baire!) normed space which is barrelled (see Exer. 6).

Example 3. Let $E = c_{00}$. Let $B = \{\{t_j\} : |t_j| \le 1/j \ \forall j\}$. Then B is a barrel in E but is not a norm neighborhood of 0 in E since $2e_k/k \notin B$ but $2e_k/k \to 0$ in E.

Proposition 4. Let (E, τ) be a Hausdorff LCS. Then $B \subseteq E$ is a barrel if and only if B is the polar of a $\sigma(E', E)$ bounded subset of E'.

Proof: The polar of a $\sigma(E', E)$ bounded subset of E' is absolutely convex and $\sigma(E, E')$ (hence, τ) closed (15.3) and absorbing (15.4). Hence, such a polar is a barrel.

If B is a barrel, $B = (B^0)_0$ by the Bipolar Theorem and B^0 is $\sigma(E', E)$ bounded (15.4).

Corollary 5. Let E be a Hausdorff LCS. Then E is barrelled if and only if every subset of E' which is $\sigma(E', E)$ bounded is equicontinuous.

Proof: If $A \subseteq E'$ is $\sigma(E', E)$ bounded, then A_0 is a barrel by Proposition 4 and, therefore, a neighborhood of 0 if E is barrelled. Hence, $(A_0)^0$ is equicontinuous (17.6), and $A \subseteq (A_0)^0$ is equicontinuous.

Let $B \subseteq E$ be a barrel. Therefore, $B = A_0$, where A is $\sigma(E', E)$ bounded, by Proposition 4. But, A is equicontinuous so $B = A_0$ is a neighborhood of 0 in E (17.7).

Corollary 6. Let E be a Hausdorff barrelled LCS. The following collections of subsets of E' are identical:

(i) equicontinuous sets,

(ii) relatively $\sigma(E', E)$ compact sets,

(iii) $\beta(E', E)$ bounded sets (i.e., strongly bounded sets),

(iv) $\sigma(E', E)$ bounded sets (i.e., weak* bounded sets).

Proof: (i) ⇒ (iii) by 19.5; (iii) ⇒ (iv) by 19.3; (iv) ⇒ (i) by Corollary 5; (i) ⇒ (ii) by the Banach-Alaoglu Theorem and 17.6; (ii) ⇒ (iv) always.

Corollary 7. Let E be a Hausdorff LCS. Then E is barrelled if and only if E carries the strong topology $\beta(E, E')$.

Proof: 17.7 and the definition of the strong topology along with Corollary 6.

We have the following criterion for reflexivity (§19).

Theorem 8. Let (E, τ) be a Hausdorff LCTVS. Then E is reflexive if and only if E is semi-reflexive and barrelled.

Proof: Since a barrelled space is quasi-barrelled (19.13), a barrelled, semi-reflexive space is reflexive (19.14). On the other hand, if E is reflexive, then $\beta(E'', E') = \beta(E, E') = \tau$ by 19.12, 19.14. E is barrelled by Corollary 7.

Proposition 9. The strong dual, E'_b, of a semi-reflexive space is barrelled.

Proof: Let B be a barrel in E'_b. Then B^0 is

$$(E'', \sigma(E'', E')) = (E, \sigma(E, E'))$$

bounded (Proposition 4). Hence, $(B^0)_0 = B$ is a strong neighborhood of 0 in E'_b.

The converse is false; take the dual of a non-reflexive NLS.

Proposition 10. The strong dual, E'_b, of a reflexive LCS is reflexive.

Proof: E'_b is barrelled by Proposition 9. Let $M \subseteq E'$ be $\sigma(E', E'')$ closed and bounded. Since E is reflexive, M is $\sigma(E', E)$ closed and bounded. M is $\sigma(E', E)$ compact by Corollary 6 and Theorem

8. Hence, E'_b is semi-reflexive (19.11), and E'_b is reflexive by Theorem 8.

UBP:

One version of the UBP established in 9.8 asserted that a pointwise bounded family of continuous linear operators on a quasi-normed $\mathcal{K}$-space is equicontinuous. It follows from Corollary 5 that if a Hausdorff LCS E is to satisfy this conclusion for continuous linear functionals it must be necessarily barrelled. We show that barrelled spaces satisfy this version of the UBP and so are the natural class of LCS for which this version of the UBP holds.

Theorem 11 (UBP). Let E be barrelled and F a LCS. If $\mathcal{F} \subseteq L(E, F)$ is pointwise bounded on E, then $\mathcal{F}$ is equicontinuous.

Proof: Let V be a closed, absolutely convex neighborhood of 0 in F. Then $U = \cap\{T^{-1}V : T \in \mathcal{F}\}$ is absolutely convex and closed. U is also absorbing since if $x \in E$, $\exists\ t > 0$ such that $\{Tx : T \in \mathcal{F}\} \subseteq tV$ so $x \in tU$. Hence, U is a neighborhood of 0 and $\mathcal{F}$ is equicontinuous.

As a corollary of this UBP, we obtain a version of the Banach-Steinhaus Theorem (9.11).

Theorem 12 (Banach-Steinhaus). Let E be barrelled and F a LCS. If $\{T_k\} \subseteq L(E, F)$ is such that $\lim T_k x = Tx$ exists $\forall\ x \in E$, then $T \in L(E, F)$ and $\lim T_k x = Tx$ converges uniformly for x in any precompact subset of E.

Proof: $\{T_k\}$ is equicontinuous by Theorem 11 so if V is a closed neighborhood of 0 in F, ∃ a neighborhood of 0, U, in E such that $T_k U \subseteq V$ ∀ k. Hence, $TU \subseteq V$ and T is continuous. The second statement follows from 23.6.

Corollary 13. A quasi-barrelled space E is barrelled if and only if $(E', \sigma(E', E))$ is sequentially complete.

Proof: If E is barrelled, $\sigma(E', E)$ is sequentially complete by Theorem 12.

If $\sigma(E', E)$ is sequentially complete, $(E', \sigma(E', E))$ is an $\mathscr{A}$-space (13.18) so $\beta^*(E, E') = \beta(E, E')$ by 20.5. If E is quasi-barrelled, its topology is $\beta^*(E, E')$ (20.1) so E has the strong topology, $\beta(E, E')$, and is barrelled by Corollary 7.

Bilinear Maps

In 9.1.6 we established a general hypocontinuity result for bilinear maps. We now establish a similar hypocontinuity result for barrelled spaces. Let E, F, G be TVS and $B : E \times F \to G$ a separately continuous bilinear map. Recall that B is $\mathscr{B}_F$-hypocontinuous, where $\mathscr{B}_F$ is the family of all bounded subsets of F, if ∀ neighborhoood of 0, W, in G and every $A \in \mathscr{B}_F$, ∃ a neighborhood of 0, U, in E such that

$B(U, A) \subseteq W$ (9.1.5). We say that B is <u>hypocontinuous</u> if B is both $\mathscr{B}_F$ and $\mathscr{B}_E$-hypocontinuous.

Theorem 14. Let E be a TVS, F and G LCS with F barrelled. If $B : E \times F \to G$ is separately continuous, then B is $\mathscr{B}_E$-hypocontinuous.

Proof: Let W be a closed, absolutely convex neighborhood of 0 in G and $A \in \mathscr{B}_E$. Put $T = \{y \in F : B(x, y) \in W \ \forall x \in A\}$. We claim that T is a barrel and this will establish the result.

First, T is absorbing. Let $y \in F$. The linear map $B(\cdot, y)$ is continuous so $\exists$ a neighborhood of 0, U, in E such that $x \in U \Rightarrow B(x, y) \in W$. $\exists t > 0$ such that $tA \subseteq U$ so if $x \in A$, $B(x, ty) = B(tx, y) \in W$ and $ty \in T$.

Next, T is balanced since if $y \in T$ and $|t| \le 1$, then $\forall x \in A$ $B(x, ty) = tB(x, y) \in tW \subseteq W$ so $ty \in T$. Similarly, T is convex.

Finally, T is closed. Let $y \in \bar{T}$. Then $\exists$ a net $\{y_\delta\} \subseteq T$ such that $y_\delta \to y$ in F. For $x \in A$, $B(x, \cdot)$ is continuous so $B(x, y_\delta) \to B(x, y)$. Since $B(x, y_\delta) \in W \ \forall \delta$, $B(x, y) \in W$. Thus, $y \in T$ and T is a barrel.

As promised in §9.1, we can now give an example of a hypocontinuous bilinear map which is not continuous.

Example 15. Let E be a Hausdorff LCS and consider the bilinear pairing between E'_b and E. If $A \subseteq E$ is $(\sigma(E, E'))$ bounded, then A^0 is a strong neighborhood of 0 so if $\varepsilon > 0$, $|\langle x', x\rangle| \le \varepsilon$ for $x' \in \varepsilon A^0$, $x \in A$

and $<\,,\,>$ is $\mathscr{B}_E$-hypocontinuous.

Let $\mathscr{B}'$ be the strongly bounded subsets of E'_b. We claim that $<\,,\,>$ is $\mathscr{B}'$-hypocontinuous if and only if E is quasi-barrelled. First suppose that E is quasi-barrelled, and let $B \subseteq E'_b$ be strongly bounded. Then B is equicontinuous (19.12) so B_0 is a neighborhood of 0 in E. Thus, if $\varepsilon > 0$, $|<x', x>| \le \varepsilon$ for $x' \in B$, $x \in \varepsilon B_0$ and $<\,,\,>$ is $\mathscr{B}'$-hypocontinuous. Conversely, assume that $<\,,\,>$ is $\mathscr{B}'$-hypocontinuous. If $B \subseteq E'_b$ is strongly bounded, there is a neighborhood, U, of 0 in E such that $|<B, U>| \le 1$. Hence, $B \subseteq U^0$ and B is equicontinuous. E is quasi-barrelled by 19.12.

Example 16. If the pairing above between E'_b and E is continuous, we claim that E must be normable. If $<\,,\,>$ is continuous, there is a neighborhood of 0, U, in E and a closed, absolutely convex bounded set $A \subseteq E$ such that $|<x', x>| \le 1$ for $x' \in A^0$, $x \in U$. But, then $U \subseteq (A^0)_0 = A$ so U must be bounded, and E is normable by Kolmogorov's Theorem.

From Examples 15 and 16, we see that if E is a quasi-barrelled, Hausdorff LCS which is not normable (e.g., s), then $<\,,\,> : E'_b \times E \to \mathbb{F}$ furnishes an example of a hypocontinuous bilinear map which is not continuous.

Exercise 1. Give an example of a quasi-barrelled space which is not barrelled.

Exercise 2. Prove the converse of the result in Example 16.

Exercise 3. If E is barrelled, show that $(E', \sigma(E', E))$ is quasi-complete.

Exercise 4. Show the inductive limit of barrelled spaces is barrelled.

Exercise 5. Show the quotient of barrelled spaces is barrelled.

Exercise 6. Show that $(m_0, \| \|_\infty)$ is barrelled. (Hint: 9.2.5) Show that this space is not a Baire space ([Sa]). The space $c_{oo} = m_0 \cap c_0$ furnishes an example of a closed subspace of a barrelled space which is not barrelled (Wilansky).

Exercise 7. If X is an infinite dimensional B-space, show that $(X, \sigma(X, X'))$ is not barrelled.

Exercise 8. A continuous linear map $T : E \to F$, E, F LCS, is <u>almost open</u> if $\overline{TU}$ is a neighborhood of 0 when U is a neighborhood of 0 in E (10.1). If F is barrelled and T is onto F, show that T is almost open.

Exercise 9. Let E be barrelled and $D \subseteq E$ a dense linear subspace. Show that D is barrelled if and only if every $\sigma(E', D)$ bounded subset is $\sigma(E', E)$ bounded.

Exercise 10. If E has a barrelled dense subspace, show that E is barrelled.

Exercise 11. Show a subset $A \subseteq E'$ is $\sigma(E', E)$ bounded if and only if A is contained in the polar of a barrel in E.

Exercise 12. Let E be a vector space with its strongest LC topology (Exercise 13.15). Show E is barrelled. Show c_{00} with this topology is not a Baire space.

Exercise 13. Give an example of a non-complete, Baire NLS. [Hint: Exer. 2.19.]

Exercise 14 (UBP for Quasi-barrelled Spaces). Let E be quasi-barrelled. If $\mathcal{F}$ is bounded in $L_b(E, F)$, then $\mathcal{F}$ is equicontinuous.

Exercise 15. Let X be a Sargent TVS (Def. 9.14) and Y a LCS. If $\mathcal{F} \subseteq L(X, Y)$ is pointwise bounded, show $\mathcal{F}$ is equicontinuous. [Hint: Let V be an absolutely convex neighborhood of 0 in Y and consider $E_k = \{x : Tx \in 2^k V, T \in \mathcal{F}\}$.]

Exercise 16. Show a LC Sargent space is barrelled. Show the converse is false. [Hint: Exercise 12.]

25
The UBP and Equicontinuity

In §9 and 20.7 we established several versions of the UBP which required no completeness or barrelledness assumptions on the range space. We showed that any pointwise bounded family of continuous linear operators between TVS is uniformly bounded on certain subfamilies of bounded subsets of the domain space (9.4, 20.7). These results were then used to show that any pointwise bounded family of continuous linear operators defined on a quasi-normed $\mathscr{K}$-space is always equicontinuous (9.8). It was further shown in Example 9.9 that a pointwise bounded family of continuous linear operators needn't be equicontinuous. Since any equicontinuous family of continuous linear operators is always uniformly bounded on bounded sets (Exer. 9.2), it is desirable to obtain equicontinuity versions of the UBP whenever possible (e.g., 9.7 and 24.11). In this section we use our earlier UBP to show that in the locally convex case there is a natural locally convex topology on the domain space with the property that any pointwise bounded family of continuous linear operators is always equicontinuous with respect to this topology. We begin by describing the topology.

Let (E, τ) be a Hausdorff LCS. Let $\mathscr{B}$ be the family of all τ bounded subsets of E. For any subfamily $\mathscr{A}$ of $\mathscr{B}$ with the property that $\cup\{A : A \in \mathscr{A}\} = E$, consider the family of subsets, V, of E satisfying:

(1) $V \subseteq E$ is absolutely convex and absorbs every member of $\mathscr{A}$.

The family of subsets satisfying (1) forms a neighborhood base at 0 for a locally convex topology, $\tau^{\mathscr{A}}$, on E (1.18). In particular, if $\mathscr{B} = \mathscr{A}$, then $\tau^{\mathscr{B}} = \tau^{b}$ is the bornological topology of E (§21). We have immediately

Proposition 1. (i) If $\mathscr{A} \subseteq \mathscr{C} \subseteq \mathscr{B}$, then $\tau^{\mathscr{A}} \supseteq \tau^{\mathscr{C}} \supseteq \tau^{b}$. (ii) $\tau^{\mathscr{A}}$ is the strongest locally convex topology on E such that all members of $\mathscr{A}$ are bounded.

We describe equicontinuity with respect to $\tau^{\mathscr{A}}$. In what follows F is a LCS.

Theorem 2. Let $\mathscr{F} \subseteq L(E, F)$. The following are equivalent.

(i) $\mathscr{F}$ is uniformly bounded on the members of $\mathscr{A}$,

(ii) $\mathscr{F}$ is $\tau^{\mathscr{A}}$ equicontinuous.

Proof. Suppose (i) holds. Let V be an absolutely convex neighborhood of 0 in F. Set $U = \cap\{T^{-1}V : T \in \mathscr{F}\}$. Then U is absolutely convex and absorbs every member of $\mathscr{A}$. Hence, U is a $\tau^{\mathscr{A}}$

neighborhood of 0, and $\mathscr{F}$ is $\tau^{\mathscr{A}}$ equicontinuous.

Suppose (ii) holds. Let $A \in \mathscr{A}$ and V be a neighborhood of 0 in F. There exists a $\tau^{\mathscr{A}}$ basic neighborhood of 0, U, such that $TU \subseteq V$ $\forall\, T \in \mathscr{F}$ and (i) holds.

For uniform boundedness on the family of all bounded sets and equicontinuity with respect to the original topology of E, we have

Theorem 3. Let (E, τ) be quasi-barrelled and $\mathscr{F} \subseteq L(E, F)$. The following are equivalent.

(iii) $\mathscr{F}$ is uniformly bounded on $\mathscr{B}$,

(iv) $\mathscr{F}$ is τ equicontinuous.

Proof: That (iv) implies (iii) follows from the proof of (ii) implies (i) above (or Exer. 9.2).

Also, (iii) implies (iv) by the proof of (i) implies (ii) above since we can take V to be a closed neighborhood and then U is τ closed by the continuity of the elements of $\mathscr{F}$. Thus, U is a bornivorous barrel and, therefore, a neighborhood of 0 by the quasi-barrel assumption (19.13).

In particular, Theorem 5 applies to barrelled and bornological spaces (Exer. 21.4). The quasi-barrelled spaces form the largest class of LCS for which (iii) and (iv) are equivalent since if F is the scalar field (iii) and (iv) are equivalent if and only if E is quasi-barrelled (19.12).

We can now obtain general equicontinuity forms of the UBP from

Theorem 2 and our earlier results.

Theorem 4. If $\mathscr{F} \subseteq L(E, F)$ is pointwise bounded on E, then $\mathscr{F}$ is $\tau^{\mathscr{A}}$ equicontinuous, where $\mathscr{A}$ is either the family of all $w(L(E, F))$ - $\mathscr{K}$ bounded or all $\beta(E, E')$ bounded subsets of E.

Proof: Theorem 2 and Theorems 9.5 and 20.7.

Since the original topology of E is stronger than $w(L(X, Y))$, any pointwise bounded family $\mathscr{F} \subseteq L(E, F)$ is also equicontinuous with respect to $\tau^{\mathscr{K}}$ when $\mathscr{K}$ is the family of $\mathscr{K}$ bounded subsets of E with respect to the original topology of E. Thus,

Corollary 5. Let (E, τ) be an $\mathscr{A}$-space. If $\mathscr{F} \subseteq L(E, F)$ is pointwise bounded, then $\mathscr{F}$ is τ^b equicontinuous.

The example given in 9.9 shows that the family $\mathscr{F}$ in Corollary 5 may fail to be equicontinuous with respect to the original topology of E even when E is an $\mathscr{A}$-space (14.19). In this example, $\sigma(\ell^2, \ell^2)^b = \| \|_2$ and the family $\{e_k\}$ is certainly equicontinuous with respect to $\| \|_2$.

26
The Transpose of a Linear Operator

For any linear mapping T between LCS E and F there is a natural linear mapping T', called the transpose of T, from a linear subspace of Y' into X', and many of the important properties of T are reflected through corresponding properties of T'. In this section we define and study properties of the transpose operator.

Let E, F be Hausdorff LCS. If T is a linear map whose domain is a linear subspace of E and has range in F, we write $\mathscr{D}(T)$ for its domain and $T : \mathscr{D}(T) \subseteq E \to F$. We define the transpose of T.

Lemma 1. Let $T : \mathscr{D}(T) \subseteq E \to F$ be linear. If $\forall\, y' \in F', \exists\, x' \in E'$ such that

(1) $$\langle y', Tx\rangle = \langle x', x\rangle \quad \forall\, x \in \mathscr{D}(T),$$

then x' is uniquely determined by y' if and only if $\mathscr{D}(T)$ is dense in E.

Proof: $\Leftarrow$: From the linearity of (1), it suffices to show that $\langle x', x\rangle = 0 \ \forall x \in \mathscr{D}(T) \Rightarrow x' = 0$. This follows from the continuity of x'.

$\Rightarrow$: Suppose $\overline{\mathscr{D}(T)} \subsetneq X$. By 13.9 $\exists x'_0 \in X'$ such that $\langle x'_0, x\rangle = 0 \ \forall x \in \mathscr{D}(T)$ and $x'_0 \neq 0$. This contradicts the hypothesis and establishes the result.

Definition 2. Let $T : \mathscr{D}(T) \subseteq E \to F$ be linear with $\mathscr{D}(T)$ dense in E. The <u>transpose</u> (<u>adjoint</u>, <u>dual</u>, <u>conjugate</u>) of T, T', is defined by:

(i) $\mathscr{D}(T') = \{y' \in Y' : y'T$ is continuous on $\mathscr{D}(T)\}$,

(ii) $T' : \mathscr{D}(T') \subseteq F' \to E' = (\mathscr{D}(T))'$ is given by $T'y' = y'T$.

Several remarks are in order:

(2) $T'y' = y'T$ is well-defined by Lemma 1 and this is the reason for the dense range assumption on T.

(3) T' is clearly linear.

(4) Since $\mathscr{D}(T)$ is dense in E, $E' = (\mathscr{D}(T))'$, and we can (and shall) consider T' to have range in E'.

(5) It is not clear that $\mathscr{D}(T')$ contains anything other than 0, and as the following example shows this phenomena can occur.

Example 3 (Berberian). Let $E = F = \ell^2$. Let $\mathscr{D}(T) = \text{span}\{e_k : k \in \mathbb{N}\}$ and let $\{u_{kj} : k, j \in \mathbb{N}\}$ be any double indexing of $\{e_k\}$. For each j, k set $T(u_{kj}) = e_k$ and extend T to $\mathscr{D}(T)$ by linearity. Suppose $y' = (y'_1, y'_2, \ldots) \in \mathscr{D}(T') \subseteq \ell^2$. Then

$$\forall\, k \, \langle T'y', u_{kj} \rangle = \langle y', Tu_{kj} \rangle = y'_k .$$

Now

$$\sum_{j=1}^{\infty} |\langle T'y', u_{kj} \rangle|^2 \leq \|T'y'\|^2$$

by Bessel's Inequality so $\lim_j \langle T'y', u_{kj} \rangle = y'_k = 0$ and $y' = 0$.

We next show that the phenomena of Example 3 cannot occur for closed operators.

Throughout the remainder of this section, we assume that $\mathscr{D}(T)$ is dense in E.

Theorem 4. Let $T : \mathscr{D}(T) \subseteq E \to F$ be linear. If T is closed, $\mathscr{D}(T')$ separates the points of F (and is, therefore, weak* dense in F' (Exer. 14.3)).

Proof: Let $y_0 \in F$, $y_0 \neq 0$. Now $(0, y_0) \notin G(T)$ and $G(T)$ is closed in $E \times F$ so by the Hahn-Banach Theorem,

$$\exists\, z' = (x', y') \in (E \times F)' = E' \times F'$$

such that $\forall\, x \in \mathscr{D}(T)$ $\langle z', (x, Tx) \rangle = \langle x', x \rangle + \langle y', Tx \rangle = 0$ and $\langle z', (0, y_0) \rangle = \langle y', y_0 \rangle \neq 0$. Then

$$-\langle x', x \rangle = \langle y', Tx \rangle \ \forall\, x \in \mathscr{D}(T) \Rightarrow y' \in \mathscr{D}(T')$$

with $T'y' = -x'$ so $\mathscr{D}(T')$ separates the points of F.

The converse holds; see Exer. 13.

We give a characterization of the situation when $\mathscr{D}(T') = F'$.

Theorem 5. Let $T : \mathscr{D}(T) \subseteq E \to F$ be linear. Then $\mathscr{D}(T') = F'$ if and only if T is continuous with respect to the original topology of E and $\sigma(F, F')$.

Proof: $\Rightarrow$: Let $\{x_\delta\}$ be a net in $\mathscr{D}(T)$ which is convergent to 0 in E. If $y' \in F'$, $\langle y', Tx_\delta \rangle = \langle T'y', x_\delta \rangle \to 0$ so $Tx_\delta \to 0$ in $\sigma(F, F')$.

$\Leftarrow$: Let $y' \in F'$ and $\{x_\delta\}$ be a net in $\mathscr{D}(T)$ which converges to 0 in E. Then $Tx_\delta \to 0$ in $\sigma(F, F')$ so $\langle y', Tx_\delta \rangle$. Hence, $y'T$ is continuous with respect to the original topology of E. So $y'T \in (\mathscr{D}(T))' = E'$.

For the case when $\mathscr{D}(T) = E$, we have

Theorem 6. Let $T : E \to F$ be linear. Then $\mathscr{D}(T') = F'$ if and only if T is $\sigma(E, E')$ - $\sigma(F, F')$ continuous.

Proof: The argument is like that in Theorem 5.

If an operator $T : E \to F$ is $\sigma(E, E')$ - $\sigma(F, F')$ continuous, we say briefly that T is weakly continuous. Recall that if T is continuous with respect to the original topologies of E and F, then T is weakly continuous (14.11) so Theorem 6 is applicable in this case.

Continuity Properties of the Transpose:

First, the transpose is always a closed operator.

Proposition 7. T' is a closed linear operator with respect to the weak* topologies, $\sigma(F', F)$ and $\sigma(E', E)$.

Proof: Let $\{y'_\delta\}$ be a net in $\mathscr{D}(T')$ which converges to a point $y' \in F$ and $\{T'y'_\delta\}$ converges to $x' \in E'$. For $x \in \mathscr{D}(T)$,

$$<y'_\delta, Tx> = <T'y'_\delta, x> \to <y', Tx> = <x', x> .$$

Thus, $y' \in \mathscr{D}(T')$ with $T'y' = x'$.

We next consider the continuity properties of the transpose. For operators whose domains are proper linear subspaces, the transpose operator can fail to even be bounded with respect to the respective weak* topologies.

Example 8. Let $E = \ell^1$, $F = c_{00}$ with the ℓ^1-norm, $\mathscr{D}(T) = c_{00}$ and $T : \mathscr{D}(T) \to F$ the identity operator. Then T is continuous with respect to the relative $\sigma(\ell^1, \ell^\infty)$ topology on $\mathscr{D}(T)$ and $\sigma(c_{00}, \ell^\infty)$. T' is just the identity operator on ℓ^∞ and $T' = I$ is not continuous with respect to $\sigma(\ell^\infty, c_{00})$ and $\sigma(\ell^\infty, \ell^1)$; indeed, I is not even bounded with respect to these topologies [consider $\{ke_k\}$].

We next consider continuity and boundedness properties of the transpose operator for operators whose domain is all of E. First we establish a sequential continuity property for any such transpose operator.

This result furnishes another example of where $\mathscr{K}$ convergence can be used as a substitute for completeness.

Theorem 9. Let $T : E \to F$ be linear. Then $T' : \mathscr{D}(T') \subseteq F' \to E'$ is sequentially continuous with respect to the relative $\sigma(F', F)$ topology on $\mathscr{D}(T')$ and the topology on E' of uniform convergence on $\sigma(E, E')$ - $\mathscr{K}$ convergent sequences in E.

Proof: Let $\{y'_k\} \subseteq \mathscr{D}(T')$ be $\sigma(F', F)$ convergent to 0, and let $\{x_j\} \subseteq E$ be $\sigma(E, E')$ - $\mathscr{K}$ convergent to 0. Consider the matrix $M = [<T'y'_i, x_j>] = [<y'_i, Tx_j>]$. It is easily checked that M is a $\mathscr{K}$-matrix so by the Basic Matrix Theorem, $\lim_i <T'y'_i, x_j> = 0$ uniformly for $j \in \mathbb{N}$.

The transpose operator also has a boundedness property for the appropriate topologies.

Theorem 10. If $T : E \to F$ is linear, then T' is bounded with respect to the relative $\sigma(F', F)$ topology of $\mathscr{D}(T')$ and the topology $\beta^*(E', E)$.

Proof: Let $B \subseteq \mathscr{D}(T')$ be $\sigma(F', F)$ bounded. Then for $x \in E$, $\{<T'y', x> : y' \in B\} = \{<y', Tx> : y' \in B\}$ is bounded so $T'B$ is $\sigma(E', E)$ bounded. By 20.6, $T'B$ is $\beta^*(E', E)$ bounded.

For $\mathscr{A}$-spaces, we have

Corollary 11. Let E be an $\mathscr{A}$-space and $T : E \to F$ linear. Then T' is bounded with respect to $\sigma(F', F)$ and $\beta(E', E)$.

Proof: For an $\mathscr{A}$-space $\beta^*(E', E) = \beta(E', E)$ by 20.5. The result follows from Theorem 10.

If in Corollary 11 E is a NLS, then $\beta(E', E)$ is just the dual norm topology so T' carries weak* bounded subsets of $\mathscr{D}(T')$ into norm bounded subsets of E'. Thus, for a NLS E which is a $\mathscr{K}$-space, the transpose operator T' carries norm bounded subsets of $\mathscr{D}(T')$ into norm bounded subsets of F'. This result for inner product $\mathscr{K}$-spaces is due to Pap ([P]); it was extended to NLS in [AS] and to LCS in [PS].

We next consider continuity properties of the transpose operator when T is weakly continuous. In this case the transpose operator has many nice continuity properties.

Lemma 12. Let $T : E \to F$ be weakly continuous. If $A \subseteq E$, then $(T')^{-1}A^0 = (TA)^0$.

We leave the proof as an easy exercise.

Lemma 13. Let $T : E \to F$ be weakly continuous. Let $\mathscr{A}(\mathscr{B})$ be a family of $\sigma(E, E')$ bounded subsets of E ($\sigma(F, F')$ bounded subsets of

F). Assume that $\mathscr{A}$, $\mathscr{B}$ satisfy conditions (i) and (ii) of 17.9. If $T\mathscr{A} \subseteq \mathscr{B}$, then T' is $\tau_{\mathscr{B}}$ - $\tau_{\mathscr{A}}$ continuous.

Proof: This is immediate from Lemma 12.

Using the fact that a weakly continuous linear operator preserves weakly bounded sets, absolutely convex weakly compact sets and finite sets, we have from Lemma 13.

Theorem 14. Let $T : E \to F$ be weakly continuous. Then T' is $\sigma(F', F)$ - $\sigma(E', E)$, $\beta(F', F)$ - $\beta(E', E)$, and $\tau(F', F)$ - $\tau(E', E)$ continuous.

If $T : E \to F$ is weakly continuous, then T' is $\sigma(F', F)$ - $\sigma(E', E)$ continuous so we may apply Theorem 14 to T' and the duality between F', F and E', E, noting that the transpose of T' in this duality is just T, to obtain

Corollary 15. Let $T : E \to F$ be weakly continuous. Then T is $\beta(E, E')$ - $\beta(F, F')$ and $\tau(E, E')$ - $\tau(F, F')$ continuous.

A classic result of Hellinger and Toeplitz for linear operators on Hilbert space can be obtained from Corollary 15. Let H be a Hilbert space and $T : H \to H$ linear. Then T is said to be <u>symmetric</u> if $Tx \cdot y = x \cdot Ty$ for all $x, y \in H$, where $x \cdot y$ is the inner product of x and y. The Hellinger-Toeplitz Theorem asserts that such a symmetric operator on a Hilbert space is continuous with respect to the norm topology of H. Such a symmetric operator is clearly weakly continuous by the self-duality of a Hilbert space so by Corollary 15 a symmetric operator is strongly (i.e., norm) continuous; that is, the Hellinger-Toeplitz Theorem is a direct consequence of Corollary 15.

Uniform Boundedness and Equicontinuity:

We now show that in the case of locally convex spaces the general UBP derived in 20.7 can be used to establish the equicontinuity of the transpose operators with respect to certain polar topologies on the dual spaces. First we require a lemma which follows directly from the identity in Lemma 12. Let $\mathcal{A}(\mathcal{B})$ be a family of bounded subsets of E (F) which satisfies conditions (i) and (ii) of 17.9.

Lemma 16. Let $\mathcal{F} \subseteq L(E, F)$. If for each $A \in \mathcal{A}$ there exists $B \in \mathcal{B}$ such that $\mathcal{F}A \subseteq B$, then the family of transpose operators $\mathcal{F}' = \{T' : T \in \mathcal{F}\}$ is $\tau_{\mathcal{B}}$-$\tau_{\mathcal{A}}$ equicontinuous.

Let $\mathcal{A}$ be the family of all $\beta(E, E')$ bounded (strongly bounded) subsets of E. Then the polar topology $\tau_{\mathcal{A}}$ is just $\beta^{*}(E', E)$ of §20. Let $\mathcal{B}$ be the family of all bounded ($\sigma(F, F')$ bounded) subsets of F. Then the polar topology $\tau_{\mathcal{B}}$ is just the strong topology $\beta(F', F)$. From the UBP given in 20.7, we obtain an equicontinuity version of the UBP for transpose operators.

Theorem 17. If $\mathcal{F} \subseteq L(E, F)$ is pointwise bounded on E, then $\mathcal{F}'$ is $\beta(F', F)$ - $\beta^*(E', E)$ equicontinuous.

A similar version of Theorem 17 can be formulated by using the version of the UBP given in 9.4 (see Exer. 12).

Algebraic Properties of the Transpose:

Proposition 18. Let G be a Hausdorff LCS, $T_1, T_2 \in L(E, F)$ and $S \in L(F, G)$. Then

(i) $(sT_1 + tT_2)' = sT_1' + tT_2' \ \forall s, t \in \mathbb{F}$,

(ii) $(ST_1)' = T_1'S'$.

The proof is left to the reader.

Corollary 19. If $T : E \to F$ is linear, continuous and has a continuous inverse T^{-1} with domain F, then $(T')^{-1}$ exists, has domain E' and $(T^{-1})' = (T')^{-1}$.

If L is a linear subspace of E, then

$$L^0 = \{x' \in E' : |\langle x', x\rangle| \le 1 \ \forall x \in L\} = \{x' : \langle x', x\rangle = 0 \ \forall x \in L\}$$

and this latter set is often called the annihilator of L (in E'). More generally if $A \subseteq E$, the <u>annihilator</u> of A (in E') is defined to be

$$A^{\perp} = \{x' \in E' : \langle x', x\rangle = 0 \ \forall x \in A\} ,$$

and if $B \subseteq E'$, the annihilator of B (in E) is defined to be

$$B_{\perp} = \{x \in E : \langle x', x\rangle = 0 \ \ \forall x' \in B\} .$$

Proposition 20. Let $T : E \to F$ be weakly continuous with kernel $\mathcal{N}(T)$ and range $\mathcal{R}T$.

(i) $\mathcal{N}(T') = (\mathcal{R}T)^{\perp}$ and $\mathcal{N}(T) = (\mathcal{R}T')_{\perp}$.

(ii) $\mathcal{N}(T')_{\perp} = \overline{\mathcal{R}T}$ and $\mathcal{N}(T)^{\perp} = \overline{\mathcal{R}T'}$, where the latter closure is in the weak* topology of F'.

Proof: (i): If $y' \in \mathcal{R}T^{\perp}$, then $\langle T'y', x\rangle = \langle y', Tx\rangle = 0 \ \forall\, x \in E$ so $y' \in \mathcal{N}(T')$. On the other hand, if $y' \in \mathcal{N}(T')$,

$$\langle T'y', x\rangle = \langle y', Tx\rangle = 0 \quad \forall\, x \in E$$

so $y' \in \mathcal{R}T^{\perp}$. The second statement follows by applying the first part to the transpose, T'.

(ii) follows from (i) and the Bipolar Theorem.

Corollary 21. T is 1-1 if and only if $\mathcal{R}T'$ is weak* dense in E'; T has dense range if and only if T' is 1-1.

The Case of NLS:

Let X and Y be NLS.

Theorem 22. Let $T \in L(X, Y)$. Then $T' \in L(Y', X')$ and $\|T\| = \|T'\|$.

Proof: $T' \in L(Y', X')$ by Theorem 14, and

$$\|T\| = \sup\{\|Tx\| : \|x\| \le 1\} = \sup\{|\langle y', Tx\rangle| : \|x\| \le 1, \|y'\| \le 1\} =$$
$$\sup\{|\langle T'y', x\rangle| : \|x\| \le 1, \|y'\| \le 1\} = \sup\{\|T'y'\| : \|y'\| \le 1\} = \|T'\|$$

by 8.1.5.

Thus the transpose map $T \to T'$ is a linear isometry from $L(X, Y)$ into $L(Y', X')$. We consider the problem as to when the map is onto. For this the following observation is useful: Let $T \in L(X, Y)$.

$$\begin{array}{ccccc} & X'' & \xrightarrow{T''} & Y'' & \\ J_X & X' & \xleftarrow{T'} & Y' & J_Y \\ & X & \xrightarrow{T} & Y & \end{array}$$

It is easy to check that $T''J_X = J_Y T$ or $J_Y^{-1}T''J_X = T$; for this reason, we say that T'' "extends" T (here we are identifying X and $J_X X$, Y and $J_Y Y$).

Theorem 23. The transpose map $T \to T'$ from $L(X, Y)$ into $L(Y', X')$ is onto if and only if Y is reflexive (here $X \neq \{0\}$).

Proof: $\Leftarrow$: Let $A \in L(Y', X')$. Define $T \in L(X, Y)$ by $T = J_Y^{-1}A'J_X$ (check the diagram above). Then $T' = A$ since for $y' \in Y', x \in X$,

$$\begin{aligned} \langle T'y', x\rangle &= \langle y', Tx\rangle = \langle y', J_Y^{-1}A'J_X x\rangle = \langle A'J_X x, y'\rangle \\ &= \langle J_X x, Ay'\rangle = \langle Ay', x\rangle . \end{aligned}$$

$\Rightarrow$: Let $y'' \in Y''$. Choose $x_0' \in X'$, $x_0 \in X$ such that $\langle x_0', x_0\rangle = 1$. Define $A: Y' \to X'$ by $Ay' = \langle y'', y'\rangle x_0'$. By hypothesis, $\exists\, T \in L(X, Y)$ such that $A = T'$. Hence,

$$\langle y', Tx_0\rangle = \langle T'y', x_0\rangle = \langle Ay', x_0\rangle = \langle y'', y'\rangle\langle x_0', x_0\rangle = \langle y'', y'\rangle$$

or $y'' = J_Y Tx_0$.

We give an application of Theorem 23 to the representation of operators on particular function spaces.

Example 24. Let $1 < p < \infty$ and $1/p + 1/q = 1$. Let $T \in L(\ell^1, \ell^p)$ and let $[t_{ij}]$ be the matrix representation of T with respect to $\{e_k\}$ (§10.1). We give a necessary condition that the matrix $[t_{ij}]$ must satisfy:

$$\|T\| = \sup\{\|Tx\|_p : \|x\|_1 \le 1\} = \sup\{|\langle y', Tx\rangle| : \|x\|_1 \le 1, \|y'\|_q \le 1\}$$

$$= \sup\{\|T'y'\|_\infty : \|y'\|_q \le 1\} =$$

$$\sup\{|\langle T'y', e_j\rangle| : \|y'\|_q \le 1, j \in \mathbb{N}\}$$

$$= \sup\{\|Te_j\|_p : j \in \mathbb{N}\} = \sup\{\|\sum_{i=1}^{\infty} t_{ij}e_i\|_p : j \in \mathbb{N}\}$$

$$= \sup\{(\sum_{i=1}^{\infty} |t_{ij}|^p)^{1/p} : j \in \mathbb{N}\} < \infty .$$

We show, conversely, that if the matrix $[t_{ij}]$ satisfies

$$M = \sup\{(\sum_{i=1}^{\infty} |t_{ij}|^p)^{1/p} : j\} < \infty ,$$

then it induces a continuous linear operator $T \in L(\ell^1, \ell^p)$. For each j the jth column $(t_{1j}, t_{2j}, \ldots) \in \ell^p$ induces a continuous linear functional $\varphi_j \in \ell^p = (\ell^q)'$ with $\|\varphi_j\| = (\sum_{i=1}^{\infty} |t_{ij}|^p)^{1/p} \le M < \infty$. Define $S : \ell^q \to \ell^\infty$ by $S(x) = \{\langle\varphi_j, x\rangle\}$. Note $S \in L(\ell^q, \ell^\infty)$ since

$$\|Sx\|_\infty = \sup\{|\langle\varphi_j, x\rangle| : j\} \le M\|x\|_q.$$

ℓ^p is reflexive so by Theorem 23 $\exists\, T \in L(\ell^1, \ell^p)$ such that $T' = S$. We claim the matrix representation of T is $[t_{ij}]$. For $i, j \in \mathbb{N}$, we have

$$<e_i, Te_j> = <Se_i, e_j> = <\sum_{k=1}^{\infty} <\varphi_k, e_i>e_k, e_j>$$

$$= \sum_{k=1}^{\infty} <\varphi_k, e_i><e_k, e_j> = <\varphi_j, e_i> = t_{ij}$$

so $Te_j = \sum_{i=1}^{\infty} t_{ij}e_i$.

Given any pair of sequence spaces X and Y, each operator $T \in L(X, Y)$ is represented by a matrix $[t_{ij}]$ as noted in §10.1. One of the principal problems in this area is to give necessary and sufficient conditions for a matrix $[t_{ij}]$ to represent an operator $T \in L(X, Y)$. For some pairs of spaces such conditions are known and for some there are no known necessary and sufficient conditions. For further examples and results see [TL], [Ma] and [DS].

More generally if X_0 is a specific function space (e.g., $C(S)$, $L^p(\mu)$), it is desirable to give a representation for a general operator $T \in L(X_0, Y)$ ($T \in L(Y, X_0)$) for an arbitrary B-space Y. There is a table of such operator representations given in [DS], V1.12.

Exercise 1. Let $k \in L^2([a, b] \times [a, b])$. Show $Kf(s) = \int_a^b k(s, t)f(t)dt$ defines a continuous linear operator from $L^2[a, b]$ into $L^2[a, b]$. Describe K'.

Exercise 2. Let $k \in C([a, b] \times [a, b])$. Show the formula in Exercise 1 defines a continuous linear operator from C[a, b] into C[a, b]. Describe K′.

Exercise 3. Let $\varphi \in L^\infty(\mu)$ and $1 \le p < \infty$. Define $T : L^p(\mu) \to L^p(\mu)$ by $Tf = \varphi f$. Show that T is continuous and describe T′.

Exercise 4. Define $L_n, R_n \in L(\ell^2)$ by

$$L_n(\{t_j\}) = \sum_{j=1}^{\infty} t_{j+n} e_j,\ R_n(\{t_j\}) = \sum_{j=1}^{\infty} t_j e_{n+j}.$$

Compute R_n', T_n'.

Exercise 5. If I is the identity operator on X, show I′ is the identity operator on X′.

Exercise 6. Show the map $T \to T'$ from $L_s(X, Y)$ into $L_s(Y', X')$ is not, in general, continuous. [Hint: Shift operators.]

Exercise 7. Let X, Y be Hausdorff LCS. If $S : Y' \to X'$ is weak* continuous, show $\exists\, T \in L(X, Y)$ such that $T' = S$. Conversely?

Exercise 8. Let X, Y be NLS and $\mathscr{D}(T)$ dense in X. Show T′ is continuous if and only if $\mathscr{D}(T')$ is norm closed.

Exercise 9. Let $T \in L(X, Y)$. Show that $\mathscr{R}T''$ is weak* dense if and only if $\mathscr{R}T$ is dense. Show that T″ is onto if and only if T is onto.

Exercise 10. If T'' is 1-1, show T is 1-1, but not conversely. [Hint: $T(\{t_j\}) = \{t_j - t_{j+1}\}$ on c_0.]

Exercise 11. Establish the converse of Lemma 16 when the elements of $\mathscr{B}$ are absolutely convex and closed.

Exercise 12. Formulate a result analogous to Theorem 17 using 9.4 and the family of $w(L(X,Y))$ - $\mathscr{K}$ bounded sets.

Exercise 13. Establish the converse of Theorem 4.

Exercise 14. If F is a Banach-Mackey space and $T : E \to F$ is weakly continuous, show T' is $\beta^*(F', F)$ - $\beta(E', E)$ continuous.

26.1 Banach's Closed Range Theorems

In this section we establish Banach's Closed Range Theorem for operators between B-spaces. Let X, Y be B-spaces and $T \in L(X, Y)$.

Lemma 1. Suppose that T is onto. $\exists\, c > 0$ such that $\forall\, y \in Y\ \exists\, x \in X$ with $Tx = y$ and $\|x\| \le c\|y\| = c\|Tx\|$.

Proof. Let φ be the quotient map from X onto $X/\mathcal{N}(T) = \dot{X}$ and let $\dot{T}$ be the induced map from $\dot{X}$ onto Y ($\dot{T}\varphi = T$, Exer. 6.3). $\dot{T}$ is onto and so has a continuous inverse by the OMT. Let $c = \|\dot{T}^{-1}\| + 1$. Then, for $y \ne 0$, $c\|y\| > \|\dot{T}^{-1}\|\|y\| \ge \|\dot{T}^{-1}y\|$ so the existence of x follows from the definition of the quotient norm (§6).

Theorem 2 (Banach). The following are equivalent.

(i) $\mathcal{R}T$ is (norm) closed.

(ii) $\mathcal{R}T'$ is norm closed.

(iii) $\mathcal{R}T'$ is weak* closed.

(iv) $\mathcal{R}T = \mathcal{N}(T')_{\perp}$.

(v) $\mathcal{R}T' = \mathcal{N}(T)^{\perp}$.

Proof: By 26.20 (iii) and (v) [(i) and (iv)] are equivalent. Clearly (iii) implies (ii).

(i) implies (v): Let $x' \in \mathcal{N}(T)^{\perp}$. Define a linear functional f_1 on $\mathcal{R}T = \overline{\mathcal{R}T}$ by $\langle f_1, Tx\rangle = \langle x', x\rangle$; f_1 is well-defined since if

$Tx_1 = Tx_2$, then $x_1 - x_2 \in \mathcal{N}(T)$ and

$$\langle x', x_1\rangle = \langle x', x_2\rangle = \langle f_1, Tx_1\rangle = \langle f_1, Tx_2\rangle.$$

f_1 is clearly linear, and we claim that f_1 is continuous. For if $y_n \in \mathcal{R}T$ and $\|y_n\| \to 0$, choose $x_n \in X$ such that $Tx_n = y_n$ as in Lemma †. (Here, we are considering T as a linear operator from X onto $\mathcal{R}T$.) By Lemma †, $\|x_n\| \to 0$ so $\langle f_1, y_n\rangle = \langle x', x_n\rangle \to 0$ since x' is continuous. Extend f_1 to a linear functional $y' \in Y'$ (Hahn-Banach). Then $\langle y', Tx\rangle = \langle f_1, Tx\rangle = \langle x', x\rangle$ which implies that $x' = T'y'$, i.e., $x' \in \mathcal{R}T'$. Hence, $\mathcal{N}(T)^\perp \subseteq \mathcal{R}T'$ and the reverse inclusion always holds (26.20).

(ii) implies (i): Consider T as a continuous linear operator, T_1, from X into $Y_1 = \overline{\mathcal{R}T}$. Now $T_1' : Y_1' \to X'$ is defined by $\langle T_1'y_1', x\rangle = \langle y_1', Tx\rangle$ for $x \in X$, $y_1' \in Y_1'$. By the Hahn-Banach Theorem y_1' can be extended to $y' \in Y'$ so

$$\langle y_1', Tx\rangle = \langle y', Tx\rangle = \langle T'y', x\rangle = \langle T_1'y_1', x\rangle$$

and $T'y' = T_1'y_1'$ so $\mathcal{R}T' = \mathcal{R}T_1'$. T_1' has an inverse (26.21) and has norm closed range so $(T_1')^{-1}$ is a continuous linear operator from $\mathcal{R}T'$ onto $Y_1' = (\mathcal{R}T)'$ by the OMT. Let $\delta = 1/\|(T_1')^{-1}\|$.

To show $\mathcal{R}T$ is closed it suffices to show that T_1 is an almost open map from X to Y_1 (10.2). To show that T_1 is almost open, we show that

(1) $\|y_1\| \le \delta$, $y_1 \in Y_1$ implies $y_1 \in \overline{T_1S}$, where

$$S = \{x \in X : \|x\| \le 1\}.$$

Suppose $y_1 \notin \overline{T_1S}$. By the Hahn-Banach Corollary 13.13, $\exists\ y_1' \in Y_1'$, $\|y_1'\| = 1$, satisfying $|\langle y_1', y_1\rangle| > \sup\{|\langle y_1', z\rangle| : z \in T_1S = TS\} = a$. For $x \in S$,

$$|\langle T_1'y_1', x\rangle| = |\langle y_1', T_1x\rangle| \le a < |\langle y_1', y_1\rangle| \le \|y_1\|$$

so $\|T_1'y_1'\| \le a < \|y_1\|$. But

$$1 = \|y_1'\| = \|(T_1')^{-1}T_1'y_1'\| \le (1/\delta)\|T_1'y_1'\|$$

so $\|y_1\| > \|T_1'y_1'\| \ge \delta$, and (1) holds.

We now give an important application of the closed range theorem to the existence of bounded inverses for an operator and its transpose.

Theorem 3. (i) T is onto if and only if T' has a continuous inverse.

(ii) T' is onto if and only if T has a continuous inverse.

Proof: (i): $\Rightarrow$: T' is 1-1 (26.21) and has closed range (Theorem 2). The result follows from the OMT.

$\Leftarrow$: T has dense range (26.21), and $\mathscr{R}T'$ is closed (10.14) so T has closed range by Theorem 2. Hence, $\mathscr{R}T = Y$.

(ii): $\Rightarrow$: T has closed range by Theorem 2 and T is 1-1 by 26.21. T has a bounded inverse by the OMT.

$\Leftarrow$: T' has weak* dense range by 26.21, and $\mathscr{R}T$ is closed (10.14) so $\mathscr{R}T'$ is weak* closed by Theorem 2. Hence, $\mathscr{R}T' = X'$.

Exercise 4 shows that the completeness assumptions in Theorems 2 and 3 are important.

As an application of Theorem 3, we prove a result due to Banach and Mazur.

Corollary 4. Let Y be a separable B-space. Then there exists a continuous linear map $T : \ell^1 \to Y$ which is onto Y and Y' is linearly isometric to a subspace of ℓ^∞.

Proof: Pick $\{y_k\}$ dense in $\{y \in Y : \|y\| \le 1\}$. Define $T : \ell^1 \to Y$ by $T\{t_k\} = \sum_{k=1}^{\infty} t_k y_k$; note that T is well-defined since the series defining T is absolutely convergent. Moreover, since $\|T\{t_k\}\| \le \sum_{k=1}^{\infty} |t_k|$, T is continuous. For $y' \in Y'$,

$$\|T'y'\| = \sup\{|\langle T'y', e_k\rangle| : k\} = \sup\{|\langle y', y_k\rangle| : k\} = \|y'\|$$

so T' is an isometry, and T is onto by Theorem 3.

Remark 5. The equivalence of (i), (iii), (iv) and (v) also holds in a locally convex F-space. The proof in this case is more difficult, and we do not develop the necessary machinery for the proof. We do, however, use the metric version of Theorem 2 in section 26.7. For the proof in this case, see [T1].

Exercise 1. If T' is 1-1 and $\mathcal{R}T'$ is weak* closed, show T is onto.

Exercise 2. If T is 1-1 and $\mathcal{R}T$ is closed, show T' is onto.

Exercise 3. If T is 1-1 and has closed range, show that T'' is 1-1. (Compare with Exer. 26.10.)

Exercise 4. Let L be the left shift operator $L\{t_k\} = (t_2, t_3, ...)$ from c_{00} with the ℓ^2-norm into ℓ^2. Show L' has a bounded inverse but L is not onto. Similarly, if R is the right shift, show R' has closed range but R does not. This example shows that completeness is important in Theorems 2 and 3.

Exercise 5. Show that every separable B-space is topologically isomorphic to a quotient of ℓ^1. Contrast this with the situation in ℓ^2.

26.2 The Closed Graph and Open Mapping Theorems for LCS

In §10, we established versions of the Closed Graph Theorem (CGT) and Open Mapping Theorem (OMT) for quasi-normed spaces. In this section we establish versions of these theorems for LCS. Throughout this section let E and F be Hausdorff LCS and $T : E \to F$ linear. We first consider the CGT. For this we introduce a new class of LCS.

Definition 1. A Hausdorff LCS F is an infra-Pták (B_r-complete) space if every $\sigma(F', F)$ dense linear subspace $D \subseteq F'$ which is such that $D \cap U^0$ is $\sigma(F', F)$ closed $\forall$ neighborhood of 0, U, in F is $\sigma(F', F)$ closed. If the requirement that D is $\sigma(F', F)$ dense is dropped, a LCS satisfying the condition above is called a Pták (B-complete, fully complete) space.

The existence of an infra-Pták space which is not a Pták space was an open problem for some time, but an example of such a space has recently been given by Valdivia ([V]). From §23.1 it follows that a complete metrizable LCS is a Pták space. For further examples of Pták spaces and their permanence properties see [K2].

Theorem 2. Let E be barrelled and F infra-Pták. If T is closed, then T is weakly continuous.

Proof: Since $\mathscr{D}(T')$ is $\sigma(F', F)$ dense in F' (26.4), it suffices to show that $\mathscr{D}(T')$ is $\sigma(F', F)$ closed (26.6). For this we use the infra-Pták assumption. Let V be a neighborhood of 0 in F. Let $\{y'_\delta\}$ be a net in $\mathscr{D}(T') \cap V^0$ which is $\sigma(F', F)$ convergent to some $y' \in V^0$. It suffices to show that $y' \in \mathscr{D}(T')$, i.e., $y'T$ is continuous. Now $\mathscr{D}(T') \cap V^0$ is $\sigma(F', F)$ bounded so $T'(\mathscr{D}(T') \cap V^0)$ is $\sigma(E', E)$ bounded (26.10) and, therefore, equicontinuous since E is barrelled (24.6). If $x \in E$, $\langle T'y'_\delta, x\rangle = \langle y'_\delta, Tx\rangle \to \langle y', Tx\rangle$ so the net $\{y'_\delta T\}$ converges pointwise to $y'T$ and is contained in an equicontinuous subset of E' and, hence, $y'T$ must be continuous (Exer. 5.14). Hence, $\mathscr{D}(T') \cap V^0$ is $\sigma(F', F)$ closed, and $\mathscr{D}(T')$ is $\sigma(F', F)$ closed by the infra-Pták assumption.

Theorem 3 (CGT). Let E be barrelled and F infra-Pták. If T is closed, then T is $\beta(E, E') - \beta(F, F')$ continuous and, hence, continuous with respect to the original topologies of E and F.

Proof: The first statement holds by Theorem 2 and 26.15. If E is barrelled, then the original topology of E is $\beta(E, E')$ (24.7) and the original topology of F is weaker than $\beta(F, F')$ so the second statement follows immediately from the first.

The class of barrelled spaces in the CGT cannot be replaced by a larger class. Makowald has shown that if a LCS E is such that every closed linear operator from E into an arbitrary B-space is continuous, then E must be barrelled (see [RR] 6.2.11). The class of infra-Pták spaces is not the largest class for which Theorem 3 holds (see Supplement (2) of

Chapter 6 of [RR] and 34.9 of [K2]).

There are other versions of the CGT for LCS in which the barrelled assumption on the domain space is strengthened and the infra-Pták assumption on the range space is relaxed. For such results, see [K2] §34 and §35.

We now consider the OMT.

Proposition 4. If E is a Pták space and $T : E \to F$ is a continuous linear open map from E onto F, then F is a Pták space.

Proof: Let $G' \subseteq F'$ be a linear subspace with the property that $G' \cap V^0$ is $\sigma(F', F)$ closed for every neighborhood of 0, V, in F. It suffices to show that $T'G'$ is $\sigma(E', E)$ closed since if this is the case, $G' = (T')^{-1}T'G'$ is $\sigma(F', F)$ closed by the weak* continuity of T' (26.14). Let U be a neighborhood of 0 in E. By the Pták hypothesis on E it suffices to show that $T'G' \cap U^0 = T'(G' \cap (T')^{-1}U^0)$ is $\sigma(E', E)$ closed. Since $(T')^{-1}U^0 = (TU)^0$ (26.12) and TU is a neighborhood of 0, $G' \cap (TU)^0$ is $\sigma(F', F)$ compact by the Banach-Alaoglu Theorem and the hypothesis so $T'G' \cap U^0 = T'(G' \cap (TU)^0)$ is $\sigma(E', E)$ closed (compact) by the weak* continuity of T' (26.14) and the result follows.

Corollary 5. If E is a Pták space and H is a linear subspace of E, then E/H is a Pták space.

We can now easily derive the OMT from the CGT.

Theorem 6 (Open Mapping Theorem). If E is a Pták space, F is barrelled and $T \in L(E, F)$ is onto, then T is open.

Proof: Let $G = E/\mathcal{N}(T)$, let φ be the quotient map from E onto G and let S be the induced map from G onto F. The inverse map $S^{-1} : F \to G$ is closed and G is a Pták space from Corollary 5 so S^{-1} is continuous from the CGT (Theorem 3). Thus, S is open and since φ is open, $T = S\varphi$ is open.

For more general forms of the CGT and OMT see the discussions in [RR] V1.2 and [K2].

