

Graduate Texts in Mathematics

Paul R. Halmos

A Hilbert Space Problem Book



Springer-Verlag
New York Heidelberg Berlin

Graduate Texts in Mathematics 19

Managing Editors: P. R. Halmos
C. C. Moore

Paul R. Halmos

A Hilbert Space Problem Book

Springer-Verlag New York · Heidelberg · Berlin

Managing Editors

P. R. Halmos

Indiana University
Department of Mathematics
Swain Hall East
Bloomington, Indiana 47401

C. C. Moore

University of California
at Berkeley
Department of Mathematics
Berkeley, California 94720

AMS Subject Classification (1970)

Primary: 46Cxx

Secondary: 46Axx, 47Axx

Library of Congress Cataloging in Publication Data

Halmos, Paul Richard, 1914-

A Hilbert space problem book.

(Graduate texts in mathematics, v. 19)

Reprint of the ed. published by Van Nostrand,
Princeton, N.J., in series: The University series
in higher mathematics.

Bibliography: p.

1. Hilbert space—Problems, exercises, etc.

I. Title. II. Series.