Exercise 1. When E and F are LCS, show that Theorems 3 and 6 imply the versions of the theorems given in §10.

Exercise 2. Let E be a reflexive Frechet space. Show that E_b' is a Pták space.

26.3 Vector Integration

In this section we consider the integration of vector-valued functions with respect to a positive measure. As was noted in §9.4 there are often two approaches to extending a concept to the vector case, the weak and strong approaches. In §9.4 we showed that in the case of analytic vector-valued functions, these two approaches led to the same class of functions. The theory of vector integration is quite different and the two approaches lead to very different theories. We begin with the weak theory.

Let (S, Σ, μ) be a σ-finite measure space and X a B-space. A function $f : S \to X$ is said to be scalarly μ-integrable (weakly μ-integrable) if $x'f = x' \circ f$ is μ-integrable $\forall\, x' \in X'$. Suppose that f is scalarly μ-integrable and define $F : X' \to L^1(\mu)$ by $Fx' = x'f$. It is easily checked that F is closed and is, therefore, continuous by the CGT (Exer. 10.6). The transpose $F' : L^\infty(\mu) \to X''$ is given by

$$<F'g, x'> = <g, Fx'> = \int_S g(x'f)d\mu,$$

where $g \in L^\infty(\mu)$, $x' \in X'$. For any $E \in \Sigma$ we define the integral of f over E (with respect to μ), denoted by $\int_E f d\mu$, to be the element $F'C_E \in X''$, where C_E is the characteristic function of E. Thus, $<\int_E f d\mu, x'> = \int_E x' f d\mu$. This integral is called the Gelfand integral of f; if $\int_E f d\mu \in X$ $\forall\, E \in \Sigma$, then f is said to be Pettis integrable and the integral is called the Pettis integral of f. Of course, if X is reflexive, then any Gelfand integrable function is Pettis integrable, but this is not the case in general.

Example 1. Let μ be counting measure on $\mathbb{N}$. Define $f : \mathbb{N} \to c_0$ by $f(n) = e_n$. Let $x' = \{t_n\} \in \ell^1$. Then $x'f(n) = t_n$ so $x'f$ is μ-integrable with $\int_E x'fd\mu = \sum_{n\in E} t_n = \langle x', C_E\rangle$. Hence, $\int_E fd\mu = C_E$ and when E is infinite $\int_E fd\mu \notin c_0$ so f is Gelfand integrable but not Pettis integrable.

From continuity properties of the transpose map, we obtain additivity properties of the integral as a set function.

Proposition 2. Let $f : S \to X$ be scalarly μ-integrable. The set function $E \to \int_E fd\mu$ is countably additive with respect to $\sigma(X'', X')$. If f is Pettis μ-integrable, the set function is norm countably additive.

Proof: If (Y, τ) is a TVS, a set function $\mu : \Sigma \to Y$ is countably additive if for each disjoint sequence $\{E_j\}$ from Σ, the series $\sum_{j=1}^{\infty} \mu(E_j)$ is τ-convergent to $\mu(E)$, where $E = \bigcup_{j=1}^{\infty} E_j$. Now the series $\sum_{j=1}^{\infty} C_{E_j}$ is weak* convergent to C_E in $L^\infty(\mu)$ so $\sum_{j=1}^{\infty} F'C_{E_j} = \sum_{j=1}^{\infty} \int_{E_j} fd\mu$ is $\sigma(X'', X')$ convergent to $F'C_E = \int_E fd\mu$ by 26.14.

If f is Pettis integrable, it follows from the first part that $\int_{\cdot} fd\mu$ is countably additive by the Orlicz-Pettis Theorem.

We give some of the elementary properties of the Gelfand and Pettis integrals in the exercises. For a more detailed study of these integrals we refer the reader to the excellent exposition in [DU].

We now consider a strong theory of integration. Here we try to extend the approach taken in the scalar case to the vector case ("replace $| \ |$ by $\| \ \|$"). First we define strong measurability for a vector-valued function. A function $\varphi : S \to X$ is <u>simple</u> (Σ-simple) if $\varphi = \sum_{k=1}^{n} x_k C_{E_k}$, where $x_k \in X$, $E_k \in \Sigma$. A function $f : S \to X$ is <u>scalarly measurable</u> if $x'f$ is measurable $\forall\, x' \in X'$. A simple function is clearly scalarly measurable and its Pettis integral is $\int_E \varphi d\mu = \sum_{k=1}^{n} x_k \mu(E \cap E_k)$, provided $\mu(E \cap E_k) < \infty \ \forall\, k$.

Definition 3. A function $f : S \to X$ is <u>strongly measurable</u> (with respect to μ) if $\exists$ a sequence of simple functions $\varphi_k : S \to X$ such that $\varphi_k \to f$ μ-a.e.

Obviously, any simple function is strongly measurable, and a strongly measurable function is scalarly measurable. The converse is false (Example 5).

Proposition 4. If $f : S \to X$ is strongly measurable, then the scalar function $\|f(\cdot)\|$ is measurable.

Proof: If $\varphi = \sum_{k=1}^{n} x_k C_{E_k}$ is a simple function with $\{E_k\}$ pairwise disjoint, then $\|\varphi(\cdot)\| = \sum_{k=1}^{n} \|x_k\| C_{E_k}$ is measurable, and if the notation is as in Definition 3, $\|\varphi_k(\cdot)\| \to \|f(\cdot)\|$ μ-a.e. so $\|f(\cdot)\|$ is measurable.

We give an example of a scalarly measurable function which is not strongly measurable.

Example 5. Let $S = [0, 1]$ and let λ be counting measure on S. Set $H = \ell^2(S) = L^2(\lambda)$. Then H is a non-separable Hilbert space and $\{e_t : t \in S\}$, $e_t(s) = 1$ if $t = s$ and $e_t(s) = 0$ if $t \neq s$, is a complete orthonormal set in H.

Let μ be Lebesgue measure on S and let P be a non-Lebesgue measurable subset of S. Define $f : S \to H$ by $f(t) = 0$ if $t \notin P$ and $f(t) = e_t$ if $t \in P$. Now f is scalarly measurable since if $g \in H' = H$, then g has a Fourier expansion $g = \sum_{t \in S} (g \cdot e_t) e_t$ with $\{t : g \cdot e_t \neq 0\}$ countable and $g \cdot f(s) = 0$ if $s \notin P$ and $g \cdot f(s) = g \cdot e_s$ if $s \in P$ so $g \cdot f(\cdot)$ is 0 μ-a.e. However, $\|f(\cdot)\| = C_P$ is not μ-measurable so f is not strongly measurable by Proposition 4.

Note the range of f is non-separable. We next establish an important result of Pettis relating strong and scalar measurability.

Proposition 6. Suppose that X' has a countable norming set A for X (Definition 8.1.9). If $f : S \to X$ is such that $x'f$ is measurable $\forall\, x' \in A$, then $\|f(\cdot)\|$ is measurable.

Proof: For each $t \in S$, $\|f(t)\| = \sup\{\, |\langle x', f(t)\rangle| : x' \in A\}$ and each $|\langle x', f(\cdot)\rangle|$ is measurable so $\|f(\cdot)\|$ is measurable.

Recall that any separable NLS has a countable norming set (8.1.11) so Proposition 6 is applicable to scalarly measurable functions with range in a separable B-space. A function $\varphi : S \to X$ is <u>countably valued</u> if $\varphi = \sum_{k=1}^{\infty} x_k C_{E_k}$, with $x_k \in X$, $E_k \in \Sigma$. A countably valued function is clearly strongly measurable.

Proposition 7. Let X be separable and $f : S \to X$ scalarly measurable. Then $\exists$ a sequence of countably valued functions $\{\varphi_k\}$ such that $\|\varphi_k(t) - f(t)\| \le 1/k$ $\forall\, t \in S$.

Proof: Since X is separable, $\forall\, n$ X can be covered by a countable number of spheres, $S_{in} = S(x_{in}, 1/n)$, $i \in \mathbb{N}$. By Proposition 6, the function $t \to \|f(t) - x_{in}\|$ is measurable so $B_{in} = \{t \in S : f(t) \in S_{in}\} \in \Sigma$ and $S = \bigcup_{i=1}^{\infty} B_{in}$. For each n, disjointify the sequence $\{B_{in} : i \in \mathbb{N}\}$ by setting $A_{1n} = B_{1n}$, $A_{in} = B_{in} \setminus \bigcup_{j=1}^{i-1} B_{jn}$ for $i \ge 2$. Define $\varphi_n : S \to X$ by $\varphi_n(t) = x_{in}$ if $t \in A_{in}$, i.e., $\varphi_n = \sum_{i=1}^{\infty} x_{in} C_{A_{in}}$. Then φ_n is certainly countably valued and $\|\varphi_n(t) - f(t)\| \le 1/n \ \forall\, t \in S$.

Proposition 8. Let $f_n : S \to X$ be strongly measurable and $f : S \to X$. If $f_n \to f$ μ-a.e., then f is strongly measurable.

Proof: Pick $g : S \to \mathbb{R}$ such that $g \in L^1(\mu)$ and $g(t) > 0 \ \forall\, t \in S$. Fix m. Now $\|f_m(t) - f_n(t)\| \to \|f_m(t) - f(t)\|$ μ-a.e. so $\|f_m(\cdot) - f(\cdot)\|$ is measurable (Proposition 4). Hence,

$$\varepsilon_m = \int_S g(t)\|f(t) - f_m(t)\|/(1 + \|f(t) - f_m(t)\|)d\mu(t) \to 0$$

by the Dominated Convergence Theorem. For each m $\exists$ a sequence of simple functions $\{\varphi_{mk}\}_{k=1}^{\infty}$ such that $\lim_k \varphi_{mk} = f_m$ μ-a.e. By the argument above $\exists$ a simple function ψ_m such that

$$\int_S g(t)\|f_m(t) - \psi_m(t)\|/(1 + \|f_m(t) - \psi_m(t)\|)d\mu(t) < 1/m .$$

Hence,

$$\int_S g(t)\|f(t) - \psi_m(t)\|/(1 + \|f(t) - \psi_m(t)\|)d\mu(t) \le \varepsilon_m + 1/m \to 0 ,$$

and $\exists$ a subsequence $\{\psi_{m_k}\}$ such that $\psi_{m_k} \to f$ μ-a.e. ([AB] p. 207).

We now have the machinery in place to establish the Pettis measurability result. A function $f : S \to X$ is <u>μ-almost separable valued</u> if $\exists$ a μ-null set $E \in \Sigma$ such that $f(S \setminus E)$ is separable.

Theorem 9 (Pettis). Let $f : S \to X$. Then f is strongly measurable if and only if f is scalarly measurable and μ-almost separable valued.

Proof: $\Rightarrow$: $\exists$ a μ-null set E and a sequence $\{\varphi_k\}$ of simple functions such that $\varphi_k(t) \to f(t)\ \forall\, t \in S\backslash E$. Since $f(S\backslash E) \subseteq \overline{\bigcup_{k\in\mathbb{N}} \mathcal{R}\varphi_k}$, f is μ-almost separable valued; f is clearly scalarly measurable.

$\Leftarrow$: Without loss of generality, we may assume that X is separable. The result then follows from Propositions 7 and 8.

Bochner Integral:

A simple function $\varphi = \sum_{k=1}^{n} x_k C_{E_k}$ is Bochner integrable over $E \in \Sigma$ if $\mu(E \cap E_k) < \infty\ \forall\, k$ with $x_k \neq 0$. Its integral is defined to be

$$\int_E \varphi d\mu = \sum_{k=1}^{n} x_k \mu(E \cap E_k)$$

which agrees with the Pettis integral of φ. Note that a simple function φ is Bochner integrable if and only if $\|\varphi(\cdot)\|$ is μ-integrable and in this case $\|\int_E \varphi d\mu\| \le \int_E \|\varphi(\cdot)\| d\mu$.

Definition 10. A strongly measurable function $f : S \to X$ is <u>μ-Bochner integrable</u> if and only if

(i) $\exists$ a sequence of Bochner integrable simple functions $\{\varphi_k\}$ such that $\varphi_k \to f$ μ-a.e.,

(ii) $\lim \int_S \|f(\cdot) - \varphi_k(\cdot)\| d\mu = 0$.

The <u>Bochner integral</u> of f with respect to μ is defined to be

(1) $$\int_S f d\mu = \lim \int_S \varphi_k d\mu .$$

For $E \in \Sigma$, we say that f is Bochner integrable over E if and only if $C_E f$ is Bochner integrable and define $\int_E f d\mu = \int_S C_E f d\mu$.

We must first show that the limit in (1) exists. Note the integral in (ii) makes sense since $\|f(\cdot) - \varphi_k(\cdot)\|$ is measurable (Proposition 4). Now

$$\|\int_S \varphi_k d\mu - \int_S \varphi_j d\mu\| \le \int_S \|\varphi_k(\cdot) - \varphi_j(\cdot)\| d\mu$$

$$\le \int_S \|\varphi_k(\cdot) - f(\cdot)\| d\mu + \int_S \|f(\cdot) - \varphi_j(\cdot)\| d\mu$$

so (ii) implies that $\{\int_S \varphi_k d\mu\}$ is Cauchy in X and, therefore, converges. Second, it is easy to see that the value of the integral is independent of the sequence $\{\varphi_k\}$ [if $\{\psi_k\}$ is another sequence satisfying (i) and (ii), consider the interlaced sequence, $\varphi_1, \psi_1, \varphi_2, \ldots$].

If both the Bochner and Pettis integrals are being considered, it is necessary to distinguish between the integrals, but for the present time we consider only the Bochner integral.

We have a very useful scalar criterion for Bochner integrability.

Theorem 11. Let $f : S \to X$ be strongly measurable. Then f is μ-Bochner integrable if and only if $\|f(\cdot)\|$ is μ-integrable. In this case, $\|\int_S f d\mu\| \le \int_S \|f(\cdot)\| d\mu$.

Proof: $\Rightarrow$: Let $\{\varphi_k\}$ satisfy (i) and (ii). Then

$$\|f(\cdot)\| \le \|\varphi_k(\cdot)\| + \|f(\cdot) - \varphi_k(\cdot)\|$$

so $\|f(\cdot)\|$ is μ-integrable.

$\Leftarrow$: Let $\{\varphi_k\}$ be a sequence of simple functions such that $\varphi_k \to f$ μ-a.e. Put $\psi_k(t) = \varphi_k(t)$ if $\|\varphi_k(t)\| \le 2\|f(t)\|$ and $\psi_k(t) = 0$ if $\|\varphi_k(t)\| > 2\|f(t)\|$. Then each ψ_k is a simple function, $\psi_k \to f$ μ-a.e., and $\|\psi_k(\cdot)\| \le 2\|f(\cdot)\|$. Since $\|\psi_k(\cdot) - f(\cdot)\| \le 3\|f(\cdot)\|$ and $\|f(\cdot)\|$ is μ-integrable, the Dominated Convergence Theorem implies that

$$\lim \int_S \|\psi_k(\cdot) - f(\cdot)\| d\mu = 0$$

so f is μ-Bochner integrable.

Also, the Dominated Convergence Theorem implies that

$$\lim \int_S \|\psi_k(\cdot)\| d\mu = \int_S \|f(\cdot)\| d\mu$$

so $\|\int_S f d\mu\| = \lim \|\int_S \psi_k d\mu\| \le \lim \int_S \|\psi_k(\cdot)\| d\mu = \int_S \|f(\cdot)\| d\mu.$

We now compare the Bochner and Pettis integrals. We denote the Bochner Integral by $B\int_S f d\mu$ and the Pettis integral by $P\int_S f d\mu$.

Proposition 12. Let $f : S \to X$ be strongly measurable. If f is μ-Bochner integrable, then f is μ-Pettis integrable and the two integrals agree.

Proof: If $x' \in X'$, $|x'f(\cdot)| \le \|x'\| \|f(\cdot)\|$ so Theorem 11 implies that $x'f \in L^1(\mu)$. Let $\{\psi_k\}$ be as in Theorem 11. Then the Dominated Convergence Theorem implies

$$\lim\langle x', P\int_S \psi_k d\mu\rangle = \lim\langle x', B\int_S \psi_k d\mu\rangle = \langle x', B\int_S f d\mu\rangle$$

$$= \lim\int_S \langle x', \psi_k\rangle d\mu = \int_S x' f d\mu .$$

Hence, f is μ-Pettis integrable and its Pettis integral is $B\int_S f d\mu$.

The converse does not hold.

Example 13. Let μ be counting measure on $\mathbb{N}$. Define $f : \mathbb{N} \to c_0$ by $f(k) = e_k/k$. If $x' = \{t_k\} \in \ell^1$, $x'f(k) = t_k/k$ so $x'f \in L^1(\mu)$ and $\int_{\mathbb{N}} x' f d\mu = \sum_{k=1}^{\infty} t_k/k$. Hence, f is Pettis integrable with $P\int_{\mathbb{N}} f d\mu = \{1/k\}$. Now f is also strongly measurable, but $\|f(k)\| = 1/k$ so $\|f(\cdot)\|$ is not μ-Bochner integrable (Theorem 11).

Proposition 12 and Example 13 show that the Pettis (and the Gelfand) integral is more general than the Bochner integral. However, it is difficult to give criteria which guarantee that a function is Pettis integrable whereas Theorem 11 supplies a very useful condition for Bochner integrability. The Bochner integral also has many other nice properties (such as the Dominated Convergence Theorem given below) which make it very useful. This is one of the reasons that is often employed.

We now establish the Dominated Convergence Theorem for the Bochner integral. Some of the other elementary properties of the integral are given in the exercises. The reader is encouraged to consult the monograph of Diestel and Uhl ([DU]) for more information.

Theorem 14. Let $f_k : S \to X$ be μ-Bochner integrable and $f : S \to X$ be such that $f_k \to f$ μ-a.e. If $\exists\, g \in L^1(\mu)$ such that $\|f_k(t)\| \le g(t)$ $\forall$ μ-almost all $t \in S$ and all $k \in \mathbb{N}$, then f is μ-Bochner integrable with $\int_S f d\mu = \lim \int_S f_k d\mu$.

Proof: Since $\|f_k(\cdot)\| \to \|f(\cdot)\|$ μ-a.e., the Dominated Convergence Theorem implies $\|f(\cdot)\|$ is μ-integrable so Proposition 8 and Theorem 11 imply that f is Bochner integrable. Also, $\|f_k(\cdot) - f(\cdot)\| \to 0$ μ-a.e. and $\|f_k(\cdot) - f(\cdot)\| \le 2g$ imply that

$$\lim \left\| \int_S (f_k - f) d\mu \right\| \le \lim \int_S \|f_k(\cdot) - f(\cdot)\| d\mu = 0$$

so

$$\lim \int_S f_k d\mu = \int_S f d\mu.$$

For a history of the integration of vector-valued functions, see [Hi].

Exercise 1. Show each of the integrals (Gelfand, Pettis, Bochner) is linear.

Exercise 2. State and prove Egoroff's Theorem for strongly measurable functions.

Exercise 3. If $f : S \to X$ is Pettis integrable and $T \in L(X, Y)$, where Y is a B-space, show Tf is Pettis integrable with $T\int_S fd\mu = \int_S Tfd\mu$. Repeat for the Bochner integral.

Exercise 4. Let $f : S \to X$ be μ-Bochner integrable. Let $m(E) = \int_E fd\mu$ for $E \in \Sigma$. Show that m is countably additive (norm) and absolutely continuous with respect to μ in the sense that $\lim_{\mu(E)\to 0} \|m(E)\| = 0$.

Exercise 5. Let $L^1(\mu, X)$ be all μ-Bochner integrable functions. Define a semi-norm on $L^1(\mu, X)$ by $\|f\|_1 = \int_S \|f(\cdot)\|d\mu$. Show $\| \|_1$ is complete.

Exercise 6. Show that $\|f\| = \sup\{\int_S |x'f|d\mu : \|x'\| \le 1\}$ defines a semi-norm on the μ-Pettis integrable functions. [This semi-norm is not complete ([DU]).]

Exercise 7. If $\lambda : \Sigma \to X$ is finitely additive, its variation is defined to be $v(\lambda)(E) = \sup\{\sum_{i=1}^{n} \|\lambda(E_i)\| : \{E_i\}$ a measurable partition of $E\}$, $E \in \Sigma$. If $f : S \to X$ is Bochner μ-integrable and $\lambda(E) = \int_E fd\mu$, show $v(\lambda)(E) = \int_E \|f\|d\mu < \infty$.

Exercise 8. Let $\lambda : \Sigma \to L^1[0, 1]$ be given by $\lambda(E) = C_E$, where Σ is the σ-algebra of Lebesgue measurable subsets of $[0, 1]$. Show that λ has finite variation. Show that there is no Bochner μ-integrable function

$f : [0, 1] \to X$ such that $\lambda(E) = \int_E f d\mu$, where μ is Lebesgue measure. That is, λ has no Radon-Nikodym derivative. For a discussion of the Radon-Nikodym property for vector-valued measures, see [DU].

Exercise 9. Compute the variation of the set function $\int_E f d\mu$ in Examples 1 and 13.

26.4 The Space of Schwartz Distributions

Recall (§22) that $\mathscr{D}(\Omega)$ is the space of all infinitely differentiable functions defined on the open subset $\Omega \subseteq \mathbb{R}^n$ and having compact support. We have $\mathscr{D}(\Omega) = \cup \mathscr{D}_K$, where K runs through the family of compact subsets of Ω, and we give $\mathscr{D}(\Omega)$ the inductive limit topology from the subspaces $\mathscr{D}_K$. If $\{K_m\}$ is an increasing sequence of compact subsets of Ω with non-void interiors such that $\cup K_m = \Omega$ and each compact subset of Ω is contained in some K_m, then $\mathscr{D}(\Omega) = \text{ind } \mathscr{D}_{K_m}$ is a strict inductive limit of Frechet spaces so $\mathscr{D}(\Omega)$ is an LF-space. The elements of $\mathscr{D}(\Omega)$ are called <u>test functions</u>. The dual of $\mathscr{D}(\Omega)$ is called the space of <u>Schwartz distributions</u> or simply <u>distributions</u>. The space of distributions has many important applications in analysis, particularly in partial differential equations, and we now study some of the basic properties of distributions.

From previous results on inductive limits in section 22, we have

Theorem 1. (i) $\mathscr{D}(\Omega)$ is a quasi-complete, Hausdorff, LF, bornological, barrelled, non-metrizable LCS.

(ii) A sequence $\{\varphi_k\}$ in $\mathscr{D}(\Omega)$ converges to 0 if and only if $\exists$ compact $K \subseteq \Omega$ such that $\{\varphi_k\} \subseteq \mathscr{D}_K$ and $D^\alpha \varphi_k \to 0$ uniformly on K $\forall$ multi-index α.

Proof: 22.14, Exer. 24.4.

The space $\mathscr{D}(\Omega)$ has one further important property which we now establish.

Proposition 2. A closed, bounded subset $B \subseteq \mathscr{D}(\Omega)$ is compact.

Proof: It suffices to show that any closed, bounded subset B of $\mathscr{D}_K$ is compact (22.12). Let $J = \{0, 1, 2, \ldots\}$, and let I be the product of J with itself n-times. The map $\varphi \to \{D^\alpha \varphi\}_{\alpha \in I}$ imbeds $\mathscr{D}_K$ linearly into $\prod_I C(K)$; the topology induced on $\mathscr{D}_K$ by the product topology is exactly the original topology of $\mathscr{D}_K$ so $\mathscr{D}_K$ is closed in $\prod_I C(K)$ since it is complete.

For each $\alpha \in I$, $\exists\, c_\alpha$ such that

$$(1) \qquad |D^\alpha \varphi(x)| \le c_\alpha \ \forall\ \varphi \in B,\ x \in K\,.$$

The projection $\Pi_\alpha : B \to C(K)$ defined by $\Pi_\alpha(\varphi) = D^\alpha \varphi$ has equicontinuous range since if $x, y \in K$ and $x(s) = (1 - s)x + sy$, $0 \le s \le 1$, then

$$\begin{aligned} |D^\alpha \varphi(x) - D^\alpha \varphi(y)| &= \left| \int_0^1 \frac{d}{ds} D^\alpha \varphi(x(s))\,ds \right| \\ &\le \int_0^1 \sum_{j=1}^n \left| \frac{\partial}{\partial x_j} D^\alpha \varphi(x(s)) \right| |y_j - x_j|\,ds \\ &\le \max\{c_\beta : |\beta| \le |\alpha| + 1\} n \|x - y\|\,. \end{aligned}$$

By the Arzela-Ascoli Theorem, it follows that $\Pi_\alpha(B)$ is relatively compact in C(K). By Tychonoff's Theorem, B is relatively compact in $\prod_I C(K)$, and the proof is complete.

A (barrelled) Hausdorff LCS with the property in Proposition 2 is called a semi-Montel (Montel) space after the corresponding result in $\mathscr{H}(D)$ which was established by Montel ([K1] 27.3). Thus, $\mathscr{D}(\Omega)$ is also a Montel space. Note that any semi-Montel space is semi-reflexive (19.11) so, in particular, $\mathscr{D}(\Omega)$ is reflexive since it is barrelled (24.8).

Theorem 3. $\mathscr{D}'(\Omega)_b$ is a complete, barrelled space. On bounded subsets B of $\mathscr{D}'(\Omega)_b$ the strong and weak* topologies coincide. In particular, a sequence $\{T_k\}$ in $\mathscr{D}'(\Omega)$ converges strongly if and only if it converges weak*.

Proof: The first two statements follow from 21.4 and 24.9 since $\mathscr{D}(\Omega)$ is reflexive as observed above. If $B \subseteq \mathscr{D}'(\Omega)$ is bounded, then B is equicontinuous (24.6) so the topologies of pointwise convergence and precompact convergence coincide on B (23.6). But, the precompact and bounded subsets of $\mathscr{D}(\Omega)$ are the same (Proposition 2).

Corollary 4. Let $\{T_k\} \subseteq \mathscr{D}'(\Omega)$ [or, $\{T_\varepsilon : 0 < \varepsilon \le a\} \subseteq \mathscr{D}'(\Omega)$] and suppose that $\lim \langle T_k, \varphi\rangle = \langle T, \varphi\rangle$ exists $\forall\, \varphi \in \mathscr{D}(\Omega)$

$$[\lim_{\varepsilon\to 0^+} \langle T_\varepsilon, \varphi\rangle = \langle T, \varphi\rangle].$$

Then $T \in \mathscr{D}'(\Omega)$ and $T_k \to T$ strongly.

Proof: Theorem 3 and the Banach-Steinhaus Theorem 24.12.

We now give examples of distributions.

Proposition 5. Let T be a linear functional on $\mathscr{D}(\Omega)$. The following are equivalent.

(i) T is a distribution.

(ii) T is sequentially continuous.

(iii) For each compact $K \subseteq \Omega$ $\exists$ $C > 0$ and a non-negative integer N such that

(2) $$|\langle T, \varphi\rangle| \le C\|\varphi\|_N \ \forall\ \varphi \in \mathscr{D}_K,$$

where

$$\|\varphi\|_N = \sup\{|D^\alpha \varphi(x)| : |\alpha| \le N,\ x \in K\}\ .$$

Proof: The equivalence of (i) and (ii) is the bornological property of $\mathscr{D}(\Omega)$; (ii) and (iii) are equivalent since T is continuous on $\mathscr{D}(\Omega)$ if and only if the restriction of T to each $\mathscr{D}_K$ is continuous (22.2).

Definition 6. If $\exists$ N such that (2) holds $\forall$ compact K (but with possibly different C), the distribution T is said to be of finite order; the smallest such N is called the order of T. (See Exer. 8 for an example of a distribution which is not of finite order.)

Example 7. Let $f : \mathbb{R}^n \to \mathbb{R}$ be locally integrable (i.e., integrable over every compact subset with respect to Lebesgue measure). Then f induces a distribution, $f^\#$, via $\langle f^\#, \varphi\rangle = \int_{\mathbb{R}^n} f\varphi$ for $\varphi \in \mathscr{D}(\mathbb{R}^n)$. Indeed, if $\varphi \in \mathscr{D}_K$,

$|\langle f^{\#}, \varphi\rangle| \le \int_K |f| \sup\{|\varphi(x)| : x \in \mathbb{R}^n\}$ so $f^{\#}$ is a distribution of order 0.

Moreover, the map $f \to f^{\#}$ is an imbedding of $L^1_{loc}(\mathbb{R}^n)$, the locally integrable functions, into $\mathscr{D}'(\mathbb{R}^n)$. This follows from

Lemma 8. Let $K \subseteq \mathbb{R}^n$ be compact and $G \supseteq K$ open. Then $\exists\, \varphi \in \mathscr{D}(\mathbb{R}^n)$ such that $\varphi(x) = 1\ \forall\, x \in K, 0 \le \varphi \le 1$ and $\varphi(x) = 0\ \forall\, x \notin G$.

Proof: For $\delta > 0$, let $S(K, \delta) = \{y : \text{distance}(y, K) < \delta\}$. Choose $\delta > 0$ such that $S(K, 2\delta) \subseteq G$. Define $\psi \in \mathscr{D}(\mathbb{R}^n)$ by

$$\psi(x) = c \exp(\delta^2/(\delta^2 - \|x\|^2))$$

if $\|x\| < \delta$ and $\psi(x) = 0$ if $\|x\| \ge \delta$, where c is chosen such that $\int_{\mathbb{R}^n} \psi = 1$ (recall example 2.27). Note the support of ψ is contained in $\overline{S(0, \delta)}$. Define

$$(3) \qquad \varphi(x) = \int_{\mathbb{R}^n} C_V(y)\psi(x - y)dy = C_V * \psi(x) = \psi * C_V(x),$$

$$\text{where } V = S(K, \delta)$$

[a sketch at this point is useful]. Clearly, $\varphi \in C^\infty(\mathbb{R}^n)$. Now $\varphi(x) = 1$ for $x \in K$ since in this case the integral in (3) is over a ball with center at x and radius δ and, hence, gives value 1. Also, $\varphi(x) = 0$ for $x \notin G$ since then the integral in (3) is 0. Clearly, $0 \le \varphi \le 1$.

Recall that this result was needed in Example 22.14 to show that $\mathscr{D}(\Omega)$ was non-metrizable.

Now suppose $f \in L^1_{loc}(\mathbb{R}^n)$ is such that $f^{\#} = 0$, i.e., $\int f\varphi = 0$

$\forall\ \varphi \in \mathscr{D}(\mathbb{R}^n)$. Let K be compact and choose a decreasing sequence of open sets, $\{G_m\}$, such that $\cap G_m = K$. Choose φ_m as in Lemma 8. Then $f\varphi_m \to fC_K$ pointwise so the Dominated Convergence Theorem implies

$$\int_{\mathbb{R}^n} f\varphi_m = \langle f^{\#}, \varphi_m \rangle = 0 \to \int_K f = 0\ .$$

Thus, $f = 0$ a.e. and the map $f \to f^{\#}$ is indeed an imbedding. From this point of view distributions generalize the notion of a function and so are sometimes referred to as generalized functions.

Not every distribution is given by a locally integrable function.

Example 9. Let μ be a bounded Borel (signed) measure on $\mathbb{R}^n$. Then μ induces a distribution, $\mu^{\#}$, via $\langle \mu^{\#}, \varphi \rangle = \int_{\mathbb{R}^n} \varphi d\mu$ for $\varphi \in \mathscr{D}(\mathbb{R}^n)$. If $\varphi \in \mathscr{D}_K$, $|\langle \mu^{\#}, \varphi \rangle| \le |\mu|(K)\sup\{|\varphi(x)| : x \in K\}$ so $\mu^{\#}$ is a distribution of order 0.

In particular, if $x \in \mathbb{R}^n$ and δ_x is the Dirac measure (point mass) concentrated at x, then $\langle \delta_x, \varphi \rangle = \varphi(x)$ defines a distribution of order 0 which is certainly not given by a locally integrable function.

Henceforth, we freely identify locally integrable functions and bounded Borel measures with the distributions that they induce and the "sharp", #, is dropped.

Example 10 (Regularization of Divergent Integrals). Consider the function $f(t) = 1/t$, $t \in \mathbb{R}$. This function is not locally integrable, but, nevertheless, f induces a distribution. The distribution, still denoted by f, is defined to be,

$\langle f, \varphi\rangle = \lim_{\varepsilon\to 0}\int_{|t|\ge\varepsilon} \varphi(t)/t\, dt$ for $\varphi \in \mathscr{D}(\mathbb{R})$. Note that this limit exists since if $\varphi \in \mathscr{D}(\mathbb{R})$ and $\varphi(t) = 0$ for $|t| \ge a$, we may choose $\psi : \mathbb{R} \to \mathbb{R}$ continuous so that $\varphi(t) = \varphi(0) + t\psi(t)$, $\psi(0) = \varphi'(0)$ and then

$$\int_{|t|\ge\varepsilon} \varphi(t)/t\, dt = \varphi(0)\int_{\varepsilon\le|t|\le a} 1/t\, dt + \int_{\varepsilon\le|t|\le a} \psi(t)dt$$

$$= \varphi(0)[\ell n(a/\varepsilon) - \ell n(a/\varepsilon)] + \int_{\varepsilon\le|t|\le a} \psi \to \int_{|t|\le a} \psi .$$

For each $\varepsilon > 0$, $\varphi \to \int_{|t|\ge\varepsilon} \varphi(t)/t\, dt$ is a distribution (Example 7) so by Corollary 4 f induces a distribution, called the <u>Cauchy Principal Value of 1/t</u> and denoted by $\text{pv}\,\frac{1}{t}$. Functions such as this are called <u>pseudo-functions</u> by Schwartz. A similar construction can be carried out for any function which has a singularity at the origin of order $1/\|x\|^m$. See [Sch] and [GS] I.1.7 for such examples.

Concerning the imbedding of the locally integrable functions in the space of distributions we have

Proposition 11. $\mathscr{D}(\mathbb{R}^n)$ is dense in $\mathscr{D}'(\mathbb{R}^n)_b$.

Proof: If this is not the case, $\exists$ a continuous linear functional L on $\mathscr{D}'(\mathbb{R}^n)_b$ such that $\langle L, \mathscr{D}'(\mathbb{R}^n)\rangle = 0$ and $L \ne 0$ (Hahn-Banach). But, $\mathscr{D}(\mathbb{R}^n)$ is reflexive so $\exists\, \varphi \ne 0$, $\varphi \in \mathscr{D}(\mathbb{R}^n)$, such that $\langle \mathscr{D}'(\mathbb{R}^n), \varphi\rangle = 0$. In particular, $\int_{\mathbb{R}^n} \varphi^2 = 0$ so $\varphi = 0$.

It is actually the case that $\mathscr{D}(\mathbb{R}^n)$ is sequentially dense in $\mathscr{D}'(\mathbb{R}^n)_b$ ([Sch]).

Differentiation in $\mathscr{D}'$:

We consider the problem of defining the derivative of a distribution. Any time we wish to extend a classical operation on functions to distributions, such as differentiation, we want to do so in a manner which preserves the operation when the function is imbedded in the distributions. Thus, suppose that $f : \mathbb{R} \to \mathbb{R}$ has a locally integrable derivative, f'. Then integration by parts implies

$$<(f')^{\#}, \varphi> = \int_{\mathbb{R}} f' \varphi = -\int_{\mathbb{R}} f\varphi' = <-f^{\#}, \varphi'> \text{ for } \varphi \in \mathscr{D}(\mathbb{R}) .$$

This motivates the following definition of the derivative of a distribution.

Definition 12. Let $T \in \mathscr{D}'(\mathbb{R}^n)$ and $\alpha = (\alpha_1, ..., \alpha_n)$ a multi-index. The (distributional) <u>derivative of order α</u> of T, $D^{\alpha}T$, is defined by

$$(4) \qquad <D^{\alpha}T, \varphi> = (-1)^{|\alpha|}<T, D^{\alpha}\varphi>, \ \varphi \in \mathscr{D}(\mathbb{R}^n) .$$

Proposition 13. (i) $D^{\alpha}T$ is a distribution.

(ii) $\forall \alpha$ the map $D^{\alpha} : \mathscr{D}'(\mathbb{R}^n) \to \mathscr{D}'(\mathbb{R}^n)$ is continuous with respect to either the strong or weak* topologies of $\mathscr{D}'(\mathbb{R}^n)$.

Proof: Note that except that for the factor $(-1)^{|\alpha|}$, the map $T \to D^{\alpha}T$ is the transpose of the map $\varphi \to D^{\alpha}\varphi$ from $\mathscr{D}(\mathbb{R}^n)$ into $\mathscr{D}(\mathbb{R}^n)$

and this map is continuous (Example 22.14). Thus, (i) and (ii) are immediate from 26.14.

It now follows that every distribution has derivatives of all orders (even Weierstrass's nowhere differentiable, continuous function) and moreover if $T_k \to T$ in $\mathscr{D}'(\mathbb{R}^n)$, either strongly or weak*, then $D^\alpha T_k \to D^\alpha T$, either strongly or weakly, respectively. This should be contrasted with the classical case (see Example 15 below).

Example 14. Let $H(t) = 1$ if $t \geq 0$, $H(t) = 0$ if $t < 0$; H is called the Heaviside function. Then $<DH, \varphi> = -<H, \varphi'> = -\int_0^\infty \varphi' = \varphi(0) = <\delta_0, \varphi>$ for $\varphi \in \mathscr{D}(\mathbb{R})$ so the (distributional) derivative of the Heaviside function is the Dirac delta measure ("function") concentrated at the origin.

Example 15. Let $f_k(t) = (\sin kt)/\sqrt{k}$. Since $\{f_k\}$ converges uniformly to 0, $f_k \to 0$ in $\mathscr{D}'(\mathbb{R})_b$ (Exer. 10). Thus, $f_k' \to 0$ in $\mathscr{D}'(\mathbb{R})_b$ although $f_k'(t) = \sqrt{k}\cos kt$ doesn't even converge pointwise.

For further examples, see the exercises.

Multiplication by C^∞ functions:

If $\alpha = (\alpha_1, ..., \alpha_n)$ and $\beta = (\beta_1, ..., \beta_n)$ are multi-indices, we define $\alpha \leq \beta$ if and only if $\alpha_i \leq \beta_i$ for $i = 1, ..., n$, $\alpha! = \alpha_1!...\alpha_n!$ and if $\alpha \leq \beta$, $\binom{\beta}{\alpha} = \beta!/\alpha!(\beta - \alpha)!$. With this notation, we have the general Leibniz

Rule for differentiation of products.

Lemma 16. If φ and ψ have continuous partials up to order $|\beta|$ in some neighborhood of x, then $D^{\beta}(\varphi\psi)(x) = \sum_{\alpha\leq\beta} \binom{\beta}{\alpha} D^{\alpha}\varphi(x) D^{\beta-\alpha}\psi(x)$.

Proposition 17. Let $\psi \in C^{\infty}(\Omega)$. The linear map $\varphi \to \psi\varphi$ from $\mathscr{D}(\Omega)_b$ into $\mathscr{D}(\Omega)_b$ is continuous.

Proof: Lemma 16 and Theorem 1.

We now want to define the product of a distribution by a C^{∞} function.

Motivation: If $\psi \in C^{\infty}(\Omega)$ and f is locally integrable on Ω, then ψf is locally integrable and therefore induces a distribution satisfying

$$<(\psi f)^{\#}, \varphi> = \int_{\Omega} \psi f \varphi = <f^{\#}, \psi\varphi> \text{ for } \varphi \in \mathscr{D}(\Omega).$$

Definition 18. Let $\psi \in C^{\infty}(\Omega)$ and $T \in \mathscr{D}'(\Omega)$. The product, ψT, is defined to be $<\psi T, \varphi> = <T, \psi\varphi>$, $\varphi \in \mathscr{D}(\Omega)$.

Since the map $T \to \psi T$ is the transpose of the map in Proposition 17, we have from 26.14

Proposition 19. For $\psi \in C^{\infty}(\Omega)$ and $T \in \mathscr{D}'(\Omega)$, $\psi T \in \mathscr{D}'(\Omega)$ and the map $T \to \psi T$ is continuous from $\mathscr{D}'(\Omega)$ into $\mathscr{D}'(\Omega)$ with respect to either the strong or weak* topologies.

The Leibniz differentation rule in Lemma 16 extends to the product of a C^∞ function and a distribution.

Proposition 20. If $\psi \in C^\infty(\Omega)$ and $T \in \mathscr{D}'(\Omega)$, then for any multi-index β,

$$D^\beta(\psi T) = \sum_{\alpha\le\beta} \binom{\beta}{\alpha} D^\alpha \psi D^{\beta-\alpha} T .$$

We leave the proof as an exercise for the reader.

Example 21. Let $\psi \in C^\infty(\mathbb{R}^n)$. Then

$$<\psi\delta_0, \varphi> = <\delta_0, \psi\varphi> = \psi(0)\varphi(0) = \psi(0)<\delta_0, \varphi>$$

so $\psi\delta_0 = \psi(0)\delta_0$. Let $D^i = \frac{\partial}{\partial x_i}$. Then

$$D^i(\psi\delta_0) = \psi D^i\delta_0 + (D^i\psi)\delta_0 = \psi D^i\delta_0 + (D^i\psi(0))\delta_0 .$$

Structure of Distributions:

If $g : \Omega \to \mathbb{R}$ is continuous, recall that the support of g is the closure of $\{x : g(x) \neq 0\}$. We first give a result which gives a local description of a distribution.

Theorem 22. Let $T \in \mathscr{D}'(\mathbb{R}^n)$ and let Ω be a bounded open subset of

$\mathbb{R}^n$. Then $\exists$ a multi-index α and a continuous function g with support in an arbitrary neighborhood of $\bar{\Omega}$ such that $\langle T, \varphi\rangle = \langle D^\alpha g, \varphi\rangle$ $\forall\ \varphi \in \mathscr{D}(\Omega)$.

Proof: By Proposition 5, $\exists$ $c > 0$ and a non-negative integer N such that $|\langle T, \varphi\rangle| \le c\|\varphi\|_N$ $\forall\ \varphi \in \mathscr{D}_{\bar{\Omega}}$. By the Mean Value Theorem, for any $\varphi \in \mathscr{D}_{\bar{\Omega}}$, $\sup\{|\varphi(x)| : x \in \Omega\} \le \mathrm{diam}(\Omega)\sup\{|\frac{\partial\varphi}{\partial x_i}(x)| : x \in \Omega\}$ so for some positive integer ℓ,

$$\sup\{|D^\alpha\varphi(x)| : |\alpha| \le N, x \in \Omega\} \le (\mathrm{diam}\ \Omega)^\ell \sup\{|D^{\bar{N}}\varphi(x)| : x \in \Omega\},$$

where $\bar{N} = (N, \ldots, N)$. For $\varphi \in \mathscr{D}_{\bar{\Omega}}$,

$$\sup\{|\varphi(x)| : x \in \Omega\} \le \int_{\mathbb{R}^n} |D^{\bar{1}}\varphi(x)|dx \text{ so } \exists\ M > 0$$

such that

$$(5) \qquad |\langle T, \varphi\rangle| \le M\int_{\mathbb{R}^n} |D^{\overline{N+1}}\varphi(x)|dx \ \ \forall\ \varphi \in \mathscr{D}_{\bar{\Omega}}.$$

Let $\Delta = \{\psi : \psi = D^{\overline{N+1}}\varphi$ for some $\varphi \in \mathscr{D}_{\bar{\Omega}}\}$. Then Δ is a linear subspace of $L^1(\bar{\Omega})$ and the linear functional $L : \Delta \to \mathbb{R}$ defined by $\langle L, \psi\rangle = \langle L, D^{\overline{N+1}}\varphi\rangle = \langle T, \varphi\rangle$ is continuous with respect to the L^1-norm by (5). [Note L is well-defined since if $\psi = 0$ and $\psi = D^{\overline{N+1}}\varphi$ with $\varphi \in \mathscr{D}(\mathbb{R}^n)$, then $\varphi = 0$.] By the Hahn-Banach Theorem, L can be called a $\underline{B^*\text{-algebra}}$. A $\underline{B^*\text{-subalgebra}}$ X_0 of a B^*-algebra is a closed subalgebra which is closed under involution and contains the identity. A $\underline{C^*\text{-algebra}}$ is a B^*-subalgebra of $L(H)$. [Note that $L(H)$ is a B^*-algebra by 34.4.]

As noted above $L(H)$ is a B^*-algebra and likewise $C(S)$ is a B^*-algebra.

Proposition 6: Let X be a B^*-algebra.

(i) If $x \in X$ is normal, then $\|x^2\| = \|x\|^2$ and $r(x) = \|x\|$ (compare 35.20 and 35.26).

(ii) If $x \in X$ is Hermitian, then $\sigma(x) \subseteq \mathbb{R}$ (compare 35.27).

Proof: (i):

$$\|x^2\|^2 = \|x^2(x^2)^*\| = \|x^2(x^*)^2\| = |xxx^*x^*\| = \|xx^*xx^*\|$$

$$= \|(xx^*)(xx^*)^*\| = \|xx^*\|^2 = \|x\|^4$$

so $\|x^2\| = \|x\|^2$. Thus, $\|x^n\| = \|x\|^n$ for $n = 2^j$ and

$$r(x) = \lim \sqrt[n]{\|x^n\|} = \|x\|$$

(42.16).

(ii): Suppose $x = x^*$ and let X_0 be the commutative B-algebra generated by x and e. Let Δ_0 be the maximal ideal space of X_0. For $\varphi \in \Delta_0$, write $\varphi(x) = r + is$, $r, s \in \mathbb{R}$. Set $y = x + ite$, where $t \in \mathbb{R}$. Note $yy^* = (x^* + ite)(x - ite) = x^2 + t^2e$. Also $y \in X_0$ and if each point $x \in \Omega$ has a neighborhood $W(x)$ which intersects only a finite number of the $\{G_i : i \in I\}$.

Definition 23. If the open cover $\{U_j : j \in \mathbb{N}\}$ of Ω is locally finite, a <u>partition of unity</u> subordinate to $\{U_j\}$ is a sequence $\{\varphi_j\} \subseteq \mathcal{D}(\Omega)$ such that

(i) $0 \le \varphi_j \le 1 \ \forall j \in \mathbb{N}$,

(ii) the support of φ_j is contained in U_j,

(iii) $\sum_{j=1}^{\infty} \varphi_j(x) = 1 \ \forall x \in \Omega$.

Note by the local finiteness of $\{U_j\}$, the series in (iii) only has a finite number of non-zero terms $\forall x \in \Omega$.

Proposition 24. If $\{U_j : j \in \mathbb{N}\}$ is a locally finite, bounded, open cover of $\mathbb{R}^n$, $\exists$ a partition of unity subordinate to $\{U_j\}$.

Proof: We first construct a locally finite cover $\{V_j\}$ of $\mathbb{R}^n$ such that $\bar{V}_j \subseteq U_j$.

Note the complement of $\bigcup_{j=2}^{\infty} U_j$ is a closed set $F_1 \subseteq U_1$. Choose V_1 to be an open set such that $F_1 \subseteq V_1 \subseteq \bar{V}_1 \subseteq U_1$. Suppose $V_1, \ldots, V_{m-1}$ have been chosen such that $\bar{V}_k \subseteq U_k$ for $k = 1, \ldots, m-1$ and

$$\{V_1, \ldots, V_{m-1}, U_m, U_{m+1}, \ldots\}$$

is a locally finite cover of $\mathbb{R}^n$. The complement of $(\bigcup_{i=1}^{m-1} V_i) \cup (\bigcup_{i=m+1}^{\infty} U_i)$ is a closed set $F_m \subseteq U_m$. Choose V_m to be an open set such that $F_m \subseteq V_m \subseteq \bar{V}_m \subset U_m$. Then $\{V_j\}$ is a locally finite cover of $\mathbb{R}^n$ such that $\bar{V}_j \subseteq U_j$.

By Lemma 8 $\exists\ \psi_j \in \mathscr{D}(\mathbb{R}^n)$ such that $\psi_j(x) = 1\ \forall\, x \in \bar{V}_j$ and $\psi_j(x) = 0\ \ \forall\, x \notin U_j$ and $0 \le \psi_j(x) \le 1$. Set $h(x) = \sum_{j=1}^{\infty} \psi_j(x)$; this series converges since $\forall\, x$ only a finite number of the terms are non-zero and also $h(x) \ge 1$ since at least one term in the series is equal to 1. Moreover, h is actually infinitely differentiable since for each x, $\exists$ a neighborhood W(x) of x such that W(x) intersects only a finite number of the $\{U_j\}$ and the series for h only involves a fixed finite number of terms for each point in W(x). If we set $\varphi_j(x) = \psi_j(x)/h(x)$, then $\{\varphi_i\}$ is the desired partition of unity.

We now give a global representation for a distribution.

Theorem 25. Let $T \in \mathscr{D}'(\mathbb{R}^n)$ and $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$. Then $\exists$ a sequence $\{\alpha_j\} \subseteq \mathbb{N}_0^n$, a sequence $\{\psi_j\} \subseteq \mathscr{D}(\mathbb{R}^n)$ and a sequence of continuous functions with compact support, $\{g_j\}$, such that

(6) $$T = \sum_{j=1}^{\infty} \psi_j D^{\alpha_j} g_j\,,$$ where the series is locally finite.

Proof: Let $\{U_j\}$ be a locally finite, bounded, open cover of $\mathbb{R}^n$

and $\{\psi_j\}$ a partition of unity subordinate to $\{U_j\}$. For each j let α_j and g_j be as in Theorem 22 with respect to $\Omega = U_j$.

If $\varphi \in \mathscr{D}(\mathbb{R}^n)$, then $\varphi = \sum_{j=1}^{\infty} \varphi\psi_j$ pointwise, and the series actually converges in $\mathscr{D}(\mathbb{R}^n)$ since the support of φ only intersects a finite number of the $\{U_j\}$.

Then

$$<T, \varphi> = <T, \sum_{j=1}^{\infty} \varphi\psi_j> = \sum_{j=1}^{\infty} <T, \psi_j\varphi>$$

$$= \sum_{j=1}^{\infty} <D^{\alpha_j} g_j, \psi_j\varphi> = \sum_{j=1}^{\infty} <\psi_j D^{\alpha_j} g_j, \varphi> .$$

Each term in the series (6) can be written as the finite sum of derivatives of continuous functions with compact support.

Lemma 26. Let $\psi \in C^\infty(\mathbb{R}^n)$ and $T = D^\alpha g$, where g is continuous. Then $\psi T = \sum_{q \le \alpha} (-1)^{|\alpha|+|q|} \binom{\alpha}{q} D^q(gD^{\alpha-q}\psi)$.

Proof:

$$<\psi T, \varphi> = <D^\alpha g, \psi\varphi> = (-1)^{|\alpha|} <g, D^\alpha(\psi\varphi)>$$

$$= (-1)^{|\alpha|} <g, \sum_{q \le \alpha} \binom{\alpha}{q} D^q\varphi D^{\alpha-q}\psi>$$

$$= \sum_{q \le \alpha} (-1)^{|\alpha|} \binom{\alpha}{q} <(D^{\alpha-q}\psi)g, D^q\varphi>$$

$$= \sum_{q \le \alpha} (-1)^{|\alpha|} \binom{\alpha}{q} (-1)^{|q|} \langle D^q[(D^{\alpha-q}\psi)g], \varphi \rangle .$$

The representation in Lemma 26 can be applied to each term in the series (6) to obtain a global representation of a distribution as an infinite series of finite derivatives of continuous functions.

For an interesting expository article on distributions, see [H2]; [L] contains a historical account of the theory of distributions.

Exercise 1. Show $\mathscr{D}(\mathbb{R}^n)$ is dense in $L^p(\mathbb{R}^n)$ for $1 \le p < \infty$.

Exercise 2. Show the injections of $L^p(\mathbb{R}^n)$, $\mathscr{E}(\mathbb{R}^n)$ and $\mathscr{D}(\mathbb{R}^n)$ into $\mathscr{D}'(\mathbb{R}^n)$ are continuous.

Exercise 3. Show the series $\sum_{k=1}^{\infty} \delta_k$ and $\sum_{k=1}^{\infty} D^k \delta_k$ define distributions in $\mathscr{D}'(\mathbb{R})$.

Exercise 4. Show $\langle D^\alpha \delta_x, \varphi \rangle = (-1)^{|\alpha|} D^\alpha \varphi(x)$.

Exercise 5. Give necessary and sufficient conditions on $\{t_j\}$ so that the series $\sum_{j=1}^{\infty} t_j D^j \delta_0$ converges in $\mathscr{D}'(\mathbb{R})$.

Exercise 6. Find the distributional derivative of $\ln|t|$.

Exercise 7. Let $T \in \mathscr{D}'(\mathbb{R}^n)$ be of finite order N. Show T can be represented as a finite sum of derivatives of bounded Borel measures on $\mathbb{R}^n$.

Exercise 8. Show the distribution $T = \sum_{k=0}^{\infty} D^k \delta_k$ is not of finite order.

Exercise 9. Fix $h \in \mathbb{R}^n$. Define the translation operator τ_h on $\mathscr{D}(\mathbb{R}^n)$ by $\tau_h \varphi(x) = \varphi(x + h)$. How should τ_h be defined on $\mathscr{D}'(\mathbb{R}^n)$? Is it continuous?

Exercise 10. Let $f_k : \mathbb{R}^n \to \mathbb{R}$ be locally integrable and suppose that $\{f_k\}$ converges uniformly to 0 on compact subsets. Show $f_k \to 0$ in $\mathscr{D}'(\mathbb{R}^n)_b$.

Exercise 11. If $f(t) = |t|$, compute the first and second derivatives of f.

Exercise 12. Let $f : \mathbb{R}^n \to \mathbb{R}$ be given by $f(x_1, \ldots, x_n) = 1$ if each $x_i \geq 0$ and 0 otherwise. Compute $\frac{\partial f}{\partial x_i}$.

Exercise 13. Let $f_\varepsilon(t) = \frac{1}{\pi} \frac{\varepsilon}{t^2 + \varepsilon^2}$ $(\varepsilon > 0)$. Show $\lim_{\varepsilon \to 0} f_\varepsilon = \delta_0$ in $\mathscr{D}'(\mathbb{R})$.

Exercise 14. Let $\mathscr{S}(\mathbb{R}^n)$ be the space of all infinitely differentiable functions $\varphi : \mathbb{R}^n \to \mathbb{R}$ such that

$$\sup\{(1 + \|x\|^2)^k |D^\alpha \varphi(x)| : |\alpha| \le k, x \in \mathbb{R}^n\} = \|\varphi\|_k < \infty$$

for all $k = 0, 1, \ldots$. $\mathscr{S}(\mathbb{R}^n)$ is called the space of rapidly decreasing functions.

(a) Show that $\mathscr{S}(\mathbb{R}^n)$ is complete when it has the locally convex topology generated by the semi-norms $\{\| \ \|_k : k \ge 0\}$.

(b) Show that $\mathscr{D}(\mathbb{R}^n)$ is dense in $\mathscr{S}(\mathbb{R}^n)$ with respect to this topology and $\mathscr{D}(\mathbb{R}^n) \ne \mathscr{S}(\mathbb{R}^n)$.

(c) Show the inclusion $\mathscr{D}(\mathbb{R}^n) \to \mathscr{S}(\mathbb{R}^n)$ is continuous.

(d) Show that every element $T \in \mathscr{S}(\mathbb{R}^n)'$ can be identified with a unique distribution in $\mathscr{D}(\mathbb{R}^n)'$. These distributions in $\mathscr{S}(\mathbb{R}^n)'$ are called tempered distributions.

(e) If $T \in \mathscr{S}(\mathbb{R}^n)'$, show $D^\alpha T$ is a tempered distribution for each multi-index α and the map $T \to D^\alpha T$ is weak* continuous.

(f) Show that the Fourier transform is a topological isomorphism from $\mathscr{S}(\mathbb{R}^n)$ onto itself.

(g) Use duality and Plancherel's formula to define the Fourier transform on $\mathscr{S}'(\mathbb{R}^n)$.

26.5 The Lax-Milgram Theorem

In this section we give an application of distribution theory to elliptic partial differential equations (PDE). The principal tool which we use is the Lax-Milgram Theorem which can be viewed as a generalization of the Riesz Representation Theorem for Hilbert space.

Let H be a Hilbert space, which for convenience we assume to be real. A bilinear form $b : H \times H \to \mathbb{R}$ is called <u>coercive</u> if there is a positive constant $c > 0$ such that $b(x, x) \geq c\|x\|^2$ for all $x \in H$.

Theorem 1 (Lax-Milgram). Suppose b is a continuous, coercive bilinear form on H. For each $y \in H$ there is a unique $u \in H$ such that $b(u, x) = y \cdot x \quad \forall\, x \in H$.

Proof: For each $z \in H$, $b(\cdot, z)$ is a continuous linear functional on H so by the Riesz Representation Theorem there is a unique $y_z \in H$ such that $b(x, z) = y_z \cdot x$ for all $x \in H$. This defines a mapping $B : H \to H$ by $Bz = y_z$ which is linear since b is bilinear. Since b is coercive, $c\|z\|^2 \leq b(z, z) = Bz \cdot z \leq \|Bz\|\|z\|$ or $c\|z\| \leq \|Bz\|$ for all $z \in H$. Hence, B has a continuous inverse (23.14). Moreover, B is continuous since

$$|b(z, Bz)| = Bz \cdot Bz = \|Bz\|^2 \leq k\|z\|\|Bz\|$$

(where $k = \sup\{|b(x, y)| : \|x\| \leq 1, \|y\| \leq 1\}$ Exer. 9.1.2). Hence, B has closed range (10.14). We claim that $\mathscr{R}B = H$. If this is not the case, $\exists$ $w \neq 0$ such that $Bz \cdot w = b(z, w) = 0 \quad \forall\, z \in H$ (8.1.2 and the Riesz Representation Theorem). Setting $z = w$ implies $0 = Bw \cdot w = b(w, w)$

and $w = 0$ by the coerciveness of b. Thus, $\mathscr{R}B = H$.

If $y \in H$, then $u = B^{-1}y$ satisfies $y \cdot x = Bu \cdot x = b(u, x)$ $\forall\, x \in H$.

We give an application of Theorem 1 to elliptic PDE. We first describe the function space which will be used. Let $\Omega \subseteq \mathbb{R}^n$ be open and let k be a positive integer. For $1 \leq p < \infty$, $W^{k,p}(\Omega)$ is the vector space of all functions $f : \Omega \to \mathbb{R}$ such that f and all distributional derivatives $D^\alpha f$ for $|\alpha| \leq k$ belong to $L^p(\Omega)$. We define a norm on $W^{k,p}(\Omega)$ by

$$\|f\|_{k,p} = \Big(\sum_{|\alpha| \leq k} \int |D^\alpha f|^p \Big)^{1/p}$$

(all integrals are over Ω). For $p = 2$, $W^{k,2}(\Omega)$ is denoted by $W^k(\Omega)$; its norm is induced by the inner product $f \cdot g = \sum_{|\alpha| \leq k} \int (D^\alpha f)(D^\alpha g)$. These spaces are called <u>Sobolev spaces</u> and are important in PDE.

Theorem 2. $W^{k,p}(\Omega)$ is a B-space; in particular, $W^k(\Omega)$ is a Hilbert space.

Proof: Let $\{f_j\}$ be Cauchy in $W^{k,p}(\Omega)$. Then $\{f_j\}$ and $\{D^\alpha f_j\}$ are Cauchy in $L^p(\Omega)$ for each $0 < |\alpha| \leq k$. Let $f_j \to f$ and $D^\alpha f_j \to f^\alpha$ in $L^p(\Omega)$ for $0 < |\alpha| \leq k$. Let $\varphi \in \mathscr{D}(\Omega)$. By Hölder's Inequality,

$$\begin{aligned}\lim \langle D^\alpha f_j, \varphi\rangle &= \lim \int (D^\alpha f_j)\varphi = \lim (-1)^{|\alpha|} \int f_j D^\alpha \varphi \\ &= \int f^\alpha \varphi = (-1)^{|\alpha|} \int f D^\alpha \varphi = \langle D^\alpha f, \varphi\rangle\end{aligned}$$

so $D^\alpha f = f^\alpha \in L^p(\Omega)$ or $f \in W^{k,p}(\Omega)$ and $f_j \to f$ in $W^{k,p}(\Omega)$.

Let $C_0^k(\Omega)$ be all functions $f : \Omega \to \mathbb{R}$ with compact support and such that $D^\alpha f$ is continuous for all $|\alpha| \le k$. We equip $C_0^k(\Omega)$ with the norm $\|f\| = \left(\sum_{|\alpha| \le k} \int |D^\alpha f|^2 \right)^{1/2}$ and denote its completion by $H_0^k(\Omega)$. $H_0^k(\Omega)$ is a closed subspace of $W^k(\Omega)$ and is, therefore, a Hilbert space. (There are situations when $H_0^k(\Omega) = W^k(\Omega)$, see [Y], I.10.)

Henceforth, we assume that Ω is bounded.

Let a_{ij} be C^1 on $\bar{\Omega}$ and let c be a non-negative continuous function on $\bar{\Omega}$. Define a partial differential operator P by

$$Pu(x) = -\sum_{i,j=1}^{n} D_i(a_{ij}(x)D_j u(x)) + c(x)u(x)$$

for $x \in \Omega$, where $D_i = \frac{\partial}{\partial x_i}$. The operator P is said to be <u>strongly elliptic</u> if there exists $\delta > 0$ such that $\sum_{i,j=1}^{n} a_{ij}(x)\xi_i\xi_j \ge \delta(\xi_1^2 + \dots + \xi_n^2)$ for all $x \in \Omega$, $\xi \in \mathbb{R}^n$. For example, the Laplace operator $P = \frac{\partial^2}{\partial x_1^2} + \dots + \frac{\partial^2}{\partial x_n^2}$ is strongly elliptic.

The operator P induces a bilinear form b on $C_0^1(\Omega)$ which can be written as

$$(1) \qquad b(u, v) = Pu \cdot v = \int (Pu)v = \sum_{i,j=1}^{n} \int a_{ij} D_i u D_j v + \int cuv$$

(Exer. 1). If P is strongly elliptic, then for $u \in C_0^1(\Omega)$

$$(2) \qquad b(u, u) = \sum_{i,j=1}^{n} \int a_{ij} D_i u D_j u + \int cu^2 \ge \delta \sum_{j=1}^{n} \int (D_j u)^2$$

since $c \geq 0$.

Lemma 3. There is a positive constant k such that $\int |u|^2 \leq k \sum_{j=1}^{n} \int |D_j u|^2$ for all $u \in H_0^1(\Omega)$.

Proof: Choose $a > 0$ such that Ω is contained in the product $\prod_{i=1}^{n} [-a, a] = \Omega'$. Let $u \in C_0^1(\Omega)$ and extend u to Ω' by setting $u = 0$ on $\Omega' \backslash \Omega$. For $x \in \Omega'$ write $x = (x', x_n)$ where $x' \in \mathbb{R}^{n-1}$, $x_n \in \mathbb{R}$. Then

$$u(x) = \int_{-a}^{x_n} D_n u(x', t)dt$$

so

$$|u(x)|^2 \leq (2a)\int_{-a}^{a} |D_n u(x', t)|^2 dt.$$

Integrating over x' gives

$$\int |u(x', t)|^2 dx' \leq 2a \int_{\Omega} |D_n u|^2$$

so

$$\int_{\Omega} |u|^2 \leq (2a)^2 \int_{\Omega} |D_n u|^2.$$

Treating the other variables in a similar fashion gives the desired inequality for $u \in C_0^1(\Omega)$ and, therefore, for $u \in H_0^1(\Omega)$.

From (2) and Lemma 3, it follows that $b(u, u) \geq k\|u\|_{1,2}^2$ for some constant k and $u \in C_0^2(\Omega)$. Since each a_{ij} is bounded on Ω, then there is a constant K such that $|b(u, v)| \leq K\|u\|_{1,2}\|v\|_{1,2}$ for $u, v \in C_0^1(\Omega)$

(Exer. 2). Thus, b is a continuous bilinear functional on $C_0^1(\Omega)$ with respect to $\| \|_{1,2}$ and since $C_0^1(\Omega)$ is dense in $H_0^1(\Omega)$, b has a continuous bilinear extension to $H_0^1(\Omega)$. When P is strongly elliptic, it follows from the inequality above that b is also coercive.

By the Lax-Milgram Theorem, for every $f \in H_0^1(\Omega)$ there is a unique $u \in H_0^1(\Omega)$ such that $b(u, \varphi) = f \cdot \varphi = Pu \cdot \varphi$ for every $\varphi \in C_0^1(\Omega)$. u is called a weak solution to the PDE $Pu = f$; u satisfies the equation

$$\int Pu(x)\varphi(x)dx = \int f(x)\varphi(x)dx$$

for every $\varphi \in C_0^1(\Omega)$.

There are theorems, called regularity theorems, in PDE which guarantee that a weak solution (in the sense above) to a PDE is actually a classical solution, i.e., are actually smooth functions. The classical theorem of this type is the Weyl regularity theorem for the Laplace operator which asserts that any continuous function u satisfying $\sum_{i=1}^{n} D_i^2 u = 0$ (in the distributional sense) is actually a harmonic function (see [Ag], 3.9).

Distribution theory is one of the most important tools of modern PDE. The modest example above, hopefully, gives some indication as to how distributions arise in PDE in a natural way. For additional information, the reader can consult the interesting expository article in [T2] or the text [T3].

Exercise 1. Derive (1).

Exercise 2. Show the bilinear form b above is bounded on $C_0^1(\Omega)$.

26.6 Distributions with Compact Support

The support of a continuous function $f : \mathbb{R}^n \to \mathbb{R}$ is the closure of $\{t : f(t) \neq 0\}$; denote this set by <u>spt(f)</u>. If $t \notin$ spt(f), $\exists$ a neighborhood U of t which is disjoint from spt(f) such that $\langle f^{\#}, \varphi \rangle = 0$ whenever $\varphi \in \mathscr{D}(U)$; we say that $f^{\#}$ vanishes in U. More generally, for distributions we have

Definition 1. Let $\Omega \subseteq \mathbb{R}^n$ be open and $T \in \mathscr{D}'(\Omega)$. If $W \subseteq \Omega$ is open, <u>T vanishes in W</u> or <u>T = 0 in W</u> if $\langle T, \varphi \rangle = 0 \ \forall \varphi \in \mathscr{D}(W)$.

If f is a continuous function, $f^{\#} = 0$ on the open set $\mathbb{R}^n \backslash$spt(f) and this is clearly the largest open set on which $f^{\#}$ vanishes. We show that there is an analogous set for any distribution.

Lemma 2. Let $\Omega = \bigcup_{i=1}^{\infty} G_i$, where each G_i is open. If $\varphi \in \mathscr{D}(\Omega)$, then $\exists$ a finite set $\varphi_1, \ldots, \varphi_p \in \mathscr{D}(\Omega)$ such that $\text{spt}(\varphi_i) \subseteq G_i$ and $\varphi = \sum_{i=1}^{p} \varphi_i$.

Proof: Let $\varphi \in \mathscr{D}(\Omega)$ and $K = \text{spt}(\varphi)$. A finite number of the $\{G_i\}$, say $G_1, \ldots, G_p$, cover K. Therefore, $G_1, \ldots, G_p$ and $V = \mathbb{R}^n \backslash K$ cover $\mathbb{R}^n$. Any unbounded G_i, $i = 1, \ldots, p$, and V can be replaced by a countable union of bounded open sets which together with the other G_i, $i = 1, \ldots, p$, form a locally finite, bounded, open cover of $\mathbb{R}^n$. Let $\{\psi_j\}$ be a partition of unity subordinate to this locally finite cover. Then

$\varphi = \sum_{j=1}^{\infty} \psi_j \varphi$, where the series is actually finite and the terms in the series corresponding to subsets of V are zero because φ vanishes on them. Combining the terms corresponding to the subsets of $G_1, \ldots, G_p$ gives the desired expansion.

Theorem 3. Let $T \in \mathscr{D}'(\Omega)$. Let $\Gamma = \{W : W \subseteq \Omega$ open, $T = 0$ on $W\}$ and set $\Omega' = \cup\{W : W \in \Gamma\}$. Then T vanishes on Ω'.

Proof: Let $\varphi \in \mathscr{D}(\Omega')$ and $K = \text{spt}(\varphi)$. $\exists$ a finite $W_1, \ldots, W_p \in \Gamma$ covering K. Let $\varphi_1, \ldots, \varphi_p$ be as in Lemma 2 with respect to $W_1, \ldots, W_p$. Then $\langle T, \varphi \rangle = \sum_{j=1}^{p} \langle T, \varphi_j \rangle = 0$.

Thus, we can make the following definition of the support of a distribution.

Definition 4. Let $T \in \mathscr{D}'(\Omega)$. The <u>support of T, spt(T)</u>, is the complement of the largest open subset of Ω on which T vanishes.

Note that this agrees with the definition of the support of a continuous function when the distribution is induced by a continuous function. For we saw above that $\text{spt}(f^{\#}) \subseteq \text{spt}(f)$, and if, conversely, $t \notin \text{spt}(f^{\#})$, then $\exists$ an open set W containing t which is disjoint from $\text{spt}(f^{\#})$ and $f = 0$ on W. That is, $\langle f^{\#}, \varphi \rangle = \int f\varphi = 0 \quad \forall\, \varphi \in \mathscr{D}(W)$.

Hence, $f = 0$ on W and $t \notin \text{spt}(f)$. This justifies the use of the notation in Definition 4.

Example 5. $\text{spt}(D^{\alpha}\delta_x) = \{x\}$.

Proposition 6. Let $T \in \mathscr{D}'(\Omega)$.

(i) If $\varphi \in \mathscr{D}(\Omega)$ is such that $\text{spt}(\varphi) \cap \text{spt}(T) = \emptyset$, then $\langle T, \varphi \rangle = 0$.

(ii) If $\psi \in \mathscr{E}(\Omega)$ and $\psi = 1$ on some open set containing $\text{spt}(T)$, then $\psi T = T$.

Proof: (i) is clear. For (ii), if $\varphi \in \mathscr{D}(\Omega)$, then $\text{spt}(\varphi - \psi\varphi) \cap \text{spt}(T) = \emptyset$ so $\langle T, \varphi - \psi\varphi \rangle = \langle T, \varphi \rangle - \langle \psi T, \varphi \rangle = 0$ by (i).

We now show that the distributions with compact support can be identified with the dual of the space $\mathscr{E}(\Omega)$ (2.28). Recall that a sequence $\{\psi_j\} \subseteq \mathscr{E}(\Omega)$ converges to 0 if and only if $\{D^{\alpha}\psi_j\}$ converges uniformly to 0 on compact subsets of Ω for each multi-index α (2.28). It follows from this fact that the differential operator $D^{\alpha} : \varphi \to D^{\alpha}\varphi$ is continuous $\mathscr{E}(\Omega)$ into itself. Additionally, we have

Proposition 7. $\mathscr{D}(\Omega) \subseteq \mathscr{E}(\Omega)$ and the injection is continuous and has dense range.

Proof: The continuity is clear. For the denseness, let $\varphi \in \mathscr{E}(\Omega)$.

Choose a sequence of open subsets Ω_j of Ω such that $\bar{\Omega}_j$ is compact, $\bar{\Omega}_j \subseteq \Omega_{j+1}$, $\Omega = \bigcup_{j=1}^{\infty} \Omega_j$ and every compact subset of Ω is contained in some Ω_j. For each j pick $\psi_j \in \mathscr{D}(\Omega)$ such that $\psi_j(x) = 1 \ \forall x \in \Omega_j$, $\psi_j(x) = 0 \ \forall x \notin \Omega_{j+1}$ and $0 \le \psi_j \le 1$ (26.4.8). Then $\psi_j \to 1$ in $\mathscr{E}(\Omega)$ so $\psi_j\varphi \to \varphi$ in $\mathscr{E}(\Omega)$ with $\psi_j\varphi \in \mathscr{D}(\Omega)$.

Since the injection of $\mathscr{D}(\Omega)$ into $\mathscr{E}(\Omega)$ is continuous and has dense range, its transpose is a 1-1 map (26.21) so the dual, $\mathscr{E}'(\Omega)$, of $\mathscr{E}(\Omega)$ can be identified with a linear subspace of $\mathscr{D}'(\Omega)$. We show that $\mathscr{E}'(\Omega)$ can be identified with the space of distributions with compact support.

Theorem 8. A distribution $T \in \mathscr{D}'(\Omega)$ has compact support if and only if $T \in \mathscr{E}'(\Omega)$.

Proof: The proof is based on the observation that a distribution T belongs to $\mathscr{E}'(\Omega)$ if and only if $\varphi \to \langle T, \varphi \rangle$ is continuous on $\mathscr{D}(\Omega)$ with the topology induced by the topology of $\mathscr{E}(\Omega)$ since if this linear form is continuous, it has a unique continuous linear extension to $\mathscr{E}(\Omega)$.

First, suppose $T \in \mathscr{E}'(\Omega)$. Then $\exists$ compact $K \subseteq \Omega$, a non-negative integer m and $c > 0$ such that $|\langle T, \varphi \rangle| \le c p_{K,m}(\varphi)$, where

$$p_{K,m}(\varphi) = \sup\{ |D^{\alpha}\varphi(x)| : |\alpha| \le m, x \in K \} ,$$

for all $\varphi \in \mathscr{E}(\Omega)$. Then $\langle T, \varphi \rangle = 0 \ \forall \varphi \in \mathscr{D}(\Omega)$ with $\mathrm{spt}(\varphi) \cap K = \emptyset$.

That is, T vanishes in $\Omega\backslash K$ and $\text{spt}(T) \subseteq K$.

Next, suppose that $K = \text{spt}(T)$ is compact. Pick $\psi \in \mathscr{D}(\Omega)$ such that $\psi = 1$ on some neighborhood of K so $\psi T = T$ (Proposition 6). All of the functions $\psi\varphi$, $\varphi \in \mathscr{D}(\Omega)$, have their supports contained in a fixed compact subset of Ω, namely, $\text{spt}(\psi) = L$. On $\mathscr{D}_L$ the topologies induced by $\mathscr{D}(\Omega)$ and $\mathscr{E}(\Omega)$ coincide so if a sequence $\{\varphi_k\}$ converges to 0 in $\mathscr{E}(\Omega)$, then $\{\psi\varphi_k\}$ converges to 0 in $\mathscr{D}(\Omega)$, and $\langle T, \psi\varphi_k\rangle = \langle \psi T, \psi_k\rangle = \langle T, \varphi_k\rangle \to 0$. Thus, the linear form $\varphi \to \langle T, \varphi\rangle$ is continuous on $\mathscr{D}(\Omega)$ with respect to the topology induced by $\mathscr{E}(\Omega)$.

Theorem 9. Let $T \in \mathscr{D}'(\Omega)$. Then $T \in \mathscr{E}'(\Omega)$ if and only if $\exists\ c > 0$ and a non-negative integer m such that

(1) $$|\langle T, \varphi\rangle| \le c\|\varphi\|_m, \text{ where}$$
$$\|\varphi\|_m = \sup\{|D^\alpha\varphi(x)| : x \in \Omega, |\alpha| \le m\}, \forall\ \varphi \in \mathscr{D}(\Omega).$$

Proof: Suppose $T \in \mathscr{E}'(\Omega)$. Pick $\psi \in \mathscr{D}(\Omega)$ such that $\psi = 1$ on an open set containing spt(T). Let $K = \text{spt}(\psi)$. Since $T \in \mathscr{D}'(\Omega)$, $\exists\ c_1$ and m such that $|\langle T, \varphi\rangle| \le c_1\|\varphi\|_m\ \forall\ \varphi \in \mathscr{D}_K$. Leibniz' Rule implies $\exists\ c_2 > 0$ such that $\|\psi\varphi\|_m \le c_2\|\varphi\|_m\ \forall\ \varphi \in \mathscr{D}(\Omega)$. Thus,

$$|\langle T, \varphi\rangle| = |\langle \psi T, \varphi\rangle| = |\langle T, \psi\varphi\rangle| \le c_1c_2\|\varphi\|_m\ \forall\ \varphi \in \mathscr{D}(\Omega).$$

The converse was observed in Theorem 8.

It follows from Theorem 9 that every distributions in $\mathscr{E}'(\Omega)$ has finite order (26.4.6). In particular, we have

Corollary 10. Let $T \in \mathcal{E}'(\Omega)$ and let V be an arbitrary neighborhood of $K = \mathrm{spt}(T)$ with $K \subseteq V \subseteq \Omega$. Then $\exists$ finitely many continuous functions $\{f_\beta\}$ on Ω with support in V such that

$$(2) \qquad \langle T, \varphi\rangle = \Sigma\langle D^\beta f_\beta, \varphi\rangle \ \forall\, \varphi \in \mathcal{E}(\Omega).$$

Proof: Choose W to be an open set with compact closure such that $K \subseteq W \subseteq \bar{W} \subseteq V$. Apply 26.4.22 with Ω as W to obtain a continuous function f on Ω with support in V such that

$$(3) \qquad \langle T, \varphi\rangle = \langle D^\alpha f, \varphi\rangle \ \forall\, \varphi \in \mathcal{D}(W).$$

Pick $\psi \in \mathcal{D}(\Omega)$ with $\mathrm{spt}(\psi) \subseteq W$ and $\psi = 1$ on some open set containing K. Then (3) implies

$$\langle T, \varphi\rangle = \langle \psi T, \varphi\rangle = \langle T, \psi\varphi\rangle = (-1)^{|\alpha|}\int_\Omega f D^\alpha(\psi\varphi)$$

$$= (-1)^{|\alpha|}\int_\Omega f \sum_{\beta\le\alpha} \binom{\alpha}{\beta} D^{\alpha-\beta}\psi D^\beta\varphi \ \forall\, \varphi \in \mathcal{D}(\Omega).$$

This gives (2) with $f_\beta = (-1)^{|\alpha|+|\beta|}\binom{\alpha}{\beta} f D^{\alpha-\beta}\psi$.

Finally, for use in a later application, we give a characterization of those distributions which have for their supports a single point.

Theorem 11. Let $T \in \mathcal{D}'(\Omega)$ and $x_0 \in \Omega$. Then $\mathrm{spt}(T) = \{x_0\}$ if and only if $\exists$ finitely many unique constants $\{c_\alpha\}$ such that

$$(4) \qquad T = \Sigma c_\alpha D^\alpha \delta_{x_0}.$$

Proof: $\Rightarrow$: For convenience assume $x_0 = 0$. Assume that (1) holds.

First, we claim that if $\varphi \in \mathscr{D}(\Omega)$ is such that $D^\alpha\varphi(0) = 0 \;\; \forall\; |\alpha| \le m$, then $\langle T, \varphi\rangle = 0$. If $\eta > 0$, $\exists$ a compact sphere $K \subseteq \Omega$ with center at 0 such that

$$(5) \qquad |D^\alpha\varphi| \le \eta \text{ in } K \;\; \forall\; |\alpha| = m.$$

From this we claim that

$$(6)\; |D^\alpha\varphi(x)| \le \eta n^{m-|\alpha|}\|x\|^{m-|\alpha|} \quad \text{for } x \in K,\ |\alpha| \le m\,.$$

For $|\alpha| = m$, (6) is just (5) so assume that (6) holds for all α with $|\alpha| = i$ and suppose $|\beta| = i - 1$. The induction hypothesis implies $\|\mathrm{grad}(D^\beta\varphi)(x)\| \le n\eta n^{m-i}\|x\|^{m-i}$, $x \in K$. Since $D^\beta\varphi(0) = 0$, the Mean Value Theorem implies that (6) holds with β in place of α and, thus, (6) holds for $|\alpha| \le m$. Choose $\psi \in \mathscr{D}(\mathbb{R}^n)$ such that $\psi = 1$ on a neighborhood of 0 and $\mathrm{spt}(\psi) \subseteq \{x : \|x\| \le 1\} = B$. Set $\psi_r(x) = \psi(x/r)$ for $r > 0$. If r is sufficiently small, $\mathrm{spt}(\psi_r) \subseteq K$. By Leibniz' Rule

$$D^\alpha(\psi_r\varphi)(x) = \sum_{\beta\le\alpha} \binom{\alpha}{\beta} D^{\alpha-\beta}\psi(x/r)D^\beta\varphi(x)r^{|\beta|-|\alpha|}$$

so from (6), $\|\psi_r\varphi\|_m \le \eta c_1\|\psi\|_m$ for r sufficiently small; the constant c_1 depends on n and m. Since $\psi_r = 1$ on some neighborhood of $\mathrm{spt}(T)$,

$$|\langle T, \varphi\rangle| = |\langle\psi_r T, \varphi\rangle| = |\langle T, \psi_r\varphi\rangle| \le c\|\psi_r\varphi\|_m \le \eta c c_1\|\psi\|_m\,.$$

Since η is arbitrary, $\langle T, \varphi\rangle = 0$ justifying our original claim.

Thus, T vanishes on the kernels of $\{D^\alpha\delta_0 : |\alpha| \le m\}$ and must be a linear combination of these distributions (14.5).

The converse is clear (Example 5). For the uniqueness apply both sides of (4) to x^α to obtain $\langle T, x^\alpha\rangle = c_\alpha\langle D^\alpha\delta_0, x^\alpha\rangle = (-1)^{|\alpha|}c_\alpha\alpha!$.

Corollary 12. Let $\Gamma \subseteq \mathcal{E}'(\mathbb{R}^n)$ be the subspace consisting of the distributions with support $\{0\}$ plus the zero distribution. Then Γ is weak* closed in $\mathcal{E}'(\mathbb{R}^n)$.

Proof: If $\{T_\delta\}$ is a net in Γ which is weak* convergent to $T \in \mathcal{E}'(\mathbb{R}^n)$, then T vanishes on $\mathbb{R}^n \backslash \{0\}$ and, therefore, belongs to Γ by Theorem 11.

Exercise 1. If $\psi \in \mathcal{E}(\mathbb{R}^n)$, $T \in \mathcal{D}'(\mathbb{R}^n)$, show $\operatorname{spt}(\psi T) \subseteq \operatorname{spt}(\psi) \cap \operatorname{spt}(T)$ with proper containment possible.

Exercise 2. Show $\operatorname{spt}(T + S) \subseteq \operatorname{spt}(T) \cup \operatorname{spt}(S)$.

Exercise 3. Let $T \in \mathcal{D}'(\mathbb{R}^n)$. Suppose $\forall\, x \in \mathbb{R}^n$ $\exists$ a neighborhood U_x of x and $\psi \in \mathcal{E}(\mathbb{R}^n)$ such that $\psi(y) \neq 0$ $\forall\, y \in \mathbb{R}^n$ satisfying $\psi T = 0$ in U_x. Show $T = 0$.

26.7 A Classical Theorem of Borel

We give one further application of distribution theory to a classical result of Borel.

Suppose that $f : \mathbb{R} \to \mathbb{R}$ is infinitely differentiable. What conditions must the sequence $\{f^{(k)}(0)\}_{k=0}^{\infty}$ satisfy? It is a remarkable result of Borel that the sequence can be an arbitrary sequence of real numbers. We now give a functional analytic proof of this result. Our proof uses Banach's Closed Range Theorem (26.1.2) for locally convex F-spaces; we only proved the result for B-spaces, but a general proof can be found in [Tr].

Theorem 1 (Borel). Let $\{a_k\}_{k=0}^{\infty}$ be an arbitrary sequence in $\mathbb{R}$. Then $\exists\ \varphi \in \mathscr{E}(\mathbb{R})$ such that $\varphi^{(k)}(0) = a_k$ for $k = 0, 1, \ldots$.

Proof: Define $T : \mathscr{E}(\mathbb{R}) \to s$ by $T\varphi = \{\varphi^{(k)}(0)\}_{k=0}^{\infty}$. To establish the result we must show that T is onto. Note T is linear and continuous.

First, $T' : c_{00} \to \mathscr{E}(\mathbb{R})$ is given by $T'\{t_k\} = \sum_{k=0}^{\infty} t_k D^k \delta_0$ and is 1-1 by 26.6.11. Hence, $\mathscr{R}T$ is (weak) dense by 26.21.

Next, $\mathscr{R}T'$ consists of all distributions with support equal to $\{0\}$ plus the zero distribution so $\mathscr{R}T'$ is weak* closed by 26.6.12. By the Closed Range Theorem (the equivalence of 26.1.2 (i) and (iii) for locally convex F-spaces), $\mathscr{R}T$ is closed and T must be onto.

The result can obviously be generalized to $\mathbb{R}^n$, but the notation becomes very clumsy. The proof given above is obviously an existence proof; for a more classical proof see [Me].

27
Projections

If M is a linear subspace of a vector space X, then there is always a linear subspace N of X such that X is the direct sum of M and N. However, what if X is a NLS and M is closed, is there always a closed subspace N of X such that X is the direct sum of M and N? Of course, if X is a Hilbert space, the answer is yes; we just take N to be the orthogonal complement of N (see the Appendix). However, we will see below that the situation is much different in a general NLS.

Let X be a B-space.

Definition 1. A projection on X is a continuous linear operator $P : X \to X$ satisfying $P^2 = P$.

Note that if P is a projection so is $I - P$.

Proposition 2. Let P be a projection. Set $M = \{x : Px = x\}$,

$$N = \{x : Px = 0\} = \{x : (I - P)x = x\}.$$

Then M and N are closed linear subspaces of X with $X = M \oplus N$.

Proof: Clearly M and N are closed linear subspaces and $M \cap N = \{0\}$.

Let $x \in X$. Then $x = Px + (I - P)x$ and $Px \in M$ since $P(Px) = P^2x = Px$ and $(I - P)x \in N$ since $P(I - P) = (P - P^2)x = 0$.

Note that $M = \mathcal{R}P$ and $N = \ker P = \mathcal{R}(I - P)$ so the result in Proposition 2 can be written $X = \mathcal{R}P \oplus \ker P$. Thus, if P is a projection, X can be written as the direct sum of closed linear subspaces. We consider the converse.

Definition 3. Two closed linear subspaces M and N of X are <u>complementary</u> if $X = M \oplus N$.

Proposition 4. Let M and N be complementary subspaces of X. For $x = m + n$, $m \in M$, $n \in N$ define $Px = m$. Then P is a projection on X with $\mathcal{R}P = M$, $\ker P = N$.

Proof: P is clearly linear and $P^2 = P$, $\mathcal{R}P = M$, $\ker P = N$. Thus, we only need to show that P is continuous. By the CGT it suffices to show that P is closed. Suppose $x_k \to x$ and $Px_k \to y$. Then $Px_k \in M \Rightarrow y \in M$. Now $x_k = Px_k + (I - P)x_k$ and since $\{x_k\}$ and $\{Px_k\}$ converge, $\{(I - P)x_k\}$ converges to some $z \in X$. $(I - P)x_k \to z$ implies $z \in N$. Hence $x = y + z$ with $y \in M, z \in N$ and $Px = y$.

From Propositions 2 and 4 we have a 1-1 correspondence between projections and complementary subspaces. Not every closed linear subspace of a B-space has a complementary subspace. We now use Phillips' Lemma 16.15 to show in particular that there is no continuous projection from ℓ^∞ onto c_0. Let $J : c_0 \to \ell^\infty$ be the canonical imbedding of c_0 into its bidual. Then $J' : ba \to \ell^1$ ($ba = (\ell^\infty)'$, Example 5.15) is given by $J'v = \{v(i)\}$. Phillips' Lemma asserts that if $v_i \to 0$ in $\sigma(ba, m_0)$, then $\|J'v_i\|_1 \to 0$. In particular, if $v_i \to 0$ weak*, then $\|J'v_i\|_1 \to 0$.

Theorem 5 (Phillips). There is no continuous projection of ℓ^∞ onto c_0.

Proof: If P were such a projection, $P : \ell^\infty \to c_0$, then for $y \in \ell^\infty$, $Py \in c_0$ so $\langle e_k, Py\rangle = \langle P'e_k, y\rangle \to 0$. Hence, $P'e_k \to 0$ weak* in ba. By the observation above,

$$\|J'P'e_k\|_1 = \sup\{|\langle J'P'e_k, x\rangle| : x \in c_0, \|x\| \le 1\}$$
$$= \sup\{|\langle e_k, PJx\rangle| : x \in c_0, \|x\| \le 1\}$$
$$= \sup\{|\langle e_k, x\rangle| : x \in c_0, \|x\| \le 1\} = \|e_k\|_1 = 1 \nrightarrow 0,$$

an obvious contradiction.

It follows from Theorem 5 that the identity operator on c_0 has no continuous linear extension to an operator from ℓ^∞ to c_0 (see the remark on the Hahn-Banach Theorem following 8.1.1). For Hilbert space, the situation is much better; see Exercise 5 of the Appendix.

For an elementary proof of Theorem 5 see [Wh2].

For separable spaces containing c_0, we have an interesting positive result due to Sobzćzyk.

Theorem 6 (Sobzczyk). Let X be separable. If E is a linear subspace of X which is linearly isomorphic to c_0, then E is complemented in X.

Proof: We identify E and c_0 and consider the unit vectors $e_k \in (c_0)' = \ell^1$. By the Hahn-Banach Theorem, each e_k has a continuous linear extension $u_k' \in X'$ with $\|u_k'\| = \|e_k\| = 1$. Let S' be the closed unit ball of X' with the weak* topology. By 16.11 S' is metrizable; let d be such a metric. Let $L = \{x' \in S' : x'$ restricted to c_0 is zero$\}$ and set $d_k = \inf\{d(x', u_k') : x' \in L\}$. Since every subsequence of $\{u_k'\}$ has a subsequence which is weak* convergent to an element of X' which belongs to L, every subsequence of $\{d_k\}$ has a subsequence which converges to 0. Thus, $d_k \to 0$ and there is a sequence $\{v_k'\} \subseteq L$ such that $d(u_k', v_k') \to 0$ so $\{u_k' - v_k'\}$ is weak* convergent to 0. Now define $P : X \to c_0$ by $Px = \{<u_k' - v_k', x>\}$. It is easily checked that P is a projection onto c_0 with $\|P\| \le 2$.

If H is a Hilbert space, it is known that every closed linear subspace M of H has a complementary subspace, namely, $M^\perp = \{y : x \cdot y = 0 \ \forall x \in M\}$ so the phenomena in Theorem 5 does not occur in Hilbert space. Indeed, a B-space is linearly homeomorphic to a Hilbert space if and only if every closed linear subspace has a complementary subspace ([LT]).

For finite dimensional subspaces the situation is better.

Proposition 7. Let M be a finite dimensional subspace of X. Then $\exists$ a continuous projection P of X onto M.

Proof: Let $\{x_1, \ldots, x_n\}$ be a basis for M. We construct $x_i' \in X'$ such that

$$<x_i', x_j> = \delta_{ij} \ (i, j = 1, \ldots, n).$$

Fix i and set $M_i = \text{span}\{x_j : j \neq i\}$. Then M_i is closed and $x_i \notin M_i$. By the Hahn-Banach Theorem $\exists\ x_i' \in X'$ such that $<x_i', M_i> = 0$ and $<x_i', x_i> = 1$. Now define $P : X \to M$ by $Px = \sum_{i=1}^{n} <x_i', x>x_i$.

Proposition 8. If $P : X \to X$ is a projection, then $P' : X' \to X'$ is a projection with $\mathscr{R}P' = (\ker P)^{\perp}$ and $\ker P' = \mathscr{R}(I - P') = (\mathscr{R}P)^{\perp}$.

Proof: By 26.18, $(P')^2 = (P^2)' = P'$. The other statements follow from 26.20.

Thus, if $X = M \oplus N$, where M and N are closed, then $X' = M^{\perp} \oplus N^{\perp}$.

Exercise 1. Give an example of a projection from c onto c_0 and compute its norm.

Exercise 2. Show that any projection from c onto c_0 has norm > 1.

Exercise 3. Let P_1, P_2 be projections on X. Set $M_i = \mathscr{R}P_i$, $N_i = \ker P_i$. Show $P_1P_2 = P_2 \Leftrightarrow M_2 \subseteq M_1 \Leftrightarrow N_1 \subseteq N_2$.

Exercise 4. Show $P_1 \leq P_2$ if and only if $P_1P_2 = P_2P_1 = P_1$ defines a partial order on the projections in $L(X, X)$.

Exercise 5. If X is a B-space, show there is always a projection of norm 1 from X''' onto X'.

Exercise 6. Use Exer. 5 to show that c_0 does not have a predual.

28
Compact Operators

In this section we consider the class of compact operators. These operators arose from studies in integral equations and have important applications to this subject. We consider only operators between NLS.

Let X, Y be NLS. A linear map $T : X \to Y$ is compact (precompact) if T carries bounded sets to relatively compact (precompact) sets. Such a map is obviously bounded and so continuous (5.4). Of course, if Y is complete any precompact operator is compact. We denote the class of compact (precompact) operators by K(X, Y) (PC(X, Y)); it is easily checked that both K(X, Y) and PC(X, Y) are linear subspaces of L(X, Y).

Proposition 1. Let Z be a NLS, $T \in L(X, Y)$, $S \in L(Y, Z)$.

(i) If T is compact (precompact), then ST is compact (precompact).

(ii) If S is compact (precompact), then ST is compact (precompact).

Proposition 2. $PC(X, Y)$ is a closed linear subspace of $L(X, Y)$ in the operator norm.

Proof: Let $T_k \in PC(X, Y)$, $T \in L(X, Y)$ and $\|T_k - T\| \to 0$. Let $\varepsilon > 0$. $\exists$ N such that $\|T_N - T\| < \varepsilon/3$. Since T_N is precompact, $\exists$ $x_1, ..., x_k$ of norm 1 such that $\{T_N x_i : i = 1, ..., k\}$ is an $\varepsilon/3$-net for $\{T_N x : \|x\| \le 1\}$. If $\|x\| \le 1$, $\exists$ j such that $\|T_N x_j - T_N x\| < \varepsilon/3$ so

$$\|Tx - Tx_j\| \le \|Tx - T_N x\| + \|T_N x - T_N x_j\| + \|T_N x_j - Tx_j\| < \varepsilon$$

and $\{Tx_j : j = 1, ..., k\}$ is an ε-net for $\{Tx : \|x\| \le 1\}$. Hence, T is precompact.

Corollary 3. If Y is a B-space, $K(X, Y)$ is closed in the operator norm of $L(X, Y)$.

Example 4. Completeness in Corollary 3 cannot be dropped. Define $T : c_0 \to \ell^2$ by $T(\{t_i\}) = \{t_i/i\}$. Then $T \in L(c_0, \ell^2)$. If $s_k = \sum_{j=1}^{k} e_j$, then $\|s_k\| = 1$ and $Ts_k = \sum_{j=1}^{k} e_j/j \to y = \sum_{j=1}^{\infty} e_j/j$.

Now consider T as a continuous linear operator from c_0 onto $\mathcal{R}T = Y_0$. Then $T \in L(c_0, Y_0)$, but T is not compact since if $\{s_k\}$ is as above, then $\{Ts_k\}$ has no convergent subsequence in Y_0 since $y \notin Y_0$.

Define $T_k \in K(c_0, Y_0)$ by $T_k(\{t_i\}) = \sum_{j=1}^{k} (t_j/j)e_j$ [Exer. 1]. Now $\|T_k - T\| \to 0$ since $\sup\{\|(T_k - T)\{t_j\}\| : \|\{t_j\}\| \le 1\} \le (\sum_{i=k+1}^{\infty} 1/i^2)^{1/2}$.

Note that T furnishes an example of a precompact operator which is not compact.

Corollary 5. If X is a B-space, then $K(X, X)$ is a closed 2-sided ideal in $L(X, X)$.

There are B-spaces in which $K(X, X)$ is the only closed 2-sided ideal in $L(X, X)$. For example, $X = \ell^p (1 \le p < \infty)$ and c_0 ([Go1], p.84).

We now give some examples of compact operators.

Example 6. Let $I = [0, 1]$ and $k \in C(I \times I)$. Define $K : L^2(I) \to L^2(I)$ by $Kf(t) = \int_0^1 k(t, s)f(s)ds$. Then $K \in L(L^2(I), L^2(I))$ since

$$(1) \qquad \|Kf\|_2^2 = \int_0^1 \left|\int_0^1 k(t, s)f(s)ds\right|^2 dt$$

$$\le \int_0^1 \left(\int_0^1 |k(t, s)|^2 ds\right)\left(\int_0^1 |f|^2\right) = \|f\|_2^2 \|k\|_2^2$$

[and $\|K\| \le \|k\|_2$]. We show that K is compact by showing that $\{Kf_n\}$ has a uniformly convergent subsequence for any bounded sequence $\{f_n\}$. Note that each Kf is continuous since k is continuous

$$[|Kf(t) - Kf(\tau)| \le \|f\|_2 \sup\{|k(t, s) - k(\tau, s)| : s \in I\}].$$

Thus, it suffices to show by the Arzelà-Ascoli Theorem, that $\{Kf_n\}$ is equicontinuous when $\|f_n\|_2 \le 1$. Let $\varepsilon > 0$. $\exists\, \delta > 0$ such that $|t - \tau| < \delta$ implies $|k(t, s) - k(\tau, s)| < \varepsilon$. Then

$$|Kf_n(t) - Kf_n(\tau)| \le \sup\{|k(t, s) - k(\tau, s)| : s \in I\} < \varepsilon$$

so $\{Kf_n\}$ is equicontinuous.

Example 7. Let $k \in L^2(I \times I)$. Define $K \in L(L^2(I), L^2(I))$ by

$$Kf(t) = \int_0^1 k(t, s)f(s)ds;$$

K is continuous by the computation in (1). We show that K is compact. Choose $k_n \in C(I \times I)$ such that $\|k_n - k\|_2 \to 0$. Let K_n be the integral operator induced by the kernel k_n. From (1), $\|K_n - K\| \to 0$ so K is compact by Corollary 3.

Example 8. Let Y be a B-space. Let $\{t_k\} \in \ell^1$, $\{x_k'\}$ be bounded in X' and $\{y_k\}$ be bounded in Y. Define $T : X \to Y$ by

$$Tx = \sum_{k=1}^{\infty} t_k \langle x_k', x \rangle y_k.$$

If $\|x_k'\| \le M$, $\|y_k\| \le M$ $\forall\, k$, then $\|Tx\| \le M^2\|x\| \sum_{k=1}^{\infty} |t_k|$ so the series defining T is absolutely convergent and, therefore, convergent and $T \in L(X, Y)$. Moreover, T is compact since if $T_k = \sum_{j=1}^{k} t_j \langle x_j', x \rangle y_j$, then

$\|T - T_k\| \le M^2 \sum_{j=k+1}^{\infty} |t_j|$ and each T_k is compact (Exer. 1 and Corollary 3). Operators of this form are called <u>nuclear operators</u>.

In 26.24 we gave necessary and sufficient conditions for an infinite matrix to represent a continuous linear operator from ℓ^1 into ℓ^p $(1 < p < \infty)$. When any such operator representation is obtained it is important to give a characterization of the compact operators (and other classes of operators) in this representation. We give such a characterization. Recall that a matrix $[t_{ij}]$ defines an operator $T \in L(\ell^1, \ell^p)$ if and only if $\sup\{(\sum_{i=1}^{\infty} |t_{ij}|^p)^{1/p} : j\} = \|T\| < \infty$. For compact operators, we have additionally

Example 9. T is compact if and only if

$$(2) \qquad \lim_n (\sum_{i=n}^{\infty} |t_{ij}|^p)^{1/p} = 0 \text{ uniformly for } j \in \mathbb{N}.$$

Proof: If T is compact, then $\{Te_j : j\}$ is relatively compact so by 10.1.15 $\lim_n \|\sum_{i=n}^{\infty} t_{ij}e_i\|_p = 0$ uniformly for $j \in \mathbb{N}$ and (2) holds.

If (2) holds, define a sequence of compact operators, T_k, by the matrix

$$T_k = \begin{bmatrix} t_{11} & t_{12} & \cdots \\ \vdots & & \\ t_{k1} & t_{k2} & \cdots \\ 0 & 0 & \cdots \\ 0 & & \\ \vdots & & \end{bmatrix}$$

(Exer. 1). Since

$$\|T_k - T\| = \sup\{(\sum_{i=k+1}^{\infty} |t_{ij}|^p)^{1/p} : j\} \to 0,$$

T is compact by Corollary 3.

Each of the compact operators in Examples 8 and 9 was the limit of a sequence of operators with finite dimensional range. This property holds in general if the range space has a Schauder basis.

Theorem 10. Let Y be a B-space with a Schauder basis $\{b_k\}$. If $T \in K(X, Y)$, then $\exists$ a sequence $\{T_k\}$ of compact operators with finite dimensional ranges such that $\|T - T_k\| \to 0$.

Proof: Let $S = \{x \in X : \|x\| \le 1\}$ and let $\{f_k\}$ be the coordinate functionals associated with $\{b_k\}$. For $x \in X$, $k \in \mathbb{N}$, set

$$T_k x = \sum_{j=1}^{k} \langle f_j, Tx\rangle b_j.$$

Then T_k has finite dimensional range. Let $\varepsilon > 0$. Since TS is relatively compact, $\exists$ N such that $k \ge N$ implies

$$\|\sum_{j=k}^{\infty} \langle f_j, Tx\rangle b_j\| < \varepsilon$$

for $\|x\| \le 1$ (10.1.15) so $\|T_k - T\| < \varepsilon$ when $k \ge N$.

A B-space X is said to have the Approximation Property if every compact operator $T : X \to X$ is the limit in the operator norm of a sequence of operators with finite dimensional ranges. By Theorem 10 every B-space with a Schauder basis has the approximation property. It was an open question as to whether every B-space has the approximation property. Enflo's example of a separable B-space without a Schauder basis also fails to satisfy the approximation property settling the question in the negative. Szankowski also showed that L(H), where H is a Hilbert space, does not have the approximation property ([Sz]).

Concerning the transpose of compact operators, we have a result of Schauder.

Theorem 11 (Schauder). Let $T \in L(X, Y)$. Then T is precompact if and only if T' is compact.

Proof: Let $T \in PC(X, Y)$ and $\varepsilon > 0$. We show T' is precompact and then T' is compact since Y' is complete. $\exists$ $x_1, ..., x_k$ in $S = \{x \in X : \|x\| \le 1\}$ such that $Tx_1, ..., Tx_k$ is an $\varepsilon/3$-net for TS. Define $A : Y' \to \mathbb{F}^k$ by $Ay' = (y'Tx_1, ..., y'Tx_k)$. Then A is compact so $\exists$ $y_1', ..., y_m'$ in $S' = \{y' \in Y' : \|y'\| \le 1\}$ such that $Ay_1', ..., Ay_m'$ is an $\varepsilon/3$-net for AS'. If $y' \in S'$, $\exists\, i$ such that $\|Ay' - Ay_i'\| < \varepsilon/3$ and, hence, $|\langle y', Tx_j\rangle - \langle y_i', Tx_j\rangle| < \varepsilon/3$ for $j = 1, \ldots, k$. Hence, if $x \in S$, $\exists\, j$ such that $\|Tx_j - Tx\| < \varepsilon/3$ and

$$|\langle T'y', x\rangle - \langle T'y_i', x\rangle| \le |\langle y', Tx - Tx_j\rangle| + |\langle y' - y_i', Tx_j\rangle| + |\langle y_i', Tx_j - Tx\rangle| < \varepsilon$$

so $T'y_1', \ldots, T'y_m'$ is an ε-net for $T'S'$.

If T' is compact, then T'' is compact by the first part. If J_X (J_Y) is the canonical imbedding of X into X'' (Y into Y''), then $T''J_X = J_YT$ is compact. Hence, T is precompact since J_Y has a bounded inverse.

It follows from Theorem 11 that if Y is a B-space then an operator $T \in L(X, Y)$ is compact if and only if its transpose T' is compact. The operator in Example 4 shows that the completeness cannot be dropped from this statement.

Exercise 1. Show that a continuous linear operator with finite dimensional range is compact.

Exercise 2. Show the identity operator on a NLS X is compact if and only if X is finite dimensional.

Exercise 3. Let X be a B-space. Suppose $\exists$ a B-space Y and a compact operator $T \in L(Y, X)$ which is onto. Show X is finite dimensional.

Exercise 4. Show every compact operator has separable range.

Exercise 5. Show that $T \in L(X, X)$ is compact if and only if the map $S \to STS$ from $L(X, X)$ into itself is compact.

Exercise 6. Let X, Y be B-spaces. If $T \in K(X, Y)$ has closed range, show $\mathcal{R}T$ is finite dimensional.

Exercise 7. If P is a projection on a B-space, show P is compact if and only if $\mathcal{R}P$ is finite dimensional.

Exercise 8. Show $T : C[0, 1] \to C[0, 1]$ defined by $Tf(t) = \int_0^t f$ is compact.

Exercise 9. Let $k \in C(I \times I), I = [0, 1]$. Show $K : C(I) \to C(I)$ defined by $Kf(t) = \int_0^1 k(t, s)f(s)ds$ is compact.

Exercise 10. Let H be a Hilbert space and $\{f_k\}$ an orthonormal subset. If $\{t_k\} \in c_0$, show that $Tx = \sum_{k=1}^{\infty} t_k(f_k \cdot x)f_k$ defines a compact operator on H.

Exercise 11. Let $\sum_{i,j} |t_{ij}|^2 < \infty$. Show that $T\{x_j\} = \{\sum_{j=1}^{\infty} t_{ij}x_j\}_{i=1}^{\infty}$ defines a compact operator on ℓ^2.

Exercise 12. Let $\varphi : [a, b] = I \to \mathbb{R}$ be continuous. Define $T : L^2(I) \to L^2(I)$ by $Tf = \varphi f$. Show T is continuous and if there exists $t \in [a, b]$ such that $\varphi(t) \neq 0$, then T is not compact.

28.1 Continuity Properties of Compact Operators

Compact operators have special continuity properties which we now consider. In this section X and Y will again be NLS. We first establish a sequential continuity property of compact operators which was originally employed by Hilbert in his study of compact operators on ℓ^2.

Theorem 1. Let $T \in K(X, Y)$. Then T carries weakly convergent sequences into norm convergent sequences.

Proof: Suppose $x_k \to 0$ weakly but $\{Tx_k\}$ isn't norm convergent to 0. Then $\exists\ \varepsilon > 0$ and a subsequence $\{x_{n_k}\}$ such that $\|Tx_{n_k}\| \geq \varepsilon$. Now $\{x_k\}$ is weakly bounded and, therefore, norm bounded so $\{Tx_{n_k}\}$ is relatively compact. Hence, $\{Tx_{n_k}\}$ has a subsequence which is norm convergent to some $y \in Y$; for convenience, assume that $Tx_{n_k} \to y$. But, T is weakly continuous (14.11) so $Tx_{n_k} \to 0$ weakly and $y = 0$ so $\|Tx_{n_k}\| \to 0$ which is the desired contradiction.

Operators with the property in Theorem 1 are called <u>completely continuous</u> (vollstetig by Hilbert); the definition of compact operator which we now use is due to F. Riesz. The converse of Theorem 1 is false; consider the identity operator $I : \ell^1 \to \ell^1$ (recall Schur's Theorem 16.14). We do have a partial converse.

Proposition 2. Let X be reflexive. $T \in L(X, Y)$ is compact if and only if T is completely continuous.

Proof: Assume T is completely continuous. Let $\{x_k\}$ be bounded in X. Then $\{x_k\}$ has a subsequence, $\{x_{n_k}\}$, which is weakly convergent to some $x \in X$ (16.6, 16.26). Then $\|Tx_{n_k} - Tx\| \to 0$ so $\{Tx_k\}$ is relatively compact, and T is compact.

We have another similar result.

Proposition 3. Suppose that X' is separable and Y is a B-space. If $T \in L(X, Y)$ is completely continuous, then T is compact.

Proof: Let $x_k \in X$, $\|x_k\| \le 1$. Consider $\{Jx_k\}$ in X''. Since the closed unit ball in X'' is metrizable in the weak* topology (16.11), ∃ a subsequence $\{Jx_{n_k}\}$ which is weak* convergent to some $x'' \in X''$ (Banach-Alaoglu). Therefore, $\{x_{n_k}\}$ is weak Cauchy in X. By Exer. 1, $\{Tx_{n_k}\}$ is norm Cauchy in Y, and T is compact.

We have a continuity characterization of compact operators.

Theorem 4. Let X, Y be B-spaces and $T \in L(X, Y)$. Then T is compact if and only if T' carries bounded nets which converge in the $\sigma(Y', Y)$

topology of Y' into nets which are norm convergent in Y'.

Proof: Let S and S' be the closed unit balls of X and Y', respectively. TS is isometric to a bounded subset of $C(S')$, where S' has the weak* topology (15.9). The condition in the theorem is equivalent to the assertion that if $\{y'_\delta\}$ is a net in S' which is weak* convergent to $y' \in S'$, then $\|T'y'_\delta - T'y'\| \to 0$ or $\langle T'y'_\delta, x\rangle = \langle y'_\delta, Tx\rangle \to \langle y', Tx\rangle$ uniformly for $x \in S$. By identifying TS with a subset of $C(S')$, this is equivalent to TS being an equicontinuous subset of $C(S')$ (Exer. 2), and by the Arzelà-Ascoli Theorem, this is equivalent to TS being relatively norm compact.

For separable spaces we have a sequential version of one part of Theorem 4.

Proposition 5. Let X, Y be B-spaces with Y separable. If $T \in L(X, Y)$ is such that T' carries weak* convergent sequences in Y' to norm convergent sequences, then T' (and, hence, T) is compact.

Proof: Let $\{y'_k\} \subseteq Y'$ be bounded. By the Banach-Alaoglu Theorem and the metrizability of the unit ball of S' in the weak* topology (16.11), $\exists$ a subsequence $\{y'_{n_k}\}$ which is weak* convergent to some $y' \in Y'$. By hypothesis, $\|T'y'_{n_k} - T'y'\| \to 0$ and T' is compact.

Exercise 1. If $T \in L(X, Y)$ is completely continuous, then T carries weak Cauchy sequences to norm Cauchy sequences.

Exercise 2. Show F is an equicontinuous subset of $C(S)$, S compact Hausdorff, if and only if whenever $\{y_\delta\}$ is a net in S which converges to $y \in S$, then $\lim f(y_\delta) = f(y)$ uniformly for $f \in F$.

Exercise 3. Let H be a Hilbert space and $T \in K(H, H)$. If $\{\varphi_k\}$ is an orthonormal sequence in H, show $\|T\varphi_k\| \to 0$.

Exercise 4. Give an example of a completely continuous operator whose transpose is not completely continuous.

28.2 Fredholm Alternative

The Fredholm Alternative from integral equations considers the possibilities for solving the integral equation

$$f(t) - \lambda\int_a^b k(t, s)f(s)ds = g(t)$$

for the unknown function f. We establish an abstract version of the alternative for compact operators which is due to F. Riesz.

Let X be a complex B-space and $T \in K(X, X)$. Let $\lambda \in \mathbb{C}, \lambda \neq 0$. We write $\lambda - T$ for $\lambda I - T$, where I is the identity operator on X.

Theorem 1. If $\mathscr{R}(\lambda - T) = X$, then $\lambda - T$ is 1-1.

Proof: Suppose $\exists\, x_1 \neq 0$ such that $Tx_1 = \lambda x_1$. Set $S = \lambda - T$ so $Sx_1 = 0$. Now $\mathscr{N}(S) \subseteq \mathscr{N}(S^2) \subseteq \ldots \subseteq \mathscr{N}(S^n) \subseteq \ldots$ and each of these subspaces is closed. We claim that the containments are each proper. Since S is onto, $\exists\, x_2$ such that $Sx_2 = x_1$ and similarly $\exists\, x_3$ such that $Sx_3 = x_2$, etc. Thus, we have a sequence $\{x_n\}$ with $Sx_{n+1} = x_n$ and $x_n \neq 0$ since $x_1 \neq 0$. Now $x_n \in \mathscr{N}(S^n)$ since

$$S^n x_n = S^{n-1}(Sx_n) = S^{n-1}x_{n-1} = \ldots = Sx_1 = 0.$$

But $x_n \notin \mathscr{N}(S^{n-1})$ since $S^{n-1}x_n = Sx_2 = x_1 \neq 0$.

By Riesz's Lemma (7.6) $\forall\, n\ \exists\ y_{n+1} \in \mathscr{N}(S^{n+1})$ such that $\|y_{n+1}\| = 1$ and $\|y_{n+1} - x\| \geq 1/2\ \ \forall\, x \in \mathscr{N}(S^n)$. Consider $\{Ty_n\}$; for $n > m$,

$$\|Ty_n - Ty_m\| = \|\lambda y_n - (\lambda y_m - Sy_n - Sy_m)\|$$

$$= |\lambda| \|y_n - (y_m - S(y_n/\lambda) - S(y_m/\lambda))\| \ge |\lambda|/2$$

since $\lambda y_m - S(y_n/\lambda) - S(y_m/\lambda) \in \mathcal{N}(S^{n-1})$. Thus, $\{y_n\}$ is a bounded sequence such that $\{Ty_n\}$ has no convergent subsequence which implies that T is not compact.

Lemma 2. If $\lambda - T$ is 1-1, then $\mathcal{R}(\lambda - T)$ is closed.

Proof: Let $y \in \overline{\mathcal{R}(\lambda - T)}$ and $y_n = (\lambda - T)x_n \to y$. If $\{x_n\}$ contains a bounded sequence, then since T is compact, $\{x_n\}$ contains a subsequence $\{x_{n_k}\}$ such that $\{Tx_{n_k}\}$ converges. Since

$$x_{n_k} = (y_{n_k} + Tx_{n_k})/\lambda,$$

the sequence $\{x_{n_k}\}$ must converge to some x with $(\lambda - T)x = y$. If $\{x_n\}$ contains no bounded subsequence, then $\|x_n\| \to \infty$. Put $z_n = x_n/\|x_n\|$ so that $(\lambda - T)z_n \to 0$ and $\|z_n\| = 1$. Since T is compact, $\exists$ a subsequence $\{z_{n_k}\}$ such that $\{Tz_{n_k}\}$ converges. Since $\{(z_{n_k} - Tz_{n_k})/\lambda\} \to 0$, it follows that $\{z_{n_k}\}$ converges to, say, z. Then $\|z\| = 1$ and $(\lambda - T)z = 0$ so $\lambda - T$ is not 1-1.

Theorem 3. If $\lambda - T$ is 1-1, $\lambda - T$ is onto.

Proof: Since $\mathcal{R}(\lambda - T)$ is closed (Lemma 2) and $\lambda - T$ is 1-1, $\mathcal{R}(\lambda - T') = \mathcal{N}(\lambda - T)^{\perp} = X'$ (26.20). Since T' is compact, Theorem 1 implies that $\mathcal{N}(\lambda - T') = \{0\}$ and 26.20 implies

$$\mathscr{R}(\lambda - T) = \mathscr{N}(\lambda - T')_{\perp} = X.$$

Combining Theorems 1 and 3, we obtain the Fredholm alternative for the operator $\lambda - T$.

Theorem 4 (Fredholm Alternative). $\lambda - T$ is 1-1 if and only if $\lambda - T$ is onto.

If $I = [a, b]$ and $k \in C(I \times I)$, then the Fredholm equation $f(t) - \lambda\int_a^b k(t, s)f(s)ds = g(t)$ has a solution $\forall\, g \in C(I)$ if and only if the homogeneous equation $f(t) - \lambda\int_a^b k(t, s)f(s)ds = 0$ has only the trivial solution $f = 0$. Similar remarks apply to L^2-kernels k.

Concerning the solutions of the homogeneous equation $(\lambda - T)x = 0$, we have

Theorem 5. $\mathscr{N}(\lambda - T)$ is finite dimensional.

Proof: It suffices to show $\{x \in \mathscr{N}(\lambda - T) : \|x\| \le 1\} = S$ is compact (7.8). Suppose $x_n \in \mathscr{N}(\lambda - T)$, $\|x_n\| \le 1$. Then $x_n = Tx_n/\lambda$, and since T is compact, $\{x_n\}$ must have a convergent subsequence, i.e., S is compact.

Corollary 6. $\forall\, n \in \mathbb{N}$, $\mathscr{N}((\lambda - T)^n)$ is finite dimensional.

Proof: $(\lambda - T)^n = \lambda^n - n\lambda^{n-1}T + \ldots + (-1)^n T^n = \lambda^n - TA$, where A is bounded. Since TA is compact, Theorem 5 is applicable.

Exercise 1. Show the conclusion of Theorem 3 is false if $\lambda = 0$. [Hint: $\int_0^x f = Tf(x)$.]

Exercise 2. Show Theorems 1 and 3 are false if compactness is dropped.

Exercise 3. Show $\mathcal{R}((\lambda - T)^n)$ is closed $\forall\, n \in \mathbb{N}$.

Exercise 4. Show Theorem 5 is false if $\lambda = 0$. [Hint: Consider the integral operator induced by the kernel $k(t, s) = t \sin s$.]

28.3 Factoring Compact Operators

In this section we establish several theorems which show that compact operators can be factored through subspaces of c_0. We begin by establishing a representation theorem for linear operators with range in c_0. Throughout this section X and Y will be NLS.

Theorem 1. (i) $T \in L(X, c_0)$ if and only if $\exists$ a (norm) bounded sequence, $\{x'_k\}$, in X' which is weak* convergent to 0 such that

$$Tx = \sum_{k=1}^{\infty} \langle x'_k, x\rangle e_k \ \forall\, x \in X.$$

Moreover, $\|T\| = \sup\{\|x'_k\| : k\}$.

(ii) $T \in L(X, c_0)$ is compact if and only if $\|x'_k\| \to 0$.

Proof: Let $T \in L(X, c_0)$. Set $x'_k = T'e_k$. Then

$$\|x'_k\| \le \|T'\| = \|T\|,$$

and if $x \in X$, $Tx = \sum_{k=1}^{\infty} \langle e_k, Tx\rangle e_k = \sum_{k=1}^{\infty} \langle x'_k, x\rangle e_k$. $\{x'_k\}$ is weak* convergent to 0 since $\{e_k\}$ is weak* convergent to 0 in ℓ^1 (26.14).

Conversely, suppose $\{x'_k\}$ is bounded and weak* convergent to 0. Define $T : X \to c_0$ by $Tx = \{\langle x'_k, x\rangle\}$. Then

$$\|Tx\| = \sup\{|\langle x'_k, x\rangle| : k\} \le \|x\| \sup\{\|x'_k\| : k\}$$

so T is continuous and $\|T\| \le \sup\|x'_k\|$.

If T is compact, then $\|T'e_k\| = \|x'_k\| \to 0$ by 28.1.3.

If $\|x_k'\| \to 0$, then $\|\sum_{k=N}^{\infty} \langle x_k', x\rangle e_k\| \le \|x\|\sup\{\|x_k'\| : k \ge N\} \to 0$ so T is compact by 28.3.

We also need a sharper form of 23.1.5.

Proposition 2. Let Z be a dense linear subspace of X and $K \subseteq X$ be precompact. There exists a sequence $\{x_k\} \subseteq Z$ converging to 0 such that each $x \in K$ has a representation $x = \sum_{k=1}^{\infty} t_k x_k$ with $\sum_{k=1}^{\infty} |t_k| \le 1$.

Proof: For convenience assume that $\|x\| \le 1\ \forall x \in K$. K is precompact and has a finite $1/2\cdot 2^3$ net, $F_1 \subseteq Z$. Thus, $\forall x \in K\ \exists\ z^1(x) \in F_1$ with $\|x - z^1(x)| < 1/2\cdot 2^3$. Now $K - F_1$ is precompact and, therefore, has an $1/3\cdot 2^4$-net, $F_2 \subseteq X$. Thus, $\forall x \in K\ \exists\ z^2(x) \in F_2$ with $\|x - z^1(x) - z^2(x)\| < 1/3\cdot 2^4$. Continuing produces a sequence of finite subsets of Z, $F_1, F_2, \dots$, such that $\forall x \in K\ \exists\ z^i(x) \in F_i$ satisfying

(1) $$\|x - z^1(x) - \dots - z^i(x)\| < 1/(i + 1)2^{i+2}.$$

Thus,

(2) $$\|z^i(x)\| \le \|x - z^1(x) - \dots - z^i(x)\| + \|x - z^1(x) - \dots - z^{i-1}(x)\| < 1/i2^i .$$

Set $x^i(x) = 2^i z^i(x)$ for $x \in K$, $i \in \mathbb{N}$. Arrange the elements of $2F_1, 2^2F_2, \dots$ in a sequence with the elements of F_1 first, those of F_2 second, etc.

By (2), this sequence converges to 0 and by (1)

$$x = z^1(x) + z^2(x) + \dots = \frac{1}{2}x^1(x) + \frac{1}{2^2}x^2(x) + \dots$$

which gives the desired representation.

We now give a characterization of precompact operators.

Theorem 3 (Terzioglu). Let $T \in L(X, Y)$. The following are equivalent.

(i) T is precompact.

(ii) $\exists \{x_k'\} \subseteq X'$, $\|x_k'\| \to 0$, such that

(3) $\|Tx\| \le \sup\{|\langle x_k', x\rangle| : k\} \ \forall x \in X$.

(iii) $\exists$ a linear subspace $Z \subseteq c_0$, $A \in PC(X, Z)$, $B \in L(Z, Y)$ such that $T = BA$.

Proof: (i) $\Rightarrow$ (ii): Assume that T is precompact. Let S' be the closed unit ball of Y'. Since T' is compact (Schauder's Theorem), $T'S'$ is relatively compact. Apply Proposition 2 to $T'S'$ to obtain a sequence $\{x_k'\}$, $\|x_k'\| \to 0$, satisfying the conclusion of Proposition 2. Each $T'y'$, $\|y'\| \le 1$, has a representation $T'y' = \sum_{k=1}^{\infty} t_k x_k'$ with $\sum_{k=1}^{\infty} |t_k| \le 1$. Hence, if $x \in X$, $\|Tx\| = \sup\{|\langle y', Tx\rangle| : \|y'\| \le 1\} \le \sup\{|\langle x_k', x\rangle| : k\}$ and (3) holds.

(ii) $\Rightarrow$ (iii): Define $A \in K(X, c_0)$ by $Ax = \{\langle x_k', x\rangle\}$ (Theorem 1). Set $Z = AX$. Then A is a precompact operator from X onto Z. Define $B : Z \to Y$ by $B(Ax) = Tx$ for $Ax \in Z$; note that B is well-defined since $\|B(Ax)\| = \|Tx\| \le \sup\{|\langle x_k', x\rangle| : k\} = \|Ax\|$ by (3). This computation also shows that $B \in L(Z, Y)$ with $\|B\| \le 1$. Obviously, we have $T = BA$.

(iii) ⇒ (i) by 28.1.

By a slight alteration in the proof above we can obtain a factorization of a precompact operator into the product of two precompact operators. In the proof of (ii) ⇒ (iii), choose $r_k \to \infty$ such that $\|r_k x_k'\| \to 0$. Set $u_k' = r_k x_k'$ and define $A_1 \in K(X, c_0)$ by $A_1 x = \{\langle u_k', x\rangle\}$ (Theorem 1). Define $A_2(\{t_k\}) = \{d_k t_k\}$, where $d_k = 1/r_k$. A_2 is a compact operator from c_0 into c_0 and is a precompact operator from A_1X onto $AX = Z$. $T = (BA_2)A_1$ is now the product of two precompact operators. We have

Theorem 4. $T \in L(X, Y)$ is precompact if and only $\exists$ a linear subspace Z of c_0 and precompact operators $B_1 \in PC(X, Z)$, $B_2 \in PC(Z, Y)$ such that $T = B_2 B_1$.

It is reasonable to ask what operators have precompact factorizations through c_0 instead of a linear subspace of c_0. We proceed to give a characterization of such operators. For this we require some preliminary results on series in NLS.

A series $\sum_{k=1}^{\infty} x_k$ in a NLS X is said to be <u>weakly unconditionally Cauchy</u> (<u>w.u.c.</u>) if $\sum_{k=1}^{\infty} |\langle x', x_k\rangle| < \infty \ \forall\, x' \in X'$. A w.u.c. series needn't converge; for example, $\sum_{k=1}^{\infty} e_k$ in c_0. We give a characterization of w.u.c.

series.

Theorem 5. The following are equivalent.

(i) Σx_k is w.u.c.,

(ii) $\{\sum_{i\in\sigma} x_i : \sigma \subseteq \mathbb{N}$ finite$\}$ is (norm) bounded,

(iii) $\sup\{\sum_{k=1}^{\infty} |\langle x', x_k\rangle| : \|x'\| \le 1\} = M < \infty$,

(iv) $\exists\ c > 0$ such that $\|\sum_{i=1}^{n} t_i x_i\| \le c\|\{t_i\}\|_\infty\ \forall\, n \in \mathbb{N}, \{t_i\} \in c_0$.

Proof: (i) ⇒ (ii): Let $\mathcal{F}$ be the family of all finite subsets of $\mathbb{N}$. For $x' \in X'$, $|\langle x', \sum_{i\in\sigma} x_i\rangle| \le \sum_{i=1}^{\infty} |\langle x', x_i\rangle|\ \forall\, \sigma \in \mathcal{F}$ so $\{\sum_{i\in\sigma} x_i : \sigma \in \mathcal{F}\}$ is weakly bounded and, therefore, norm bounded.

(ii) ⇒ (iii): Let $N > 0$ be such that $\|\sum_{i\in\sigma} x_i\| \le N\ \forall\, \sigma \in \mathcal{F}$. For $x' \in X'$, $\|x'\| \le 1$, $|\sum_{i\in\sigma} \langle x', x_i\rangle| \le N\ \forall\, \sigma \in \mathcal{F}$ which implies

$$\sum_{i=1}^{\infty} |\langle x', x_i\rangle| \le 4N$$

(Exer. 9.5.1).

(iii) ⇒ (iv): Let $\{t_i\} \in c_0$. Then

$$\|\sum_{i=1}^{n} t_i x_i\| = \sup\{|\langle x', \sum_{i=1}^{n} t_i x_i\rangle| : \|x'\| \le 1\}$$

$$\le \sup\{\sum_{i=1}^{\infty} |\langle x', t_i x_i\rangle| : \|x'\| \le 1\} \le M\|\{t_i\}\|_\infty.$$

(iv) ⇒ (i): Let $x' \in X'$, $\|x'\| \le 1$. Then

$$\sum_{i=1}^{n} |\langle x', t_i x_i\rangle| \le c\|\{t_i\}\|_\infty$$

implies $\sum_{i=1}^{\infty} |t_i||\langle x', x_i\rangle| < \infty \ \forall \{t_i\} \in c_0$ so $\{\langle x', x_i\rangle\} \in \ell^1$.

Theorem 6. If X is complete, then (i)-(iv) are equivalent to

(v) $\sum_{i=1}^{\infty} t_i x_i$ converges $\forall \{t_i\} \in c_0$.

Proof: (iv) ⇒ (v): If $\{t_i\} \in c_0$ and $n > m$, then

$$\|\sum_{i=m}^{n} t_i x_i\| \le c \sup\{|t_i| : m \le i \le n\}$$

so the partial sums of the series $\sum_{i=1}^{\infty} t_i x_i$ are Cauchy.

(v) ⇒ (i): For $x' \in X'$, $\sum_{i=1}^{\infty} t_i \langle x', x_i\rangle$ converges $\forall \{t_j\} \in c_0$ so $\{\langle x', x_i\rangle\} \in \ell^1$.

From (iv) and (v), we obtain

Corollary 7. If X is complete and Σx_i is w.u.c., then $T\{t_i\} = \sum_{i=1}^{\infty} t_i x_i$ defines a continuous linear operator $T : c_0 \to X$.

Before proceeding onto the question about factoring compact operators through c_0, we show that the appearance of c_0 in the example of a w.u.c. series which was not convergent is no accident.

Lemma 8. Let X be complete and $x_{ij} \in X$ be such that $\lim_i x_{ij} = 0 \ \forall j$ and $\lim_j x_{ij} = 0 \ \forall i$. Given $\varepsilon > 0$ $\exists$ an increasing sequence of positive integers $\{m_i\}$ such that $\|\sum_{i=1}^{\infty} \sum_{\substack{j=1 \\ j \neq i}}^{\infty} x_{m_i m_j}\| < \varepsilon$.

Proof: From 9.1 $\exists$ a subsequence $\{m_i\}$ such that $\|x_{m_i m_j}\| < \varepsilon/2^{i+j}$ for $i \neq j$. This subsequence satisfies the desired conclusion.

Theorem 9 (Bessaga-Pelczynski). A B-space X is such that every w.u.c. series in X is convergent if and only if X contains no subspace (topologically) isomorphic to c_0.

Proof: $\Rightarrow$: Consider Σe_j in c_0.

$\Leftarrow$: Suppose that X contains a series Σx_i which is w.u.c. but is not convergent. $\exists\ \delta > 0$ and an increasing sequence of positive integers $\{p_i\}$ such that $\|z_j\| = \|\sum_{i=p_j+1}^{p_{j+1}} x_i\| > \delta\ \forall\ j$. By Theorem 5 (ii), the series Σz_j is w.u.c.

By replacing X by the closed linear span of $\{z_j\}$, if necessary, we may assume that X is separable. For each j, pick $z_j' \in X'$, $\|z_j'\| = 1$, such that $<z_j', z_j> = \|z_j\|$. By 16.11 and the Banach-Alaoglu Theorem $\{z_j'\}$ has a subsequence, $\{z_{n_j}'\}$, which is weak* convergent to some $z' \in X'$. For notational convenience, assume $z_j' \to z'$ weak*. Now

$$|<z_i' - z', z_i>| \geq \delta - |<z', z_i>| \geq \delta/2$$

for large i; again for notational convenience, assume that

$$|<z_i' - z', z_i>| \geq \delta/2$$

for all i.

The matrix $[<z_i' - z', z_j>]$ satisfies the conditions of Lemma 8 so let $\{m_i\}$ be the sequence associated with $\varepsilon = \delta/4$.

The operator $T : c_0 \to X$ defined by $T\{t_i\} = \sum_{i=1}^{\infty} t_i z_{m_i}$ is linear and continuous (Corollary 7). We show that T has a bounded inverse and this will establish the result. Setting $x_i' = z_{m_i}' - z'$ and using the conclusion of Lemma 8, we obtain

$$2\|T\{t_i\}\| \geq |\langle x_i', T\{t_j\}\rangle| \geq |t_i\langle x_i', z_{m_i}\rangle| - \sum_{j\neq i} |t_j|\,|\langle x_i', z_{m_j}\rangle|$$

$$\geq |t_i|\,\delta/2 - \|\{t_j\}\|\delta/4 \text{ for each } i.$$

Thus, $2\|T\{t_i\}\| \geq (\delta/4)\|\{t_j\}\|$ which implies that T has a bounded inverse (23.14).

We introduce the class of compact operators which factor through c_0.

Definition 10. An operator $T \in L(X, Y)$ is <u>infinite nuclear</u> if it has a representation $Tx = \sum_{k=1}^{\infty} \langle x_k', x\rangle y_k$, where $x_k' \in X'$, $\lim\|x_k'\| \to 0$ and Σy_k is w.u.c.

Proposition 11. Every infinite nuclear operator is precompact.

Proof: With the notation as above, we have

$$\|Tx\| \leq \sup_k |\langle x_k', x\rangle| \sup\{ \sum_{k=1}^{\infty} |\langle y', y_k\rangle| : \|y\| \leq 1\}$$

so the result follows from Theorem 3.

We now derive our factorization theorem for compact operators.

Theorem 12. Let Y be complete. A compact operator $T \in K(X, Y)$ is infinite nuclear if and only if T has a compact factorization through c_0.

Proof: Assume that $T = AB$, where $B \in K(X, c_0)$ and $A \in L(c_0, Y)$ [note that we do not need that A is compact]. By Theorem 1 $\exists \{x_k'\} \subseteq X'$ such that $\|x_k'\| \to 0$ and $Bx = \sum_{k=1}^{\infty} \langle x_k', x\rangle e_k$. Thus,

$$Tx = A(\sum_{k=1}^{\infty} \langle x_k', x\rangle e_k) = \sum_{k=1}^{\infty} \langle x_k', x\rangle Ae_k$$

and since ΣAe_k is w.u.c. (Exer. 1), this shows that T is infinite nuclear.

Conversely, let T be infinite nuclear and let the notation be as in Definition 10. Pick $r_k \to \infty$ such that $\|r_k x_k'\| \to 0$ and set $z_k' = r_k x_k'$. Define $B : X \to c_0$ by $Bx = \{\langle z_k', x\rangle\}$; B is compact by Theorem 1. Define $A \in L(c_0, Y)$ by $A\{t_j\} = \sum_{j=1}^{\infty} t_j(y_j/r_j)$ (Corollary 7). Obviously, $T = AB$ and it only remains to show that A is compact. Define $A_n : c_0 \to Y$ by $A\{t_j\} = \sum_{j=1}^{n} t_j y_j/r_j$. By Theorem 5,

$$\|(A - A_n)\{t_j\}\| \le M \sup\{|t_j/r_j| : j > n\}$$

so A is precompact by 28.3.

For examples of spaces which admit compact operators which are not infinite nuclear see [K2], 42.8.

Exercise 1. Show a continuous linear operator carries w.u.c. series to w.u.c. series.

Exercise 2. Let $T : X \to Y$ be infinite nuclear. Suppose Z is a NLS containing X as a linear subspace. Show that T has an infinite nuclear extension from Z into Y.

Exercise 3. Let $S, T \in L(X, Y)$ with T precompact. If $\|Sx\| = \|Tx\|$ $\forall\, x \in X$, show S is precompact.

28.4 Projecting the Bounded Operators onto the Compact Operators

Let X and Y be B-spaces. The space of compact operators, K(X, Y), is a closed subspace of L(X, Y) (in the operator norm) so it is natural to ask if $\exists$ a continuous projection from L(X, Y) onto K(X, Y). There are instances where L(X, Y) = K(X, Y) so the answer in this case is clear. However, it is not known, in general, if

(1) $L(X, Y) \neq K(X, Y)$ and

(2) $\exists$ no continuous projection from L(X, Y) onto K(X, Y)

are equivalent (see Exer. 3). There are many results which show that (1) and (2) are equivalent in special cases. We give one such result in this section; our result, in particular, shows that (1) and (2) are equivalent when X = Y = Hilbert space, an important case. For general results and references, see [Jo].

To establish our result we need to consider a special class of Schauder bases. For this we consider several types of convergence for series in a TVS which are of interest in their own right.

Let E be a TVS and $\sum x_k$ a series in E.

Definition 1. The series $\sum x_k$ is <u>unconditionally convergent</u> if $\forall$ permutation $\pi : \mathbb{N} \to \mathbb{N}$, the series $\sum_k x_{\pi(k)}$ converges.

Definition 2. The series $\sum x_k$ is <u>subseries convergent</u> if $\forall$ subsequence $\{x_{n_k}\}$, the subseries $\sum_k x_{n_k}$ converges.

Definition 3. The series Σx_k is <u>unordered convergence</u> if the net $\{\sum_{n\in\sigma} x_k : \sigma \in \mathcal{F}\}$ converges, where $\mathcal{F}$ is the set of finite subsets of $\mathbb{N}$ ordered by inclusion.

Definition 4. The series Σx_k is <u>bounded multiplier convergent</u> if the series $\Sigma t_k x_k$ converges $\forall \{t_k\} \in \ell^\infty$.

Theorem 5. Let Σx_k be a series in X. The following are equivalent

(i) Σx_k is unconditionally convergent,

(ii) Σx_k is unordered convergent,

(iii) $\lim_N \sum_{k=N}^{\infty} |\langle x', x_k\rangle| = 0$ uniformly for $\|x'\| \le 1$,

(iv) Σx_k is bounded multiplier convergent,

(v) Σx_k is subseries convergent.

Proof: (i) ⇒ (ii): Let $x = \sum_{k=1}^{\infty} x_k$. Assume (ii) fails. Then $\exists\, \varepsilon > 0$ such that $\forall\, \sigma \in \mathcal{F}$, $\exists\, \sigma' \in \mathcal{F}$ with $\sigma' \supseteq \sigma$ and $\|x - \sum_{\sigma'} x_k\| \ge \varepsilon$. $\exists$ N such that $\|x - \sum_{k=1}^{n} x_k\| < \varepsilon/2$ $\forall\, n \ge N$. Let $d_1 = \{1, ..., N\}$ and let d_1' be as above. Set $d_2 = \{1, ..., \max d_1'\}$ and let d_2' be as above. Continue in this way to obtain a sequence $d_1, d_1', d_2, d_2', ...$. Define a permutation π of $\mathbb{N}$ by enumerating the elements of $d_1, d_1'\backslash d_1, d_2\backslash d_1', d_2'\backslash d_2, ...$. The series $\sum_{k=1}^{\infty} x_{\pi(k)}$ is not convergent since

$$\|\sum_{d_n'\backslash d_n} x_k\| = \|x - \sum_{d_n'} x_k + x - \sum_{d_n} x_k\| \geq \varepsilon - \varepsilon/2 = \varepsilon/2 .$$

(ii) ⇒ (iii): Let $x = \lim_{\mathcal{F}} \sum_{\sigma} x_k$ and let $\varepsilon > 0$. $\exists\ \sigma_0 \in \mathcal{F}$ such that $\sigma \supseteq \sigma_0$ implies $\|x - \sum_{\sigma} x_k\| < \varepsilon$. Let $N = \max \sigma_0$. If σ is a finite subset of $\{N + 1, N + 2, \ldots\}$, then

$$\|\sum_{\sigma} x_k\| \leq \|x - \sum_{\sigma_0 \cup \sigma} x_k\| + \|\sum_{\sigma_0} x_k - x\| < 2\varepsilon$$

so if $\|x'\| \leq 1$, $|\sum_{\sigma} \langle x', x_k\rangle| < 2\varepsilon$ and, therefore, $\sum_{k=N+1}^{\infty} |\langle x', x_k\rangle| \leq 8\varepsilon$ for $\|x'\| \leq 1$ (Exer. 9.5.1).

(iii) ⇒ (iv): Let $\{t_k\} \in \ell^{\infty}$. Then

$$\|\sum_{N}^{N+P} t_k x_k\| = \sup\{|\sum_{N}^{N+P} \langle x', t_k x_k\rangle| : \|x'\| \leq 1\}$$

$$\leq \|\{t_j\}\|_{\infty} \sup\{\sum_{N}^{N+P} |\langle x', x_k\rangle| : \|x'\| \leq 1\} \to 0$$

by (iii).

(iv) ⇒ (v): If $\{n_k\}$ is an increasing sequence of positive integers, define $t_i = 0$ if i is not one of the n_k and $t_i = 1$ otherwise. Then $\sum_k x_{n_k} = \sum_k t_k x_k$.

(v) ⇒ (i): If (i) fails, $\exists\ \varepsilon > 0$, a permutation π of $\mathbb{N}$ and an increasing sequence of positive integers $\{m_n\}$ such that

$$\|\sum_{i=m_n+1}^{m_{n+1}} x_{\pi(i)}\| \geq \varepsilon.$$

Choose a subsequence $\{m_{n_j}\}$ of $\{m_n\}$ such that

$$\min\{\pi(i) : m_{n_{j+1}} + 1 \leq i \leq m_{n_{j+1}+1}\} > \max\{\pi(i) : m_{n_j} + 1 \leq i \leq m_{n_j+1}\} .$$

Arrange the integers $\pi(i)$, $m_{n_j} + 1 \leq i \leq m_{n_j+1}$, $j \in \mathbb{N}$, into an increasing sequence $\{i_j\}$. Then $\sum_{j=1}^{\infty} x_{i_j}$ doesn't converge.

A Schauder basis $\{b_k\}$ for X is said to be <u>unconditional</u> if the series $\sum_{k=1}^{\infty} \langle f_k, x\rangle b_k = x$ is unconditionally convergent $\forall\, x \in X$, where $\{f_k\}$ are the coordinate functionals associated with $\{b_k\}$. For example, the Schauder basis $\{e_k\}$ in c_0 or ℓ^p, $1 \leq p < \infty$, is unconditional. The (original) Schauder basis in $C[0, 1]$ is not unconditional, and, in fact, $C[0, 1]$ has no unconditional Schauder basis (see [Si]).

We assume henceforth that $\{b_k\}$ is an unconditional Schauder basis for X with associated coordinate functionals $\{f_k\}$ satisfying $\|b_k\| \leq 1$ and $\|f_k\| \leq c\ \ \forall\, k$. Let P_j be the projection $P_j(x) = \langle f_j, x\rangle b_j$. From Theorem 5, for each $\{t_j\} \in \ell^\infty$ the series $\sum_{j=1}^{\infty} t_j P_j$ converges pointwise on X and by the Banach-Steinhaus Theorem defines an element of $L(X, X)$.

Lemma 6. Define $A : \ell^\infty \to L(X, X)$ by $A\{t_j\} = \sum_{j=1}^{\infty} t_j P_j$.

(i) A is linear and continuous.

(ii) $A(c_0) \subseteq K(X, X)$.

Proof: A is clearly linear and to show A is continuous it suffices to show that A has a closed graph (CGT). If $z_k \to z_0$ in ℓ^∞ and $A(z_k) \to T$ in $L(X, X)$, then, with $z_k = \{t_{kj}\}_{j=1}^{\infty}$, we have $(Az_k)b_j = t_{kj}b_j \to Tb_j$ and $(Az_k)b_j = t_{kj}b_j \to t_{0j}b_j$ as $k \to \infty$ $\forall j$. Hence, $Tb_j = t_{0j}b_j = (Az_0)b_j$ $\forall j$ and $T = Az_0$ so A has a closed graph.

For $x \in X$, $\{\sum_{j\in\sigma} P_j x : \sigma \in \mathscr{F}\}$ is bounded so by the Uniform Boundedness Principle $\exists M > 0$ such that $\|\sum_{j\in\sigma} P_j\| \le M$ $\forall \sigma \in \mathscr{F}$. For $\|x\| \le 1$, $\|x'\| \le 1$ and $\sigma \in \mathscr{F}$, $|\sum_{j\in\sigma} \langle x', P_j x\rangle| \le M$ which implies $\sum_{j=1}^{\infty} |\langle x', P_j x\rangle| \le 4M$ (Exer. 9.5.1). Now if $\{t_j\} \in c_0$,

$$\|\sum_{j=N}^{N+P} t_j P_j\| \le 4M \sup\{|t_j| : N \le j \le N+P\},$$

and since $\sum_{j=1}^{N} t_j P_j$ is compact $\forall N$, $A\{t_j\}$ is compact and (ii) holds.

Next define $B : L(X, X) \to \ell^\infty$ by $B(T) = \{\langle f_j, Tb_j\rangle\}$. Note that

B is linear, continuous and $\|B\| \le c$.

Lemma 7.

(i) $B(K(X, X)) \subseteq c_0$.

(ii) $BA = I$, the identity on ℓ^∞.

Proof: (i): Let $T \in K(X, X)$. Since $\|P_j\| \le c$ and $\|P_j x\| \to 0$ $\forall x \in X$, $\lim_j P_j(Tx) = 0$ uniformly for $\|x\| \le 1$ (9.11 or 23.6). In particular, if we put $x = b_j$, we obtain $\lim_j P_j(Tb_j) = 0$ or $B(T) \in c_0$.

(ii) is clear.

Theorem 8. There is no continuous linear projection from $L(X, X)$ onto $K(X, X)$.

Proof: Suppose that P were such a projection. From Lemmas 6 and 7, BPA would be a projection from ℓ^∞ onto c_0 which is impossible by 27.5.

In particular it follows from Theorem 8 that there is no continuous projection from $L(\ell^p, \ell^p)$ $(1 \le p < \infty)$ or $L(c_0, c_0)$ onto $K(\ell^p, \ell^p)$ or $K(c_0, c_0)$, respectively. The following result expands the range of applicability of Theorem 8.

Proposition 9. Suppose that X_1 and Y_1 are complemented subspaces of

X and Y, respectively. If K(X, Y) is complemented in L(X, Y), then $K(X_1, Y_1)$ is complemented in $L(X_1, Y_1)$.

We leave the proof to Exer. 1.

If S is uncountable and μ is counting measure on S, set $\ell^p(S) = L^p(\mu)$, $1 \le p < \infty$. Then ℓ^p is a complemented subspace of $\ell^p(S)$ so from Proposition 9 and Theorem 8, there is no continuous projection from $L(\ell^p(S), \ell^p(S))$ onto $K(\ell^p(S), \ell^p(S))$. If H is a Hilbert space, then H is isometrically isomorphic to $\ell^2(S)$ for some S, so if H is infinite dimensional, there is no continuous projection from L(H, H) on K(H, H). [See [Go] for another proof of the Hilbert space case; the proof above is adapted from [C_0] who gave the proof for Hilbert space.]

Exercise 1. Prove Proposition 8.

Exercise 2. Show that any absolutely convergent series in a B-space is subseries convergent.

Exercise 3. Show that if X is reflexive, then $K(X, \ell^1) = L(X, \ell^1)$.

Exercise 4. Show that a series Σx_k in a Hausdorff TVS is subseries convergent if and only if $\{\sum_{k \in \sigma} x_k : \sigma \text{ finite}\}$ is relatively compact.

29
Weakly Compact Operators

Let X, Y be NLS. A linear operator $T \in L(X, Y)$ is said to be <u>weakly compact</u> if T carries bounded sets to relatively weakly compact sets. A compact operator is weakly compact; the identity on a B-space is weakly compact if and only if the B-space is reflexive so, for example, the identity operator on an infinite dimensional reflexive space is weakly compact but not compact. More generally, we have

Proposition 1. If either X or Y is reflexive, then any $T \in L(X, Y)$ is weakly compact.

Proof: 16.5 and 14.11.

Theorem 2. $T \in L(X, Y)$ is weakly compact if and only if $T''X'' \subseteq J_Y Y$.

Proof: Let S, S'' be the closed unit balls of X and X'', respectively. From Goldstine's Theorem S'' is the weak* closure of $J_X S$. Since T'' is weak* continuous,

$$(1) \qquad T''(S'') \subseteq \overline{T''(J_X S)} = \overline{J_Y(TS)} \subseteq \overline{J_Y(\overline{TS})}\,,$$

where the closures in X'' are in the weak* topology and the closure $\overline{TS}$ is in the weak topology.

Now assume that T is weakly compact. Then $\overline{TS}$ is weakly compact. Since J_Y is a homeomorphism with respect to the weak topology of Y and the weak* topology of Y'', $J_Y(\overline{TS})$ is weak* compact. From (1), $T''S'' \subseteq J_Y(\overline{TS})$ so $T''X'' \subseteq J_Y Y$.

Assume that $T''X'' \subseteq J_Y Y$. Since S'' is weak* compact and T'' is weak* continuous, $T''S'' \subseteq J_Y Y$ is weak* compact in Y''. Since $J_Y(TS) = T''(J_X S) \subseteq T''S''$, $\overline{J_Y(TS)}$ is weak* compact and $\overline{J_Y(TS)} \subseteq T''S''$. Since J_Y is a homeomorphism from Y with the weak topology onto $J_Y Y$ with the weak* topology, $\overline{TS}$ is weak compact in Y, and T is weakly compact.

We denote the class of all weakly compact operators from X into Y by $W(X, Y)$.

Proposition 3. (i) $W(X, Y)$ is a vector space.

(ii) If Z is a NLS and $A \in L(Z, X)$, $B \in L(Y, Z)$, $T \in W(X, Y)$, then $TA \in W(Z, Y)$, $BT \in W(X, Z)$,

(iii) If Y is complete, $W(X, Y)$ is closed in $L(X, Y)$ with respect to the operator norm.

Proof: (i) and (ii) are easily checked. For (iii) suppose $T_k \in W(X, Y)$, $T \in L(X, Y)$ and $\|T_k - T\| \to 0$. For $x'' \in X''$, $T_k''x'' \in J_Y Y$ and $T_k''x'' \to T''x''$ so $T''x'' \in J_Y Y$ since $J_Y Y$ is closed in Y''. $T \in W(X, Y)$ by Theorem 2.

Thus, if X is a B-space, $W(X, X)$ is a closed two-sided ideal in $L(X, X)$.

For weakly compact operators we have a strong continuity condition (compare 28.1.4).

Theorem 4. $T \in L(X, Y)$ is weakly compact if and only if T' is continuous with respect to $\sigma(Y', Y)$ and $\sigma(X', X'')$.

Proof: ⇒: Suppose T is weakly compact, and let $\{y'_\delta\}$ be a net in Y' which is weak* convergent to 0. Let $x'' \in X''$. By Theorem 2 $\exists\, y \in Y$ such that $T''x'' = J_Y y$. Hence,

$$\langle x'', T'y'_\delta \rangle = \langle y'_\delta, y \rangle \to 0.$$

⇐: Let $x'' \in X''$. Let $\{y'_\delta\}$ be a net in Y' which is weak* convergent to 0. Then $T'y'_\delta \to 0$ weakly by hypothesis so $\langle T''x'', y'_\delta \rangle = \langle x'', T'y'_\delta \rangle \to 0$ and $T''x''$ is weak* continuous. Hence, $T''x'' \in J_Y Y$ so T is weakly compact by Theorem 2.

The analogue of the Schauder Theorem for compact operators (28.11) holds the weakly compact operators.

Theorem 5 (Gantmacher). Let $T \in L(X, Y)$.

(i) If T is weakly compact, then T' is weakly compact.

(ii) If Y is a B-space and T' is weakly compact, then T is weakly compact.

Proof: (i): The closed unit ball, S', of X' is weak* compact so $T'S'$ is weakly compact by Theorem 4.

(ii): Let S, S'' be the closed unit balls of X, X'', respectively. By Goldstine's Theorem S'' is the weak* closure of $J_X S$ so by Theorem 4 $T''S'' \subseteq \overline{T''J_X S} = \overline{J_Y TS}$, where the closure is in the weak topology of Y''. But $J_Y TS$ has the same closure in the weak and norm topologies and $J_Y Y$ is norm closed so $T''S'' \subseteq J_Y Y$ and T is weakly compact by Theorem 2.

We now proceed to establish a factorization theorem for weakly compact operators due to Davis, Figiel, Johnson and Pelczynski.

If $\{X_k\}$ is a sequence of B-spaces, let $\ell^2(\{X_k\})$ be all sequences $x = \{x_k\}$ with $x_k \in X_k$ and $|||x||| = (\sum_{k=1}^{\infty} \|x_k\|^2)^{1/2} < \infty$. Then $\ell^2(\{X_k\})$ is a vector space under coordinatewise addition and scalar multiplication and is a B-space under the norm $|||\ |||$ (Exer. 6). If $P_i : \ell^2(\{X_k\}) \to X_i$ is defined by $P_i(\{x_k\}) = x_i$, then P_i is linear and continuous with $\|P_i\| = 1$.

Let X be a B-space and W an absolutely convex, bounded subset of X. For $n \in \mathbb{N}$ set $U_n = 2^n W + 2^{-n} S$, where $S = \{x \in X : \|x\| \le 1\}$, and let p_n be the Minkowski functional of U_n. Each p_n is a norm which is equivalent to the original norm, $\| \|$, of X. [It is clear that $p_n \le \| \|$; on the other hand W is bounded so $\exists\, M > 0$ such that $\|x\| \le M$ $\forall\, x \in U_n$ and if $x \in U_n$, $p_n(x) \le 1$ so $\| \| \le Mp_n$.]

Let $R = \{x : |||x||| = (\sum_{n=1}^{\infty} p_n(x)^2)^{1/2} < \infty\}$ and set $C = \{x \in R : |||x||| \le 1\}$.

Proposition 6. (i) $W \subseteq C$.

(ii) $(R, ||| \, |||)$ is a B-space and the inclusion map $j : R \to X$ is continuous.

(iii) $j'' : R'' \to X''$ is 1-1 and $(j'')^{-1}X = R$.

(iv) R is reflexive if and only if W is relatively weakly compact.

Proof: (i): If $x \in W$, $p_n(x) \le 2^{-n}$ $\forall\, n$ so $|||x||| \le 1$.

(ii): Let X_n be X with the norm p_n and set $Z = \ell^2(\{X_n\})$. The map $\varphi : R \to Z$ defined by $\varphi(x) = (x, x, \ldots)$ is a linear isometry and $\varphi(R) = \{\{x_k\} \in Z : x_k = x_1 \ \forall\, k\}$ is a closed subspace so R is complete. The inclusion j is the composition of φ and the projection onto the first coordinate, P_1.

(iii): $\varphi'' : R'' \to Z'' = \ell^2(\{X_k''\})$ is given by $\varphi''(y'') = (j''y'', j''y'', \ldots)$. Since φ is 1-1 and has closed range, φ'' is 1-1 (Exer. 26.1.3) and, therefore, j'' is 1-1. If $y'' \in (j'')^{-1}X$, then $j''y'' \in X$ so $\varphi''y'' = (j''y'', j''y'', \ldots) \in \varphi(R)$ and since φ'' is

1-1, $y'' \in R$.

(iv): If R is reflexive, W is relative weakly compact from (i) and 16.6.

Assume W is relatively weakly compact. Let $\overline{W}$ be the weak closure of W and S'' be the closed unit ball of X''. For each n the set $2^n\overline{W} + 2^{-n}S''$ contains C and is $\sigma(X'', X')$ compact so each contains $j''(S_{R''})$ where $S_{R''}$ is the closed unit ball of R''. Since

$$j''(S_{R''}) \subseteq \bigcap_n (2^n W + 2^{-n}S'') \subseteq \bigcap_n (X + 2^{-n}S'') = X,$$

$j''(R'') \subseteq X$ and $R'' \subseteq R$ by (iii) so R is reflexive.

Theorem 7. If X and Z are B-spaces and $T \in W(Z, X)$, then $\exists$ a reflexive space R and $A \in L(Z, R)$, $B \in L(R, X)$ such that $T = BA$.

Proof: Put $W = T(S)$, where S is the closed unit ball of Z, and let R be as in Proposition 6. The operators $j^{-1}T = A$ and $j = B$ give the desired factorization.

Exercise 1. Show that if $T \in L(X, \ell^1)$ is weakly compact, then T is compact.

Exercise 2. If Y is a B-space and $T \in L(c_0, Y)$ is weakly compact, show T is compact.

Exercise 3. Let X be a NLS and Y a B-space and $T \in W(X, Y)$. If L

is a closed linear subspace such that $L \subseteq TX$, show L is reflexive.

Exercise 4. Show the "summing operator" $S : \ell^1 \to \ell^\infty$, $S(\{t_j\}) = \{\sum_{j=1}^{n} t_j\}$ is not weakly compact.

Exercise 5. Show that completeness can't be dropped in either Proposition 3 or Theorem 5.

Exercise 6. Show $\ell^2(\{X_k\})$ is complete under $|||\ |||$.

30
Absolutely Summing Operators

In this section we consider an interesting class of operators associated with the different types of convergent series in NLS that we have encountered. Let X and Y be real NLS. An operator $T \in L(X, Y)$ is <u>absolutely summing</u> if T carries w.u.c. series in X into absolutely convergent series in Y. We also have a characterization in terms of unconditionally convergent series.

Proposition 1. Let X be complete. Then $T \in L(X, Y)$ is absolutely summing if and only if T carries unconditionally convergent series in X to absolutely convergent series in Y.

Proof: $\Rightarrow$: Clear. $\Leftarrow$: Let $\sum x_k$ be w.u.c in X. Let $\{t_i\} \in c_0$. Then $\sum t_i x_i$ is unconditionally convergent (28.3.6) so $\sum_i |t_i| \|Tx_i\| < \infty$. Since $\{t_i\} \in c_0$ is arbitrary, $\sum \|Tx_i\| < \infty$.

Let $\ell^1_w(X)$ be all sequences $\{x_i\}$ in X such that the series $\sum x_i$ is w.u.c. Define a norm on $\ell^1_w(X)$ by $\|\{x_i\|\} = \sup\{\sum|<x', x_i>| : \|x'\| \le 1\}$ (28.3.5). If X is complete, $\ell^1_w(X)$ is complete under this norm (Exer. 1). Similarly, let $\ell^1_s(X)$ be all sequences $\{x_i\}$ in X such that $\sum\|x_i\| < \infty$ and define a norm on $\ell^1_s(X)$ by $\|\{x_i\}\| = \sum_{i=1}^{\infty} \|x_i\|$. If X is complete, $\ell^1_s(X)$ is also complete under this norm (Exer. 1).

Proposition 2. Let $T \in L(X, Y)$ be absolutely summing. Define $\tilde{T} : \ell^1_w(X) \to \ell^1_s(X)$ by $\tilde{T}\{x_i\} = \{Tx_i\}$. Then $\tilde{T}$ is continuous.

Proof: Suppose $\tilde{T}$ is not bounded. Then $\exists\ z_n = \{x_i^n\}_{i=1}^{\infty} \in \ell^1_w(X)$ such that $\|z_n\| \le 1$ and finite $\sigma_n \subseteq \mathbb{N}$ such that $\sum_{i \in \sigma_n} \|Tx_i^n\| > 2^n$. The series $\sum_{n=1}^{\infty} \sum_{i \in \sigma_n} x_i^n/2^n$ is w.u.c but $\sum_{n=1}^{\infty} \sum_{i \in \sigma_n} \|T(x_i^n/2^n)\| = \infty$ so T is not absolutely summing.

Corollary 3. $T \in L(X, Y)$ is absolutely summing if and only if $\exists\ r > 0$ such that for every finite subset $x_1, ..., x_n$ of X,

$$\sum_{i=1}^{n} \|Tx_i\| \le r \sup\{ \sum_{i=1}^{n} |<x', x_i>| : \|x'\| \le 1\}.$$

Proof: $\Rightarrow$: Let r be the operator norm of the operator $\tilde{T}$ in

Proposition 2.

$\Leftarrow$: Let $\sum x_i$ be w.u.c. in x. Then for finite $\sigma \subseteq \mathbb{N}$,

$$\sum_{i\in\sigma} \|Tx_i\| \le r \sup\{ \sum_{i\in\sigma} |<x', x_i>| : \|x'\| \le 1\}$$

$$\le r \sup\{ \sum_{i=1}^{\infty} |<x', x_i>| : \|x'\| \le 1\}$$

so $\sum\|Tx_i\| < \infty$.

Let $\mathcal{AS}(X, Y)$ be all the absolutely summing operators from X into Y. $\mathcal{AS}(X, Y)$ is a linear subspace of L(X, Y) and the operator norm, $\|\tilde{T}\|$, of the induced operator defines a norm $\pi(T) = \|\tilde{T}\|$ on $\mathcal{AS}(X, Y)$ called the absolutely summing norm. $\pi(T)$ is the infimum of all r satisfying the inequality in Corollary 3; so, in particular, $\pi(T) \ge \|T\|$.

If Y is complete, $\mathcal{AS}(X, Y)$ is complete under π (Exer. 3).

We now give some examples.

Example 4. Let S be a compact Hausdorff space and μ a positive, finite Borel measure on S. The injection $j : C(S) \to L^1(\mu)$ is absolutely summing. For if $\varphi_1, \ldots, \varphi_n \in C(S)$, then

$$\sum_{i=1}^{n} \|\varphi_i\|_1 = \int_S \sum_{i=1}^{n} |\varphi_i(t)| d\mu(t) = \int_S \sum_{i=1}^{n} |<\delta_t, \varphi_i>| d\mu(t)$$

$$\le \int_S \sup\{ \sum_{i=1}^{n} |<v, \varphi_i>| : v \in C(S)', \|v\| \le 1\} d\mu(t) =$$

$$\mu(S) \sup\{ \sum_{i=1}^{n} |<v, \varphi_i>| : \|v\| \le 1\}.$$

Apply Corollary 3.

Example 5. The injection $\ell^2 \to c_0$ is not absolutely summing. For the series $\sum e_i/i$ is w.u.c. in ℓ^2 but not absolutely convergent in c_0.

Example 6. As in Example 5 the injection $\ell^2 \to \ell^\infty$ is not absolutely summing.

Example 7. We now show the injection $\ell^1 \to \ell^2$ is absolutely summing. For this we need the Rademacher functions. The k^{th} Rademacher function is defined by $r_k(t) = \text{sign}\,\sin(2^k\pi t)$, $0 \le t \le 1$. The Rademacher functions $\{r_k\}_{k=1}^{\infty}$ form an orthonormal sequence in $L^2[0, 1]$ which is not complete ($\cos 2\pi t$ is orthogonal to each r_k). For the Rademacher functions, we have the important Khintchine Inequality:

$$(1) \qquad \Big(\sum_{i=1}^{n} t_i^2\Big)^{1/2} \le \sqrt{3}\int_0^1 \Big|\sum_{i=1}^{n} t_i r_i(t)\Big|\,dt \quad \text{for } t_1, \dots, t_n \in \mathbb{R}$$

[see [LT] p. 66].

For $n \in \mathbb{N}$, $t \in [0, 1]$, define a continuous linear functional on ℓ^1 by $\rho_{n,t} : \{s_i\} \to \sum_{i=1}^{n} s_i r_i(t)$. Note $\|\rho_{n,t}\| \le 1$. Now suppose $x_1, \dots, x_k \in \ell^1$ with $x_i = \{x_j^i\}_{j=1}^{\infty}$. Then for arbitrary n we obtain from (1),

$$\sum_{i=1}^{k}\Big(\sum_{j=1}^{n} |x_j^i|^2\Big)^{1/2} \le \sum_{i=1}^{k}\sqrt{3}\int_0^1\Big|\sum_{j=1}^{n} x_j^i r_j(t)\Big|\,dt = \sqrt{3}\sum_{i=1}^{k}\int_0^1 |\langle x_i, \rho_{n,t}\rangle|\,dt$$

$$= \sqrt{3} \int_0^1 \sum_{i=1}^{k} |\langle x_i, \rho_{n,t} \rangle| dt$$

$$\leq \sqrt{3} \int_0^1 \sup\{ \sum_{i=1}^{k} |\langle x_i, a \rangle| : a \in \ell^\infty, \|a\| \leq 1\} dt$$

$$= \sqrt{3} \sup\{ \sum_{i=1}^{k} |\langle x_i, a \rangle| : \|a\| \leq 1\},$$

and since n is arbitrary,

$$\sum_{i=1}^{k} \|x_i\|^2 \leq \sqrt{3} \sup\{ \sum_{i=1}^{k} |\langle x_i, a \rangle| : \|a\| \leq 1\}.$$

Hence, the injection of ℓ^1 into ℓ^2 is absolutely summing and the absolutely summing norm is $\leq \sqrt{3}$ (Corollary 3).

Remark 8. It is actually the case that $L(\ell^1, \ell^2) = \mathscr{AS}(\ell^1, \ell^2)$ ([LT], p. 69). From Example 5, 6, and 7, it follows that there is no "Schauder-type" theorem with respect to the transpose of an absolutely summing operator (28.11). It is, however, true that if $T : X \to Y$ is absolutely summing, then T'' is absolutely summing ([Pi], p. 87).

We next establish the domination theorem of Pietsch which characterizes absolutely summing operators. We give the very beautiful proof of this theorem due to B. Maurey.

Theorem 9. Let $T \in L(X, Y)$ and let U be the closed unit ball of X.

Then $T \in \mathscr{AS}(X, Y)$ if and only if $\exists$ a regular probability measure μ on U^0 such that

$$\|Tx\| \le \pi(T)\int_{U^o} |\langle x', x\rangle| d\mu(x') \quad \forall x \in X.$$

Proof: $\Leftarrow$: For $x_1, ..., x_k \in X$,

$$\sum_{i=1}^{k} \|Tx_i\| \le \pi(T)\int_{U^0} \sum_{i=1}^{k} |\langle x', x_i\rangle| d\mu(x')$$

$$\le \pi(T)\sup\{\sum_{i=1}^{k} |\langle x', x_i\rangle| : x' \in U^0\}$$

so $T \in \mathscr{AS}(X, Y)$ by Corollary 3.

$\Rightarrow$: For $x_1, ..., x_n \in X$ define $f_{x_1,...,x_n} : U^0 \to \mathbb{R}$ by

$$f_{x_1,...,x_n}(x') = \pi(T)\sum_{i=1}^{n} |\langle x', x_i\rangle| - \sum_{i=1}^{n} \|Tx_i\|.$$

Note each $f_{x_1,...,x_n}$ is continuous with respect to the weak* topology of U^0, i.e., belongs to $C(U^0)$. Let $C = \{f_{x_1,...,x_n} : x_1, ..., x_n \in X\}$. Note C is a convex cone in $C(U^0)$ each member of which is somewhere non-negative (by definition of $\pi(T)$). [A convex subset K of a vector space is a convex cone if $tx \in K$ when $t > 0$.] Now

$$N = \{f \in C(U^0) : f(x') < 0 \;\; \forall\, x' \in U^0\}$$

is a convex cone in $C(U^0)$ with non-empty interior such that $C \cap N = \emptyset$. By 12.2 $\exists\, \mu \in C(U^0)' = rca(S)$ such that

$$\int_{U^0} f d\mu \le t \le \int_{U^0} g d\mu$$

for $f \in N$, $g \in C$; moreover, because N and C are cones we may assume that $t = 0$. Since $\int_{U^0} f d\mu \le 0 \ \forall f \in N$, μ is a positive measure which we can assume to be a probability measure. If $x \in X$, setting $g = f_x$ gives $\pi(T)\int_{U^0} |\langle x', x\rangle| d\mu(x') - \|Tx\| \ge 0$.

Corollary 10. If $T \in \mathcal{AS}(X, Y)$, then T carries weakly convergent sequences to norm convergent sequences, i.e., T is completely continuous (§28.1).

Proof: Let $x_k \to 0$ weakly in X. Let μ be as in Theorem 9. The sequence is uniformly bounded on U^0 (i.e., norm bounded) and converges pointwise on U^0 so by the Dominated Convergence Theorem

$$\|Tx_k\| \le \pi(T)\int_{U^0} |\langle x', x_k\rangle| d\mu(x') \to 0.$$

We next establish a factorization theorem for absolutely summing operators.

Theorem 11. Let Y be complete. Let $T \in \mathcal{AS}(X, Y)$ and μ be as in Theorem 9. Then $\exists\ A \in L(X, C(U^0))$, $B \in L(L^2(U^0, \mu), Y)$ such that $T = BjA$, where j is the inclusion of $C(U^0)$ into $L^2(U^0, \mu)$.

Proof: For $x \in X$, let $\hat{x} : U^0 \to \mathbb{R}$ be defined by $\hat{x}(x') = \langle x', x\rangle$. Then $Ax = \hat{x}$ defines a linear isometry from X into $C(U^0)$ (15.9). Since

$$\|Tx\| \le \pi(T)\int_{U^0} |\langle \hat{x}, x'\rangle|\, d\mu(x') \le \pi(T)\Big(\int_{U^0} |\langle \hat{x}, x'\rangle|^2 d\mu(x')\Big)^{1/2},$$

the map B_0 from $X_0 = \{\hat{x} : x \in X\} \subseteq L^2(U^0, \mu)$ into Y defined by $B_0\hat{x} = Tx$ is linear and continuous. Since $L^2(U^0, \mu)$ is a Hilbert space and Y is complete, B_0 can be extended to a continuous linear operator $B : L^2(U^0, \mu) \to Y$ (Appendix: Exercise 5). Clearly, we now have $T = BjA$.

Corollary 12. If Y is complete and $T \in \mathscr{AS}(X, Y)$, then T is weakly compact.

Exercise 1. If X is complete, show $\ell^1_w(X)$ and $\ell^1_s(X)$ are complete.

Exercise 2. Show the inclusion of $\ell^1_s(X)$ into $\ell^1_w(X)$ is continuous.

Exercise 3. If Y is complete, show $\mathscr{AS}(X, Y)$ is complete under π.

Exercise 4. If T is absolutely summing and A, B are continuous linear operators between appropriate NLS, show AT and TB are absolutely summing.

Exercise 5. $T \in L(X, Y)$ is <u>unconditionally converging</u> if and only if T

carries w.u.c. series into subseries convergent series. Show any weakly compact operator is unconditionally converging. (Hint: Exer. 28.4.4.). Give an example of an operator which is not unconditionally converging.

30.1 The Dvoretsky-Rogers Theorem

It is established in a beginning analysis course that a series Σt_i of real numbers is subseries convergent (or unconditionally convergent) if and only if the series is absolutely convergent ([DeS]). It follows that a series $\mathbb{R}^n$ is subseries convergent if and only if it is absolutely convergent. The Dvoretsky-Rogers Theorem is the converse of this statement.

Theorem 1. Let X be a real B-space. X is finite dimensional if and only if every subseries convergent series in X is absolutely convergent.

Proof: If every subseries convergent series in X is absolutely convergent, then the identity operator I on X is absolutely summing and by 30.12 weakly compact. Hence, X is reflexive. But, I is also completely continuous (30.10) so it is compact (28.1.2) and, hence, X is finite dimensional.

The converse was observed above.

For a geometric proof of Theorem 1 see [Da].

Exercise 1. Give examples of series in c_0, c, ℓ^∞ and ℓ^p $(1 < p < \infty)$ which are unconditionally convergent but not absolutely convergent. Such examples in ℓ^1 are more difficult to construct.

Part V
Spectral Theory

31

The Spectrum of an Operator

The spectrum of a linear operator on a finite dimensional space is just the set of eigenvalues of the operator. For operators defined on infinite dimensional spaces the situation is much more complicated for even compact operators can fail to have eigenvalues. For example, the Volterrra operator $Tf(t) = \int_0^t f$ on $C[0, 1]$ has no eigenvalue. We begin by defining the spectrum of a linear operator.

Let $X \neq \{0\}$ be a complex NLS; it is important that we consider spaces over the complex numbers in order to develop a reasonable general theory. Let $T : \mathscr{D}(T) \subseteq X \to X$ be linear. Let I be the identity on X.

Definition 1. If $\lambda \in \mathbb{C}$ is such that $\lambda - T = \lambda I - T$ has range dense in X and a bounded inverse (on $\mathscr{R}T$), then λ is said to belong to the <u>resolvent set</u> of T. The resolvent set of T is denoted by $\rho(T)$. The set $\mathbb{C}\backslash\rho(T) = \sigma(T)$ is called the <u>spectrum</u> of T.

Theorem 2. If X is a B-space and $T \in L(X)$, then $\lambda \in \rho(T)$ if and only if $(\lambda - T)^{-1} \in L(X)$.

Proof: $\Leftarrow$: clear. $\Rightarrow$: If $\lambda - T$ has a bounded inverse, then $\mathcal{R}(\lambda - T)$ is closed since X is complete (10.14). But, $\mathcal{R}(\lambda - T)$ is dense so $\lambda - T$ is onto.

Thus, if X is finite dimensional and $T : X \to X$ is linear, then the spectrum of T is just the set of eigenvalues of T. We compute the spectrum of several operators on infinite dimensional spaces in the next section. We now establish several general properties of the spectrum.

Lemma 3. Suppose $\mu \in \mathbb{C}$ is such that $\mu - T$ has a bounded inverse and let $M(\mu) = \|(\mu - T)^{-1}\|$ (the norm as a bounded linear operator on $\mathcal{R}(\mu - T)$). If $\lambda \in \mathbb{C}$ is such that $|\lambda - \mu| M(\mu) < 1$, then $\lambda - T$ has a bounded inverse and $\overline{\mathcal{R}(\lambda - T)}$ is not a proper subset of $\overline{\mathcal{R}(\mu - T)}$.

Proof: Let $x \in \mathcal{D}(T)$. Then $(\lambda - T)x = (\lambda - \mu)x + (\mu - T)x$ implies $\|(\lambda - T)x\| \geq \|(\mu - T)x\| - |\lambda - \mu| \, \|x\|$. Now $\|x\| = \|(\mu - T)^{-1}(\mu - T)x\| \leq M(\mu)\|(\mu - T)x\|$ so

$$\begin{aligned} M(\mu)\|(\lambda - T)x\| &\geq M(\mu)\|(\mu - T)x\| - M(\mu)|\lambda - \mu| \, \|x\| \\ &\geq \|x\|(1 - M(\mu)|\lambda - \mu|). \end{aligned}$$

Since $1 - M(\mu)|\lambda - \mu| > 0$, this shows that $\lambda - T$ has a bounded inverse (23.14).

Suppose $\overline{\mathcal{R}(\lambda - T)}$ is a proper subset of $\overline{\mathcal{R}(\mu - T)}$. Choose θ

such that $|\lambda - \mu|M(\mu) < \theta < 1$. By Riesz's Lemma (7.6) $\exists$ $y_\theta \in \overline{\mathscr{R}(\mu - T)}$, $\|y_\theta\| = 1$, such that $\|y - y_\theta\| \geq \theta$ $\forall y \in \overline{\mathscr{R}(\lambda - T)}$. Choose $y_n \in \mathscr{R}(\mu - T)$ such that $\|y_n - y_\theta\| \to 0$. Let $x_n = (\mu - T)^{-1} y_n$. But $(\lambda - T)x_n \in \mathscr{R}(\lambda - T)$ so

$$\theta \leq \|(\lambda - T)x_n - y_\theta\| \leq \|(\mu - T)x_n - y_\theta\|$$
$$+ \|(\lambda - T)x_n - (\mu - T)x_n\| = \|y_n - y_\theta\| + |\lambda - \mu|\,\|x_\lambda\|$$
$$\leq \|y_n - y_\theta\| + |\lambda - \mu|M(\mu)\|y_n\|.$$

Letting $n \to \infty$ implies $\theta \leq |\lambda - \mu|M(\mu)\|y_\theta\| = |\lambda - \mu|M(\mu)$ which is a contradiction.

Theorem 4. The resolvent set $\rho(T)$ is open and the spectrum $\sigma(T)$ is closed.

Proof: If $\mu \in \rho(T)$, $(\mu - T)^{-1}$ is bounded and $\overline{\mathscr{R}(\mu - T)} = X$. Lemma 3 implies that if λ is close to μ, then $\lambda - T$ has a bounded inverse and $\overline{\mathscr{R}(\lambda - T)}$ is not a proper subspace of X so $\mathscr{R}(\lambda - T)$ must be dense, i.e., $\lambda \in \rho(T)$.

Notation: For $\lambda \in \rho(T)$, we write $R_\lambda(T) = R_\lambda = (\lambda - T)^{-1}$; R_λ is called the <u>resolvent operator</u> of T.

Theorem 5. Let X be a B-space and $T \in L(X)$. For $\lambda, \mu \in \rho(T)$,

(i) $R_\lambda - R_\mu = (\mu - \lambda)R_\lambda R_\mu$ (Resolvent Equation)

(ii) $R_\lambda R_\mu = R_\mu R_\lambda$.

Proof: $(\mu - T)(\lambda - T)(R_\lambda - R_\mu) = \mu - T - (\lambda - T) = (\mu - \lambda)I.$ Multiplying this equation by $R_\lambda R_\mu$ gives (i). By symmetry $R_\mu - R_\lambda = (\lambda - \mu)R_\mu R_\lambda$ and adding this to (i) gives (ii).

Theorem 6. Let X be a B-space and $T \in L(X)$. If $|\lambda| > \lim \sqrt[n]{\|T^n\|}$ (recall that this limit exists, 23.10), then $(\lambda - T)^{-1} \in L(X)$ with

$$R_\lambda = (\lambda - T)^{-1} = \sum_{n=1}^{\infty} (1/\lambda^n)T^{n-1} \quad \text{(norm convergence).}$$

Proof: Since $\lambda - T = \lambda(I - T/\lambda)$ and

$$\lim \sqrt[n]{\|(T/\lambda)^n\|} = (1/|\lambda|)\lim \sqrt[n]{\|T^n\|} < 1,$$

it follows from 23.12 that the Neumann series for $(I - T/\lambda)^{-1} = \sum_{n=0}^{\infty} T^n/\lambda^n$ is norm convergent and so $(\lambda - T)^{-1} = \sum_{n=1}^{\infty} (1/\lambda^n)T^{n-1}$.

Corollary 7. Let X be a B-space and $T \in L(X)$. Then the spectrum of T is compact with $\sigma(T) \subseteq \{\lambda : |\lambda| \le \lim \sqrt[n]{\|T^n\|}\}$.

Proof: $\sigma(T)$ is closed by Theorem 4 and bounded by $\lim \sqrt[n]{\|T^n\|}$ from Theorem 6.

We now show that the spectrum of a bounded linear operator on a B-space is non-void.

Theorem 8. Let X be a B-space and $T \in L(X)$. Then $\sigma(T) \neq \emptyset$.

Proof: Let $z \in L(X)'$ and define $f = f_z : \rho(T) \to \mathbb{C}$ by $f(\lambda) = \langle z, R_\lambda \rangle = \langle z, (\lambda - T)^{-1} \rangle$. By the Resolvent Equation, $(f(\lambda) - f(\mu))/(\lambda - \mu) = \langle z, (R_\lambda - R_\mu)/(\lambda - \mu) \rangle = \langle z, -R_\mu R_\lambda \rangle$. By 23.13 the map $\lambda \to R_\lambda$ is continuous from $\rho(T)$ to $L(X)$ so letting $\lambda \to \mu$ we have $\lim_{\lambda \to \mu}(f(\lambda) - f(\mu))/(\lambda - \mu) = -\langle z, R_\mu^2 \rangle$. That is, f is analytic on $\rho(T)$ with $f'(\mu) = -\langle z, R_\mu^2 \rangle$.

Moreover, f is bounded for large λ since

$$|f(\lambda)| \leq \|z\| \|R_\lambda\| \leq \|z\|/(|\lambda| - \|T\|)$$

for $|\lambda| > \|T\|$ by Theorem 6. Hence, if $\sigma(T) = \emptyset$, then f is a bounded entire function and must be a constant by Liouville's Theorem. Since $|f(\lambda)| \to 0$ as $|\lambda| \to \infty$, f would then be identically zero, but since $z \in L(X)'$ is arbitrary, this means that $R_\lambda = 0$ which is impossible since R_λ is an inverse. Hence, $\sigma(T) \neq \emptyset$.

From the proof we also obtain

Corollary 9. The function $\lambda \to R_\lambda$ is an analytic map from $\rho(T)$ into $L(X)$ with $\frac{d}{d\lambda} R_\lambda = - R_\lambda^2$.

From Theorem 8 and Corollary 7, the spectrum of a bounded linear operator on a B-space is a non-empty compact set. The Example 32.4 shows conversely that any compact subset of $\mathbb{C}$ is the spectrum of some bounded linear operator on a Hilbert space.

Example 32.7 shows that the completeness assumption in Theorem 8 is important.

A further property of the analytic function R_λ is given by

Theorem 10. Let X be a B-space and $T \in L(X)$. For $\mu \in \rho(T)$ set $d(\mu) = \text{distance}(\mu, \sigma(T)) > 0$. Then $\|R_\mu\| \geq 1/d(\mu)$ so $\|R_\mu\| \to \infty$ as $d(\mu) \to 0$. Thus, the resolvent set is the natural domain of analyticity of R_μ.

Proof: From Lemma 3, if $\mu \in \rho(T)$ and λ is such that $|\lambda| < 1/\|R_\mu\|$, then $\lambda + \mu \in \rho(T)$ so $d(\mu) \geq 1/\|R_\mu\|$.

Definition 11. For $T \in L(X)$, the <u>spectral radius</u> of T, <u>$r(T)$</u>, is $r(T) = \sup\{|\lambda| : \lambda \in \sigma(T)\}$.

By Corollary 7, we have $r(T) \leq \lim \sqrt[n]{\|T^n\|}$ and we now show that this inequality is actually an equality, giving a formula for the spectral radius. Example 33.7 shows the inequality $r(T) < \|T\|$ can hold.

Theorem 12. Let X be a B-space and $T \in L(X)$. Then $r(T) = \lim \sqrt[n]{\|T^n\|}$.

Proof: The series $\sum_{n=0}^{\infty} (1/\lambda^{n+1})T^n$ is norm convergent to R_λ for

$|\lambda| > \|T\|$ by Theorem 6, and the function $\lambda \to R_\lambda$ is analytic for $|\lambda| > r(T)$ by Corollary 9. Hence, for every $z' \in L(X)'$ the series $\sum_{n=0}^{\infty} (1/\lambda^{n+1})<z', T^n>$ representing the analytic function $<z', R_\lambda>$ must converge for $|\lambda| > r(T)$. Therefore, for $|\lambda| > r(T)$, $\sup\{|<z', T^n>/\lambda^{n+1}| : n\} < \infty$, and by the UBP, $\sup\{\|T^n/\lambda^{n+1}\| : n\} < \infty$. Hence, $\lim \sqrt[n]{\|T^n\|} \le |\lambda|$ for $|\lambda| > r(T)$ and $\lim \sqrt[n]{\|T^n\|} \le r(T)$. The reverse inequality was observed above.

We give two further results on the spectrum of a bounded linear operator.

Theorem 13 (Spectral Mapping Theorem: Junior Grade). Let X be a B-space and $T \in L(X)$. If p is a (complex) polynomial, then $\sigma(p(T)) = p(\sigma(T))$.

Proof: Fix μ and let $p(\lambda) - \mu = c \prod_{j=1}^{n} (\lambda - \beta_j)$ so

(1) $$p(T) - \mu = c \prod_{j=1}^{n} (T - \beta_j) .$$

If $\mu \in \sigma(p(T))$, then by (1) some $\beta_j \in \sigma(T)$ since otherwise $p(T) - \mu$ would be invertible. But $p(\beta_j) = \mu$ so $\mu \in p(\sigma(T))$ and $\sigma(p(T)) \subseteq p(\sigma(T))$.

Suppose some $\beta_j \in \sigma(T)$, say β_1. If $T - \beta_1$ has an inverse, then the range of $T - \beta_1$ is not X and (1) implies $\mathcal{R}(p(T) - \mu)$ is not X. Then $\mu \in \sigma(p(T))$. If $T - \beta_1$ doesn't have an inverse, then (1) implies

$p(T) - \mu$ has no inverse so in this case $\mu \in \sigma(p(T))$. Hence, $p(\sigma(T)) \subseteq \sigma(p(T))$.

Riesz has established a much more general form of the Spectral Mapping Theorem. If X is a B-space and $T \in L(X)$, then Riesz associates with each function f which is analytic on an open set D containing the spectrum of T an operator $f(T) \in L(X)$ defined by the integral $f(T) = \frac{1}{2\pi i}\int_C f(\lambda)(\lambda - T)^{-1}d\lambda$, where C is a positively oriented contour in D whose interior contains $\sigma(T)$. The map $f \to f(T)$ is called the Riesz operational calculus and has many interesting and useful properties. In particular, the Spectral Mapping Theorem $f(\sigma(T)) = \sigma(f(T))$ holds. For details see [RN], §151 or [DS] §VII.3.

We give an extended version of the Spectral Mapping Theorem for Hermitian operators on a Hilbert space in 41.4.

Finally concerning the spectrum of an operator and its transpose, we have

Theorem 14. Let X be a B-space and $T \in L(X)$. Then

(i) $\sigma(T) = \sigma(T')$

(ii) $R_\lambda(T)' = R_\lambda(T')$ for $\lambda \in \rho(T) = \rho(T')$.

Proof: Note $(\lambda - T)' = (\lambda - T')$ and apply 26.19.

Exercise 1. Let X be a B-space with $X = M_1 \oplus M_2$; M_i closed. Suppose $TM_i \subseteq M_i$ where $T \in L(X)$. Set $T_i = T|M_i$. Show that

$$\sigma(T) = \sigma(T_1) \cup \sigma(T_2).$$

32
Subdivisions of the Spectrum and Examples

Let X be a complex NLS and $T : \mathscr{D}(T) \subseteq X \to X$ linear.

Definition 1. All $\lambda \in \sigma(T)$ such that $\exists\, x \neq 0$ with $Tx = \lambda x$ is called the <u>point spectrum</u> of T and is denoted by $P\sigma(T)$. Such a λ is called an <u>eigenvalue</u> of T and x is called an <u>eigenvector</u> associated with λ. All $\lambda \in \sigma(T)$ such that $\lambda - T$ is 1-1 and $\mathscr{R}(\lambda - T)$ is dense in X is called the <u>continuous spectrum</u> of T and is denoted by $C\sigma(T)$. All $\lambda \in \sigma(T)$ such that $\lambda - T$ is 1-1 but $\mathscr{R}(\lambda - T)$ is not dense in X is called the <u>residual spectrum</u> of T and is denoted by $R\sigma(T)$.

Thus, the spectrum, $\sigma(T)$, is the disjoint union of the point spectrum, the continuous spectrum and the residual spectrum.

We compute and classify the points of the spectrum of several linear operators.

Example 2. If $X = \mathbb{C}^n$ and $T : X \to X$ is linear, then $\sigma(T) = P\sigma(T)$

consists of the eigenvalues of T.

Example 3. Let $X = C[0, 1]$, where we are using complex-valued functions. Define $T : X \to X$ by $Tx(t) = tx(t)$. Then $\|T\| = 1$ so $\sigma(T) \subseteq \{\lambda : |\lambda| \le 1\}$.

Suppose $\lambda = a + bi$ with $|\lambda| \le 1$ and $b \ne 0$. Then

$$|(\lambda - T)x(t)|^2 = |(a - t)x(t) + bx(t)i|^2$$
$$= (a - t)^2|x(t)|^2 + b^2|x(t)|^2 \ge b^2|x(t)|^2$$

implies $\|(\lambda - T)x\| \ge b\|x\|$ so $\lambda \in \rho(T)$ since $\lambda - T$ is clearly onto X (23.14). Hence, $\sigma(T) \subseteq [-1, 1]$.

Similarly, for $-1 \le \lambda < 0$, $|(\lambda - T)x(t)| = |\lambda - t|\,|x(t)| \ge |\lambda|\,|x(t)|$ implies $\|(\lambda - T)x\| \ge |\lambda|\,\|x\|$ so $\lambda \in \rho(T)$ since $\lambda - T$ is onto (23.14). Hence $\sigma(T) \subseteq [0, 1]$.

Let $\lambda \in [0, 1]$ and $\varepsilon > 0$ be such that either $[\lambda, \lambda + \varepsilon]$ or $[\lambda - \varepsilon, \lambda]$ is contained in $[0, 1]$. For definiteness, assume the former. Construct x_ε to be 1 on $[\lambda + \varepsilon/3, \lambda + 2\varepsilon/3]$, 0 on $[0, \lambda]$ and $[\lambda + \varepsilon, 1]$ and linear on $[\lambda, \lambda + \varepsilon/3]$, $[\lambda + 2\varepsilon/3, \lambda + \varepsilon]$. Then $\|x_\varepsilon\| = 1$ and $\varepsilon \ge \|(\lambda - T)x_\varepsilon\|$ so $\|(\lambda - T)x_\varepsilon\| \to 0$ as $\varepsilon \to 0$. Hence, $\lambda \in \sigma(T)$ and $\sigma(T) = [0, 1]$.

Note that $P\sigma(T) = \emptyset$ since $(\lambda - T)x = 0$ implies $x(t) = 0$ except possibly at $t = \lambda$ and, therefore, $x = 0$. Note also that $\mathcal{R}(\lambda - T)$ is not dense in X since any function $y \in \mathcal{R}(\lambda - T)$ vanishes at $t = \lambda$. Thus, $\sigma(T) = R\sigma(T) = [0, 1]$.

A point $\lambda \in \mathbb{C}$ is called an <u>approximate eigenvalue</u> of T if to each $\varepsilon > 0$ $\exists x \in \mathscr{D}(T)$ such that $\|x\| = 1$ and $\|(\lambda - T)x\| < \varepsilon$. The set of all approximate eigenvalues is called the <u>approximate point spectrum</u> of T. From 23.14 a point λ is an approximate eigenvalue of T if and only if $\lambda - T$ does not have a continuous inverse so any approximate eigenvalue is a point of the spectrum. In Example 3 every point in the spectrum is an approximate eigenvalue.

In Exercise 1 the reader is asked to repeat Example 3 with C[0, 1] replaced by $L^2[0, 1]$.

Example 4. Let $X = \ell^2$, where we are considering complex-valued sequences. Let $D \subseteq \mathbb{C}$ be compact with $\{d_k : k \in \mathbb{N}\}$ dense in D. Define $T : \ell^2 \to \ell^2$ by $T\{t_k\} = \{d_k t_k\}$.

If $\lambda \in \mathbb{C}\backslash D$, there exists $d > 0$ such that $|\lambda - d_k| > d$ $\forall k$. Thus, $\|(\lambda - T)\{t_k\}\|^2 = \sum_{k=1}^{\infty} |(\lambda - d_k)t_k|^2 \geq d^2 \sum_{k=1}^{\infty} |t_k|^2$ implies $\lambda - T$ has a bounded inverse (23.14) and since $\lambda - T$ is onto, $\lambda \in \rho(T)$. Hence, $D \supseteq \sigma(T)$.

Now $Te_k = d_k e_k$ so $d_k \in P\sigma(T) \subseteq \sigma(T)$. Hence, $\overline{\{d_k : k \in \mathbb{N}\}} = D \subseteq \sigma(T)$ and $D = \sigma(T)$.

For $\lambda \in D\backslash\{d_k : k \in \mathbb{N}\}$, we have $(\lambda - T)\{t_k\} = 0$ implies $(\lambda - d_k)t_k = 0$ $\forall k$ so $t_k = 0$ $\forall k$ and $\{t_k\} = 0$. Thus, $\lambda - T$ is 1-1 and $P\sigma(T) = \{d_k : k \in \mathbb{N}\}$. Note $\lambda - T$ has dense range since

$$\mathscr{R}(\lambda - T) \supseteq \{e_k : k \in \mathbb{N}\} \; [(\lambda - T)(e_k/(\lambda - d_k)) = e_k]$$

and, therefore, $C\sigma(T) = D \setminus \{d_k : k \in \mathbb{N}\}$, $R\sigma(T) = \emptyset$.

This example shows that any compact subset of the complex plane is the spectrum of a continuous linear operator on a Hilbert space.

We next consider the left shift operator $L : (t_1, t_2, \ldots) \to (t_2, t_3, \ldots)$ on the sequence spaces ℓ^p for $1 \le p \le \infty$.

Example 5. Since $\|L\| = 1$ for any $1 \le p \le \infty$, $\sigma(L) \subseteq \{\lambda : |\lambda| \le 1\}$.

First consider the case $p = \infty$. If $|\lambda| \le 1$, then λ is an eigenvalue of L with associated eigenvector $(1, \lambda, \lambda^2, \ldots)$. In this case, $\sigma(L) = P\sigma(L) = \{\lambda : |\lambda| \le 1\}$.

Now assume $1 \le p < \infty$. If $|\lambda| < 1$, then λ is an eigenvalue of L with associated eigenvector $(1, \lambda, \lambda^2, \ldots)$ [note when $p < \infty$ and $|\lambda| = 1$, the sequence $(1, \lambda, \lambda^2, \ldots)$ is not in ℓ^p]. Thus, $P\sigma(L) = \{\lambda : |\lambda| < 1\}$ and $\sigma(L) = \{\lambda : |\lambda| \le 1\}$.

If $|\lambda| = 1$, then $\lambda - L$ is 1-1 since $(1, \lambda, \lambda^2, \ldots) \notin \ell^p$ for $1 \le p < \infty$. The range of $\lambda - L$ is dense in ℓ^p since

$$(\lambda - L)(e_k - (\lambda - 1)e_{k+1}) = e_k.$$

Hence, $C\sigma(L) = \{\lambda : |\lambda| = 1\}$.

The reader is asked to consider the right shift operator on ℓ^p in Exercise 2.

The operators in the examples above are continuous linear operators on B-spaces. We now show that the extreme cases of the spectrum, $\sigma(T) = \emptyset$ or $\mathbb{C}$, can occur for general linear operators.

Example 6. Let $X = C[0, 1]$, where we are considering complex-valued functions. Let $\mathscr{D}(T) = C^1[0, 1]$ and define $T : \mathscr{D}(T) \to X$ by $Tf = f'$. Then $\sigma(T) = P\sigma(T) = \mathbb{C}$ since for each $\lambda \in \mathbb{C}$, $e^{\lambda t}$ is an eigenvector associated with λ.

Example 7. Let $X = C[0, 1]$ and $\mathscr{D}(T) = \{f \in C^1[0, 1] : f(0) = 0\}$. Define $T : \mathscr{D}(T) \to X$ by $Tf = f'$. For each $\lambda \in \mathbb{C}$ the operator R_λ defined by $R_\lambda f(t) = e^{-\lambda t}\int_0^t f(s)e^{\lambda s}ds$ belongs to $L(X)$ and $R_\lambda(\lambda - T) = (\lambda - T)R_\lambda = I$ so R_λ is the resolvent operator. Hence, each $\lambda \in \mathbb{C}$ is in the resolvent set and $\sigma(T) = \emptyset$.

This example shows that the completeness assumption in 31.8 is important.

Exercise 1. Repeat Example 3 using $X = L^2[0, 1]$. Show that $\sigma(T) = C\sigma(T) = [0, 1]$.

Exercise 2. Define $R : (t_1, t_2, \ldots) \to (0, t_1, t_2, \ldots)$ on ℓ^p for $1 \le p \le \infty$. Show $\sigma(R) = \{\lambda : |\lambda| \le 1\}$ and classify the points of the spectrum.

Exercise 3. In Exercise 31.1 show $P\sigma(T) = P\sigma(T_1) \cup P\sigma(T_2)$.

33
The Spectrum of a Compact Operator

The examples in §32 show that the spectrum of a general linear operator can be quite complex. However, the case of a compact operator is much simpler, and we now describe the spectrum of a compact operator.

Let $X \neq \{0\}$ be a complex B-space and let $T \in K(X)$.

Theorem 1. Let $0 \neq \lambda \in \mathbb{C}$. If $\lambda \in \sigma(T)$, then λ is an eigenvalue of T.

Proof: Suppose $\lambda \in \mathbb{C} \setminus P\sigma(T)$. Then $\lambda - T$ is 1-1 and, therefore, onto by the Fredholm Alternative (28.2.4). Then $(\lambda - T)^{-1} \in L(X)$ by the Open Mapping Theorem so $\lambda \in \rho(T)$.

Theorem 2. Let $0 \neq \lambda \in P\sigma(T)$. Then $\mathcal{N}(\lambda - T)$ is finite dimensional.

Proof: It suffices to show that $S = \{x \in \mathcal{N}(\lambda - T) : \|x\| \leq 1\}$ is compact (7.9). Let $\{x_k\} \subseteq S$. Then $x_k = (1/\lambda)Tx_k$ and since T is compact, $\{Tx_k\}$ has a convergent subsequence so the same is true for $\{x_k\}$. Thus, S is compact.

Lemma 3. Let $\{\lambda_k\}$ be a distinct sequence of eigenvalues of T with x_k an eigenvector associated with λ_k. Then $\lambda_k \to 0$.

Proof: Suppose $\{\lambda_k\}$ doesn't converge to 0. Then $\exists\, \varepsilon > 0$ and a subsequence satisfying $|\lambda_{n_k}| \geq \varepsilon$. Assume for simplicity that $|\lambda_k| \geq \varepsilon$. Let $H_k = \text{span}\{x_1, ..., x_k\}$. So H_k is a closed subspace of X. We claim that $H_k \subsetneq H_{k+1}$. For this it suffices to show that $\{x_1, ..., x_k\}$ is linearly independent $\forall$ k. Suppose $\{x_1, ..., x_{n-1}\}$ is linearly independent but $x_n = \sum_{k=1}^{n-1} t_k x_k$. Then $0 = (\lambda_n - T)x_n = \sum_{k=1}^{n-1} t_k(\lambda_n - T)x_k = \sum_{k=1}^{n-1} t_k(\lambda_n - \lambda_k)x_k$ which implies that $t_k = 0$ for $k = 1, ..., n - 1$ since $\lambda_n \neq \lambda_k$ for $k = 1, ..., n - 1$. Then $x_n = 0$ which is impossible. Therefore, $\{x_1, ..., x_k\}$ is linearly independent $\forall$ k.

By Riesz's Lemma (7.6) $\forall\, n \geq 1\ \exists\, y_n \in H_n$, $\|y_n\| = 1$ such that $\|y_n - x\| \geq 1/2\ \forall\, x \in H_{n-1}$. Each y_n has the form $y_n = \sum_{k=1}^{n} \alpha_k x_k$ so that $(\lambda_n - T)y_n \in H_{n-1}$. Thus, if $n > m$,

$$z_{nm} = (y_n - \frac{1}{\lambda_n} Ty_n) + \frac{1}{\lambda_m} Ty_m \in H_{n-1}$$

implies $\|T(\frac{1}{\lambda_n} y_n) - T(\frac{1}{\lambda_m} y_m)\| = \|y_n - z_{mm}\| \geq 1/2.$ Hence, no subsequence of $\{T(\frac{1}{\lambda_n} y_n)\}$ converges while $\|\frac{1}{\lambda_n} y_n\| \leq 1/\varepsilon\ \forall$ n. This contradicts the compactness of T.

We now give a precise description of the form of the spectrum of a compact operator.

Theorem 4. The spectrum of T is at most countable and has no point of accumulation in $\mathbb{C}$ except possibly 0.

Proof: Since $\sigma(T)$ is compact, it suffices to show that every $0 \neq \lambda \in \sigma(T)$ is isolated. If $0 \neq \lambda \in \sigma(T)$ is not an isolated point of $\sigma(T)$, $\exists\, 0 \neq \lambda_k \in \sigma(T)$ such that $\lambda_k \to \lambda$ with $\{\lambda_k\}$ distinct. By Theorem 1 each $\lambda \in P\sigma(T)$ and Lemma 3 implies that $\lambda_k \to 0$. This contradiction yields the desired result.

The following example shows that, conversely, the range of any sequence which converges to 0 plus 0 is the spectrum of a compact operator on a Hilbert space.

Example 5. Let $\{\lambda_k\} \subseteq C$, $\lambda_k \neq 0$, be such that $\lambda_k \to 0$. Define $T : \ell^2 \to \ell^2$ by $T\{t_k\} = \{\lambda_k t_k\}$. Then $Te_k = \lambda_k e_k\ \forall\, k$ so $\lambda_k \in P\sigma(T)$, and $\sigma(T) = \{\lambda_{k_n} : k \in \mathbb{N}\} \cup \{0\}$ (32.4). Note T is compact since if $T_n\{t_k\} = \sum_{k=1}^{n} \lambda_k t_k e_k$, then

$$\|(T_n - T)\{t_k\}\|^2 \leq \sup\{|\lambda_k|^2 : k \geq n+1\} \sum_{k=1}^{\infty} |t_k|^2.$$

Theorem 6. If X is infinite dimensional, then $0 \notin \rho(T)$ (or $0 \in \sigma(T)$).

Proof: If $0 \in \rho(T)$, $T^{-1} \in L(X)$ so $I = T^{-1}T$ is compact and X is finite dimensional.

The following example shows that the spectrum of a compact operator on an infinite dimensional space can consist only of 0.

Example 7. Let $X = C[0, 1]$ and define $T \in K(X)$ by $Tf(t) = \int_0^t f$. If $\lambda \neq 0$ and $f \in X$ satisfy $Tf = \lambda f$, then $\lambda f(t) = \int_0^t f$ so $f(t) = \lambda f'(t)$ for $0 \le t \le 1$. Therefore, $f(t) = ce^{t/\lambda}$ and $f(0) = 0$ implies that $c = 0$ so $f = 0$. By Theorems 1 and 6, $\sigma(T) = \{0\}$. Note that $\sigma(T) = R\sigma(T)$. This operator furnishes an example where $r(T) = 0 < \|T\| = 1$ (see also Exercise 23.5).

Exercise 1. Repeat Example 7 with $C[0, 1]$ replaced by $L^2[0, 1]$.

Exercise 2. Let $a_k \to 0$. Define $T : \ell^2 \to \ell^2$ by

$$T\{t_k\} = (0, a_1 t_2, a_2 t_2, \dots).$$

Show that T is compact and find $\sigma(T)$, $\|T\|$ and $r(T)$.

33.1 Invariant Subspaces and Lomonosov's Theorem

If X is a NLS and $T \in L(X)$, then a linear subspace M of X is invariant under T if $TM \subseteq M$. If λ is an eigenvalue of T, then $\mathcal{N}(\lambda - T)$ is clearly an invariant subspace of T. One of the longstanding open problems in operator theory was whether any continuous linear operator on a B-space has a non-trivial invariant subspace. Per Enflo has given an example of a continuous linear operator on a B-space which has no non-trivial invariant subspace ([E2]); however, it is still an open question as to whether such an operator exists on a Hilbert space. For compact operators the situation is much better; Aronszajn and Smith proved that every compact operator has a non-trivial invariant subspace and Bernstein and Robinson showed, more generally, that p(T) has a non-trivial invariant subspace for any operator T such that p(T) is compact for the polynomial p.

The Russian mathematician V.I. Lomonosov has given a remarkable generalization of this result; his proof uses the Schauder fixed point theorem but a somewhat less general form of his theorem with an elementary proof has been given by M. Hilden. We now present his proof; for the general form of Lomonosov's result, the reader can consult [C1].

Theorem 1. Let X be a B-space and $0 \neq T \in K(X)$. Then $\exists$ a non-trivial closed subspace of X that is invariant under any element of L(X) which commutes with T.

Proof: If $\sigma(T)$ contains a non-zero point λ, then λ is an eigenvalue of T (33.1) so $\mathcal{N}(\lambda - T)$ is a non-trivial closed subspace of X which is invariant under T. It is easily checked that $\mathcal{N}(\lambda - T)$ is also invariant under any operator in L(X) which commutes with T.

It remains to check the case where $\sigma(T) = \{0\}$. In this case the spectral radius, r(T), is 0 so $\lim \sqrt[n]{\|T^n\|} = 0$ (31.12). Let $\mathscr{C} = \{S \in L(X) : ST = TS\}$. For any non-zero $x \in X$, $\mathscr{C}x$ is a non-zero linear subspace of X and since $AB \in \mathscr{C}$ when $A, B \in \mathscr{C}$, any such $\mathscr{C}x$ is an invariant subspace for every operator in $\mathscr{C}$. Therefore, $\overline{\mathscr{C}x}$ is also a closed invariant subspace for $\mathscr{C}$. If $\exists\, x \neq 0$ such that $\overline{\mathscr{C}x} \neq X$, then the theorem is proven. Suppose that $\overline{\mathscr{C}x} = X$ for each $x \neq 0$. Pick $x_0 \in X$ such that $x_0 \neq 0$, $Tx_0 \neq 0$. There is an open sphere $S(x_0, r) = S$ about x_0 such that $0 \notin \overline{S}$ and $0 \notin \overline{TS}$. Given $y \in \overline{TS}$, $x_0 \in \overline{\mathscr{C}y}$ by assumption so $\exists\, A \in \mathscr{C}$ (depending on y) such that $Ay \in S$. There is an open neighborhood, U_y, of y such that $A(U_y) \subseteq S$. $\overline{TS}$ is compact so a finite number, $U_{y_1}, \ldots, U_{y_m}$, cover $\overline{TS}$. Let $A_1, \ldots, A_m$ be operators in $\mathscr{C}$ associated with $U_{y_1}, \ldots, U_{y_m}$. For each $y \in \overline{TS}$ $\exists\, i \in \{1, \ldots, m\}$ such that $A_i y \in S$. There exists i_1, $1 \le i_1 \le m$, such that $A_{i_1}(Tx_0) \in S$. Since $T(A_{i_1}Tx_0) \in TS$, $\exists\, i_2$, $1 \le i_2 \le m$, such that $A_{i_2}TA_{i_1}Tx_0 \in S$. Continuing, we obtain a sequence of operators $\{A_{i_n}\}$ satisfying

$$x_n = A_{i_n}T \ldots A_{i_1}Tx_0 \in S.$$

Let $M = \max\{\|A_1\|, \ldots, \|A_m\|\}$. Since T commutes with each A_i,

$\|x_n\| \le M^n\|T^n\|\|x_0\|$ and, therefore, $\|x_n\|^{1/n} \to 0$. Hence, $0 \in \bar{S}$ which gives the desired contradiction.

In particular, if $T \in K(X)$, then there is a non-trivial subspace which is invariant for T and $p(T)$ for any polynomial p. We also have the result of Bernstein and Robinson.

Corollary 2. Let X be infinite dimensional and $T \in L(X)$ be such that $p(T)$ is compact for some non-zero polynomial p. Then T has a non-trivial invariant subspace.

Proof: If $p(T) \neq 0$, Theorem 1 applies. Suppose $p(T) = 0$ and $p(t) = \sum_{k=0}^{n} a_k t^k$. For $x \neq 0$, set $M = \text{span}\{x, Tx, \ldots, T^{n-1}x\}$. Then $M \neq \{0\}$ and $M \neq X$ since X is infinite dimensional. Since $T^n = -a_n^{-1} \sum_{k=0}^{n-1} a_k T^k$, M is invariant under T.

Exercise 1. Give an example of non-trivial invariant subspace for the Volterra operator in Example 33.7.

Exercise 2. Let $I = [a, b]$ and $k \in L^2(I \times I)$. If K is the integral operator induced by the kernel k, show that $M_t = \{f \in L^2(I) : f = 0 \text{ on } [a, t]\}$ is invariant under K for each $a < t < b$.

Exercise 3. Let $T \in L(\ell^2)$ have the matrix representation $[t_{ij}]$ (§10.1). If $[t_{ij}]$ is upper triangular ($t_{ij} = 0$ for $i > j$), show $\text{span}\{e_1, ..., e_k\}$ is invariant under T for each k.

Exercise 4. Let $T_1, T_2 \in K(X)$ commute. Show T_1 and T_2 have a common non-trivial invariant subspace.

34

Adjoints in Hilbert Space

We are going to now consider linear operators in Hilbert space and establish the spectral theorem for several different classes of operators. For this we require the Hilbert space adjoint for operators on a Hilbert space.

Let H be a Hilbert space with inner product $x \cdot y$. Any $x \in H$ induces a continuous linear functional $x^* \in H'$ defined by $\langle x^*, y \rangle = y \cdot x$ for $y \in H$ and $\|x\| = \|x^*\|$. The Riesz Representation Theorem for Hilbert space, then guarantees that the linear isometry $U_H = U : H \to H'$, $Ux = x^*$, is onto, i.e., every continuous linear functional has this form. The map $U : H \to H'$ is additive but is only conjugate homogeneous, i.e., $U(x + y) = Ux + Uy$ and $U(tx) = \bar{t}Ux$.

If H_1 and H_2 are Hilbert spaces and $T : \mathscr{D}(T) \subseteq H_1 \to H_2$ is a linear map with dense domain, we may "identify" H_1' and H_2' under the isometries U_{H_1} and U_{H_2} and then consider the transpose operator $T' : \mathscr{D}(T') \subseteq H_2' \to H'$ to be a linear operator from a subspace of H_2

into H_1. The operator which results from this identification is called the Hilbert space adjoint of T and is denoted by T^*. Thus,

$$\mathscr{D}(T^*) = U_{H_2}^{-1}(\mathscr{D}(T'))$$

and $T^* = U_{H_1}^{-1}T'U_{H_2}$. For $x \in \mathscr{D}(T^*)$ and $y \in \mathscr{D}(T)$, we have

$$<T'U_{H_2}x, y> = <U_{H_2}x, Ty> = Ty\cdot x$$

and

$$<T'U_{H_2}x, y> = <U_{H_1}T^*x, y> = y\cdot T^*x$$

so

(1) $$Ty\cdot x = y\cdot T^*x \quad \forall\, x \in \mathscr{D}(T^*),\ y \in \mathscr{D}(T),$$

characterizes the Hilbert space adjoint of T.

From the previously established properties of the transpose operator, we have the following properties of the adjoint.

Proposition 1. Let $T, S \in L(H_1, H_2)$. Then $T^*, S^* \in L(H_2, H_1)$ and

(i) $(T + S)^* = T^* + S^*$,

(ii) $(TS)^* = S^*T^*$ if $H_1 = H_2$,

(iii) $(tT)^* = \bar{t}T'$ $\forall\, t \in \mathbb{F}$,

(iv) $T^{**} = T$,

(v) $I^* = I$ when $H_1 = H_2$,

(vi) $\|T^*\| = \|T\|$,

(vii) if either T^{-1} or $(T^*)^{-1}$ exists and is in $L(H)$, then the other exists and $(T^{-1})^* = (T^*)^{-1}$.

From (iii) and 31.14, we have

Proposition 2. If $T \in L(H)$, then $\sigma(T^*) = \overline{\sigma(T)}$.

To illustrate the difference between the transpose and the adjoint, consider

Example 3. Let $T \in L(\ell^2)$ and let $T = [t_{ij}]$ be the matrix representation of T with respect to $\{e_k\}$. Suppose $T^* = [\tau_{ij}]$ is the matrix representation of T^* with respect to $\{e_k\}$. Then $T^* e_k = \sum_j \tau_{jk} e_j$ so

$$Te_\ell \cdot e_k = e_\ell \cdot T^* e_k = \bar{\tau}_{\ell k} = (\sum_j t_{j\ell} e_j) \cdot e_k = t_{k\ell} \text{ or } \tau_{\ell k} = \bar{t}_{k\ell}.$$

From Example 26.24, the matrix representation of T' with respect to $\{e_k\}$ is $[s_{ij}]$ with $s_{ij} = t_{ji}$.

Proposition 4. Let $T \in L(H)$. Then $\|T^*T\| = \|T\|^2 = \|TT^*\|$.

Proof: $\|T^*T\| \le \|T^*\|\|T\| = \|T\|^2$ by (vi). On the other hand,

$$\begin{aligned} \|T\|^2 &= (\sup\{\|Tx\| : \|x\| \le 1\})^2 \\ &= \sup\{\|Tx\|^2 : \|x\| \le 1\} = \sup\{Tx \cdot Tx : \|x\| \le 1\} \\ &= \sup\{T^*Tx \cdot x : \|x\| \le 1\} \\ &\le \sup\{\|T^*Tx\| : \|x\| \le 1\} = \|T^*T\|. \end{aligned}$$

The other equality is obtained in a similar manner using $\|T^*\|^2 = \|T\|^2$.

Exercise 1. Let $\varphi \in L^{\infty}[a, b]$. Define $T_{\varphi} = T : L^2[a, b] \to L^2[a, b]$ by $Tf = \varphi f$. Describe T^*.

Exercise 2. Let $k \in L^2[a, b] \times [a, b]$. Define $K : L^2[a, b] \to L^2[a, b]$ by $Kf(t) = \int_a^b k(t, s)f(s)ds$. Describe K^*. Compare with Exercise 26.1.

Exercise 3. Let H_1, H_2 be Hilbert spaces and $T \in L(H_1, H_2)$. Show that if T^*T is compact, then T is compact. [Hint: $\|Tx\|^2 \le \|T^*Tx\| \|x\|$.]

Exercise 4. If $T \in L(H)$ and M is a closed linear subspace of H which is invariant under T, show that $M^{\perp}$ is invariant under T^*.

Exercise 5. If $T \in L(H)$, show $T = 0$ if and only if $T^*T = 0$.

Exercise 6. If H_1 and H_2 are Hilbert spaces and M_i is a closed subspace of H_i and $T \in L(H_1, H_2)$, show that $TM_1 \subseteq M_2$ if and only if $T^*M_2^{\perp} \subseteq M_1^{\perp}$.

35
Symmetric, Hermitian and Normal Operators

In this section we describe the basic properties of the symmetric, Hermitian and normal operators. Ultimately we will establish the spectral theorem for the Hermitian and normal operators.

Let X be a complex inner product space with inner product $x \cdot y$ and let $A : \mathscr{D}(A) \subseteq X \to X$ be linear.

Definition 1. A is <u>symmetric</u> if and only if $Ax \cdot y = x \cdot Ay \;\; \forall\, x, y \in \mathscr{D}(A)$.

Example 2. Let $k \in C([a, b] \times [a, b])$, where $C[a,b]$ is equipped with the inner product of $f \cdot g = \int_a^b f\bar{g}$. Define $K \in L(C[a, b])$ by

$$Kf(t) = \int_a^b k(t, s)f(s)ds.$$

Then K is symmetric if and only if $k(t, s) = \overline{k(s, t)} \;\; \forall\, s, t$.

Example 3. Let $k \in L^2([a, b] \times [a, b])$. Define $K \in L^2[a, b]$ as in Example 2. Then K is symmetric if and only if $k(t, s) = \overline{k(s, t)}$ for almost all t, s.

Example 4. Let $p, p', q, w \in C[a, b]$ be real-valued with $p(t) > 0$, $w(t) > 0$ for $a \le t \le b$. Let H be the complex Hilbert space which consists of all functions $u : [a, b] \to \mathbb{C}$ satisfying $\int_a^b |u^2| w < \infty$; the inner product is given by $u \cdot v = \int_a^b u\bar{v}w$ [i.e., $H = L^2(\mu)$ where $\mu(E) = \int_E w$ is the measure induced by the <u>weight function</u> w]. Let $\mathscr{D}(L)$ be all functions in $C^2[a, b]$ which satisfy the boundary conditions $B_1u = \alpha_1 u(a) + \alpha_2 u'(a) = 0$, $B_2 u = \beta_1 u(b) + \beta_2 u'(b) = 0$ where $|\alpha_1| + |\alpha_2| \neq 0$, $|\beta_1| + |\beta_2| \neq 0$ and $\alpha_i, \beta_i \in \mathbb{R}$. Define $L : \mathscr{D}(L) \to H$ by $Lu = \frac{1}{w}[-(pu')' + qu]$; L is called a <u>Sturm-Liouville</u> differential operator.

We claim that L is a symmetric operator. Let $u, v \in \mathscr{D}(L)$. Then

$$Lu \cdot v - u \cdot Lv = \int_a^b (-(pu')'\bar{v} + qu\bar{v} + (p\bar{v}')'u - qu\bar{v}).$$

Integrating by parts gives

$$(1) \qquad Lu \cdot v - u \cdot Lv = p(b)[u(b)\overline{v'(b)} - u'(b)\overline{v(b)}]$$
$$- p(a)[u(a)\overline{v'(a)} - u'(a)\overline{v(a)}].$$

Since α_i and β_i are real and $\bar{u}, \bar{v}$ belong to $\mathscr{D}(L)$, applying B_1 to u and $\bar{v}$ gives $\alpha_1 u(a) + \alpha_2 u'(a) = 0$, $\alpha_1 \bar{v}(a) + \alpha_2 \bar{v}'(a) = 0$, and since not both α_1 and α_2 are zero,

$$\begin{vmatrix} u(a) & u'(a) \\ \bar{v}(a) & \bar{v}'(a) \end{vmatrix} = 0 .$$

Thus, the second term on the right hand side of (1) is 0. Using B_2 in the same way, the first term on the right hand side of (1) is 0, and (1) implies that L is symmetric.

We will consider the problem of solving the Sturm-Liouville boundary value problem $Lu = 0, B_1u = 0, B_2 = 0$ in section 37.

Lemma 5. If A is symmetric, then $Ax \cdot x$ is real $\forall\, x \in \mathscr{D}(A)$.

Proof: $\overline{Ax \cdot x} = x \cdot Ax = Ax \cdot x.$

Definition 6. The bounds of a symmetric operator A are defined by

$$m(A) = \inf\{Ax \cdot x : \|x\| = 1, x \in \mathscr{D}(A)\},$$

$$M(A) = \sup\{Ax \cdot x : \|x\| = 1, x \in \mathscr{D}(A)\}.$$

(Note these quantities are meaningful by Lemma 5.)

Theorem 7. If A is symmetric and λ is an eigenvalue of A, then $\lambda \in \mathbb{R}$ and $m(A) \le \lambda \le M(A)$. Moreover, eigenvectors associated with distinct eigenvalues are orthogonal.

Proof: Suppose $\|x\| = 1$ and $Ax = \lambda x$. Then $Ax \cdot x = (\lambda x) \cdot x = \lambda$ which implies $\lambda \in \mathbb{R}$ by Lemma 5 and also $m(A) \le \lambda \le M(A)$.

If $Ax = \lambda x$ and $Ay = \mu y$ with $\mu \ne \lambda$, then

$$\lambda(x\cdot y) = Ax\cdot y = x\cdot Ay = x\cdot \mu y = \mu(x\cdot y)$$

so $(\lambda - \mu)x\cdot y = 0$ and $x\cdot y = 0$.

Theorem 8. If $A : X \to X$ is continuous and symmetric, then

$$\|A\| = \sup\{\,|Ax\cdot x| : \|x\| = 1\} = \max\{\,|m(A)|,\ |M(A)|\,\}.$$

Proof: For $\|x\| = 1$, by the Schwarz inequality,

$$|Ax\cdot x| \le \|Ax\|\|x\| \le \|A\|$$

so $\alpha = \sup\{\,|Ax\cdot x| : \|x\| = 1\} \le \|A\|$. For $x, y \in X$,

$$A(x + y)\cdot(x + y) = Ax\cdot x + Ay\cdot y + Ax\cdot y + Ay\cdot x$$

$$= Ax\cdot x + Ay\cdot y + Ax\cdot y + \overline{Ax\cdot y}$$

$$= Ax\cdot x + Ay\cdot y + 2\,\mathcal{R}Ax\cdot y$$

and similarly,

$$A(x - y)\cdot(x - y) = Ax\cdot x + Ay\cdot y - 2\,\mathcal{R}Ax\cdot y$$

so subtraction implies

$$4\,\mathcal{R}Ax\cdot y = A(x + y)\cdot(x + y) - A(x - y)\cdot(x - y).$$

Since $|Az\cdot z| \le \alpha\,\|z\|^2\ \forall\, z \in X$, we have

(2) $$|\,\mathcal{R}Ax\cdot y| \le (\alpha/4)(\|x + y\|^2 + \|x - y\|^2)$$
$$= (\alpha/4)(2\|x\|^2 + 2\|y\|^2) = (\alpha/2)(\|x\|^2 + \|y\|^2)$$

by the parallelogram law.

Now suppose $Ax \ne 0$ and set $y = (\|x\|/\|Ax\|)Ax$ in (2) to obtain

$$|Ax\cdot y| = Ax\cdot(\|x\|/\|Ax\|)Ax$$
$$= \mathcal{R}Ax\cdot(\|x\|/\|Ax\|)Ax \le (\alpha/2)(\|x\|^2 + \|x\|^2) = \alpha\|x\|^2$$

so $\|x\|\|Ax\| \le \alpha\|x\|^2$ and $\|Ax\| \le \alpha\|x\|$. This also holds if $Ax = 0$ so $\|A\| \le \alpha$ and $\|A\| = \alpha$.

Hermitian Operators:

Let H be a complex Hilbert space.

Definition 9. An operator $T \in L(H)$ is <u>Hermitian</u> (<u>bounded self-adjoint</u>) if $T = T^*$.

Note that if T is Hermitian, then $Tx \cdot y = x \cdot T^* y = x \cdot Ty$ $\forall$ x, $y \in H$ so T is also symmetric. A remarkable result of Hellinger and Toeplitz asserts that the converse holds for a symmetric operator whose domain is a Hilbert space, i.e., the algebraic property of symmetry implies continuity in this case. This result follows from the abstract versions of the Hellinger-Toeplitz Theorem given in §26, but we give a direct proof based on the UBP which avoids the use of the locally convex space machinery (see also Exercise 3).

Theorem 10 (Hellinger-Toeplitz). If $T : H \to H$ is linear and symmetric, then T is continuous and Hermitian.

Proof: $\{Tx : \|x\| \le 1\}$ is weakly bounded in H since $\forall\, y \in H$, $|Tx \cdot y| = |x \cdot Ty| \le \|Ty\|$. Therefore, by the UBP,

$$\sup\{\|Tx\| : \|x\| \le 1\} = \|T\| < \infty.$$

T is now clearly Hermitian.

The completeness in Theorem 10 is important for this form of the Hellinger-Toeplitz Theorem (Exercise 2).

For the algebraic and topological properties of Hermitian operators, we have

Proposition 11. The Hermitian operators in L(H) form a (norm) closed real-linear subspace which contains the identity operator I.

Proof: If T, S ∈ L(H) are Hermitian and t, s ∈ ℝ, then $(tT + sS)^* = tT^* + sS^* = tT + sS$. If T, S ∈ L(H) are Hermitian and $\|T - T_k\| \to 0$ with T ∈ L(H), then

$$\|T - T^*\| \le \|T - T_k\| + \|T_k^* - T^*\| = 2\|T - T_k\| \to 0$$

so T is Hermitian.

Proposition 12. If T, S ∈ L(H) are Hermitian, then TS is Hermitian if and only if TS = ST.

Proof: ⇒: $(TS)^* = S^*T^* = ST = TS$.
⇐: $(ST)^* = T^*S^* = TS = ST$.

Lemma 13. Let X be a complex inner product space and T ∈ L(X). If $Tx \cdot x = 0$ ∀ x ∈ X, then T = 0.

Proof: Let $x, y \in X$, $s, t \in \mathbb{C}$. Then

$$0 = T(sx + ty)\cdot(sx + ty) = |s|^2 Tx\cdot x + |t|^2 Ty\cdot y + s\bar{t}Tx\cdot y + \bar{s}tTy\cdot x$$

so

$$(3) \qquad 0 = s\bar{t}Tx\cdot y + \bar{s}tTy\cdot x.$$

Put $s = t = 1$ in (3) to obtain $Tx\cdot y + Ty\cdot x = 0$; put $s = i$ and $t = 1$ in (3) to obtain $iTx\cdot y - iTy\cdot x = 0$. Hence, $2Tx\cdot y = 0$ $\forall x, y \in X$ so $T = 0$.

It is important that complex scalars are used in Lemma 13 (see Exercise 5).

Theorem 14. Let $T \in L(H)$. Then T is Hermitian if and only if $Tx\cdot x$ is real $\forall x \in H$.

Proof: $\Rightarrow$: If T is Hermitian, it is symmetric. Apply Lemma 5.

$\Leftarrow$: If $Tx\cdot x \in \mathbb{R}$ $\forall x \in H$, $Tx\cdot x = \overline{Tx\cdot x} = \overline{x\cdot T^*x} = T^*x\cdot x$ so $(T - T^*)x\cdot x = 0$ and $T = T^*$ by Lemma 13.

In the algebra $L(H)$, the Hermitian operators play a role analogous to the real numbers in the algebra $\mathbb{C}$. To justify this statement, we have

Theorem 15. If $T \in L(H)$, $\exists$ unique Hermitian operators $A, B \in L(H)$ such that $T = A + Bi$.

Proof: Set $A = (T + T^*)/2$, $B = (T - T^*)/2i$.

For uniqueness, suppose $T = A_1 + B_1 i$ with A_1, $B_1 \in L(H)$

Hermitian. If $x \in H$, $Ax \cdot x + iBx \cdot x = A_1x \cdot x + iB_1x \cdot x$ and by Theorem 14 $Ax \cdot x = A_1x \cdot x$, $Bx \cdot x = B_1x \cdot x$ so $A = A_1$, $B = B_1$ by Lemma 13.

As in the case of real numbers, we can define a partial order on the Hermitian operators.

Definition 16. If $T, S \in L(H)$ are Hermitian, we say that $T \geq S$ if $Tx \cdot x \geq Sx \cdot x$ $\forall x \in H$. If $T \geq 0$, we say that T is <u>positive</u>. Note this definition is meaningful by Theorem 14.

Proposition 17. Let $U, T, S \in L(H)$ be Hermitian. Then

(i) $T \leq S$ implies $T + U \leq S + U$,

(ii) $T \leq S$ implies $tT \leq tS$ $\forall t \geq 0$.

This order is a partial order on the Hermitian operators.

Proof: (i) and (ii) are clear. Clearly $T \leq T$ and if $T \leq S$, $S \leq U$, then $T \leq U$. If $T \leq S$ and $S \leq T$, then $(T - S)x \cdot x = 0$ $\forall x \in H$ so $T = S$ by Lemma 13.

As in the case of real numbers, it is also the case that any positive operator has a unique square root. We give the proof of this fact in 41.3.

Normal Operators:

Definition 18. An operator $T \in L(H)$ is <u>normal</u> if and only if $TT^* = T^*T$.

A Hermitian operator is clearly normal. The multiplication operator in Exercise 34.1 is normal but not necessarily Hermitian (see Exercise 1).

In $\mathbb{C}^2$, the operator $T = \begin{bmatrix} i & 1 \\ i & 1 \end{bmatrix}$ is not normal.

Theorem 19. Let $T \in L(H)$. Then T is normal if and only if $\|T^*x\| = \|Tx\| \ \forall\, x \in H$.

Proof: For $x \in H$, we have $\|Tx\|^2 = Tx \cdot Tx = T^*Tx \cdot x$ and $\|T^*x\|^2 = T^*x \cdot T^*x = TT^*x \cdot x$. If T is normal, these equations imply $\|Tx\|^2 = \|T^*x\|^2$. On the other hand, if $\|T^*x\| = \|Tx\|$, they imply $(T^*T - TT^*)x \cdot x = 0$ and $T^*T = TT^*$ by Lemma 13.

Corollary 20. If $T \in L(H)$ is normal, then $\|T^2\| = \|T\|^2$.

Proof: Replace x by Tx in Theorem 19 to obtain $\|T^*Tx\| = \|T^2x\|$ so $\|T^*T\| = \|T^2\|$. By 34.4, $\|T^*T\| = \|T\|^2$ for any $T \in L(H)$.

We now describe some of the spectral properties of normal operators. If $M \subseteq H$, we write $M^\perp = \{x \in H : x \cdot y = 0 \ \forall\, y \in M\}$. From 26.20 and the Riesz Representation Theorem for H, we have

Lemma 21. If $T \in L(H)$, then $\overline{\mathscr{R}T} = \mathscr{N}(T^*)^\perp$.

From this we show that the range and null space of any normal operator gives an orthogonal decomposition of H.

Theorem 22. If $T \in L(H)$ is normal, then $\overline{\mathcal{R}T}$ and $\mathcal{N}(T)$ are orthogonal complements so, in particular, $H = \overline{\mathcal{R}T} \oplus \mathcal{N}(T)$.

Proof: By Theorem 19, $\mathcal{N}(T) = \mathcal{N}(T^*)$ and by Lemma 21, $\overline{\mathcal{R}T} = \mathcal{N}(T^*)^{\perp} = \mathcal{N}(T)^{\perp}$.

We can now describe the resolvent set of a normal operator.

Theorem 23. Let $T \in L(H)$ be normal. Then $\lambda \in \mathbb{C}$ is in $\rho(T)$ if and only if $\exists\, c > 0$ such that

(4) $$\|(\lambda - T)x\| \geq c\|x\| \quad \forall\, x \in H.$$

Proof: $\Rightarrow$: 23.14.

$\Leftarrow$: If (4) holds, $\lambda - T$ has a bounded inverse by 23.14 so we need to show that $\mathcal{R}(\lambda - T)$ is dense in H. Since T is normal, $\lambda - T$ is normal so by Theorem 22,

$$H = \overline{\mathcal{R}(\lambda - T)} \oplus \mathcal{N}(\lambda - T) = \overline{\mathcal{R}(\lambda - T)} \oplus \{0\}$$

and $\mathcal{R}(\lambda - T)$ is dense in H.

Corollary 24. Let $T \in L(H)$ be normal. Then $\lambda \in \mathbb{C}$ is in $\sigma(T)$ if and only if $\forall\, \varepsilon > 0\ \exists\, x \in H$, $\|x\| = 1$, such that $\|(\lambda - T)x\| < \varepsilon$.

Recall that a point in the spectrum of an operator which satisfies the condition of Corollary 24 is called an <u>approximate eigenvalue</u> (§32). The set of all approximate eigenvalues is called the approximate spectrum. It follows from Corollary 24 that the spectrum of a normal operator coincides with its approximate spectrum.

Proposition 25. If $T \in L(H)$ is normal, then $R\sigma(T) = \emptyset$.

Proof: If $\lambda \in \sigma(T)$ is such that $\mathcal{R}(\lambda - T)$ is not dense, then either $\lambda \in P\sigma(T)$ or $\lambda \in R\sigma(T)$. We show that if $\mathcal{R}(\lambda - T)$ is not dense, then $\lambda \in P\sigma(T)$. Now $\overline{\mathcal{R}(\lambda - T)} \neq H$ implies $\mathcal{R}(\lambda - T)^{\perp} \neq \{0\}$. By Theorem 22, $\mathcal{R}(\lambda - T)^{\perp} = \mathcal{N}(\lambda - T) \neq \{0\}$ so $\lambda \in P\sigma(T)$.

Finally from Corollary 20, we obtain a simple formula for the spectral radius of a normal operator.

Theorem 26. If $T \in L(H)$ is normal, then $r(T) = \|T\|$.

Proof: By 31.12, $r(T) = \lim \sqrt[n]{\|T^n\|}$. By Corollary 20, since powers of normal operators are normal (Exercise 6), $\|T^n\| = \|T\|^n$ when n is even. The result follows.

The Spectrum of Hermitian Operators:

Theorem 27. If $T \in L(H)$ is Hermitian, then $\sigma(T) \subseteq \mathbb{R}$.

Proof: Let $\lambda = s + ti$ with $t \neq 0$. For $x \in H$, let $y = (\lambda - T)x$. Then $y \cdot x = \lambda x \cdot x - Tx \cdot x$ and $x \cdot y = \overline{y \cdot x} = \overline{\lambda x \cdot x} - \overline{Tx \cdot x} = \bar{\lambda} x \cdot x - Tx \cdot x$

so $x\cdot y - y\cdot x = (\bar{\lambda} - \lambda)x\cdot x = -2itx\cdot x$ and

$$2|t|\,\|x\|^2 = |x\cdot y - y\cdot x| \le 2\|x\|\,\|y\|$$

by the Schwarz inequality. Hence, $\|y\| = \|(\lambda - T)x\| \ge |t|\,\|x\|$ and $\lambda \in \rho(T)$ by Theorem 23.

Theorem 28. If $T \in L(H)$ is Hermitian, then $\sigma(T) \subseteq [m(T), M(T)]$.

Proof: Suppose $\lambda \in \mathbb{R}$ is such that

$$\lambda > M(T) = \sup\{Tx\cdot x : \|x\| = 1\}.$$

Let $\varepsilon = \lambda - M(T)$. Then

$$(\lambda - T)x\cdot x = \lambda x\cdot x - Tx\cdot x \ge \lambda x\cdot x - M(T)x\cdot x = \varepsilon\|x\|^2.$$

But $(\lambda - T)x\cdot x \le \|(\lambda - T)x\|\,\|x\|$ so $\|(\lambda - T)x\| \ge \varepsilon\|x\|$ and $\lambda \in \rho(T)$ by Theorem 23.

Similarly, if $\lambda < m(T)$, $\lambda \in \rho(T)$.

Lemma 29 (Generalized Schwarz Inequality). If $T \in L(H)$ is positive, then

$$(5) \qquad |Tx\cdot y|^2 \le (Tx\cdot x)(Ty\cdot y) \;\; \forall\, x, y \in H.$$

Proof: Since T is positive, the function $\{x, y\} = Tx\cdot y$, $\{\ ,\ \} : H \times H \to \mathbb{C}$ has all the properties of an inner product except possibly $\{x, x\} = 0$ if and only if $x = 0$. This property is not required in "the proof" of the Schwarz inequality so (5) is a consequence of the Schwarz inequality for the "inner product" $\{\ ,\ \}$ (see the appendix for a proof of the Schwarz inequality).

Note for $T = I$, (5) is just the usual Schwarz inequality.

Theorem 30. If $T \in L(H)$ is Hermitian, then both $m(T)$ and $M(T)$ are in $\sigma(T)$.

Proof: Consider the case $m = m(T)$. For $x \in H$, $(T - m)x \cdot x \geq 0$ so since $m \in \mathbb{R}$, $T - m$ is Hermitian and positive. Apply Lemma 29 to $T - m$, x and $y = (T - m)x$ to obtain

$$|(T - m)x \cdot (T - m)x|^2 \leq ((T - m)x \cdot x)((T - m)^2 x \cdot (T - m)x)$$

so

$$\|(T - m)x\|^2 \leq ((T - m)x \cdot x)\|T - m\|^3 \|x\|^2. \tag{6}$$

From (6), $\inf\{\|(T - m)x\| : \|x\| = 1\} = 0$ since

$$\inf\{(T - m)x \cdot x : \|x\| = 1\} = 0$$

by definition of m. Therefore, $m \in \sigma(T)$ by Corollary 24.

We have a formula for the spectral radius in terms of the bounds for a Hermitian operator.

Proposition 31. If $T \in L(H)$ is Hermitian, then

$$\|T\| = r(T) = \max\{|m(T)|, |M(T)|\} = \{|Tx \cdot x| : \|x\| = 1\}.$$

Proof: Theorems 8 and 26.

Finally, we show that the Hermitian operators have a sequential completeness property very analogous to that of real sequences.

Theorem 32. Let $\{T_k\}$ be a sequence of Hermitian operators in $L(H)$ such that

(i) $T_1 \le T_2 \le \dots$

(ii) $\exists$ a Hermitian operator B such that $T_k \le B \ \forall\, k$.

Then $\exists$ a Hermitian operator $T \in L(H)$ such that $T_k \to T$ in the strong operator topology and $T \le B$.

Proof: We may assume $0 \le T_1 \le T_2 \le \dots < B$. For $n > m$ set $T_{nm} = T_n - T_m \ge 0$. By Lemma 29, for $x \in H$,

$$(T_{nm}x \cdot T_{nm}x)^2 = \|T_{nm}x\|^4 \le (T_{nm}x \cdot x)(T_{nm}^2 x \cdot T_{nm}x).$$

By Proposition 31, $0 \le T_{nm} \le B$ implies $\|T_{nm}\| \le \|B\|$ so

$$\text{(7)} \qquad \|T_{nm}x\|^4 = \|(T_n - T_m)x\|^4 \le ((T_n x \cdot x) - (T_m x \cdot x))\|B\|^3\|x\|^2.$$

The sequence of real numbers $\{T_n x \cdot x\}$ is bounded and increasing and, therefore, convergent. By (7), $\{T_n x\}$ is a Cauchy sequence and, hence, convergent to, say, Tx. T is continuous by the Banach- Steinhaus Theorem and $Tx \cdot y = \lim T_k x \cdot y = \lim x \cdot T_k y = x \cdot Ty$ for $x, y \in H$ implies that T is Hermitian. Clearly $T_k \le T \le B \ \forall\, k$.

Theorem 32 can be used to show the existence of a square root for a positive operator, but we give a proof in 41.3 based on the spectral theorem for Hermitian operators (see [BN] 23.1, [Ri] or [RN] §104).

xercise 1. Show the multiplication operator in Exercise 34.1 is normal. ;ive necessary and sufficient conditions for the operator to be Hermitian or ositive.

Exercise 2. Let c_{00} be given the inner product from ℓ^2 (complex sequences). Show $T : c_{00} \to c_{00}$ defined by $T\{t_k\} = \{kt_k\}$ is symmetric but not continuous.

Exercise 3. Give a proof of the Hellinger-Toeplitz Theorem by using the CGT.

Exercise 4. If $T \in L(H)$, show TT^* and T^*T are Hermitian and positive.

Exercise 5. Show Lemma 13 is false for real inner product spaces. (Consider rotations about the origin in the plane).

Exercise 6. If $T \in L(H)$ is normal, show T^n, λT and $\lambda - T$ are normal.

Exercise 7. Show the normal operators are a closed subspace of $L(H)$.

Exercise 8. If $T = A + iB$ with $A, B \in L(H)$ Hermitian, show T is normal if and only if A and B commute (recall Theorem 15). Thus, the normal operators do not form a linear subspace.

Exercise 9. If $T \in L(H)$ is Hermitian and p is a real polynomial, show

$p(T)$ is Hermitian.

Exercise 10. Give an example of a normal operator and a point in its spectrum which is an approximate eigenvalue but not an eigenvalue.

Exercise 11. If $T \in L(H)$ is Hermitian and $-\varepsilon I \leq T \leq \varepsilon I$, show $\|T\| \leq \varepsilon$.

Exercise 12. Let $T \in L(H)$. Show that T is an isometry if and only if $\|Tx\| = \|x\|$ $\forall\, x \in H$ if and only if $Tx \cdot Ty = x \cdot y$ $\forall\, x, y \in H$ if and only if T carries complete orthonormal sets into orthonormal sets if and only if $T^*T = I$.

Exercise 13. An operator $T \in L(H)$ is called <u>unitary</u> if T is an isometry onto H. Show that T is unitary if and only if $T^*T = TT^* = I$ if and only if T carries complete orthonormal sets to complete orthonormal sets.

Exercise 14. If $T \in L(H)$ is Hermitian, show that $U = e^{iH}$ is unitary.

Exercise 15. If U is unitary, show that eigenvectors corresponding to distinct eigenvalues are orthogonal.

36
The Spectral Theorem for Compact Symmetric Operators

In this section we establish the spectral theorem for compact symmetric operators. This theorem is a generalization of a well-known result for real symmetric matrices. Namely, if A is a real symmetric $n \times n$ matrix and $\lambda_1, ..., \lambda_n$ are the (real) eigenvalues of A with associated orthonormal eigenvectors $x_1, ..., x_n$, then $Ax = \sum_{i=1}^{n} \lambda_i (x \cdot x_i) x_i$ for each $x \in \mathbb{R}^n$. That is, the matrix representation of A with respect to the orthonormal basis $\{x_1, ..., x_n\}$ is the diagonal matrix, $A = \text{diagonal}\{\lambda_1, ..., \lambda_n\}$. We now show that any compact symmetric operator has such an eigenvalue-eigenvector series expansion.

Let $X \neq \{0\}$ be a complex, inner product space and $0 \neq T \in K(X)$ symmetric. We establish a series representation for T in terms of its eigenvalues and eigenvectors. The space X is not assumed to be complete but rather the compactness of T replaces a completeness assumption.

Lemma 1. Either $\|T\|$ or $-\|T\|$ is an eigenvalue of T and $\exists$ a corresponding eigenvector x with $\|x\| = 1$ and $|Tx \cdot x| = \|T\|$.

Proof: By 35.8, $\|T\| = \max\{|m(T)|, |M(T)|\}$ so $\exists\ x_n \in X$, $\|x_n\| = 1$, such that $Tx_n \cdot x_n \to \lambda$ where λ is real and $|\lambda| = \|T\|$ (35.5). Now

$$0 \le \|Tx_n - \lambda x_n\|^2 = \|Tx_n\|^2 - 2\lambda Tx_n \cdot x_n + \lambda^2\|x_n\|^2 \le \|T\|^2 - 2\lambda Tx_n \cdot x_n + \lambda^2 \to 0$$

so $Tx_n - \lambda x_n \to 0$. Since T is compact, $\{Tx_n\}$ has a convergent subsequence, say $\{Tx_{n_k}\}$. Since $x_{n_k} - (1/\lambda)Tx_{n_k} \to 0$, $\{x_{n_k}\}$ also converges to, say, x. Then $\|x\| = 1$ and $Tx_{n_k} \to Tx$ so $Tx = \lambda x$. Also $|Tx \cdot x| = |\lambda| x \cdot x = \|T\|$.

The symmetry of T is important; see Exercise 33.1.

Theorem 2. There exists a possibly finite sequence of non-zero eigenvalues $\{\lambda_k\}$ of T and a corresponding sequence of orthonormal eigenvectors $\{x_k\}$ such that

(1) $$Tx = \sum_k (Tx \cdot x_k)x_k = \sum_k \lambda_k (x \cdot x_k)x_k \quad \forall\ x \in X.$$

If the sequence $\{\lambda_k\}$ is infinite, $|\lambda_k| \downarrow 0$. Every non-zero eigenvalue of T occurs in the sequence $\{\lambda_k\}$; the eigenmanifold, $\mathcal{N}(\lambda_k - T)$, corresponding to each λ_k is finite dimensional and its dimension is exactly the number of times λ_k occurs in the sequence $\{\lambda_j\}$.

Proof: Let λ_1 and x_1 be the eigenvalue and corresponding eigenvector from Lemma 1. Note $|\lambda_1| = \|T\|$. Let $X_1 = X$ and $X_2 = \{x : x \cdot x_1 = 0\} = \{x_1\}^{\perp}$. Then X_2 is a closed subspace of X and X_2 is invariant under T since $x \in X_2$ implies

$$Tx \cdot x_1 = x \cdot Tx_1 = \lambda x \cdot x_1 = 0.$$

The restriction of T to X_2, T_2, is compact and symmetric. If $T_2 \neq 0$, by Lemma 1 there exists an eigenvalue λ_2 and unit eigenvector x_2 with $\|T_2\| = |\lambda_2|$. Clearly $|\lambda_2| \leq |\lambda_1|$ and $x_1 \cdot x_2 = 0$.

Continuing in this way produces non-zero eigenvalues $\lambda_1, \ldots, \lambda_k$ and corresponding unit eigenvectors $x_1, \ldots, x_k$ and closed subspaces $X_1, \ldots, X_k$ with $X_{k+1} \subseteq X_k$, where $X_{k+1} = \{x_1, \ldots, x_k\}^{\perp}$, $x_j \in X_j$ and $|\lambda_j| = \|T_j\|$, where T_j is T restricted to X_j. Thus,

$$|\lambda_1| \geq |\lambda_2| \geq \ldots \geq |\lambda_k|$$

and $x_i \cdot x_j = 0$ if $i \neq j$. This process stops at λ_n, x_n, X_{n+1} if and only if $T_{n+1} = 0$. In this case $\mathcal{R}T$ lies in the linear subspace generated by $\{x_1, \ldots, x_n\}$; for if $x \in X$ and $y_n = x - \sum_{k=1}^{n} (x \cdot x_k)x_k$, then $y_n \cdot x_k = 0$ $(k = 1, \ldots, n)$ and so $y_n \in X_{n+1}$ which implies

$$Ty_n = 0 = Tx - \sum_{k=1}^{n} (x \cdot x_k)Tx_k = Tx - \sum_{k=1}^{n} \lambda_k(x \cdot x_k)x_k$$

and $Tx \in \text{span}\{x_1, \ldots, x_n\}$. In this case (1) is clearly satisfied. If the process does not terminate, we obtain infinite sequences $\{\lambda_k\}_{k=1}^{\infty}$, $\{x_k\}_{k=1}^{\infty}$, where $|\lambda_k| \geq |\lambda_{k+1}|$, $\|x_k\| = 1$ and $\{x_k : k \in \mathbb{N}\}$ is orthonormal.

We claim $\lambda_k \to 0$. Since $|\lambda_k| \geq |\lambda_{k+1}|$, either $\lambda_k \to 0$ or $\exists\, \varepsilon > 0$ such that $|\lambda_k| \geq \varepsilon\ \forall\, k$. Suppose the latter holds. Then $\{x_k/\lambda_k\}$ is bounded so $\{T(x_k/\lambda_k)\} = \{x_k\}$ has a convergent subsequence. But $\{x_k\}$ is orthonormal so $\|x_k - x_j\|^2 = 2$, $k \neq j$, and $\{x_k\}$ cannot have a convergent subsequence. Hence, $\lambda_k \to 0$.

We next claim that (1) holds when $\{\lambda_k\}$ is infinite. For $x \in X$, let $y_n = x - \sum_{k=1}^{n} (x\cdot x_k)x_k$. Then $\|y_n\|^2 = \|x\|^2 - \sum_{k=1}^{n} |x\cdot x_k|^2 \leq \|x\|^2$. Since $y_n \in X_{n+1}$ and $|\lambda_{n+1}| = \|T_{n+1}\|$, $\|Ty_n\| \leq |\lambda_{n+1}|\,\|y_n\| \leq |\lambda_{n+1}|\,\|x\|$ so $Ty_n \to 0$. But $Ty_n = Tx - \sum_{k=1}^{n} (x\cdot x_k)Tx_k = Tx - \sum_{k=1}^{n} \lambda_k(x\cdot x_k)x_k$ and (1) holds.

If $\lambda \neq 0$ is an eigenvalue of T which is not in the sequence $\{\lambda_k\}$, then there is a corresponding unit eigenvector x and $x\cdot x_k = 0\ \forall\, k$ by 35.7. By (1), $Tx = 0 = \lambda x$ but $\lambda x \neq 0$ so $\{\lambda_k\}$ exhausts all of the non-zero eigenvalues of T.

Suppose the eigenvalue λ_k occurs N times in the sequence $\{\lambda_j\}$. Then the eigenmanifold $\mathcal{N}(\lambda_k - T) = \mathcal{N}_k$ is at least N-dimensional. If dimension $\mathcal{N}_k > N$, $\exists\, x$ with $\|x\| = 1$, $Tx = \lambda_k x$, $x\cdot x_j = 0\ \forall\, j$ (35.7). By (1), $Tx = 0 = \lambda_k x$, which is a contradiction. Hence, $\dim \mathcal{N}_k = N$.

From Theorem 2 we can now obtain a similar eigenvalue-eigenvector expansion for the resolvent operator of T.

Theorem 3. Let the notation be as in Theorem 2. If $0 \neq \lambda \in \mathbb{C}$ and

$\lambda = \lambda_k \ \forall\ k$, then $(\lambda - T)^{-1} \in L(X)$ and is given by

$$(2) \qquad (\lambda - T)^{-1}y = \frac{1}{\lambda}y + \frac{1}{\lambda}\sum_k \lambda_k \frac{y \cdot x_k}{\lambda - \lambda_k} x_k = x \quad \forall\ y \in X.$$

Proof: First suppose that (2) converges. Then $(\lambda - T)x = y$ since

$$\begin{aligned}(\lambda - T)x &= y + \sum \lambda_k \frac{y \cdot x_k}{\lambda - \lambda_k} x_k - \frac{1}{\lambda} Ty - \frac{1}{\lambda}\sum \lambda_k \frac{y \cdot x_k}{\lambda - \lambda_k} Tx_k \\ &= y + \sum \lambda_k \frac{y \cdot x_k}{\lambda - \lambda_k} x_k - \frac{1}{\lambda}\sum \lambda_k (y \cdot x_k) x_k - \frac{1}{\lambda}\sum \lambda_k^2 \frac{y \cdot x_k}{\lambda - \lambda_k} x_k \\ &= y + \sum (y \cdot x_k) \left\{ \frac{\lambda_k}{\lambda - \lambda_k} - \frac{\lambda_k}{\lambda} - \frac{\lambda_k^2}{\lambda(\lambda - \lambda_k)} \right\} x_k = y\end{aligned}$$

so $x = (\lambda - T)^{-1}y$.

We now show that (2) converges. Let

$a = \sup\{ |\lambda_k/(\lambda - \lambda_k)| : k \in \mathbb{N}\}$, $b = \sup\{ 1/|\lambda - \lambda_k| : k \in \mathbb{N}\}$,

$u_n = \sum_{k=1}^{n} \lambda_k \frac{y \cdot x_k}{\lambda - \lambda_k} x_k$ and $v_n = \sum_{k=1}^{n} \frac{y \cdot x_k}{\lambda - \lambda_k} x_k$. If $m < n$,

$$\|u_n - u_m\|^2 = \sum_{k=m+1}^{n} \left| \frac{\lambda_k}{\lambda - \lambda_k} \right|^2 |y \cdot x_k|^2 \le a^2 \sum_{k=m+1}^{n} |y \cdot x_k|^2.$$

Hence, $\{u_n\}$ is Cauchy. Now

$$\|v_n\|^2 = \sum_{k=1}^{n} \frac{|y \cdot x_k|^2}{|\lambda - \lambda_k|^2} \le b^2 \sum_{k=1}^{n} |y \cdot x_k|^2 \le b^2 \|y\|^2$$

by Bessel's inequality so $\{v_n\}$ is bounded. Since $Tv_n = u_n$ and T is compact, $\{u_n\}$ must have a convergent subsequence and since $\{u_n\}$ is Cauchy, it must converge. Hence, (2) converges.

From (2), $\|x\| \le \|y\|/|\lambda| + a\|y\|/|\lambda|$ so $(\lambda - T)^{-1}$ is continuous, has domain X and $\|(\lambda - T)^{-1}\| \le (1 + a)/|\lambda|$.

Corollary 4. Let the notation be as in Theorem 2, and let $M = \overline{\text{span}}\{x_k : k \in \mathbb{N}\}$. Then $M^{\perp} = \mathcal{N}(T)$ so $X = M \oplus \mathcal{N}(T)$.

Proof: Let $y \in M^{\perp}$. Then $y \cdot x_k = 0$ $\forall$ k and

$$Ty = \sum \lambda_k (y \cdot x_k) x_k = 0.$$

Therefore, $M^{\perp} \subseteq \mathcal{N}(T)$.

Let $y \in \mathcal{N}(T)$. Then $y \cdot x_k = \frac{1}{\lambda_k} y \cdot Tx_k = \frac{1}{\lambda_k} Ty \cdot x_k = 0$ so $y \in M^{\perp}$. Hence, $\mathcal{N}(T) \subseteq M^{\perp}$.

Remark 5. If X is complete, then $\{x_k\}$ is a complete orthonormal set in X if and only if $\lambda = 0$ is not an eigenvalue of T.

Both eigenvector expansions in (1), (2) are applicable to integral equations. For example, if $k \in L^2[a, b] \times [a, b]$ and $K \in K(L^2[a, b])$ is the integral operator induced by the kernel k, $Kf(t) = \int_a^b k(t, s)f(s)ds$, and if $k(t, s) = \overline{k(s, t)}$ for almost all t, s, then

$$Kf = \sum_k \lambda_k (f \cdot x_k) x_k \tag{3}$$

and

$$(\lambda - K)^{-1} g = f = \frac{1}{\lambda} g + \frac{1}{\lambda} \sum_k \lambda_k \frac{g \cdot x_k}{\lambda - \lambda_k} x_k \quad (\lambda \ne 0, \lambda \ne \lambda_k \ \forall\, k), \tag{4}$$

where $\{\lambda_k\}$ and $\{x_k\}$ are the eigenvalues and eigenvectors of K as in Theorem 2, and the series in (3) and (4) are convergent in the L^2-norm.

If the kernel k is well-behaved the convergence of the series in (3) and (4) is actually uniform on the interval [a, b].

Theorem 6 (Hilbert-Schmidt). If the kernel k satisfies the condition

$$\sup_t \int_a^b |k(t, s)|^2 ds = M < \infty, \tag{5}$$

then the series in (3) and (4) converge uniformly and absolutely on [a, b].

Proof: Since $\lambda_j x_j(t) = Kx_j(t) = \int_a^b k(t, s)x_j(s)ds$ for any fixed t $\lambda_j x_j(t)$ is the j^{th} Fourier coefficient of the function $k(t, \cdot)$ with respect to $\{x_j\}$. By Bessel's Inequality,

$$\sum_j |\lambda_j x_j(t)|^2 \le \int_a^b |k(t, s)|^2 ds$$

and $\sum_j |f \cdot x_j|^2 \le \|f\|^2$ for every $f \in L^2[a, b]$ so by the Schwarz Inequality and (5),

$$\sum_{j=n}^{\infty} |\lambda_j (f \cdot x_j) x_j(t)| \le (\sum_{j=n}^{\infty} |f \cdot x_j|^2)^{1/2} (\sum_{j=n}^{\infty} |\lambda_j x_j(t)|^2)^{1/2}$$

$$\le (\sum_{j=n}^{\infty} |f \cdot x_j|^2)^{1/2} \int_a^b |k(t, s)|^2 ds$$

$$\le M(\sum_{j=n}^{\infty} |f \cdot x_j|^2)^{1/2}.$$

Hence, the series in (3) and (4) converge uniformly and absolutely.

In particular, if the kernel k is continuous, condition (5) is satisfied and the conclusion of Theorem 6 holds (Exercise 3).

Note that the series in (4) gives the solution to the integral equation $(\lambda - K)f = g$ when $\lambda \neq 0$, $\lambda \neq \lambda_k$ $\forall k$. For the solution of the equation $Kf = g$ see Exercise 2.

Finally we show that Theorem 2 can be used to give a spectral representation for a compact operator between arbitrary Hilbert spaces.

Theorem 7. Let H_1, H_2 be complex Hilbert spaces and $T \in K(H_1, H_2)$. Then $\exists$ orthonormal sequences $\{x_k\} \subseteq H_1$, $\{y_k\} \subseteq H_2$ and $\lambda_k \in \mathbb{R}$, $\lambda_k \downarrow 0$ such that $Tx = \sum_k \lambda_k (x \cdot x_k) y_k$ $\forall x \in H_1$.

Proof: The operator $T^*T \in K(H_1)$ and since

$$T^*Tx \cdot x = Tx \cdot Tx \geq 0,$$

the non-zero eigenvalues of T^*T are positive; denote these by λ_k^2 and arrange them so that $\lambda_{k+1} \leq \lambda_k$ $\forall k$. Let $\{x_k\}$ be the eigenvectors associated with $\{\lambda_k^2\}$ as in Theorem 2. Set $y_k = (1/\lambda_k)Tx_k$. Then $\{y_k\}$ is orthonormal in H_2 since

$$y_i \cdot y_j = (1/\lambda_i \lambda_j)Tx_i \cdot Tx_j = (1/\lambda_i \lambda_j)T^*Tx_i \cdot x_j = (\lambda_i/\lambda_j)x_i \cdot x_j = 0$$

for $i \neq j$ and $\|y_k\|^2 = (1/\lambda_k^2)T^*Tx_k \cdot x_k = 1$.

We claim that the series

$$\text{(6)} \qquad Tx = \sum_j \lambda_j (x \cdot x_j) y_j$$

converges to $Tx \ \forall x \in H_1$. First, the series on the right hand side of (6) converges in the norm topology of $L(H_1, H_2)$ since

$$\|\sum_{j=m}^{n} \lambda_j(x\cdot x_j)y_j\|^2 = \sum_{j=m}^{n} |\lambda_j|^2|x\cdot x_j|^2 \le \lambda_m^2 \sum_{j=m}^{n} |x\cdot x_j|^2 \le \lambda_m^2\|x\|^2$$

by Bessel's inequality. Thus, the right hand side of (6) represents a compact operator in $K(H_1, H_2)$. To establish the equality in (6) it suffices to establish the equality on a dense subset of H_1. By Corollary 4 $H_1 = \overline{\text{span}\{x_k\}} \oplus \mathcal{N}(T^*T)$. The equality holds for $x = x_k$ and since $\mathcal{N}(T^*T) = \mathcal{N}(T)$, the equality holds on H_1.

It follows from Theorem 7, that any compact operator between Hilbert spaces is the limit in the operator norm of a sequence of finite rank operators (Exercise 4).

Exercise 1. Show the converse of Theorem 7 holds.

Exercise 2. Let $T \in L(H)$ be compact, Hermitian. Given $z \in H$ show that the equation $Tx = z$ has a solution, x, if and only if $z \perp \mathcal{N}(T)$ and $\sum|z\cdot x_k|^2/|\lambda_k|^2 < \infty$, where the notation is as in Theorem 2.

Exercise 3. If the kernel k is continuous, show that condition (5) holds.

Exercise 4. Show that every compact operator on a Hilbert space is the norm limit of a sequence of operators with finite dimensional range. (Every Hilbert space has the approximation property.)

36.1 Hilbert-Schmidt Operators

In Example 28.7 we showed that an integral operator

$$Kf(t) = \int_a^b k(t, s)f(s)ds$$

induced by a kernel $k \in L^2(I \times I)$, $I = [a, b]$, is a compact operator. In this section we will give a characterization of such integral operators in terms of their spectral representation as given in Theorem 36.7.

Let H_1, H_2 be complex Hilbert spaces. In order to introduce the class of operators which we will study in this section, we require a preliminary lemma.

Lemma 1. Let $\{e_a : a \in A\}$, $\{f_b : b \in B\}$ be complete orthonormal sets in H_1, H_2, respectively. Let $T \in L(H_1, H_2)$. Then

$$\sum_{a \in A} \|Te_a\|^2 = \sum_{b \in B} \|T^* f_b\|^2.$$

Proof: By Parseval's Equality

$$\sum_{a \in A} \|Te_a\|^2 = \sum_{a \in A} \sum_{b \in B} |Te_a \cdot f_b|^2 = \sum_{b \in B} \sum_{a \in A} |e_a \cdot T^* f_b|^2 = \sum_{b \in B} \|T^* f_b\|^2.$$

Definition 2. An operator $T \in L(H_1, H_2)$ is a <u>Hilbert-Schmidt operator</u> (<u>HS-operator</u>) if and only if $\sum_{a \in A} \|Te_a\|^2 < \infty$ for every (for any by Lemma 1) complete orthonormal set $\{e_a : a \in A\}$ in H_1.

We denote the class of all HS-operators from H_1 into H_2 by

$HS(H_1, H_2)$. The Hilbert-Schmidt norm on $HS(H_1, H_2)$ is defined to be $|T| = (\sum_{a\in A} \|Te_a\|^2)^{1/2}$, where $\{e_a : a \in A\}$ is a complete orthonormal set in H_1; note from Lemma 1 that this norm is independent of the particular complete orthonormal family chosen.

Also, from Lemma 1, we have

Proposition 3. An operator $T \in L(H_1, H_2)$ is a HS-operator if and only if T^* is a HS-operator and, in this case, $|T| = |T^*|$.

Theorem 4. $HS(H_1, H_2)$ is a B-space under the Hilbert-Schmidt norm, $|\ |$, with

$$(1) \qquad \|T\| \le |T| \quad \forall\, T \in HS(H_1, H_2).$$

Proof: For $x \in H_1$ with $\|x\| \le 1$,

$$\|Tx\|^2 = \sum_{a\in A} |Tx\cdot e_a|^2 = \sum_{a\in A} |x\cdot T^*e_a|^2 \le \|x\|^2 \sum_{a\in A} \|T^*e_a\|^2 \le |T|,$$

when $\{e_a : a \in A\}$ is a complete orthonormal set in H_1. This gives (1).

That $T \to |T|$ defines a norm on the vector space $HS(H_1, H_2)$ is easily checked. Only the completeness remains to be checked. Suppose $\{T_k\} \subseteq HS(H_1, H_2)$ is Cauchy with respect to $|\ |$. By (1), $\{T_k\}$ is $\|\ \|$-Cauchy so there exists $T \in L(H_1, H_2)$ such that $\|T_k - T\| \to 0$. Set $M = \sup\{|T_k| : k\}$ and let $\varepsilon > 0$. Let N be such that $|T_k - T_j| < \varepsilon$ for $k, j \ge N$. For $A_0 \subseteq A$ finite,

$$\sum_{a\in A_0} \|(T - T_k)e_a\|^2 = \lim_j \sum_{a\in A_0} \|(T_j - T_k)e_a\|^2 \le \lim_j |T_j - T_k| \le \varepsilon$$

if $k \ge N$. Thus, $T - T_k \in HS(H_1, H_2)$ so $T \in HS(H_1, H_2)$ with $|T - T_k| \le \varepsilon$ if $k \ge N$.

Concerning the composition of HS-operators, we have

Proposition 5. Let H_3 be a complex Hilbert space.

(i) If $T \in HS(H_1, H_2)$ and $S \in L(H_2, H_3)$, then $ST \in HS(H_1, H_3)$ with $|ST| \le \|S\| |T|$.

(ii) If $S \in L(H_1, H_2)$ and $T \in HS(H_2, H_3)$, then $TS \in HS(H_1, H_3)$ with $|TS| \le |T| \|S\|$.

Proof: (i): If $\{e_a : a \in A\}$ is a complete orthonormal set in H_1, then

$$|ST|^2 = \sum_{a\in A} \|STe_a\|^2 \le \|S\|^2 \sum_{a\in A} \|Te_a\|^2 = \|S\|^2 |T|^2.$$

(ii) follows from (i) and Proposition 3 since $(TS)^* = S^*T^*$.

Thus, $HS(H_1) = HS(H_1, H_1)$ is a two-sided ideal in the algebra $L(H_1)$. In general, $HS(H_1)$ doesn't have an identity (Exercise 1).

Theorem 6. Let $T \in HS(H_1, H_2)$. Then there exists a sequence $\{T_k\}$ of operators with finite dimensional range such that $|T_k - T| \to 0$. In particular, any HS-operator is compact.

Proof: Let $\{e_a : a \in A\}$ be a complete orthonormal set in H_1. Since $|T|^2 = \sum_{a\in A} \|Te_a\|^2 < \infty$, only a countable number of

$$\{a_1, a_2, ...\} \subseteq A$$

are such that $\|Te_{a_i}\| \neq 0$. Define $T_k \in L(H_1, H_2)$ by $T_k e_a = Te_a$ if $a \in \{a_1, a_2, ..., a_k\}$ and $T_k e_a = 0$ otherwise. Then

$$|T - T_k|^2 = \sum_{a\in A} \|(T - T_k)e_a\|^2 = \sum_{j=k+1}^{\infty} \|Te_{a_j}\|^2 \to 0$$

as $k \to \infty$.

The last statement follows from (1) and 28.3.

Recall that if $T \in L(H_1, H_2)$ is compact, then there exist $\lambda_k \downarrow 0$ and orthonormal sets $\{x_k\} \subseteq H_1$, $\{y_k\} \subseteq H_2$ such that $Tx = \sum_k \lambda_k (x \cdot x_k) y_k$ $\forall\, x \in H_1$ (36.7). We give a characterization of the HS-operator in terms of the $\{\lambda_k\}$.

Theorem 7. $T \in HS(H_1, H_2)$ if and only if $\{\lambda_k\} \in \ell^2$ and, in this case, $|T|^2 = \sum_k |\lambda_k|^2$.

Proof: Extend $\{x_k : k\} = E$ to a complete orthonormal set E' in H_1. If $x \in E'$, then $Tx = \lambda_k y_k$ if $x = x_k \in E$ and $Tx = 0$ if $x \in E' \setminus E$ so that

$$|T|^2 = \sum_{x\in E'} \|Tx\|^2 = \sum_k \|\lambda_k y_k\|^2 = \sum_k |\lambda_k|^2 .$$

In particular, if $T \in L(H_1)$ is compact and symmetric, then T is a HS-operator if and only if the eigenvalues of T, $\{\lambda_k\}$, are in ℓ^2 (36.2).

We now show that when $H_1 = H_2 = L^2(I)$, $I = [a, b]$, the HS-operators are exactly the integral operators induced by L^2-kernels. First, assume that $k : I \times I \to \mathbb{C}$ belongs to $L^2(I \times I)$ and let $K : L^2(I) \to L^2(I)$ be the integral operator induced by k. Let $Kf = \sum_k \lambda_k (f \cdot x_k) y_k$ where $\lambda_k \downarrow 0$ and $\{x_k\}$, $\{y_k\}$ are orthonormal in $L^2(I)$ as in 36.7. For fixed t we have

$$Kx_k(t) = \lambda_k y_k(t) = \int_I k(t, s) x_k(s) ds$$

so $\lambda_k y_k(t)$ is the k^{th} Fourier coefficient of the function $k(t, \cdot)$ with respect to $\{x_k : k\}$. Bessel's Inequality implies that

$$\sum_k |\lambda_k y_k(t)|^2 \le \int_I |k(t, s)|^2 ds$$

so

$$\sum_k |\lambda_k|^2 = \sum_k \int_I |\lambda_k y_k(t)|^2 dt \le \int_I \int_I |k(t, s)|^2 ds\, dt < \infty$$

and K is a HS-operator by Theorem 7.

Conversely, assume that $K : L^2(I) \to L^2(I)$ is a HS-operator. Let $\{\lambda_k\} \in \ell^2$ and $\{x_k\}$, $\{y_k\}$ be orthonormal with $Kf = \sum_k \lambda_k (f \cdot x_k) y_k$ as in 36.7 and Theorem 7. Let $y_k \otimes \bar{x}_k : I \times I \to \mathbb{C}$ be the function

$$y_k \otimes \bar{x}_k(t, s) = y_k(t) \bar{x}_k(s).$$

Then $\{y_k \otimes \bar{x}_k : k\}$ is an orthonormal sequence in $L^2(I \times I)$ so the series

$\sum_k \lambda_k y_k \otimes \bar{x}_k$ is norm convergent to a function $k \in L^2(I \times I)$. Since the series $\sum_k \lambda_k y_k$ is norm convergent in $L^2(I)$, we have

$$Kf(t) = \sum_k \lambda_k \int_I f(s)\bar{x}_k(s)ds\, y_k(t) = \int_I \sum_k \lambda_k y_k(t)\bar{x}_k(s)f(s)ds$$
$$= \int_I k(t, s)f(s)ds,$$

and the HS-operator K is the integral operator induced by the L^2-kernel k.

Exercise 1. Show the identity operator on H_1 is a HS-operator if and only if H_1 is finite dimensional.

Exercise 2. Let $\{e_a : a \in A\}$ be a complete orthonormal set in H_1. Show $S \cdot T = \sum_{a \in A} Se_a \cdot Te_a$ defines an inner product on $HS(H_1)$ which induces the Hilbert-Schmidt norm.

Exercise 3. Give an example of a compact operator which is not a HS-operator.

Exercise 4. If $T : H \to H$ is a HS-operator and $U \in L(H)$ is unitary (Exercise 35.13), show $U^{-1}TU$ is HS and $|T| = |U^{-1}TU|$.

Exercise 5. Show the Volterra operator $Tf(t) = \int_0^t f$ is a Hilbert-Schmidt operator on $L^2[0, 1]$.

37
Symmetric Operators with Compact Resolvent

Differential operators are not, in general, compact but many such operators do have compact resolvents and the eigenvector expansions of §36 are applicable to such operators. In this section we first give an abstract description of the eigenvector expansion of symmetric operators with compact resolvent and then show how this expansion applies to Sturm-Liouville differential operators.

Let X be an infinite dimensional, complex inner product space and $A : \mathscr{D}(A) \subseteq X \to X$ linear and symmetric.

Theorem 1. Suppose $T = A^{-1}$ exists with $T \in K(X)$. Let $\{\lambda_k\}$, $\{x_k\}$ be the eigenvalues and eigenvectors of T as in Theorem 36.2. Set $\mu_k = 1/\lambda_k$. Then the sequence $\{\mu_k\}$ is infinite and $|\mu_k| \to \infty$. For each $x \in \mathscr{D}(A)$,

$$x = \sum_{k=1}^{\infty} (x \cdot x_k) x_k . \tag{1}$$

A point $\mu \in \sigma(A)$ if and only if μ is one of the $\{\mu_k\}$ and $Ax_k = \mu_k x_k$. If $\mu \notin \sigma(A)$, then

(2) $$(\mu - A)^{-1}y = \sum_{k=1}^{\infty} \frac{y \cdot x_k}{\mu - \mu_k} x_k \quad \forall\, y \in X$$

and $(\mu - A)^{-1} \in K(X)$.

Proof: Since A is symmetric, T is symmetric; for if x, $y \in X = \mathscr{D}(T)$ and $Ax_1 = x$, $Ay_1 = y$, then

$$x \cdot Ty = Ax_1 \cdot y_1 = x_1 \cdot Ay_1 = Tx \cdot y.$$

By 36.4, $X = \overline{\text{span}\{x_k\}} \oplus \mathscr{N}(T) = \overline{\text{span}\{x_k\}}$ so $\{x_k\}$ is infinite since X is infinite dimensional. Hence, $\{\lambda_k\}$ is infinite and $\{\mu_k\}$ is infinite with $|\mu_k| \to \infty$.

If $x \in \mathscr{D}(A)$, then $x = Ty$ for some y so by 36.2,

$$x = Ty = \sum_k (Ty \cdot x_k)x_k = \sum_k (x \cdot x_k)x_k$$

and (1) holds.

We claim that if $\mu \neq 0$ and $\mu \neq \mu_k \ \forall\, k$, then

(3) $$(\mu - A)^{-1} = \frac{1}{\mu} T(T - 1/\mu)^{-1}.$$

Suppose $x \in \mathscr{D}(A)$ and set $y = (\mu - A)x$. Then $\mu Tx - x = Ty$ so $(T - 1/\mu)x = \frac{1}{\mu} Ty$ and

(4) $$x = \frac{1}{\mu}(T - 1/\mu)^{-1}Ty = \frac{1}{\mu} T(T - 1/\mu)^{-1}(\mu - A)x$$

since $\frac{1}{\mu} \in \rho(T)$ and $(T - 1/\mu)^{-1} \in L(X)$. Suppose $y \in X$ and set $x = \frac{1}{\mu} T(T - 1/\mu)^{-1}y$ so $x \in \mathscr{D}(A)$ and

(5) $$(\mu - A)x = (\mu - A)\frac{1}{\mu}T(T - 1/\mu)^{-1}y = (T - 1/\mu)(T - 1/\mu)^{-1}y = y.$$

Equations (4) and (5) establish (3), and (3) shows that $(\mu - A)^{-1}$ is compact.

Since $0 \in \rho(A)$ by hypothesis, (3) implies that

$$\sigma(A) \subseteq \{\mu_k : k \in \mathbb{N}\}.$$

But $Tx_k = \lambda_k x_k$ implies $x_k = \lambda_k A x_k$ so $Ax_k = \mu_k x_k$ and each μ_k is an eigenvalue of A with associated eigenvector x_k. Thus, $\sigma(A) = \{\mu_k : k \in \mathbb{N}\}$.

Finally, to show (2), from 36.3,

$$(1/\mu - T)^{-1}y = \mu y + \mu\sum\frac{y\cdot x_k}{\mu_k(1/\mu - 1/\mu_k)}x_k = \mu y + \mu^2\sum\frac{y\cdot x_k}{\mu_k - \mu}x_k$$

so

$$1/\mu\, T(T - 1/\mu)^{-1}y = -Ty + \mu\sum\frac{y\cdot x_k}{\mu_k(\mu - \mu_k)}x_k$$

$$= -\sum\frac{y\cdot x_k}{\mu_k}x_k + \mu\sum\frac{y\cdot x_k}{\mu_k(\mu-\mu_k)}x_k$$

$$= \sum\frac{y\cdot x_k}{\mu - \mu_k}x_k = (\mu - A)^{-1}y.$$

Remark 2. If X is complete, it follows from 36.5 that $\{x_k\}$ is a complete orthonormal set.

We now apply Theorem 1 to the Sturm-Liouville differential operator L of Example 35.4. Recall that the domain of L consists of all functions $f \in C^2[a, b]$ satisfying $B_1 f = \alpha_1 f(a) + \alpha_2 f'(a) = 0$,

$B_2 f = \beta_1 f(b) - \beta_2 f'(b) = 0$, and L is given by $Lf = \frac{1}{w}[-(pf')' + qf]$, where p, p', q and w are real-valued and continuous with $p(t) > 0$, $w(t) > 0$ for all $t \in [a, b]$. Recall further that we are using the inner product $f \cdot g = \int_a^b f\bar{g}w$.

We describe the eigenvalues of L. Fix λ. The equation $Lu = \lambda u$ reduces to the differential equation $u'' + \frac{p'}{p}u' + \frac{\lambda w - q}{p}u = 0$. Let u_1 and u_2 be two linearly independent solutions of this differential equation. We claim that λ is an eigenvalue of L if and only if

$$\Delta(\lambda) = \begin{vmatrix} B_1 u_1 & B_1 u_2 \\ B_2 u_1 & B_2 u_2 \end{vmatrix} = 0 . \tag{6}$$

Every solution of the differential equation can be written in the form $u = c_1 u_1 + c_2 u_2$. The boundary conditions for u become

$$\begin{aligned} B_1 u &= c_1 B_1 u_1 + c_2 B_1 u_2 = 0 \\ B_2 u &= c_1 B_2 u_1 + c_2 B_2 u_2 = 0. \end{aligned} \tag{7}$$

If λ is an eigenvalue, then (7) has a non-trivial solution (c_1, c_2) so the determinant $\Delta(\lambda)$ in (6) is zero. On the other hand, if $\Delta(\lambda) = 0$, then (7) has a non-trivial solution (c_1, c_2) and $u = c_1 u_1 + c_2 u_2$ is then an eigenvector associated with the eigenvalue λ.

We show that if $\lambda = 0$ is not an eigenvalue of L, then L^{-1} can be represented as an integral operator whose kernel is called the <u>Green's function</u> of L.

Theorem 3. Assume that 0 is not an eigenvalue of L. Then ∃ a continuous, symmetric function $g : [a, b] \times [a, b] \to \mathbb{R}$, called the Green's function of L, such that $L^{-1}u(t) = \int_a^b g(t, s)u(s)ds$.

Proof: For $j = 1, 2$ let u_j be a non-trivial real-valued solution to $pu'' + p'u' - qu = 0$, $B_ju = 0$. Since 0 is not an eigenvalue of L, u_1 and u_2 are linearly independent and $B_iu_j \neq 0$ for $i \neq j$. Let $W(t) = \begin{vmatrix} u_1(t) & u_1'(t) \\ u_2(t) & u_2'(t) \end{vmatrix}$ be the Wronskian of u_1, u_2. Then the general solution of the equation $Lu = v$ is given by

$$u(t) = c_1u_1(t) + c_2u_2(t) + \int_a^t \frac{u_1(t)u_2(s) - u_1(s)u_2(t)}{p(a)W(a)} v(s)w(s)ds.$$

We want to choose the coefficients c_1 and c_2 so that $u \in \mathscr{D}(L)$ when v is continuous, i.e., u satisfies the boundary conditions $B_iu = 0$, $i = 1, 2$. If we set $c_2 = 0$ and

$$c_1 = -\int_a^b \frac{u_2(s)}{p(a)W(a)} v(s)w(s)ds\ ,$$

the solution u becomes

$$u(t) = -\int_t^b \frac{u_1(t)\,u_2(s)}{p(a)W(a)} v(s)w(s)ds - \int_a^t \frac{u_2(t)\,u_1(s)}{p(a)W(a)} v(s)w(s)ds$$

$$= \int_a^b g(t, s)v(s)w(s)ds,$$

where

$$g(t, s) = \begin{cases} -u_1(t)u_2(s)/p(a)W(a) & a \le s \le t \le b \\ -u_2(t)u_1(s)/p(a)W(a) & a \le t \le s \le b\,. \end{cases}$$

It is easily checked that g is continuous, $g(s, t) = g(t, s)$ and that the function u satisfies the desired boundary conditions.

Thus, if 0 is not an eigenvalue of L, Theorems 36.2 and 36.3 are applicable and give eigenfunction expansions for both $Lu = 0$ and for the solution of the differential equation $(\lambda - L)u = v$ when λ is not an eigenvalue of L. Moreover, it follows from 36.6 that these eigenfunction expansions are uniformly convergent.

As an interesting application of these results, consider the case where $w = p = 1$, $q = 0$ and $[a, b] = [0, \pi]$ with $B_1u = u(0)$, $B_2u = u(\pi)$. Then the eigenvalues of L are $\{k^2\}$ with associated eigenfunctions $\{\sin kt\}$. By Remark 2, $\{\sin kt\}$ forms a complete orthonormal set in $L^2[0, \pi]$.

A specific example where 0 is not an eigenvalue of L is given in Exercise 1. For applications of the spectral theorem for compact symmetric operators to partial differential operators, see [NS], 7.8.

Exercise 1. Let $w = 1$, $B_1u = u(a)$, $B_2u = u(b)$. Show

$$Lu \cdot u = \int_a^b (p|u'|^2 + q|u^2|).$$

Use this to show that all non-zero eigenvalues of L are positive and 0 is not an eigenvalue of L.

Exercise 2. Use the formula in Theorem 3 to find the Green's function for $Lu = -u'' = 0$, $u(0) = u(1) = 0$. Find the eigenvalues and corresponding eigenvectors for L.

Exercise 3. Repeat Exercise 2 for $Lu = 0$ and $u(0) = u'(1) = 0$.

38
Orthogonal Projections and the Spectral Theorem for Compact Symmetric Operators

In this section we show that the spectral representation of Theorem 36.2 can be rewritten in an integral form by using orthogonal projections. This integral representation will suggest how one should seek a spectral representation for operators which are not compact and have more complicated spectra. We first discuss orthogonal projections.

Let H be a complex Hilbert space and $P \in L(H)$ a projection. Then

$$H = \mathscr{R}P \oplus \mathscr{N}(P) = \mathscr{R}P \oplus \mathscr{R}(I - P) \quad (27.2).$$

Definition 1. P is an <u>orthogonal projection</u> if $\mathscr{R}P \perp \mathscr{N}(P)$.

In terms of the operator P, we have

Proposition 2. A projection P is orthogonal if and only if P is Hermitian.

Proof: Let $x, y \in H$ with $x = Px + u$, $y = Py + v$, $u, v \in \mathcal{N}(P)$.

Then

$$(1) \qquad \begin{aligned} Px \cdot y &= Px \cdot Py + Px \cdot v , \\ x \cdot Py &= Px \cdot Py + u \cdot Py . \end{aligned}$$

If P is orthogonal, then (1) implies $Px \cdot y = x \cdot Py$ so P is Hermitian.

If P is Hermitian and $x \in \mathcal{R}P$, $y \in \mathcal{N}(P)$, then $Px = x$, $Py = 0$ so

$$x \cdot y = Px \cdot y = x \cdot Py = x \cdot 0 = 0$$

and P is orthogonal.

Proposition 3. An orthogonal projection P is positive, $\|P\| \le 1$ and $0 \le Px \cdot x \le 1$ if $\|x\| = 1$. If $P \ne 0$, then $\|P\| = 1$, and if $0 \ne P \ne I$, then $m(P) = 0$, $M(P) = 1$.

Proof: P is positive since $Px \cdot x = P^2x \cdot x = Px \cdot Px = \|Px\|^2 \ge 0$.

If $x \in H$, $x = Px + u$, where $u \in \mathcal{N}(P)$, so

$$\|x\|^2 = \|Px\|^2 + \|u\|^2 \ge \|Px\|^2$$

which implies $\|P\| \le 1$. Moreover, if $P \ne 0$, $\exists\, x \ne 0$ such that $Px = x$ so $\|P\| = 1$.

If $\|x\| = 1$, $0 \le Px \cdot x = Px \cdot Px \le \|Px\|^2 \le \|P\|^2 \le 1$.

Since $0 \le Px \cdot x \le 1$ for $\|x\| = 1$, $0 \le m(P) \le M(P) \le 1$. If $P \ne 0$, $\exists\, x$ with $\|x\| = 1$ such that $Px = x$ so $Px \cdot x = x \cdot x = 1$ and $M(P) = 1$. If $P \ne I$, $\exists\, z$ with $\|z\| = 1$ such that $Pz = 0$ so $Pz \cdot z = 0$ and $m(P) = 0$.

We defined a partial order on the class of Hermitian operators by setting $T \le S$ if and only if $Tx \cdot x \le Sx \cdot x \quad \forall\, x \in H$ (35.16). We also defined a partial order on the family of all projections in a B-space in Exercise 27.4. We now show that these two definitions agree for the orthogonal projections.

Proposition 4. If P_1, P_2 are orthogonal projections, then $P_1 \le P_2$ if and only if $P_2P_1 = P_1$. In this case, $P_2P_1 = P_1P_2$ and $P_2 - P_1$ is an orthogonal projection.

Proof: $\Leftarrow$: If $P_2P_1 = P_1$, then

$$P_1P_2x \cdot y = P_2x \cdot P_1y = x \cdot P_2P_1y = x \cdot P_1y \quad \forall\, x, y$$

so $P_1P_2 = P_1$. Thus, $P_2 - P_1$ is an orthogonal projection and $P_2 - P_1 \ge 0$ by Proposition 3.

$\Rightarrow$: If $P_2 \ge P_1$, then $I - P_1 \ge I - P_2$, $I - P_2$ is an orthogonal projection and $(I - P_1)P_1 = 0$ implies

$$0 \le (I - P_2)P_1x \cdot (I - P_2)P_1x = (I - P_2)^2P_1x \cdot P_1x$$

$$= (I - P_2)P_1x \cdot P_1x \le (I - P_1)P_1x \cdot P_1x = 0$$

so

$$(I - P_2)P_1 = 0 \text{ or } P_1 = P_2P_1 .$$

We now write the spectral representation of Theorem 36.2 in an integral form. Let $T \in K(H)$ be compact, Hermitian and let $\{\lambda_k\}$, $\{x_k\}$ be as in 36.2. Define an orthogonal projection $P_k : H \to H$ by $P_kx = (x \cdot x_k)x_k$. For any $x \in H$ the series $\sum_k (x \cdot x_k)x_k = \sum_k P_kx$ is unconditionally convergent so the series $\sum_k P_k$ is unconditionally convergent in the strong operator topology of L(H). If $\mathscr{B}$ is the σ-algebra of Borel sets of $\mathbb{R}$, then $E(\sigma) = \sum_{\lambda_k \in \sigma} P_k$ defines a continuous linear operator on H so $E : \mathscr{B} \to L(H)$. If $x, y \in H$, then $E(\cdot)x \cdot y : \mathscr{B} \to \mathbb{C}$ is a complex-valued measure with mass $(x \cdot x_k)(x_k \cdot y)$ at the eigenvalue λ_k so $Tx \cdot y = \int_{\mathbb{R}} \lambda d(E(\lambda)x \cdot y) = \int_{\sigma(T)} \lambda d(E(\lambda)x \cdot y)$. We write

$$T = \int_{\mathbb{R}} \lambda dE(\lambda) = \int_{\sigma(T)} \lambda dE(\lambda),$$

where the integrals have the meaning above. Writing T in this form suggests that if we want to find a spectral representation for a general Hermitian operator whose spectrum can be any compact subset of $\mathbb{R}$, we should seek a projection-valued measure E as above and represent the operator as an integral with respect to this measure over the spectrum of the operator. This is exactly the Spectral Theorem for Hermitian operators which we are now going to derive.

Exercise 1. Two orthogonal projections P_1, P_2 are said to be <u>perpendicular</u>, $P_1 \perp P_2$ if and only if $P_1P_2 = 0$. Show $P_1 \perp P_2$ if and only if $P_2 \perp P_1$ if and only if $\mathscr{R}P_1 \perp \mathscr{R}P_2$.

Exercise 2. If P_1, P_2 are orthogonal projections, show $P_1 + P_2$ is an orthogonal projection if and only if $P_1 \perp P_2$.

Exercise 3. Let M be a closed subspace of H, P the orthogonal projection onto M and $T \in L(H)$. Show M is invariant under T if and only if $TP = PTP$.

Exercise 4. If P is a non-trivial orthogonal projection, show that $\sigma(P) = \{0, 1\}$.

39

Sesquilinear Functionals

In this section we establish several results on sesquilinear functionals which are needed in the proof of the spectral theorem for Hermitian operators. Let H be a complex Hilbert space.

Definition 1. A function $b : H \times H \to \mathbb{C}$ is a <u>sesquilinear functional</u> if $b(\cdot, y)$ is linear $\forall\, y \in H$ and $\overline{b(x, \cdot)}$ is linear $\forall\, x \in H$.

For example, the inner product on a Hilbert space is sesquilinear. More generally, if $B : H \to H$ is linear, then $b(x, y) = Bx \cdot y$ is sesquilinear.

Definition 2. A sesquilinear functional b is <u>bounded</u> if $\exists\, k > 0$ such that $|b(x, y)| \le k\|x\|\|y\|$ $\forall\, x, y \in H$; the <u>norm</u> of b is defined to be $\|b\| = \sup\{\,|b(x, y)| : \|x\|, \|y\| \le 1\}$. (Compare Exercise 9.1.2.)

Proposition 3. Let $B : H \to H$ be linear and set $b(x, y) = Bx \cdot y$. Then B is bounded if and only if b is bounded. In this case, $\|b\| = \|B\|$.

Proof: $\Rightarrow$: By the Schwarz inequality,

(1) $$|b(x, y)| \le \|Bx\|\|y\| \le \|B\|\|x\|\|y\|, \; x, y \in H.$$

$\Leftarrow$: If $x \in H$, $\|Bx\|^2 = Bx \cdot Bx = b(x, Bx) \le \|b\|\|x\|\|Bx\|$ so

(2) $$\|Bx\| \le \|b\|\|x\|.$$

(1) and (2) show $\|b\| = \|B\|$.

We now show that every bounded sesquilinear functional has the form of the functional in Proposition 3.

Theorem 4. If b is a bounded sesquilinear functional, $\exists$ a unique $B \in L(H)$ such that $b(x, y) = Bx \cdot y \;\forall\, x, y \in H$.

Proof: Fix $x \in H$. Then $\overline{b(x, \cdot)}$ is a bounded linear functional on H so by the Riesz Representation Theorem, there is a unique $Bx \in H$ such that $\overline{b(x, y)} = y \cdot Bx \;\forall\, y \in H$. Since $Bx \cdot y = b(x, y)$, the map $B : x \to Bx$ is linear and B is bounded by Proposition 3. Uniqueness is clear.

Definition 5. A sesquilinear functional b is <u>symmetric</u> if and only if $b(x, y) = \overline{b(y, x)} \;\forall\, x, y \in H$.

Proposition 6. Let $B \in L(H)$ and $b(x, y) = Bx \cdot y$. Then B is Hermitian if and only if b is symmetric.

Proof: $\overline{b(y, x)} = \overline{By \cdot x} = x \cdot By$ for $x, y \in H$.

40

The Gelfand Map for Hermitian Operators

In this section we establish the main result required for our proof of the Spectral Theorem for Hermitian operators. Let H be a complex Hilbert space, and let $T \in L(H)$ be Hermitian.

Theorem 1. If p is a real polynomial, then

$$\|p(T)\| = \sup\{|p(\lambda)| : \lambda \in \sigma(T)\}.$$

Proof: Since p is a real polynomial, p(T) is Hermitian (Exercise 35.9) so $r(p(T)) = \|p(T)\| = \sup\{|\mu| : \mu \in \sigma(p(T))\}$ (35.26). By the Spectral Mapping Theorem (31.13), $\sigma(p(T)) = p(\sigma(T))$ so

$$\sup\{|\mu| : \mu \in \sigma(p(T))\} = \sup\{|\mu| : \mu \in p(\sigma(T))\} = \|p(T)\|.$$

Now let $\mathscr{P} = \{p(T) : p \text{ a real polynomial}\}$ and define a map $\Phi : \mathscr{P} \to C_{\mathbb{R}}(\sigma(T))$, the real-valued continuous functions on $\sigma(T)$, by

$\Phi p(T) = p$. Then Φ satisfies

(1) $$\Phi(tp(T)) = t\Phi(p(T)) \ \forall\, t \in \mathbb{R},\ \Phi(p(T) + q(T)) = \Phi(p(T)) + \Phi(q(T)),\ \Phi(p(T)q(T)) = \Phi(p(T))\Phi(q(T)),$$

and by Theorem 1, Φ is an isometry. Let $\mathcal{A}$ be the norm closure of $\mathcal{P}$ in L(H); thus, $\mathcal{A}$ is the closed, real-subalgebra of L(H) generated by T and I. We may extend Φ uniquely to a real-linear isometry from $\mathcal{A}$ into $C_{\mathbb{R}}(\sigma(T))$ (Exercise 5.3). $\mathcal{A}$ is a closed real-subspace of the complete space L(H) so $\Phi(\mathcal{A}) \subseteq C_{\mathbb{R}}(\sigma(T))$ must also be closed. Since $\Phi(\mathcal{A})$ contains the polynomials in $C_{\mathbb{R}}(\sigma(T))$, $\Phi(\mathcal{A}) = C_{\mathbb{R}}(\sigma(T))$ by the Weierstrass Approximation Theorem. Thus, Φ is a real-linear isometry from $\mathcal{A}$ onto $C_{\mathbb{R}}(\sigma(T))$ which satisfies (1).

The map Φ is called the Gelfand map (for T).

Remark 2. Note that if $S \in L(H)$ commutes with T, then S commutes with the elements of $\mathcal{A}$.

We will now employ the Gelfand map to derive the Spectral Theorem for Hermitian operators. There is likewise a Spectral Theorem for normal operators, but it is much more difficult to define the Gelfand map for normal operators. Note the following difficulties associated with trying to extend the construction above to normal operators.

(1) $\sigma(T)$ is not in general real so the identity polynomial $p(z) = z$ associated with T will be complex-valued. That is, we must use $C_{\mathbb{C}}(\sigma(T))$ instead of $C_{\mathbb{R}}(\sigma(T))$.

(2) We need the analogue of Theorem 1 for normal operators. This is obtainable.

(3) Now the real difficulties appear. $\mathcal{P}$ will not in general be dense in $C_{\mathbb{C}}(\sigma(T))$ since $\mathcal{P}$ is not closed under conjugation (recall the complex form of the Stone-Weierstrass Theorem). Hence, $\mathcal{P}$ must be replaced by the family $p(T, T^*)$, where p is a polynomial in two variables with complex coefficients. The complex form of the Stone-Weierstrass Theorem is then used.

This process is carried out in an elementary manner in [Wh3]. We will, however, give a proof of the Spectral Theorem for normal operators by employing Banach algebra techniques; this will provide us with an introduction to the topic of Banach algebras.

41

The Spectral Theorem for Hermitian Operators

We now use the machinery developed in the previous sections to give a proof of the Spectral Theorem for Hermitian operators. Let H be a complex Hilbert space and $T \in L(H)$ Hermitian. Let $\mathcal{A}$ be the closed real-subalgebra of L(H) generated by T and I and let $\Phi : \mathcal{A} \to C_{\mathbb{R}}(\sigma(T))$ be the Gelfand map of §40. Denote the inverse of the Gelfand map from $C_{\mathbb{R}}(\sigma(T))$ onto $\mathcal{A}$ by $f \to \Psi_f$ (so $p \to \Psi_p = p(T)$ when p is a real polynomial). Let $\mathcal{B}(\sigma(T))$ denote the σ-algebra of Borel sets of $\sigma(T)$.

Theorem 1. For each $\sigma \in \mathcal{B}(\sigma(T))$ $\exists$ a unique orthogonal projection $E(\sigma) \in L(H)$ such that $\mu_{x,y}(\sigma) = E(\sigma)x \cdot y$ $\forall$ x, y $\in$ H, where $\mu_{x,y}$ is the unique regular Borel measure on $\mathcal{B}(\sigma(T))$ satisfying

$$<\mu_{x,y}, f> = \int_{\sigma(T)} f(\lambda)d\mu_{x,y}(\lambda) = \Psi_f x \cdot y$$

for $f \in C_{\mathbb{R}}(\sigma(T))$. The map $E : \mathcal{B}(\sigma(T)) \to L(H)$ satisfies:

(i) $E(\emptyset) = 0,\ E(\sigma(T)) = I$

(ii) $E(\sigma)E(\tau) = E(\sigma \cap \tau) = E(\tau)E(\sigma)\ \forall\ \sigma, \tau \in \mathscr{B}(\sigma(T))$

(iii) $\sum_{j=1}^{\infty} E(\sigma_j)x = E(\bigcup_{j=1}^{\infty} \sigma_j)x \quad \forall\ x \in H$ and pairwise disjoint sequence $\{\sigma_j\} \subseteq \mathscr{B}(\sigma(T))$.

Moreover, each $E(\sigma)$ commutes with every operator in $L(H)$ which commutes with T.

Proof: For $x, y \in H$ define $\mu_{x,y} \in C_{\mathbb{R}}(\sigma(T))'$ by $\langle\mu_{x,y}, f\rangle = \Psi_f x \cdot y$. Note $|\langle\mu_{x,y}, f\rangle| \le \|\Psi_f\|\|x\|\|y\| = \|f\|\|x\|\|y\|$ so $\mu_{x,y} \in C_{\mathbb{R}}(\sigma(T))'$ with $\|\mu_{x,y}\| \le \|x\|\|y\|$. For $\sigma \in \mathscr{B}(\sigma(T))$, set $b_\sigma(x, y) = \mu_{x,y}(\sigma)$. It is easily checked that b_σ is a bounded sesquilinear functional and since Ψ_f is Hermitian, b_σ is symmetric. By 39.4 and 39.6, $\exists$ a Hermitian operator $E(\sigma) \in L(H)$ such that

$$b_\sigma(x, y) = E(\sigma)x \cdot y = \mu_{x,y}(\sigma).$$

Now $E(\emptyset) = 0$ and $E(\sigma(T))x \cdot y = \mu_{x,y}(1) = Ix \cdot y$ so (i) holds.

For (ii), let $f, g \in C_{\mathbb{R}}(\sigma(T))$, $x, y \in H$. Then

$$\Psi_f\Psi_g x \cdot y = \Psi_{fg} x \cdot y = \int_{\sigma(T)} f(\lambda)d(E(\lambda)\Psi_g x \cdot y) = \int_{\sigma(T)} f(\lambda)g(\lambda)d(E(\lambda)x \cdot y)$$

so

(1) $$E(\sigma)\Psi_g x \cdot y = \int_\sigma g(\lambda)d(E(\lambda)x \cdot y).$$

Also $\langle\mu_{\Psi_g x,y}, f\rangle = \Psi_f\Psi_g x \cdot y = \Psi_f x \cdot \Psi_g y = \langle\mu_{x,\Psi_g y}, f\rangle$ so

(2) $$\mu_{\Psi_g x, y}(\sigma) = \mu_{x, \Psi_g y}(\sigma) = E(\sigma)x \cdot \Psi_g y = E(\sigma)\Psi_g x \cdot y$$
$$= \Psi_g E(\sigma)x \cdot y = \int_{\sigma(T)} g(\lambda) d(E(\lambda)E(\sigma)x \cdot y).$$

(1) and (2) imply $E(\tau \cap \sigma) = E(\tau)E(\sigma) \ \forall \ \sigma, \tau \in \mathcal{B}(\sigma(T))$. In particular, $E(\sigma) = E(\sigma \cap \sigma) = E(\sigma)E(\sigma)$ so $E(\sigma)$ is an orthogonal projection.

For (iii), E is clearly finitely additive so let $\sigma_n \downarrow \emptyset$, $\sigma_n \in \mathcal{B}(\sigma(T))$. For $x \in H$, $\|E(\sigma_n)x\|^2 = E(\sigma_n)x \cdot E(\sigma_n)x = E(\sigma_n)x \cdot x = \mu_{x,x}(\sigma_n) \downarrow 0$ so (iii) holds.

Suppose U commutes with T. For $x, y \in H$, $f \in C_{\mathbb{R}}(\sigma(T))$, by Remark 40.2 we have

$$\int_{\sigma(T)} f(\lambda) d(E(\lambda)Ux \cdot y) = \Psi_f Ux \cdot y = U\Psi_f x \cdot y = \Psi_f x \cdot U^* y$$
$$= \int_{\sigma(T)} f(\lambda) d(E(\lambda)x \cdot U^* y) = \int_{\sigma(T)} f(\lambda) d(UE(\lambda)x \cdot y)$$

so $E(\sigma)U = UE(\sigma) \ \forall \ \sigma \in \mathcal{B}(\sigma(T))$.

If Ω is a locally compact Hausdorff space and $\mathcal{B}(\Omega)$ is the σ-algebra of Borel sets of Ω, a set function $E : \mathcal{B}(\Omega) \to L(H)$ is called a <u>spectral measure</u> (in Ω) if

(i) each $E(\sigma)$, $\sigma \in \mathcal{B}(\Omega)$, is an orthogonal projection,

(ii) $E(\emptyset) = 0$, $E(\Omega) = I$,

(iii) $E(\sigma \cap \tau) = E(\sigma)E(\tau) \ \forall \ \sigma, \tau \in \mathcal{B}(\Omega)$,

(iv) $\forall \ x \in H$ and pairwise disjoint sequence $\{\sigma_j\} \subseteq \mathcal{B}(\Omega)$, $\sum_{j=1}^{\infty} E(\sigma_j)x = E(\bigcup_{j=1}^{\infty} \sigma_j)x$ (i.e., E is countably additive in the strong operator topology of $L(H)$),

(v) $\forall\, x, y \in H$ the map $\sigma \to E(\sigma)x \cdot y$ from $\mathscr{B}(\Omega)$ to $\mathbb{C}$ is a regular Borel measure.

Note that (v) merely postulates regularity since the measure in (v) is countably additive by (iv).

From Theorem 1, any Hermitian operator induces a spectral measure E on $\sigma(T)$ such that $\int_{\sigma(T)} p(\lambda)d(E(\lambda)x \cdot y) = p(T)x \cdot y$ for any real polynomial p and $x, y \in H$. In particular, $Tx \cdot y = \int_{\sigma(T)} \lambda d(E(\lambda)x \cdot y)$, and we write $T = \int_{\sigma(T)} \lambda dE(\lambda)$. It is easily checked that this spectral measure is unique (Exercise 1). This spectral measure is called the <u>resolution of the identity</u> for T.

The proof of the spectral theorem for Hermitian operators given above uses the representation of the dual space of $C(\sigma(T))$. There are two other proofs of the spectral theorem due to Riesz and Nagy-which do not employ the dual space. The proof of Riesz uses Theorem 35.32 to obtain an operational calculus from which the projections in the resolution of the identity are constructed. The Hermitian operator is then represented as a Riemann-Stieltjes integral. See [RN] §106, 107 for details. The proof of Nagy uses the existence of the square root of a positive operator (see Theorem 3 in this section); the operator is again represented as a Riemann-Stieltjes integral. See [RN] §108 for details. The representation as a Riemann-Stieltjes integral can also be obtained by using the representation of the dual of $C[a, b]$ as the space of normalized functions of bounded variation; see [TL], VI.6, for details.

The inverse of the Gelfand maps:

For $f \in C_{\mathbb{R}}(\sigma(T))$, define a linear operator $f(T) \in L(H)$ by $f(T)x \cdot y = \int_{\sigma(T)} f(\lambda) d(E(\lambda)x \cdot y)$ [note

(3) $|f(T)x \cdot y| \le \|f\|_\infty |E(\cdot)x \cdot y|(\sigma(T)) \le \|f\|_\infty \|x\| \|y\|$

so $f(T) \in L(H)$ by 39.4]. We write $f(T) = \int_{\sigma(T)} f(\lambda) dE(\lambda)$ for this operator. If p is a real polynomial in $C_{\mathbb{R}}(\sigma(T))$, then $p(T) = \Psi_p$ and if $f \in C_{\mathbb{R}}(\sigma(T))$ and $\{p_k\}$ is a sequence of real polynomials such that $p_k \to f$ in $C_{\mathbb{R}}(\sigma(T))$, then $p_k(T) \to T$ in the norm of $L(H)$ by (3) so $f(T) = \Psi_f$ and $\|f(T)\| = \|\Psi_f\| = \|f\|$. This gives an integral representation for the inverse of the Gelfand map in terms of the spectral measure of T. The map $f \to f(T)$ from $C_{\mathbb{R}}(\sigma(T))$ into $L(H)$ is often called an <u>operational calculus</u> associated with T. We use this operational calculus to show the existence of square roots for positive operators.

Lemma 2. Let $f, g \in C_{\mathbb{R}}(\sigma(T))$ and set $A = f(T)$, $B = g(T)$. Then

$$ABx \cdot y = \int_{\sigma(T)} f(\lambda)g(\lambda) d(E(\lambda)x \cdot y) = BAx \cdot y$$

(i.e., $\int f dE \int g dE = \int fg dE$, the product of the integrals is the integral of the product!)

Proof: For $\sigma \in \mathcal{B}(\sigma(T))$, $x, y \in H$, set

$$\mu(\sigma) = E(\sigma)Bx\cdot y = Bx\cdot E(\sigma)y = \int_{\sigma(T)} g(\lambda)d(E(\lambda)x\cdot E(\sigma)y)$$
$$= \int_{\sigma(T)} g(\lambda)d(E(\sigma\cap\lambda)x\cdot y) = \int_{\sigma} g(\lambda)d(E(\lambda)x\cdot y).$$

Hence,

$$ABx\cdot y = \int_{\sigma(T)} f(\lambda)d(E(\lambda)Bx\cdot y) = \int_{\sigma(T)} f(\lambda)d\mu(\lambda)$$
$$= \int_{\sigma(T)} f(\lambda)g(\lambda)d(E(\lambda)x\cdot y).$$

Theorem 3. If $T \in L(H)$ is a positive operator, then T has a unique positive square root denoted by $\sqrt{T}$. Moreover, if $S \in L(H)$ commutes with T, then S commutes with $\sqrt{T}$.

Proof: Since $T \geq 0$, $\sigma(T) \subseteq \mathbb{R}_+$ (35.28) and the function $\lambda \to \sqrt{\lambda}$ is continuous on $\sigma(T)$. Set $\sqrt{T} = \int_{\sigma(T)} \sqrt{\lambda}\, dE(\lambda)$. By Lemma 2, $(\sqrt{T})^2 = T$.

It follows from Theorem 1 that $T \geq 0$, and if S commutes with T then S commutes with $\sqrt{T}$.

For uniqueness, suppose $C \geq 0$ and $C^2 = T$. Set $B = \sqrt{T}$. Then

(4) $$\|\sqrt{C}\, x\|^2 + \|\sqrt{B}\, x\|^2 = Cx\cdot x + Bx\cdot x = (C+B)x\cdot x \quad \forall\, x \in H.$$

Since $TC = C^2C = CC^2 = CT$, C commutes with B by the part above.

For $y \in H$ put $x = (B - C)y$ in (4). Then

$$\|\sqrt{C}\, x\|^2 + \|\sqrt{B}\, x\|^2 = (C+B)x\cdot x = (C+B)(B-C)y\cdot x = (B^2 - C^2)y\cdot x = 0$$

so $\sqrt{C}\, x = 0$, $\sqrt{B}\, x = 0$. Hence, $Cx = \sqrt{C}\sqrt{C}\, x = 0$ and $Bx = 0$ and

$$\|(B-C)y\|^2 = (B-C)^2 y\cdot y = (B-C)x\cdot y = 0$$

so $(B - C)y = 0 \ \forall y \in H$ and $B = C$.

A proof that does not use the operational calculus can be found in [BN].

As another application of the operational calculus, we give an extension of the Spectral Mapping Theorem of 31.13 from polynomials to continuous functions.

Theorem 4. If $f \in C_{\mathbb{R}}(\sigma(T))$, then $f(\sigma(T)) = \sigma(f(T))$.

Proof: Suppose $\mu \notin f(\sigma(T))$. Then $f - \mu$ is non-zero on $\sigma(T)$ so $g = 1/(f - \mu)$ is continuous. By Lemma 2, $(f(T) - \mu)g(T) = \int_{\sigma(T)} dE(\lambda) = I$ so $\mu \notin \sigma(f(T))$. Hence, $\sigma(f(T)) \subseteq f(\sigma(T))$. Next we show that $\lambda \in \sigma(T)$ implies that $f(T) - f(\lambda)I$ is not invertible. Choose a sequence of polynomials $\{p_n\}$ converging uniformly to f on $\sigma(T)$. Then $p_n(T) - p_n(\lambda)I \to f(T) - f(\lambda)I$ in norm. Since $p_n(\lambda) \in \sigma(p_n(T))$ (31.13), $p_n(T) - p_n(\lambda)I$ is not invertible so $f(T) - f(\lambda)I$ is not invertible (23.13).

Another problem associated with Hermitian operators is the simultaneous representation of a family of Hermitian operators by means of a single spectral measure. We give a simple example of such a representation.

Definition 5. Let X be a B-space. A family $\{T_t : t \geq 0\} \subseteq L(X)$ is said to be a <u>strongly</u> (<u>uniformly</u>) <u>continuous semi-group</u> if

(i) $T_{s+t} = T_s T_t \ \forall\, s, t \geq 0$

(ii) $T_0 = I$

(iii) the map $t \to T_t$ is continuous from $[0, \infty)$ into $L(X)$ with respect to the strong operator topology (uniform operator topology) of $L(X)$.

Example 6. Let $A \in L(X)$. Then $T_t = \exp(tA)$ defines a uniformly continuous semi-group. [The converse also holds, [DS].]

Example 7. Let $X = L^2(\mathbb{R})$ (with Lebesgue measure). For $f \in X$ and $t \in \mathbb{R}$, $\tau_t f(s) = f(s + t)$. Then $T_t f = \tau_t f$ defines a strongly continuous semi-group which is not uniformly continuous.

For strongly continuous semi-groups of Hermitian operators, we have the following interesting representation theorem.

Theorem 8. Let $\{T_t : t \geq 0\}$ be a strongly continuous semi-group of Hermitian operators in $L(H)$. Then $\exists$ a compact subset $K \subseteq [0, \infty)$ and a spectral measure $E : \mathscr{B}(K) \to L(H)$ with $T_t x \cdot y = \int_K \lambda^t d(E(\lambda)x, y) \ \forall\, x, y \in H$.

Proof: Since $T_t = (T_{t/2})^2$, each T_t is a positive operator so the spectrum of T_1 is a compact subset K of $[0, \infty)$. Let E be spectral resolution of T_1. Define a strongly continuous semi-group of Hermitian operators, $\{B_t : t \geq 0\}$ by $B_t x \cdot y = \int_K \lambda^t d(E(\lambda)x \cdot y)$ (Exercise 5). Now $T_1 = B_1$ and, moreover, $T_t = B_t$ for $t = 1/2, \ldots, 1/2^n, \ldots$ since by Theorem 3 each positive operator has a unique square root. By (i), $T_t = B_t$ for $t = k/2^n$, $k, n \in \mathbb{N}$. By the continuity of the maps $t \to T_t$, $t \to B_t$, it follows that $T_t = B_t$ for all $t \geq 0$.

For further such simultaneous representations, see [RN] §130. The theory of semi-groups is an important topic with many applications particularly in differential equations. For further information, see [DS] and [HP].

Exercise 1. Show the spectral measure in Theorem 1 is unique.

Exercise 2. With notation as in Theorem 1 show

$$\|\Psi_f x\|^2 = \int_{\sigma(T)} |f(\lambda)|^2 d(E(\lambda)x \cdot x) .$$

Exercise 3. Let $T : H = L^2[0, 1] \to H$ be the multiplication operator $Tf(t) = tf(t)$. Show that $E(\sigma)f = C_\sigma f$ is the resolution of the identity for T.

Exercise 4. If $T \geq 0$ is invertible, show $\sqrt{T}$ is invertible.

Exercise 5. Show the family $\{B_t : t \geq 0\}$ in Theorem 7 defines a strongly continuous semi-group.

Exercise 6. If T is a compact, Hermitian operator, is $f(T)$ compact for every continuous function f on $\sigma(T)$?

Exercise 7. If T is Hermitian, show that T is the difference of two positive operators. [Hint: Define $|T| = \sqrt{T^2}$.]

42

Banach Algebras

We are going to prove the spectral theorem for normal operators by establishing the analogue of the Gelfand map of §40 for normal operators. We will do this by using Banach algebras. This will allow us to present an elementary introduction to the important topic of Banach algebras. The topic of Banach algebras is a subject in its own right, and we present only a very brief introduction to this subject. For further studies the reader can consult ([DS2]).

Definition 1. A (complex) Banach algebra (B-algebra) is a complex B-space X which is an algebra such that $\|xy\| \le \|x\|\|y\|$ $\forall$ $x, y \in X$. If X has an identity e, it is further assumed that $\|e\| = 1$ (this can always be obtained by renorming).

Example 2. Let S be compact, Hausdorff. $C_{\mathbb{C}}(S)$ is a B-algebra under pointwise multiplication. Moreover, $C_{\mathbb{C}}(S)$ is commutative and has an identity.

Example 3. As in Example 2 $L^{\infty}(S, \Sigma, \mu)$ and $B(S)$ are both commutative B-algebras and have identities.

Example 4. Let X be a B-space. Then $L(X)$ is a B-algebra with identity but is not commutative.

Example 5. Let $L^1(\mathbb{R}^n)$ be the Lebesgue integrable functions with the L^1-norm. Then $L^1(\mathbb{R}^n)$ is a commutative B-algebra under the convolution product, $f*g(t) = \int_{\mathbb{R}^n} f(t - s)g(s)ds$, which has no identity.

Throughout the remainder of this section we assume that X is a B-algebra with identity, e. An element $x \in X$ is <u>invertible</u> if it has an inverse $x^{-1} \in X$; otherwise x is called <u>singular</u>. We now establish the analogue of 23.13 for B-algebras.

Theorem 6. The multiplicative group $\mathcal{G}$ of invertible elements of X is open in X and the map $x \to x^{-1}$ is a homeomorphism from $\mathcal{G}$ onto itself.

Proof: First we show that $\mathcal{G}$ contains the sphere $\{x : \|x - e\| < 1\}$. If $\|e - x\| < 1$, the series $\sum_{k=0}^{\infty} (e - x)^k$ (where $a^0 = e$) is absolutely convergent and, hence, convergent. Let $y = \sum_{k=0}^{\infty} (e - x)^k$. Then

$$xy = yx = y - (e - x)y = \sum_{k=0}^{\infty} (e - x)^k - \sum_{k=0}^{\infty} (e - x)^{k+1} = e \text{ so } y = x^{-1} \text{ and}$$

$$\|x^{-1} - e\| = \|\sum_{k=1}^{\infty} (e - x)^k\| \le \sum_{k=1}^{\infty} \|e - x\|^k = \|e - x\|/(1 - \|e - x\|).$$

Thus, $\mathcal{G}$ contains a neighborhood of e and $x \to x^{-1}$ is a continuous function at e. Now let $x \in \mathcal{G}$ and $\|y - x\| < 1/\|x^{-1}\|$. Then $\|x^{-1}y - e\| = \|x^{-1}(y - x)\| \le \|x^{-1}\|\|y - x\| < 1$ so $x^{-1}y \in \mathcal{G}$ by the above and $y \in \mathcal{G}$. Hence, $\mathcal{G}$ is open.

If $y_k \to y$, where $y_k, y \in \mathcal{G}$, then $y_k y^{-1} \to e$ so

$$(y_k y^{-1})^{-1} = y y_k^{-1} \to e$$

by the above so $y_k^{-1} \to y^{-1}$ and the map $y \to y^{-1}$ is continuous on $\mathcal{G}$.

Using the analogue of 31.2, we define the spectrum of an element of X. The <u>spectrum</u> of an element $x \in X$ is $\sigma(x) = \{\lambda \in \mathbb{C} : \lambda e - x$ is singular$\}$, the <u>spectral radius</u> of x is $r(x) = \sup\{|\lambda| : \lambda \in \sigma(x)\}$, and the <u>resolvent set</u> of x is $\rho(x) = \mathbb{C}\backslash\sigma(x)$. The <u>resolvent</u> of x is the map $\lambda \to R_\lambda(x) = (\lambda e - x)^{-1}$ from $\rho(x)$ into X. We now establish the analogue of 31.5, 7, 8 and 9.

Theorem 7. For $x \in X$, $\sigma(x)$ is non-void and compact. The resolvent $\lambda \to R_\lambda(x)$ is an X-valued analytic function on $\rho(x)$ which vanishes at ∞ and satisfies the <u>resolvent equation</u> $R_\lambda(x) - R_\mu(x) = (\mu - \lambda)R_\lambda(x)R_\mu(x)$.

Proof: $\sigma(x)$ is closed by Theorem 6 and Theorem 6 also implies that for large λ, $(\lambda e - x) = \lambda(e - x/\lambda) \in \mathcal{G}$ so $\sigma(x)$ is bounded. As $|\lambda| \to \infty$, $e - x/\lambda \to e$ so by Theorem 6

$$R_\lambda(x) = (\lambda e - x)^{-1} = 1/\lambda(e - x/\lambda)^{-1} \to 0$$

as $|\lambda| \to \infty$. Hence $R_\lambda(x)$ vanishes at ∞. For $\lambda, \mu \in \rho(x)$, $R_\lambda(x)$ and $R_\mu(x)$ commute and

$$(\lambda e - x)R_\lambda(x)R_\mu(x) = R_\mu(x),$$

$$(\mu e - x)R_\mu(x)R_\lambda(x) = R_\lambda(x) \text{ so } R_\mu(x) - R_\lambda(x) = (\lambda - \mu)R_\mu(x)R_\lambda(x),$$

and

$$(R_\mu(x) - R_\lambda(x))/(\mu - \lambda) = -R_\mu(x)R_\lambda(x) .$$

Letting $\mu \to \lambda$ and using Theorem 6 implies that $R_\lambda(x)$ is analytic as a function of λ with $\frac{d}{d\lambda} R_\lambda(x) = -R_\lambda(x)^2$.

Finally, if $\sigma(x) = \emptyset$, let $x' \in X'$. Then $\langle x', R_\lambda(x) \rangle$ is an entire function of λ which vanishes at ∞ and, therefore, must be 0. Hence, $R_\lambda(x) = 0 \quad \forall \lambda \in \mathbb{C}$ so $(\lambda e - x)R_\lambda(x) = 0 = e$ which is impossible. Therefore, $\sigma(x) \neq \emptyset$.

The proof of Theorem 7 should be compared with the proofs in §31.

An element $x \in X$ is a <u>right</u> (<u>left</u>) <u>topological divisor</u> of 0 if $\exists$ a sequence $\{x_k\} \subseteq X$ with $\|x_k\| = 1$ and $x_k x \to 0$ $(x x_k \to 0)$; a <u>two-sided topological divisor of 0</u> is an element which is both a right and left topological divisor of 0.

A right (left) topological divisor of 0, x, cannot be invertible since $xy = e$ and $x_k x \to 0$ implies $x_k = x_k xy \to 0$ with $\|x_k\| \to 1$.

Example 8. The element $x(t) = t$ in $C[0, 1]$ is a two-sided topological divisor of 0. For example, define $x_k \in C[0, 1]$ by $x_k(t) = 0$ if $t > 1/k$ and $x_k(t) = -kt + 1$ for $0 \le t < 1/k$. Then $x_k x \to 0$.

Topological divisors of 0 are useful in describing the boundary of $\mathcal{G}$.

If E is a subset of a topological space, let ∂E denote its boundary.

Theorem 9. Every boundary point of $\mathcal{G}$, the group of invertible elements, is a two-sided topological divisor of 0.

Proof: Let $x \in \partial\mathcal{G}$. Then $x \notin \mathcal{G}$ by Theorem 6. Pick $x_k \in \mathcal{G}$ such that $x_k \to x$. Since $x \notin \mathcal{G}$, $x_k^{-1}x \notin \mathcal{G}$ so by the proof of Theorem 6 $1 \le \|x_k^{-1}x - e\| \le \|x_k^{-1}\|\|x - x_k\|$ which implies that $\|x_k^{-1}\| \to \infty$. Put $z_k = x_k^{-1}/\|x_k^{-1}\|$. Then $\|z_k\| = 1$ and $xz_k = (x - x_k)z_k + e/\|x_k^{-1}\| \to 0$ and similarly $z_k x \to 0$ so x is a two-sided topological divisor of 0.

Theorem 10 (Gelfand-Mazur). If X has no non-zero two-sided topological divisors of 0, then X is isometrically isomorphic to $\mathbb{C}$.

Proof: Let $x \in X$. Then $\sigma(x)$ is non-void and compact so $\exists$ a point λ belonging to the boundary of $\sigma(x)$. By Theorem 9, $\lambda e - x$ is a two-sided topological divisor of 0 so $\lambda e - x = 0$ or $x = \lambda e$. Thus, the isometric isomorphism $\lambda \to \lambda e$ from $\mathbb{C}$ into X is actually onto X.

Note that it is part of the conclusion of Theorem 10 that X is commutative!

Corollary 11. If every non-zero element of X is invertible, then X is isometrically isomorphic to $\mathbb{C}$.

Corollary 11 is the original version of the Gelfand-Mazur Theorem.

If X_0 is a subalgebra of X, then an element $x \in X_0$ has a spectrum as an element of X_0 and also as an element of X; we denote these by $\sigma_{X_0}(x)$ and $\sigma_X(x)$, respectively. Of course, we have $\sigma_{X_0}(x) \supseteq \sigma_X(x)$. We first consider an example to illustrate what can occur.

Example 12. Let $D = \{z \in \mathbb{C} : |z| < 1\}$ and $X = C(\partial D)$. Let X_0 be the subalgebra of X which consists of the restrictions to ∂D of those functions f which are analytic in D and continuous on $\bar{D}$. Consider the function $f(z) = z$ belonging to X_0. Then $\sigma_X(f) = \partial D$ while $\sigma_{X_0}(f) = \bar{D}$. Thus, the spectrum of f in X_0 "lost" points when X_0 was enlarged to X but none of the boundary points of $\sigma_{X_0}(f)$ was "lost" during this enlargement. We show that this is a general phenomena.

Proposition 13. Let X_0 be a Banach subalgebra of X with the same identity and let $z \in X_0$. Then $\sigma_X(z) \subseteq \sigma_{X_0}(z)$ and $\partial\sigma_{X_0}(z) \subseteq \partial\sigma_X(z)$.

Proof: An element which is invertible in X_0 is invertible in X so the first inclusion is clear.

If $\lambda \in \partial\sigma_{X_0}(z)$, then $\lambda e - z$ is on the boundary of the group of invertible elements of X_0 so by Theorem 9, $\lambda e - z$ is a two-sided topological divisor of 0 in X_0 and, therefore, in X. Thus, $\lambda \in \sigma_X(z)$ and since

$$\rho_{X_0}(z) \subseteq \rho_X(z), \overline{\rho_{X_0}(z)} \cap \sigma_{X_0}(z) = \partial\sigma_{X_0}(z) \subseteq \overline{\rho_X(z)} \cap \sigma_X(z) = \partial\sigma_X(z).$$

Thus, as in Example 12, the spectrum of an element shrinks when its containing B-algebra is enlarged, but its boundary points are not lost in the process; the spectrum shrinks by "hollowing out".

For later use, we give criteria which guarantee that the spectrum does not change when the containing B-algebra is enlarged.

Proposition 14. Let X_0 be a Banach subalgebra of X with the same identity and let $z \in X_0$. Then

(i) If $\sigma_{X_0}(z)$ is nowhere dense, then $\sigma_X(z) = \sigma_{X_0}(z)$.

(ii) If $\rho_X(z)$ is connected, then $\sigma_X(z) = \sigma_{X_0}(z)$.

Proof: (i): $\sigma_{X_0}(z)$ is closed and nowhere dense so

$$\sigma_{X_0}(z) = \partial\sigma_{X_0}(z) \subseteq \partial\sigma_X(z) \subseteq \sigma_X(z) \subseteq \sigma_{X_0}(z)$$

by Proposition 13.

(ii): If $\rho_X(z)$ is connected and if $\lambda \in \sigma_{X_0}(z) \cap \rho_X(z)$, then λ may be connected to ∞ by a continuous path contained entirely in $\rho_X(z)$. In this case $\exists$ a boundary point of $\sigma_{X_0}(z)$ which is in $\rho_X(z)$ and this contradicts Proposition 13. Hence, $\sigma_{X_0}(z) \cap \rho_X(z) = \emptyset$ and

$$\sigma_{X_0}(z) \subseteq \sigma_X(z) \subseteq \sigma_{X_0}(z)$$

by Proposition 13.

Remark 15. Note that if $\sigma_X(z) \subseteq \mathbb{R}$, then (ii) holds.

For more precise statements describing $\sigma_X(z)$ and $\sigma_{X_0}(z)$ see [C1] p. 210.

Spectral Radius:

We now establish the analogue of the formula for the spectral radius which was established for operators in 31.12.

Theorem 16. If $x \in X$, then $r(x) = \lim \sqrt[n]{\|x^n\|} \le \|x\|$.

Proof: For $|\lambda| > \|x\|$ the series $\sum_{k=0}^{\infty} x^k/\lambda^{k+1}$ converges and since

$$(\lambda e - x) \sum_{k=0}^{\infty} x^k/\lambda^{k+1} = \sum_{k=0}^{\infty} (x^k/\lambda^k - x^{k+1}/\lambda^{k+1}) = e\,,\ R_\lambda(x) = \sum_{k=0}^{\infty} x^k/\lambda^{k+1}$$

for $|\lambda| > \|x\|$. By Theorem 7, $R_\lambda(x)$ is analytic on $\rho(x)$ so

$\langle x', R_\lambda(x)\rangle$ is analytic in $|\lambda| > r(x)$ $\forall x' \in X'$. Hence, the series $\langle x', R_\lambda(x)\rangle = \sum_{k=0}^{\infty} \langle x', x^k\rangle/\lambda^{k+1}$ converges for $|\lambda| > r(x)$ and for such λ, $\sup\{|\langle x', x^k\rangle/\lambda^{k+1}| : k\} < \infty$. By the UBP, $\sup\{\|x^k/\lambda^{k+1}\| : k\} < \infty$ so $\overline{\lim} \sqrt[k]{\|x^k\|} \le |\lambda|$ for all $|\lambda| > r(x)$. Hence, $\overline{\lim} \sqrt[k]{\|x^k\|} \le r(x)$.

Since $\lambda e - x$ is a factor of $\lambda^n e - x^n$, $\lambda^n e - x^n$ is singular if $\lambda e - x$ is singular. Thus, if $\lambda \in \sigma(x)$, $\lambda^n \in \sigma(x^n)$ and $|\lambda^n| \le \|x^n\|$ implies that $|\lambda| \le \underline{\lim} \sqrt[n]{\|x^n\|}$. Hence, $r(x) \le \underline{\lim} \sqrt[n]{\|x^n\|}$.

Finally, we establish a result on complex-valued homomorphisms on X which will be used later.

Proposition 17. Let $\varphi : X \to \mathbb{C}$ be a homomorphism with $\varphi \neq 0$. Then

(i) $|\varphi(x)| < 1$ for $\|x\| < 1$,

(ii) φ is continuous and $\|\varphi\| = 1$.

Proof: (i): Let $\lambda \in \mathbb{C}$, $|\lambda| \ge 1$. Then $e - x/\lambda$ is invertible (Theorem 6) so

$$\varphi(e - x/\lambda) = 1 - \varphi(x)/\lambda \neq 0.$$

Hence, $\varphi(x) \neq \lambda$.

(ii) follows from (i) and $\varphi(e) = 1$.

Exercise 1. If $f \in C(S)$, show $\sigma(f) = \mathcal{R} f$.

Exercise 2. If $f \in C(S)$, show that either f is invertible or a topological divisor of 0.

Exercise 3. If $z \in X$, then $L_z : x \to zx$ defines an operator $L_z \in L(X)$ such that $\|L_z\| = \|z\|$. Show $T \in L(X)$ is such that $T = L_z$ for some $z \in X$ if and only if $(Tx)y = T(xy)$ $\forall$ $x, y \in X$.

Exercise 4. In 3, show the spectrum of L_z in $L(X)$ is the same as the spectrum of z in X.

Exercise 5. State and prove the analogue of the Spectral Mapping Theorem (31.13) for Banach algebras.

Exercise 6. Show the set of all right (left) topological divisors of 0 forms a closed subset of X.

Exercise 7. An element $x \in X$ is a <u>projection</u> if $x^2 = x$. If x is a projection which is not 0 or e, show $0, 1 \in \sigma(x)$. What are the projections in $C(S)$?

Exercise 8. Show $r(xy) \leq r(x)r(y)$.

Exercise 9. Show $C[0, 1]$ is a commutative B-algebra without an identity if the product of two functions x and y is defined by

$$x*y(t) = \int_0^t x(t - s)y(s)ds.$$

Exercise 10. Let X be an algebra which is a B-space. If the multiplication in X is separately continuous, show the multiplication is continuous.

43

Commutative Banach Algebras

In this section we begin to establish an abstract version of the Gelfand map of §40 for commutative Banach algebras.

Let X be a commutative B-algebra with identity e.

Definition 1. A subset $\mathcal{J} \subseteq X$ is an ideal if $\mathcal{J}$ is a linear subspace such that $\mathcal{J}X = X\mathcal{J} \subseteq \mathcal{J}$. If $\mathcal{J} \neq \{0\}$ and $\mathcal{J} \neq X$, then $\mathcal{J}$ is called a proper ideal. An ideal $\mathcal{J}$ is maximal if it is a proper ideal and any proper ideal containing $\mathcal{J}$ must equal $\mathcal{J}$.

Lemma 2. (i) If $\mathcal{J}$ is a proper ideal, then $\mathcal{J}$ contains no invertible elements.

(ii) The closure of an ideal is an ideal.

(iii) Every proper ideal is contained in a maximal (proper) ideal.

(iv) Every maximal ideal is closed.

(v) $z \in X$ is singular if and only if z belongs to a maximal ideal.

Proof: (i) and (ii) are clear and (iii) follows from standard Zorn's lemma arguments. (iv) follows from (ii) since if M is a maximal ideal, then $\overline{M}$ cannot contain elements of $\mathcal{G}$, the invertible elements, since $\mathcal{G}$ is open and $M \cap \mathcal{G} = \emptyset$.

For (v), assume that z is singular. Then zX is a proper ideal and is contained in a maximal ideal by (iii). The converse is clear.

Lemma 3. If $\mathcal{J}$ is a closed ideal, then $X/\mathcal{J}$ is a Banach algebra.

Proof: For $x \in X$ let $[x] = x + \mathcal{J}$ be the coset determined by x. We have

$$\|[x][y]\| = \|[xy]\| = \inf\{\|xy + z\| : z \in \mathcal{J}\}$$

$$\leq \inf\{\|x + u\|\|y + v\| : u, v \in \mathcal{J}\} = \|[x]\|\|[y]\|$$

and

$$\|[e]\| = \|[e]^2\| \leq \|[e]\|^2$$

which implies $\|[e]\| \geq 1$ since $\|[e]\| \neq 0$. But $\|[e]\| \leq \|e\| = 1$ so $\|[e]\| = 1$. Since $X/\mathcal{J}$ is a B-space, this completes the proof.

Let $\Delta = \Delta_X$ be the set of all non-zero complex homomorphisms on X. We study the relationship between the elements of Δ and the maximal ideals in X.

Theorem 4. (i) Every maximal ideal is the kernel of some $h \in \Delta$.

(ii) If $h \in \Delta$, then $\mathcal{N}(h)$ is a maximal ideal of X.

(iii) $x \in X$ is invertible if and only if $\langle h, x\rangle \neq 0 \ \forall h \in \Delta$.

(iv) $\lambda \in \sigma(x)$ if and only if $\langle h, x\rangle = \lambda$ for some $h \in \Delta$.

Proof: (i): Let M be a maximal ideal. Then M is closed and X/M is a B-algebra. Choose $x \in X \setminus M$ and set

$$J = \{ax + y : a \in X, y \in M\}.$$

Then J is an ideal which properly contains M since $x \in J$. Hence, $X = J$ and $ax + y = e$ for some $a \in X$, $y \in M$. If $\varphi : X \to X/M$ is the quotient map $\varphi(x)\varphi(a) = \varphi(e)$ so every non-zero element $\varphi(x) \in X/M$ is invertible. By the Gelfand-Mazur Theorem (42.11), $\exists$ an isomorphism $j : X/M \to \mathbb{C}$. Set $h = j\varphi$. Then $h \in \Delta$ and $M = \mathcal{N}(h)$.

(ii): If $h \in \Delta$, $\mathcal{N}(h)$ is an ideal which must be maximal since it has co-dimension 1.

(iii): If x is invertible and $h \in \Delta$, then $h(x)h(x^{-1}) = 1$ so $h(x) \neq 0$. On the other hand, if x is singular, then xX doesn't contain e and is a proper ideal which is contained in some maximal ideal M. By (i), x is annihilated by some $h \in \Delta$.

(iv) follows from applying (iii) to $\lambda e - x$.

The formula $\hat{x}(h) = \langle h, x\rangle$ defines a map $\hat{x} : \Delta \to \mathbb{C}$ for each $x \in X$. The map $\hat{x}$ is called the <u>Gelfand transform of x</u> and the map $x \to \hat{x}$ is called the <u>Gelfand transform</u> (<u>Gelfand representation</u>). Let $\hat{X} = \{\hat{x} : x \in X\}$. The <u>Gelfand topology</u> of Δ is the weakest topology on

Δ such that every element of $\hat{X}$ is continuous. Thus, $\hat{X} \subseteq C(\Delta)$ and the Gelfand transform $x \to \hat{x}$ is a map from X into $C(\Delta)$.

By Theorem 4 there is a 1-1 correspondence between the maximal ideals of X and the elements of Δ so Δ equipped with the Gelfand topology is often called the maximal ideal space of X. We now consider Δ and the Gelfand transform.

The radical of X, rad X, is the intersection of all the maximal ideals of X; X is semi-simple if and only if rad $X = \{0\}$.

Theorem 5.

(i) Δ is a compact Hausdorff space under the Gelfand topology.

(ii) The Gelfand transform $x \to \hat{x}$ from X into $C(\Delta)$ is a homomorphism of X onto a subalgebra $\hat{X}$ of $C(\Delta)$ whose kernel is rad X. Therefore, the Gelfand transform is an isomorphism if and only if X is semi-simple.

(iii) For each $x \in X$, the range of $\hat{x}$ is $\sigma(x)$. Thus, $\|\hat{x}\|_\infty = r(x) \le \|x\|$ and $x \in$ rad X if and only if $r(x) = 0$.

(iv) X is semi-simple if and only if Δ separates the points of X.

Proof: (i): Let B be the closed unit ball of X'; B is compact in the weak* topology by the Banach-Alaoglu Theorem. Note that $\Delta \subseteq B$ (42.17) and the Gelfand topology of Δ is just the relative weak* topology from B so it suffices to show that Δ is weak* closed. Let $h \in B$ be in the weak* closure of Δ, and let $\{h_\delta\}$ be a net in Δ which is weak* convergent to h. Then $\langle h_\delta, e\rangle = 1 \to \langle h, e\rangle$ so $h \neq 0$. Since

$\langle h_\delta, xy\rangle = \langle h_\delta, x\rangle\langle h_\delta, y\rangle \to \langle h, xy\rangle = \langle h, x\rangle\langle h, y\rangle$, $h \in \Delta$.

(ii): Let $x, y \in X$, $t \in \mathbb{C}$, $h \in \Delta$. Then

$$(tx)^\wedge(h) = \langle h, tx\rangle = t\langle h, x\rangle = t\hat{x}(h),$$

$$(x + y)^\wedge(h) = \langle h, x + y\rangle = \langle h, x\rangle + \langle h, y\rangle = \hat{x}(h) + \hat{y}(h),$$

$$(xy)^\wedge(h) = \langle h, xy\rangle = \langle h, x\rangle\langle h, y\rangle = \hat{x}(h)\hat{y}(h)$$

so $x \to \hat{x}$ is a homomorphism. Its kernel consists of those x such that $\langle h, x\rangle = 0 \ \ \forall h \in \Delta$; by Theorem 4 this is the intersection of all maximal ideals, rad X.

(iii): λ is in the range of $\hat{x}$ if and only if $\hat{x}(h) = \langle h, x\rangle = \lambda$ for some $h \in \Delta$; by Theorem 4 this happens if and only if $\lambda \in \sigma(x)$.

(iv): $\Rightarrow$: If $x \neq 0$, $r(x) = \|\hat{x}\|_\infty \neq 0$ by (iii) so $\exists\, h \in \Delta$ such that $\langle h, x\rangle \neq 0$.

$\Leftarrow$: If $x \in X$, $x \neq 0$, $\exists\, h \in \Delta$ such that $\langle h, x\rangle \neq 0$ so by (iii), $r(x) \neq 0$ and $x \notin \text{rad } X$.

An interesting application of (iv) is given in Exercise 4, where the reader is asked, in particular, to show that the norm on a commutative B-algebra is essentially unique.

We next consider when the Gelfand transform is an isometry. For this we require a preliminary result.

Lemma 6. If $r = \inf\{\|x^2\|/\|x\|^2 : x \neq 0\}$ and $s = \inf\{\|\hat{x}\|_\infty/\|x\| : x \neq 0\}$, then $s^2 \le r \le s$.

Proof: Since $\|\hat{x}\|_\infty \ge s\|x\|$, $\|x^2\| \ge \|\hat{x}^2\|_\infty = \|\hat{x}\|_\infty^2 \ge s^2\|x\|^2$ so $s^2 \le r$.

Since $\|x^2\| \ge r\|x\|^2$, induction gives $\|x^m\| \ge r^{m-1}\|x\|^m$ for $m = 2^n$, $n \in \mathbb{N}$. Taking m^{th} roots and letting $m \to \infty$ gives by Theorem 5 (iii) and 42.16, $\|\hat{x}\|_\infty = r(x) \ge r\|x\|$ so $r \le s$.

Theorem 7. The Gelfand transform $x \to \hat{x}$ is an isometry of X into $C(\Delta)$ if and only if $\|x^2\| = \|x\|^2$ $\forall\, x \in X$ (Compare 35.20).

Proof: In Lemma 6, $x \to \hat{x}$ is an isometry if and only if $s = 1$ (always $\|\hat{x}\| \le \|x\|$ by Theorem 5). By Lemma 6, this occurs if and only if $r = 1$.

Another result about the Gelfand transform which follows from Lemma 6 is

Theorem 8. X is semi-simple and $\hat{X}$ is closed in $C(\Delta)$ if and only if $\exists\, k > 0$ such that $\|x^2\| \ge k\|x\|^2$ $\forall\, x \in X$.

Proof: The existence of $k > 0$ is equivalent to $r > 0$ in Lemma 6 and, hence, to $s > 0$. If $s > 0$, then $x \to \hat{x}$ is 1-1 and has a continuous inverse so $\hat{X}$ is closed in $C(\Delta)$. Conversely, if $x \to \hat{x}$ is 1-1 and $\hat{X}$ is closed, the OMT implies that $s > 0$.

We now compute the maximal ideal space of some familiar function spaces.

Example 9. Let S be compact Hausdorff and $X = C(S)$. For $t \in S$, $\delta_t : f \to f(t)$ is a complex homomorphism on X. Since $C(S)$ separates the points of S, the map $t \to \delta_t$ imbeds S into $\Delta = \Delta_X$. Actually, this map is onto Δ; for if not, $\exists$ a maximal ideal M in $C(S)$ such that $\forall\, t \in S$ $\exists\, f \in M$ with $f(t) \neq 0$. The compactness of S implies that M contains functions $f_1, \ldots, f_n$ such that at least one of the $\{f_i\}$ is non-zero at each point of S. Set $g = \sum_{i=1}^{n} f_i \bar{f}_i$. Then $g \in M$ and $g(t) > 0$ $\forall\, t \in S$ so g is invertible and $M = X$.

Thus, $t \to \delta_t$ establishes a 1-1 correspondence between S and Δ; this correspondence is also topological since Δ carries the relative weak* topology or Gelfand topology. Thus, S is the maximal ideal space of $C(S)$. This fact has the following interesting consequence.

Corollary 10. If $C(S)$ and $C(S')$ are algebraically isomorphic as algebras, then S is homeomorphic to S'.

Proof: $C(S)$ and $C(S')$ have the same maximal ideal spaces.

Example 11. Let $D = \{z \in \mathbb{C} : |z| \leq 1\}$ and let X be all functions which are continuous on D and analytic in $|z| < 1$. Equip X with the sup-norm. If $z \in D$, δ_z is a complex homomorphism on X and $z \to \delta_z$ imbeds D into $\Delta = \Delta_X$. Again this map is onto; for if h is a complex homomorphism on X, let $j \in X$ be the function $j(z) = z$ and set $w = \langle h, j \rangle$. We claim that $h = \delta_w$: now $|w| \leq 1$ since $|w| = |\langle h, j \rangle| \leq \|h\|\|j\| = 1$. If $p \in X$ is a polynomial, $p = \sum_{k=0}^{n} a_k j^k$, then $\langle h, p \rangle = \Sigma a_k \langle h, j^k \rangle = \Sigma a_k w^k = \langle \delta_w, p \rangle$. But, the polynomials are dense in X since if $f \in X$ and $\varepsilon > 0$, $\exists\, r < 1$ such that $|f(rz) - f(z)| < \varepsilon/2$ for $|z| \leq 1$ and $f(rz)$ is analytic in $|z| \leq 1$ so $\exists$ a polynomial p such that $|p(z) - f(rz)| < \varepsilon/2$ for $|z| \leq 1$. Hence, $|f(z) - p(z)| < \varepsilon$ for $|z| \leq 1$. Therefore, $h = \delta_w$ on X, and we have $\Delta = D$, where as above this is a topological identification.

We now consider an important algebra, called the Wiener algebra, and give an application to Fourier series. We first describe the Wiener algebra. Let $S = \{z \in \mathbb{C} : |z| = 1\}$, and let W be the space of all $\mathbb{C}$-valued functions on S of the form $x(z) = \sum_{k=-\infty}^{\infty} a_k z^k$ with $\sum_{-\infty}^{\infty} |a_k| < \infty$ (i.e., all functions which have period 2π and absolutely convergent Fourier series). W is called the Wiener algebra. We define a norm on W by $\|x\| = \sum_{-\infty}^{\infty} |a_k|$. The product on W is the convolution product; if

$y(z) = \sum_{-\infty}^{\infty} b_k z^k$, then $xy(z) = \sum_{-\infty}^{\infty} c_n z^n = \sum_{-\infty}^{\infty} (a*b)_n z^n$, where

$$c_n = (a*b)_n = \sum_{k=-\infty}^{\infty} a_k b_{n-k}$$

is the n^{th} coefficient of the convolution product of $\{a_n\}$ and $\{b_n\}$. Under this product X is a commutative B-algebra with identity $d^0(z) = 1$ (Exercise 2).

We show that the maximal ideal space of W, Δ_W, is S. If $w \in S$, $\delta_w : W \to \mathbb{C}$, $\langle \delta_w, x \rangle = x(w)$, is a multiplicative linear functional on W so S is imbedded in Δ_W by the map $w \to \delta_w$. On the other hand, if $f \in \Delta_W$, $f \neq 0$, set $w = \langle f, d^1 \rangle$, where $d^1(z) = z$. Then

$$|w| \leq \|f\|\|d^1\| = 1,$$

and if $d^n(z) = z^n$ $(n = 0, \pm 1, \ldots)$, $(d^1)^{-1} = d^{-1}$ so that $|w| = |1/\langle f, d^{-1} \rangle| \geq 1/\|f\|\|d^{-1}\| = 1$ and $|w| = 1$ so $w \in S$. We claim that $f = \delta_w$. Now $\langle f, d^1 \rangle = d^1(w)$ and $\{d^n : n = 0, \pm 1, \ldots\}$ is a Schauder basis for W and $\langle f, d^n \rangle = \langle f, d \rangle^n = w^n = \langle \delta_w, d^n \rangle$ so that $f = \delta_w$.

We now give an application of the Wiener algebra to Fourier series.

Theorem 12 (Wiener). Let $h : [-\pi, \pi] \to \mathbb{C}$ have an absolutely convergent Fourier series $\sum_{-\infty}^{\infty} a_n e^{in\theta}$, $\theta \in [-\pi, \pi]$. If $h(\theta) \neq 0$ $\forall\, \theta \in [-\pi, \pi]$, then $1/h$ also has an absolutely convergent Fourier series.

Proof: Set $a(z) = \sum_{-\infty}^{\infty} a_n z^n$, $z \in S$, so $a \in W$. The hypothesis is that

$<\delta_z, a> \neq 0 \;\; \forall z \in S$, the maximal ideal space of W. Thus, a is invertible (Theorem 4), and the result follows.

The original proof of this theorem by Wiener used classical methods and was quite lengthy; the slick Banach algebra proof above was given by Gelfand and attracted a great deal of attention to the subject of Banach algebras which Gelfand originated.

Exercise 1. Let B be the subalgebra of C(S), $S = \{z : |z| = 1\}$, consisting of those functions which are restrictions to S of elements of the algebra in Example 11. Show $\Delta_B \neq S$.

Exercise 2. Show the product defined in W satisfies $xy = yx \in W$ with $\sum_{-\infty}^{\infty} |(a*b)_n| \leq \sum_{-\infty}^{\infty} |a_n| \sum_{-\infty}^{\infty} |b_n|$. Show $xd^0 = d^0x = x \;\; \forall x \in W$.

Exercise 3. Let S be compact, Hausdorff. Show there is a 1-1 correspondence between closed subsets of S and closed ideals of C(S).

Exercise 4. Let X be a semi-simple commutative B-algebra with a unit. Let Y be a B-algebra with a unit. If $\Phi : Y \to X$ is a homomorphism, show Φ is continuous. [Hint: CGT.] If X is a commutative B-algebra under another norm show that the two norms are equivalent.

Exercise 5. If X and Y are commutative B-algebras with identities such that X and Y are isomorphic as algebras, show Δ_X and Δ_Y are homeomorphic.

Exercise 6. Show the image of X under the Gelfand transform is closed in $C(\Delta)$ if and only if $\exists\, k$ such that $\|x\|^2 \le k\|x^2\|$ $\forall\, x \in X$.

Exercise 7. Prove the converse of Corollary 10.

44

Banach Algebras with Involutions

We now consider a class of B-algebras for which the Gelfand transform is onto $C(\Delta)$, where Δ is the maximal ideal space. Let X be a B-algebra with identity e.

Definition 1. A map $* : X \to X$ $(x \to x^*)$ is an involution if

(i) $(x+y)^* = x^* + y^* \quad \forall\, x, y \in X$,

(ii) $(\lambda x)^* = \bar{\lambda} x^* \quad \forall\, x \in X, \lambda \in \mathbb{C}$,

(iii) $(xy)^* = y^* x^* \quad \forall\, x, y \in X$,

(iv) $(x^*)^* = x^{**} = x \quad \forall\, x \in X$.

An element $x \in X$ is called Hermitian if $x = x^*$; x is called normal if $x^* x = x x^*$.

Example 2. $C(S)$ has the involution $f^* = \bar{f}$.

Example 3. Let H be a complex Hilbert space. The map $T \to T^*$ is an involution on L(H). The definitions of Hermitian and normal elements in a B-algebra with involution are obviously carried over from L(H).

Proposition 4. Let X have an involution and $x \in X$. Then

(i) $x + x^*$, $i(x - x^*)$ are Hermitian,

(ii) x has a unique representation $x = u + iv$ with u, v Hermitian,

(iii) x is invertible if and only if x^* is invertible and in this case $(x^{-1})^* = (x^*)^{-1}$,

(iv) $\lambda \in \sigma(x)$ if and only if $\bar{\lambda} \in \sigma(x^*)$.

Proof: (i) is easily checked. (ii): Put $u = (x + x^*)/2$, $v = i(x^* - x)/2$. If $x = u_1 + iv_1$ is another such representation, set $w = v_1 - v$. Then both w and iw are Hermitian so that $iw = (iw)^* = -iw^* = -iw$ and $w = 0$. Hence, $v_1 = v$ and $u_1 = u$.

(iii): $(xy)^* = y^*x^*$ and $e = e^*$ give (iii).

(iv) follows from applying (iii) to $\lambda e = x$.

We now introduce the important class of B-algebras for which the Gelfand transform maps onto $C(\Delta)$.

Definition 5. A B-algebra with involution that satisfies $\|xx^*\| = \|x\|^2$ is alled a <u>B^*-algebra</u>. A <u>B^*-subalgebra</u> X_0 of a B^*-algebra is a closed ıbalgebra which is closed under involution and contains the identity. A <u>*-algebra</u> is a B^*-subalgebra of L(H). [Note that L(H) is a B^*-algebra y 34.4.]

As noted above L(H) is a B^*-algebra and likewise C(S) is a *-algebra.

'roposition 6: Let X be a B^*-algebra.

(i) If $x \in X$ is normal, then $\|x^2\| = \|x\|^2$ and $r(x) = \|x\|$ (compare 35.20 and 35.26).

(ii) If $x \in X$ is Hermitian, then $\sigma(x) \subseteq \mathbb{R}$ (compare 35.27).

Proof: (i):

$$\|x^2\|^2 = \|x^2(x^2)^*\| = \|x^2(x^*)^2\| = |xxx^*x^*\| = \|xx^*xx^*\|$$

$$= \|(xx^*)(xx^*)^*\| = \|xx^*\|^2 = \|x\|^4$$

ɔo $\|x^2\| = \|x\|^2$. Thus, $\|x^n\| = \|x\|^n$ for $n = 2^j$ and

$$r(x) = \lim \sqrt[n]{\|x^n\|} = \|x\|$$

'42.16).

(ii): Suppose $x = x^*$ and let X_0 be the commutative B-algebra generated by x and e. Let Δ_0 be the maximal ideal space of X_0. For $\varphi \in \Delta_0$, write $\varphi(x) = r + is$, $r, s \in \mathbb{R}$. Set $y = x + ite$, where $t \in \mathbb{R}$. Note $yy^* = (x^* + ite)(x - ite) = x^2 + t^2e$. Also $y \in X_0$ and

$$\varphi(y) = \varphi(x) + it = r + i(s + t).$$

Since $\|\varphi\| \leq 1$, $r^2 + (s+t)^2 = |\varphi(y)|^2 \leq \|y\|^2 = \|yy^*\| \leq \|x^2\| + t^2$ and $r^2 + s^2 + 2st \leq \|x^2\|$. This inequality cannot hold for all t unless $s = 0$, i.e., unless $\varphi(x)$ is real. But, $\sigma_{X_0}(x) = \mathscr{R}\hat{x} = \{\varphi(x) : \varphi \in \Delta_0\} \subseteq \mathbb{R}$ by 43.4. Since $X_0 \subseteq X$, $\sigma_X(x) \subseteq \sigma_{X_0}(x) \subseteq \mathbb{R}$.

Lemma 7. Let X be a B-algebra with involution. Then X is a B^*-algebra if and only if $\|x^*\| = \|x\|$ and $\|xx^*\| = \|x\|\|x^*\|$ $\forall\, x \in X$.

Proof: $\Rightarrow$: $\|x\|^2 = \|xx^*\| \leq \|x\|\|x^*\|$ so $\|x\| \leq \|x^*\|$ and, therefore, $\|x^*\| \leq \|x^{**}\| = \|x\|$. Then $\|xx^*\| = \|x\|^2 = \|x\|\|x^*\| = \|x^*\|^2$.

$\Leftarrow$: $\|xx^*\| = \|x\|\|x^*\| = \|x\|^2$.

Given two B-algebras, X_1 and X_2, with involutions, a homomorphism $h : X_1 \to X_2$ is a <u>*-homomorphism</u> if and only if $h(x^*) = h(x)^*$ $\forall\, x \in X_1$; <u>*-isomorphisms</u> are defined similarly. We now have the major result of this section.

Theorem 8 (Gelfand-Naimark). Let X be a commutative B^*-algebra with maximal ideal space Δ. The Gelfand transform $x \to \hat{x}$ of X into $C(\Delta)$ is an isometric *-isomorphism which is onto $C(\Delta)$.

Proof: Every $x \in X$ is normal since X is commutative so $\|x\|^2 = \|x^2\|$ by Proposition 6. Thus, the Gelfand transform $x \to \hat{x}$ is an isometry by 43.7. Given $x \in X$, let $x = u + iv$ with u, v Hermitian (Proposition 4). For $\varphi \in \Delta$, $\varphi(u)$ and $\varphi(v)$ are real by Proposition 6. Hence, $\overline{\varphi(x)} = \varphi(u) - i\varphi(v) = \varphi(u - iv) = \varphi(x^*) = \overline{\hat{x}(\varphi)} = \hat{x}^*(\varphi)$ and $\hat{x}^* = \overline{\hat{x}}$ so $x \to \hat{x}$ is a *-isomorphism.

$\hat{X}$ is a closed subalgebra of $C(\Delta)$ which is closed under conjugation, $\hat{e} = 1 \in \hat{X}$ and $\hat{X}$ separates the points of Δ. Hence, $C(\Delta) = \hat{X}$ by the Stone-Weierstrass Theorem.

We state the analogue of Theorem 8 except for the <u>inverse</u> of the Gelfand map.

Theorem 9. Let X be a commutative B^*-algebra which contains an element $x \in X$ such that the complex polynomials in x and x^* are dense in X. Then

$$(\Psi f)^\wedge = f \circ \hat{x} \tag{1}$$

defines an isometric *-isomorphism Ψ of $C(\sigma(x))$ onto X. Moreover, if $f(\lambda) = \lambda \ \forall \lambda \in \sigma(x)$, then $\Psi f = x$.

Proof: Let Δ be the maximal ideal space of X. Then $\hat{x}$ is a continuous function on Δ with range $\sigma(x)$ (43.4). Suppose $\varphi_1, \varphi_2 \in \Delta$ and $\hat{x}(\varphi_1) = \hat{x}(\varphi_2)$ or $\varphi_1(x) = \varphi_2(x)$. Theorem 8 implies $\varphi_1(x^*) = \varphi_2(x^*)$ so if p is a complex polynomial in two variables, $\varphi_1(p(x, x^*)) = \varphi_2(p(x, x^*))$ and $\varphi_1 = \varphi_2$ by hypothesis. Hence, $\hat{x}$ is 1-1 and since Δ is compact, Hausdorff, $\hat{x}$ is a homeomorphism from Δ onto $\sigma(x)$. The map $f \to f \circ \hat{x}$ is, therefore, an isometric *-isomorphism of

$C(\sigma(x))$ onto $C(\Delta)$.

By Theorem 8, each $f\circ\hat{x}$ is the Gelfand transform of a unique element of X which we denote by Ψf and which satisfies $\|\Psi f\| = \|f\|_\infty$.

If $f(\lambda) = \lambda \ \forall \lambda \in \sigma(x)$, $f\circ\hat{x} = \hat{x} = (\Psi f)^\wedge$ so $\Psi f = x$.

As in §41 we would like to obtain the spectral theorem for a normal operator T from an isometric *-isomorphism between $C(\sigma(T))$ and the B^*-algebra, $\mathscr{A}$, generated by T and I. For this we need to know that the spectrum of such a normal operator is the same whether it is computed in $L(H)$ or in the subalgebra $\mathscr{A}$. This fact follows from the following interesting general theorem on B^*-algebras.

Proposition 10. Let X be a B^*-algebra and X_0 a B^*-subalgebra of X. Then $x_0 \in X_0$ has an inverse in X_0 if and only if x has an inverse in X. Hence, $\sigma_X(x) = \sigma_{X_0}(x)$.

Proof: $\Rightarrow$: Clear. $\Leftarrow$: Let $x \in X_0$ be invertible in X. Then x^* is invertible in X (Proposition 4) and so is x^*x. Now $\sigma_X(x^*x)$ is real (Exercise 1 and Proposition 6) so $\sigma_{X_0}(x^*x) = \sigma_X(x^*x)$ (42.14). Hence, $0 \notin \sigma_{X_0}(x^*x)$ or $(x^*x)^{-1} \in X_0$. But, $e = (x^*x)^{-1}(x^*x)$ implies $x^{-1} = (x^*x)^{-1}x^* \in X_0$.

The last statement is immediate.

Combining Theorem 9 and Proposition 10, we obtain the desired result.

Theorem 11. Let $T \in L(H)$ be normal and let $\mathcal{A}$ be the B^*-algebra generated by T and I. Then $\exists$ an isometric *-isomorphism from $C(\sigma(T))$ onto $\mathcal{A}$.

Proof: $\mathcal{A}$ is the norm closure of the set of all $p(T, T^*)$, where p is a complex polynomial in two variables. By Proposition 10 the spectrum of T is the same in $\mathcal{A}$ and $L(H)$ so the result follows by Theorem 9.

Note in Theorem 11 that if p is a complex polynomial in two variables, then $p \to p(T, T^*)$, $1 \to I$, $p(\lambda) = \lambda \to T$ and $\bar{p}(\lambda) = \bar{\lambda} \to T^*$ under the *-isomorphism.

As in §41, we use Theorem 11 to derive the spectral theorem for normal operators. We give the details in §45, but first we give the following application of Theorem 8.

Stone-Čech Compactification:

Theorem 8 can be used to show the existence of the Stone-Čech compactification. Let S be a completely regular space and let $BC(S)$ be the B-space of all bounded continuous complex-valued functions on S equipped with the sup-norm. Then $BC(S) = X$ is a B^*-algebra under pointwise multiplication and the involution $f \to \bar{f}$. If Δ is the maximal ideal space of X, the Gelfand transform $f \to \hat{f}$ is an isometric *-isomorphism of X onto $C(\Delta)$ (Theorem 8). For each $t \in S$, $\delta_t \in \Delta$ and the map $t \to \delta_t$ is 1-1 since S is completely regular so S is imbedded in

Δ. Moreover, $<\hat{f}, \delta_t> = <\delta_t, f> = f(t)$ so every $f \in X$ is the restriction of $\hat{f}$ to S (actually the image of S which we identify with S). Since the topology of any completely regular space T is just the weak topology from $C(T)$ ([Sm] p. 134), it follows that the topology of S is just the induced topology from Δ.

We claim that S is dense in Δ. If this is not the case, there is a neighborhood

$$\{h \in \Delta : |<h, f_i> - <h_0, f_i>| < \varepsilon, f_1, ..., f_m \in X\}$$

of some $h_0 \in \Delta$ which does not intersect S. If $g_i = f_i - <f_i, h_0> \cdot 1$, then $\forall\, t \in S$ there is a g_i with $|g_i(t)| \geq \varepsilon$. Therefore, if $g(t) = \sum_{i=1}^{n} |g_i(t)|^2$, $g(t) \geq \varepsilon^2$ $\forall\, t \in S$. Hence, $g^{-1} \in X$ so g doesn't belong to any maximal ideal in X. But, $<h_0, g> = 0$ and this contradicts Theorem 43.4 (iii).

Thus, Δ is a compact Hausdorff space which contains S as a dense subset and, furthermore, has the property that every $f \in BC(S)$ has a unique continuous extension to Δ (namely, $\hat{f}$). Δ is called the Stone-Čech compactification of S and is usually denoted by $\beta(S)$. For a brief historical sketch of the evolution of the Stone-Čech compactification, see [TL].

Exercise 1. Show x^*x, xx^* and e are Hermitian.

Exercise 2. Show a B-algebra with involution is a B^*-algebra if and only if $\|x^*x\| = \|x\|^2$ $\forall\, x \in X$.

Exercise 3. Establish the converse of Proposition 6 (ii) for x normal.

Exercise 4. Show an involution on a commutative, semi-simple B-algebra is continuous. [Hint: For $h \in \Delta$, let $\varphi(x) = \bar{h}(x^*)$. Then $\varphi \in \Delta$. Use the CGT and the semi-simplicity to show $x \to x^*$ is continuous.]

45

The Spectral Theorem for Normal Operators

Let H be a complex Hilbert space and $T \in L(H)$ normal. Let $\mathscr{A}$ be the B^*-subalgebra of $L(H)$ generated by T and I; $\mathscr{A}$ is the norm closure of the set of all $p(T, T^*)$, where p is a complex polynomial in two variables. Then $\mathscr{A}$ and $C(\sigma(T))$ are isometrically *-isomorphic by 44.11 (note that $\sigma(T)$ is unambiguous by 44.10 since T has the same spectrum in $\mathscr{A}$ and L(H)). Denote the *-isomorphism from $C(\sigma(T))$ onto $\mathscr{A}$ by $f \to S_f$ (the inverse of the Gelfand map). As in §41, we have

Theorem 1. Let $\mathscr{B} = \mathscr{B}(\sigma(T))$ be the Borel sets of $\sigma(T)$. For each $\sigma \in \mathscr{B}$ $\exists$ a unique orthogonal projection $E(\sigma) \in L(H)$ such that $\mu_{x,y}(\sigma) = E(\sigma)x \cdot y$ $\forall$ x, y $\in$ H, where $\mu_{x,y}$ is the unique regular Borel measure in $C(\sigma(T))'$ satisfying $<\mu_{x,y}, f> = S_f x \cdot y$. The map $E(\cdot) : \mathscr{B} \to L(H)$ satisfies

(i) $E(\emptyset) = 0$, $E(\sigma(T)) = I$,

(ii) $E(\sigma)E(\tau) = E(\tau)E(\sigma) = E(\sigma \cap \tau)$ $\forall$ $\sigma, \tau \in \mathscr{B}$,

(iii) $\sum_{i=1}^{\infty} E(\sigma_i)x = E(\bigcup_{i=1}^{\infty} \sigma_i)x$ $\forall$ $x \in X$ and pairwise disjoint sequence $\{\sigma_i\} \subseteq \mathscr{B}$.

Moreover, each $E(\sigma)$ commutes with every operator in $L(H)$ which commutes with T and T^* (or the elements of $\mathscr{A}$).

The reader should check that the proof of Theorem 41.1 carries over to this situation.

The last statement in Theorem 1 can be strengthened by using a result of Fuglede which states that an operator $S \in L(H)$ which commutes with a normal operator T also commutes with its adjoint, T^* (see [Rs] for an elementary proof of this result).

Recalling §41, a normal operator induces a spectral measure E on the Borel sets of $\sigma(T)$ such that $Tx \cdot y = \int_{\sigma(T)} \lambda d(E(\lambda)x \cdot y)$ $\forall$ $x, y \in H$, or, more generally, $p(T, T^*)x \cdot y = \int_{\sigma(T)} p(\lambda, \bar{\lambda}) d(E(\lambda)x \cdot y)$ for every complex polynomial p. Again, one customarily writes $T = \int_{\sigma(T)} \lambda dE(\lambda)$. This spectral measure, called the <u>resolution of the identity for T</u>, is unique as we now show.

Lemma 2. Let E be the spectral measure for T and let f, g be bounded Borel functions on $\sigma(T)$. Define $A, B \in L(H)$ by

$$Ax \cdot y = \int_{\sigma(T)} f(\lambda) d(E(\lambda)x \cdot y),\ Bx \cdot y = \int_{\sigma(T)} g(\lambda) d(E(\lambda)x \cdot y)$$

for $x, y \in X$. Then

(i) $$A^* x \cdot y = \int_{\sigma(T)} \overline{f(\lambda)} d(E(\lambda)x \cdot y)$$

(ii) $$ABx \cdot y = \int_{\sigma(T)} f(\lambda) g(\lambda) d(E(\lambda)x \cdot y).$$ [Recall 41.2.]

We leave the proof to the reader who should consult 41.2.

Corollary 3. The resolution of the identity for T is unique.

Proof: Let F be another resolution of the identity for T. By Lemma 2,

$$p(T, T^*) = \int_{\sigma(T)} p(\lambda, \bar{\lambda})dF(\lambda) = \int_{\sigma(T)} p(\lambda, \bar{\lambda})dE(\lambda) = S_p$$

for every complex polynomial p. By the Stone-Weierstrass Theorem

$$S_f x \cdot y = \int_{\sigma(T)} f(\lambda)d(E(\lambda)x \cdot y) = \int_{\sigma(T)} f(\lambda)d(F(\lambda)x \cdot y)$$

$$\forall\, f \in C(\sigma(T)),\ x, y \in H.$$

Therefore $E(\sigma) = F(\sigma)\ \forall\, \sigma \in \mathscr{B}$.

Let $B(\sigma(T), \mathscr{B}) = B$ be all the bounded $\mathscr{B}$-measurable, $\mathbb{C}$-valued functions defined on $\sigma(T)$. Then B is a B^*-algebra under the involution $f^* = \bar{f}$. For $f \in B$ define an operator $f(T) \in L(H)$ by

$$f(T)x \cdot y = \int_{\sigma(T)} f(\lambda)d(E(\lambda)x \cdot y)$$

as in Lemma 2 (see Exercise 1). The map $f \to f(T)$ from B into $L(H)$ is a norm decreasing $*$-homomorphism which extends the inverse of the Gelfand map, $f \to S_f$ (Exercise 5). This map extends the operational calculus for Hermitian operators which was constructed in §41.

The spectral theorem for normal operators can also be obtained directly from the spectral theorem for Hermitian operators. If $T \in L(H)$ is normal, then T has a unique representation as $T = A + iB$, where A and B are Hermitian and commute (35.15 and Exercise 35.8). Let E (F) be the spectral resolution for A (B). Then $A = \int \lambda dE(\lambda)$, $B = \int \mu dF(\mu)$ and $I = \int dE(\lambda) = \int dF(\mu)$. Formally, this gives

$$T = A + iB = \int \lambda dE(\lambda) \int dF(\mu) + i \int dE(\lambda) \int \mu dF(\mu) = \int\int (\lambda + i\mu) dE(\lambda) dF(\mu).$$

This formal computation suggests that a resolution of the identity for T can be constructed by constructing a product measure from E and F. This construction is carried out in [Be]; see also [RN] where a similar construction using Riemann-Stieltjes integrals is carried out.

There is another version of the spectral theorem for normal operators which essentially asserts that any normal operator is just (isomorphic to) a multiplication operator as in Exercise 34.1. For this result see [TL] VII.7.4 or [DS2].

There are also versions of the spectral theorem for unbounded operators. These versions of the spectral theorem have applications to differential operators and quantum theory. For expositions of these versions of the spectral theorem see [DS2], [RN] or [C1].

For an interesting historical account of the history of spectral theory, see [St].

Exercise 1. Show the formulas in Lemma 2 define linear operators A, $B \in L(H)$.

Exercise 2. Let $T \in L(H)$ be normal. Show that T is unitary (Exercise 35.13) if and only if $\sigma(T) \subseteq \{z \in \mathbb{C} : |z| = 1\}$.

Exercise 3. Let $T \in L(H)$ be normal with resolution of the identity E. Show that $\lambda \in \sigma(T)$ is an eigenvalue if and only if $E(\{\lambda\}) \neq 0$.

Exercise 4. Let $T \in L(H)$ be normal. Show T is a projection if and only if $\sigma(T) \subseteq \{0, 1\}$.

Exercise 5. Show the map $f \to f(T)$ is a norm-reducing, *-homomorphism from $B(\sigma(T), \mathscr{B})$ to $L(H)$ and $\|f(T)x\|^2 = \int_{\sigma(T)} |f(\lambda)|^2 d(E(\lambda)x \cdot x)$.

Notation

Appendix: Hilbert Space

In this appendix we set down the basic properties of Hilbert spaces for those readers not familiar with Hilbert space.

Let X be a vector space.

Definition 1. An inner product (scalar product, dot product) on X is a function $\cdot : X \times X \to \mathbb{F}$, $(x, y) \to x \cdot y$, satisfying

(i) $(x + y)\cdot z = x\cdot z + y\cdot z \quad \forall\, x, y, z \in X$,

(ii) $\lambda(x\cdot y) = (\lambda x)\cdot y \quad \forall\, x, y \in X, \lambda \in \mathbb{F}$,

(iii) $x\cdot y = \overline{y\cdot x} \quad \forall\, x, y \in X$,

(iv) $x\cdot x \geq 0 \quad \forall\, x \in X$

(v) $x\cdot x = 0$ if and only if $x = 0$.

A vector space X with an inner product defined on it is called an inner product space.

It follows easily from the axioms that $0\cdot x = x\cdot 0 = 0$,

$x\cdot(\lambda y) = \bar{\lambda}(x\cdot y)$ and $x\cdot(y + z) = x\cdot y + x\cdot z$. We have the important inequality.

Theorem 2 (Schwarz Inequality). If X is an inner product space, then

$$(1) \qquad |x\cdot y| \le \sqrt{x\cdot x}\sqrt{y\cdot y} \;\; \forall\, x, y \in X.$$

Proof: If $y = 0$, the result is trivial so assume $y \ne 0$. In this case, (1) is equivalent to $|x\cdot y/\sqrt{y\cdot y}| \le \sqrt{x\cdot x}$ so we may assume $y\cdot y = 1$. Then

$$(2) \qquad 0 \le (x - (x\cdot y)y)\cdot(x - (x\cdot y)y)$$
$$= x\cdot x + |x\cdot y|^2 - (x\cdot y)(y\cdot x) - (\overline{x\cdot y})(x\cdot y)$$
$$= x\cdot x + |x\cdot y|^2 - (x\cdot y)(\overline{x\cdot y}) - |x\cdot y|^2 = x\cdot x - |x\cdot y|^2.$$

Remark 3. Equality holds in (1) if x and y are linearly dependent. The converse also holds for if equality holds in (1) with $y \ne 0$, then (2) implies that $x - (x\cdot y)y = 0$. Note that axiom (v) was not used in the proof of (1).

Proposition 4. If X is an inner product space, the map $x \to \sqrt{x\cdot x} = \|x\|$ defines a norm on X.

Proof: Only the triangle inequality needs to be checked. For $x, y \in X$,

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

by Theorem 2.

If X is an inner product space, we always assume that X is equipped with the norm induced by the inner product.

Proposition 5. The inner product is a continuous function from $X \times X \to \mathbb{F}$.

Proof: If $x_k \to x$ and $y_k \to y$, then

$$|x_k \cdot y_k - x \cdot y| \le |x_k \cdot y_k - x_k \cdot y| + |x_k \cdot y - x \cdot y|$$
$$\le \|x_k\| \|y_k - y\| + \|x_k - x\| \|y\|$$

by the Schwarz Inequality.

We have the following important property of the norm in an inner product space.

Proposition 6 (Parallelogram Law). If X is an inner product space, then

$$\|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2 \quad \forall\, x, y \in X.$$

Proof:

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

Remark 7. The parallelogram law characterizes inner product spaces among the class of NLS. That is, if X is a NLS whose norm satisfies the parallelogram law, then the norm of X is induced by an inner product. If X is real, the inner product is defined by $4x \cdot y = \|x + y\|^2 - \|x - y\|^2$, while if X is complex, the inner product is defined by

$$4x \cdot y = \|x + y\|^2 - \|x - y\|^2 + i\|x + iy\|^2 - i\|x - iy\|^2.$$

We leave the (tedious!) verification to the reader.

Definition 8. An inner product space which is complete under the induced norm is called a Hilbert space.

Example 9. $\mathbb{R}^n$ or $\mathbb{C}^n$ is a Hilbert space under the usual inner product $x \cdot y = \sum_{i=1}^{n} x_i \overline{y_i}$, $x = (x_1, \ldots, x_n)$, $y = (y_1, \ldots, y_n)$. The norm induced by the inner product is the Euclidean norm.

Example 10. ℓ^2 under the inner product $x \cdot y = \sum_{i=1}^{\infty} x_i \overline{y_i}$, $x = \{x_i\}$, $y = \{y_i\}$, is a (the original!) Hilbert space.

Example 11. If (S, Σ, μ) is a measure space, then $L^2(\mu)$ is a Hilbert space with the inner product $f \cdot g = \int_S f\overline{g}d\mu$.

Example 12. $C[a, b]$ with the inner product $f \cdot g = \int_a^b f(t)\overline{g}(t)dt$ is an inner product space which is not a Hilbert space.

We establish an important geometric property of Hilbert space.

Theorem 13. Let K be a non-void, closed, convex subset of a Hilbert space H. If $x \in H$, then there is a unique $y \in K$ such that

$$\|x - y\| = \min\{\|x - z\| : z \in K\} = \text{dist}(x, K).$$

Furthermore, y can be characterized by:

$$(3) \qquad y \in K, \ \mathcal{R}(x - y)\cdot(z - y) \le 0 \ \text{ for all } z \in K.$$

Proof: Set $d = \text{dist}(x, K) \ge 0$. If $w, z \in K$, applying the parallelogram law to $(x - z)/2$ and $(x - w)/2$ gives

$$(4) \qquad d^2 \le \|(w + z)/2 - x\|^2 = \|w - x\|^2/2 + \|z - x\|^2/2 - \|(w - z)/2\|^2$$

since $(w + z)/2 \in K$.

If $\|z - x\| = d$ and $\|w - x\| = d$, then (4) implies that $w = z$ so uniqueness holds.

Pick $\{y_k\} \subseteq K$ such that $\|x - y_k\| \to d$. Set $w = y_k$, $z = y_j$ in (4) to obtain

$$\|(y_k - y_j)/2\|^2 \le \|y_k - x\|^2/2 + \|y_j - x\|^2/2 - d^2 \to 0 .$$

Thus, $\{y_k\}$ is a Cauchy sequence in H and converges to some $y \in K$ with $\|x - y\| = d$.

If $y \in K$ satisfies $\|y - x\| = d$, then for $z \in K$ and $0 < t < 1$,

$$\|x - y\| \le \|x - tz - (1 - t)y\| = \|x - y - t(z - y)\|$$

so

$$\|x - y\|^2 \le \|x - y\|^2 + t^2\|z - y\|^2 - t(x - y)\cdot(z - y) - t(z - y)\cdot(x - y)$$

or

$$2\,\mathcal{R}(x - y)\cdot(z - y) \le t\|z - y\|^2 .$$

Letting $t \to 0$ gives (3).

On the other hand if y satisfies (3) and $z \in K$,

$$\|x - y\|^2 - \|x - z\|^2 = 2\,\mathcal{R}(x - y)\cdot(z - y) - \|z - y\|^2 \le 0$$

so $\|x - y\| = d$.

Let $P_K : H \to H$ be the "projection" map which sends x to y in Theorem 13. If $H = \mathbb{R}^2$, inequality (3) means that the vector from x to $P_k x$ makes an obtuse angle with the vector from z to $P_K x$. Moreover, this map is uniformly continuous on H since $\|P_K u - P_K v\| \le \|u - v\|$. [From (3), $\mathcal{R}(u - P_K u)\cdot(P_K v - P_K u) \le 0$ and

$$\mathcal{R}(v - P_K v)\cdot(P_K u - P_K v) \le 0$$

so adding gives $\mathcal{R}(u - v - (P_K u - P_K v))\cdot(P_K v - P_K u) \le 0$ so

$$\|P_K u - P_K v\|^2 \le \mathcal{R}(u - v).(P_K v - P_K u) \le \|u - v\|\|P_K v - P_K u\|$$

by the Schwarz Inequality.]

If X is an inner product space, then two elements $x, y \in X$ are <u>orthogonal</u>, written $x \perp y$, if $x\cdot y = 0$. A subset $E \subseteq X$ is said to be <u>orthogonal</u> if $x \perp y \quad \forall\, x, y \in E,\ x \ne y$. If $E \subseteq X$ is orthogonal and $\|x\| = 1 \ \forall\, x \in E$, then E is said to be <u>orthonormal</u>.

Example 14. In ℓ^2, $\{e_k : k \in \mathbb{N}\}$ is orthonormal.

Example 15. In $L^2[-\pi, \pi]$, $\{e^{int}/\sqrt{2\pi} : n = 0, \pm 1, \ldots\}$ is orthonormal.

Proposition 16 (Pythagorean Theorem). If $x \perp y$, then

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

Proof: $(x + y)\cdot(x + y) = \|x + y\|^2 = \|x\|^2 + \|y\|^2 = (x - y)\cdot(x - y)$.

If X is an inner product space and $M \subseteq X$, we set

$$M^\perp = \{x \in X : x \perp y \;\; \forall\, y \in M\};$$

$M^\perp$ is called the orthogonal complement of M. Note that $M^\perp$ is a closed linear subspace of X. We now use Theorem 13 to show that any closed linear subspace of a Hilbert space is complemented.

Theorem 17. If M is a closed linear subspace of a Hilbert space H, then $H = M \oplus M^\perp$.

Proof: For $x \in H$ let $y = P_M x \in M$ be as in Theorem 13. Then $x - y \in M^\perp$ since by (3) $\mathcal{R}(x - y)\cdot w \le 0$ for every $w \in M$.

Hence, $x = (x - y) + y$ with $x - y \in M^\perp$ and $y \in M$. Since $M \cap M^\perp = \{0\}$, we have $H = M \oplus M^\perp$.

Thus, every closed linear subspace of a Hilbert space has a projection onto it. This property actually characterizes Hilbert spaces among the class of B-spaces ([LT]).

We can now establish the Riesz Representation Theorem for Hilbert space.

Proposition 18. Let H be a Hilbert space and $y \in H$. If $f_y : H \to \mathbb{F}$ is defined by $\langle f_y, x\rangle = x \cdot y$, then $f_y \in H'$ and $\|f_y\| = \|y\|$.

Proof: f_y is clearly linear and since $|\langle f_y, x\rangle| = |x \cdot y| \le \|x\|\|y\|$, $f_y \in H'$ with $\|f_y\| \le \|y\|$. Since $\langle f_y, y\rangle = y \cdot y = \|y\|^2$, $\|f_y\| = \|y\|$.

Thus, the map $y \to f_y$ is an isometry from H into its dual space H'. We show that this map is onto.

Theorem 19 (Riesz Representation Theorem). If H is a Hilbert space and $f \in H'$, then $\exists$ a unique $y \in H$ such that $f = f_y$.

Proof: If $f = 0$, put $y = 0$. Suppose $f \ne 0$. Set $M = \mathcal{N}(f)$ so M is a proper closed subspace of H and $M^{\perp} \ne \{0\}$. Choose $z \in M^{\perp}$, $z \ne 0$. Then $\langle f, z\rangle \ne 0$. Set $y = (\overline{\langle f, z\rangle}/\|z\|^2)z$ so $y \in M^{\perp}$, $y \ne 0$, and $\langle f, y\rangle = |\langle f, z\rangle|^2/\|z\|^2 = y \cdot y$. For $x \in H$, let $x_1 = x - (\langle f, x\rangle/\|y\|^2)y$, $x_2 = (\langle f, x\rangle/\|y\|^2)y$ so $x = x_1 + x_2$ and $\langle f, x_1\rangle = 0$ so $x_1 \in M$ and $x_1 \cdot y = 0$. Hence, $x \cdot y = x_2 \cdot y = \langle f, x\rangle = \langle f_y, x\rangle$, and $f = f_y$.

The map $\Phi : y \to f_y$ is an isometry from H onto H' which is additive but is only conjugate homogeneous in the sense that $\Phi(ty) = \bar{t}\Phi(y)$. From this it follows that Hilbert spaces are always reflexive.

If X is an inner product space and $E = \{x_a : a \in A\}$ is an orthonormal subset of X, then $\forall\, x \in X$ the scalars $\hat{x}(a) = x \cdot x_a$, $a \in A$, are called the Fourier coefficients of x with respect to E. We establish several important properties of the Fourier coefficients.

Proposition 20. Let X be an inner product space and $\{x_1, ..., x_n\}$ an orthonormal set in X. Then for each $x \in X$,

(i) $$\sum_{i=1}^{n} |x \cdot x_i|^2 = \sum_{i=1}^{n} |\hat{x}(i)|^2 \le \|x\|^2,$$

(ii) $$(x - \sum_{i=1}^{n} (x \cdot x_i)x_i) \perp x_i \quad \forall i.$$

Proof: (i):

$$0 \le \|x - \sum_{i=1}^{n} (x \cdot x_i)x_i\|^2 = (x - \sum_{i=1}^{n} (x \cdot x_i)x_i, x - \sum_{i=1}^{n} (x \cdot x_i)x_i)$$

$$= \|x\|^2 - \sum_i (x \cdot x_i)(\overline{x \cdot x_i}) - \sum_i (\overline{x \cdot x_i})(x \cdot x_i)$$

$$+ \sum_i \sum_j (x \cdot x_i)(\overline{x \cdot x_j}) x_i \cdot x_j = \|x\|^2 - \sum_i |x \cdot x_i|^2.$$

(ii): $$(x - \sum_i (x \cdot x_i)x_i, x_j) = x \cdot x_j - \sum_i (x \cdot x_i)(x_i \cdot x_j) = x \cdot x_j - x \cdot x_j = 0.$$

We generalize the inequality in (i) to infinite orthonormal sets.

Proposition 21. Let $E = \{x_a : a \in A\}$ be an orthonormal set in an inner product space X. For each $x \in X$ the set $E_x = \{a \in A : x \cdot x_a \ne 0\}$ is at most countable.

Proof: For $n \in \mathbb{N}$, let $S_n = \{a \in A : |x \cdot x_a|^2 > \|x\|^2/n\}$. By Proposition 20 S_n contains at most $n - 1$ elements. Since $E_x = \bigcup_{n=1}^{\infty} S_n$, the result follows.

Theorem 22 (Bessel's Inequality). Let $E = \{x_a : a \in A\}$ be an orthonormal subset of an inner product space X. For each $x \in X$,

$$(5) \qquad \sum_{a \in A} |x \cdot x_a|^2 = \sum_{a \in A} |\hat{x}(a)|^2 \le \|x\|^2.$$

Proof: If A is finite, this is Proposition 20. If A is infinite, we must assign a meaning to the series in (5). Let $S = \{a \in A : x \cdot x_a \neq 0\}$. If $S = \emptyset$, we set $\sum_{a \in A} |x \cdot x_a|^2 = 0$, and if S is finite, we set

$$\sum_{a \in A} |x \cdot x_a|^2 = \sum_{a \in S} |x \cdot x_a|^2$$

and (5) follows from Proposition 20. If S is infinite, S is countable by Proposition 21 so the elements $\{x_a : a \in S\}$ can be arranged in a sequence, say $y_1, y_2, \ldots$. By Proposition 20, $\forall$ n

$$\sum_{i=1}^{n} |x \cdot y_i|^2 \le \|x\|^2$$

so the series $\sum_{i=1}^{\infty} |x \cdot y_i|^2$ is absolutely convergent and its sum is independent of the ordering of the elements $\{x_a : a \in S\}$. Therefore, we may define $\sum_{a\in A} |x\cdot x_a|^2 = \sum_{i=1}^{\infty} |x\cdot y_i|^2$ and $\sum_{a\in A} |x\cdot x_a|^2 \le \|x\|^2$ by Proposition 20.

We will now show that equality holds in Bessel's Inequality for certain orthonormal sets in a Hilbert space. An orthonormal subset E of a Hilbert space H is said to be complete (or a complete orthonormal set) if $E_1 \subseteq H$ orthonormal and $E_1 \supseteq E$ implies that $E = E_1$ (i.e., E is a maximal orthonormal set with respect to set inclusion). We give several criteria for an orthonormal set to be complete. First, we require a lemma.

Lemma 23. Let $\{x_1, ..., x_n\}$ be an orthonormal set in an inner product space X.

(i) If $x = \sum_{k=1}^{n} c_k x_k$, then $c_k = x\cdot x_k = \hat{x}(k)$ and

$$\|x\|^2 = \sum_{k=1}^{n} |c_k|^2.$$

(ii) For $\{c_1, ..., c_n\} \subseteq \mathbb{F}$, $\|x - \sum_{k=1}^{n} c_k x_k\|$ attains its minimum (as a function of $(c_1, ..., c_n)$) at $c_k = x\cdot x_k = \hat{x}(k)$, $k = 1, ..., n$.

Proof: (i): That $c_k = x\cdot x_k$ is immediate;

$$\|x\|^2 = x\cdot x = \sum_k \sum_j c_k \bar{c}_j x_k \cdot x_j = \sum_{k=1}^{n} |c_k|^2.$$

(ii):

$$0 \le \|x - \sum_{k=1}^{n} c_k x_k\|^2 = (x - \sum_{k=1}^{n} c_k x_k)\cdot(x - \sum_{k=1}^{n} c_k x_k)$$

$$= \|x\|^2 - \sum_{k=1}^{n} c_k(\overline{x\cdot x_k}) - \sum_{k=1}^{n} \bar{c}_k(x\cdot x_k) + \sum_{k=1}^{n} |c_k|^2$$

$$= (\|x\|^2 - \sum_{k=1}^{n} |x\cdot x_k|^2) + \sum_{k=1}^{n} (x\cdot x_k - c_k)(\overline{x\cdot x_k - c_k})$$

and the expression on the right is clearly minimal at $c_k = x\cdot x_k$.

Theorem 24. Let $E = \{x_a : a \in A\}$ be an orthonormal set in a Hilbert space H. The following are equivalent:

(i) E is complete,

(ii) $x \perp x_a \ \forall a \in A$ implies $x = 0$,

(iii) span E is dense in H

(iv) If $x \in H$, $\|x\|^2 = \sum_{a\in A} |x\cdot x_a|^2$ (equality in Bessel's Inequality)

(v) $x = \sum_{a\in A} (x\cdot x_a)x_a \ \forall x \in H$,

(vi) if $x, y \in H$, $x\cdot y = \sum_{a\in A} (x\cdot x_a)(\overline{y\cdot x_a})$ (Parseval's Equality).

Proof: (i) ⇒ (ii): If (ii) is false, $\exists\, x \neq 0$ such that $x \perp x_a \ \forall a \in A$. Set $z = x/\|x\|$ so $\{z\} \cup E$ is an orthonormal set which properly contains E so (i) does not hold.

(ii) ⇒ (iii): Let M be the closure of span E. If $M \neq H$, $H = M \oplus M^{\perp}$ with $M^{\perp} \neq \{0\}$. If $x \neq 0$, $x \in M^{\perp}$, then $x \perp x_a$ $\forall a \in A$ so (ii) fails.

(iii) ⇒ (iv): Let $\varepsilon > 0$ and $x \in H$. $\exists\, x_{a_1}, \ldots, x_{a_n} \in E$ and $c_1, \ldots, c_n \in \mathbb{F}$ such that $\|x - \sum_{k=1}^{n} c_k x_{a_k}\| < \varepsilon$. By Lemma 23 (ii),

$$\|x - \sum_{k=1}^{n} (x \cdot x_{a_k}) x_{a_k}\| < \varepsilon . \tag{6}$$

By Lemma 23 (i) and (6),

$$(\|x\| - \varepsilon)^2 \leq \|\sum_{k=1}^{n} (x \cdot x_{a_k}) x_{a_k}\|^2 = \sum_{k=1}^{n} |x \cdot x_{a_k}|^2 \leq \sum_{a \in A} |x \cdot x_a|^2 .$$

Bessel's Inequality gives the reverse inequality.

(iv) ⇒ (v): As in Theorem 22 let $S = \{x_a : x_a \cdot x \neq 0\}$ and arrange the elements of S into a sequence $y_1, y_2, \ldots$. Then

$$\|x - \sum_{k=1}^{n} (x \cdot y_k) y_k\|^2 = \|x\|^2 - \sum_{k=1}^{n} |x \cdot y_k|^2$$

$$= \sum_{k=1}^{\infty} |x \cdot y_k|^2 - \sum_{k=1}^{n} |x \cdot y_k|^2 = \sum_{k=n+1}^{\infty} |x \cdot y_k|^2$$

by (iv). Hence, $x = \sum_{k=1}^{\infty} (x \cdot y_k) y_k = \sum_{a \in A} (x \cdot x_a) x_a$.

(v) ⇒ (vi): By Proposition 5,

$$x \cdot y = \sum_{a \in A} \sum_{b \in A} (x \cdot x_a)(\overline{y \cdot x_b}) x_a \cdot x_b = \sum_{a \in A} (x \cdot x_a)(\overline{y \cdot x_a}).$$

(vi) ⇒ (i): If (i) fails, $\exists\ z \in H$ with $\|z\| = 1$ and $z \perp x_a\ \forall\ a \in A$. Then $z \cdot z = 1$ while $\sum_{a \in A} |z \cdot x_a|^2 = 0$ so (vi) fails.

Theorem 25. Let H be a Hilbert space with E and F complete orthonormal subsets. Then E and F have the same cardinality.

Proof: Since orthonormal sets are linearly independent, we may assume that E and F are infinite.

For $e \in E$, let $F_e = \{f \in F : f \cdot e \neq 0\}$. By Theorem 24 (ii), $F = \bigcup_{e \in E} F_e$ and by Proposition 21 each F_e is at most countable. Hence, the cardinality of F is less than or equal to the cardinality of E. Symmetry gives the reverse inequality.

The cardinality of a (any) complete orthonormal set is called the <u>orthonormal dimension</u> of the Hilbert-space.

Example 26. $\{e_k : k \in \mathbb{N}\}$ is a complete orthonormal subset of ℓ^2.

Example 27. $\{e^{int}/\sqrt{2\pi} : n = 0, \pm 1, \ldots\}$ is a complete orthonormal subset of $L^2[-\pi, \pi]$ ([HS] IV.16.32).

Exercise 1. Show that any orthonormal subset of a Hilbert space is contained in a complete orthonormal set.

Exercise 2. Show a Hilbert space is separable if and only if its orthonormal dimension is countable.

Exercise 3. In Theorem 24, the map $x \to \{x \cdot x_a : a \in A\}$ is an isometry from H into $\ell^2(A)$. Show this map is onto (this is the Riesz-Fischer Theorem). Thus, any Hilbert space is isomorphic to some $\ell^2(A)$.

Exercise 4. If D is a dense subset of an inner product space and $x \perp D$, show $x = 0$.

Exercise 5. If E is a linear subspace of a Hilbert space H, Y is a B-space and $T : E \to Y$ is a continuous linear operator, show T has a continuous linear extension $\hat{T} : H \to Y$. [Compare with the remark following 27.5.]

Exercise 6 (Gram-Schmidt). Let $x_1, \ldots, x_n$ be linearly independent. Set $y_1 = x_1$, $y_k = x_k - \sum_{j=1}^{k-1} (x_k \cdot y_j) y_j$ for $k > 1$ and $z_k = y_k / \|y_k\|$. Show $\{z_k\}$ is orthonormal and $\text{span}\{x_k\} = \text{span}\{z_k\}$.

References

[Ag] S. Agmon, Elliptic Boundary Value Problems, Van Nostrand, Princeton, 1965.

[AB] C. Aliprantis and O. Burkinshaw, Principles of Real Analysis, Academic Press, N.Y., 1990.

[A] P. Antosik, On uniform boundedness of families of mappings, Proc. Conf. Conv. Structures, Szczyrk, 1979, 2-16.

[AS] P. Antosik and C. Swartz, Matrix Methods in Analysis, Springer-Verlag, Heidelberg, 1985.

[BN] G. Bachman and L. Narici, Functional Analysis, Academic Press, N.Y., 1966.

[B] S. Banach, Oeuvres II, Acad. Pol. Sci., Warsaw, 1979.

[BK] G. Bennett and N. Kalton, FK-Spaces Containing c_o, Duke Math. J., 39 (1972), 561-582.

[Be] S. Berberian, Notes on Spectral Theory, Van Nostrand, Princeton, 1966.

[Bo] N. Bourbaki, Topological Vector Spaces, Springer-Verlag, Berlin, 1987.

[BM] J. Burzyk and P. Mikusiuski, On normability of semigroups, Bull. Acad. Polon. Sci., 28 (1980), 33-35.

[BKL] J. Burzyk, C. Klis and Z. Lipecki, On metrizable Abelian groups with a completeness-type property, Colloq. Math., 49 (1984), 33-39.

[Co] J. Connor, A Short Proof of Steinhaus' Theorem on Summability, Amer. Math. Monthly, 92 (1985), 420-421.

[C1] J. Conway, A Course in Functional Analysis, Springer-Verlag, N.Y., 1985.

[C2] J. Conway, The Inadequacy of Sequences, Amer. Math. Monthly, 76 (1969), 68.

[C3] J. Conway, The Compact Operators are not Complemented in B(H), Proc. AMS, 32 (1982), 549-550.

[Da] M. Day, Normed Linear Spaces, Springer-Verlag, N.Y., 1962.

[DeS] J. DePree and C. Swartz, Introduction to Real Analysis, Wiley, N.Y., 1987.

[DU] J. Diestel and J. Uhl, Vector Measures, Amer. Math. Soc. Surveys #15, Providence, 1977.

[Dr] L. Drewnowski, Equivalence of Brooks-Jewett, Vitali-Hahn-Saks and Nikodym Theorems, Bull. Pol. Acad. Sci., 20 (1972), 725-731.

[DS] N. Dunford and J. Schwartz, Linear Operators I, Interscience, N.Y., 1953.

[DS2] N. Dunford and J. Schwartz, Linear Operators II, Interscience, N.Y. 1963.

[E1] P. Enflo, A counterexample to the approximation problem in Banach spaces, Acta Math., 130 (1973), 309-317.

[E2] E. Enflo, On the invariant subspace problem for Banach spaces, Acta Math., 158 (1987), 213-313.

[FW] K. Floret and J. Wloka, Einführung in die Theorie lokalkonvexer Räume, Lecture Notes in Math. 56, Springer-Verlag, Heidelberg, 1968.

[Fr] J. Franklin, Methods of Mathematical Economics, Springer-Verlag, N.Y., 1980.

[GS] I. M. Gelfand and G. E. Shilov, Generalized Functions, Academic Press, N.Y., 1964.

[Go1] S. Goldberg, Unbounded Linear Operators, McGraw-Hill, N.Y., 1966.

[Go2] S. Goldberg, Some Properties of the Space of compact Operators on a Hilbert Space, Math. Ann., 138 (1959), 329-331.

[Gr] A. Grothendieck, Topological Vector Spaces, Gordan and Breach, N.Y., 1973.

[Ha] H. Hahn, Uber Folgen linearen Operationen, Monatsch. für Math und Phys., 32 (1922), 1-88.

[HT] E. Hellinger and O. Toeplitz, Grundlagen für eine Theorie der unendlichen Matrizen, Math. Ann., 69 (1910), 289-330.

[HL] H. Hermes and J. P. LaSalle, Functional Analysis and Time Optional Control, Academic Press, N.Y., 1969.

[HS] E. Hewitt and K. Stromberg, Real and Abstract Analysis, Springer-Verlag, N.Y., 1965.

[Hi] T. H. Hildebrandt, Integration in Abstract Spaces, Bull. Amer. Math. Soc., 59 (1953), 111-139.

[HP] E. Hille and R. Phillips, Functional Analysis and Semi-Groups, Amer. Math. Soc., Providence, 1957.

[Ho] H. Hochstadt, Eduard Helly, Father of the Hahn-Banach Theorem, Math. Intell., 2 (1979), 123-125.

[H1] J. Horváth, Topological Vector Spaces and Distributions, Addison-Wesley, Reading, 1966.

[H2] J. Horváth, An Introduction to Distributions, Amer. Math. Monthly, 77 (1970), 227-240.

[J1] R. C. James, A non-reflexive Banach space isometric with its second conjugate space, Proc. Nat. Acad. Sci. USA, 37 (1951), 174-177.

[J2] R. C. James, Reflexivity and the sup of linear functionals, Israel J. Math., 13 (1972), 289-300.

[J3] R. C. James, Bases in Banach Spaces, Amer. Math. Monthly 82 (1989), 625-640.

[Jo] J. Johnson, Remarks on Banach Spaces of Compact Operators, J. Functional Anal., 32 (1979), 304-311.

[Jor] O. G. Jorsboe, Set Transformations and Invariant Measures, Aarhus Universitet, 1966.

[Ke] J. Kelley, General Topology, Van Nostrand, Princeton, 1955.

[Kl] C. Klis, An Example of noncomplete normed (K)-space, Bull. Pol. Acad. Sci., 26 (1978), 415-420.

[K] A. Kolmogorov, Zur Normierbarkeit eines allgemeinen topologischen linearen Raumes, Studia Math., 5 (1934), 29-33.

[Kom] Y. Komura, Some examples on linear topological spaces, Math. Ann., 153 (1964), 150-162.

[K1] G. Köthe, Topological Vector Spaces I, Springer-Verlag, N.Y., 1969.

[K2] G. Köthe, Topological Vector Spaces II, Springer-Verlag, N.Y., 1979.

[LL] I. Labuda and Z. Lipecki, On subseries convergent series and m-quasi-bases in topological linear spaces, Manuscripta Math., 38 (1982), 87-98.

[LT] J. Lindenstrauss and L. Tzafriri, Classical Banach Spaces I, Springer-Verlag, N.Y., 1977.

[L] J. Lützen, The Prehistory of the theory of distributions, Springer-Verlag, N.Y., 1982.

[Ma] I. Maddox, Infinite Matrices of Operators, Springer-Verlag, Heidelberg, 1980.

[MO] S. Mazur and W. Orlicz, Uber Folgen linearen Operationen, Studia Math., 4 (1933), 152-157.

[Me] M. D. Meyerson, Every power series is a Taylor series, Amer. Math. Monthly, 88 (1981), 51-52.

[M] M. Munroe, Introduction to Measure and Integration, Addison-Wesley, Reading, Mass., 1953.

[NB] L. Narici and E. Beckenstein, Topological Vector Spaces, Marcel Dekker, N.Y., 1985.

[NS] A. W. Naylor and G. R. Sell, Linear Operator Theory, Holt-Rinehart-Winston, N.Y., 1971.

[vN] J. von Neumann, On complete topological spaces, Trans. Amer. Math. Soc., 37 (1935), 1-20.

[P] E. Pap, Functional analysis with K-convergence, Proceedings of Conference on Convergence Spaces, Bechyne, Czech. 1984, Akademie-Verlag, 245-250.

[PS] E. Pap and C. Swartz, Pap's Adjoint Theorem in Locally Convex Spaces.

[Ph] R. Phillips, On Linear Transformations, Trans. Amer. Math. Soc.,

48 (1940), 516-541.
[Pi] A. Pietsch, Quasi-nukleare Abbildungen in normierten Räumen, Math. Ann., 165 (1966), 76-90.
[Ri] N. M. Rice, On nth Roots of Positive Operators, Amer. Math. Monthly, 89 (1982), 313-314.
[RN] F. Riesz and B. Nagy, Functional Analysis, Unger, N.Y., 1955.
[RR] W. Robertson and A. Robertson, Topological Vector Spaces, Cambridge University Press, Cambridge, 1964.
[Rs] M. Rosenblum, On a Theorem of Fuglede and Putnam, J. London Math. Soc., 33 (1958), 376-377.
[Ro] R. L. Royden, Real Analysis, MacMillan, N.Y., 1988.
[Ru1] W. Rudin, Real and Complex Analysis, McGraw-Hill, N.Y., 1966.
[Ru2] W. Rudin, Functional Analysis, McGraw-Hill, N.Y., 1973.
[S] W. L. C. Sargent, On Some Theorems of Hahn, Banach and Steinhaus, J. London M. S., 28 (1953), 438-451.
[Sa] S. Saxon, Some normed barrelled spaces which are not Baire, Math. Ann., 209 (1974), 153-160.
[Sc] J. Schur, Uber lineare Transformation in der Theorie die unendlichen Reihen, J. Reine Angew. Math., 151 (1920), 79-111.
[Sch] L. Schwartz, Theorie des Distributions I, II, Hermann, Paris, 1951.
[Sm] G. Simmons, Introduction to Topology and Modern Analysis, McGraw-Hill, N.Y., 1963.
[Si] I. Singer, Bases in Banach Spaces I, Springer-Verlag, N.Y., 1970.
[St] L. Steen, Highlights in the History of Spectral Theory, Amer. Math. Monthly, 80 (1973), 359-381.
[Sw] C. Swartz, The Evolution of the Uniform Boundedness Principle, Math. Chron., 19 (1990), 1-18.
[Sz] Szankowski, B(H) does not have the approximation property, Acta Math., 147 (1981), 89-108.
[Ta] A. Taylor, Notes on the History and Uses of Analyticity in Operator Theory, Amer. Math. Monthly, 78 (1971), 331-342.
[TL] A. Taylor and D. Lay, Introduction to Functional Analysis, Wiley, N.Y., 1980.
[T1] F. Treves, Topological Vector Spaces, Distributions and Kernels, Academic Press, N.Y., 1967.
[T2] F. Treves, Applications of Distributions to PDE Theory, Amer. Math. Monthly, 77 (1970), 227-240.
[T3] F. Treves, Basic Linear Partial Differential Equations, Academic Press, N.Y., 1975.
[V] M. Valdivia, B_r-complete spaces which are not B-complete, Math. Zeit. 185 (1984), 253-259.
[Wh1] R. Whitley, An Elementary Proof of the Eberlein-Smulian Theorem, Math. Ann., 172 (1967), 116-118.
[Wh2] R. Whitley, Projecting m onto c_o, Amer. Math. Monthly, 73 (1966), 285-286.

[Wh3] R. Whitley, The Spectral Theorem for a Normal Operator, Amer. Math. Monthly, 75 (1968), 856-861.
[W] A. Wilansky, Modern Methods in Topological Vector Spaces, McGraw-Hill, N.Y., 1978.
[Y] K. Yosida, Functional Analysis, Springer-Verlag, Berlin, 1966.

Index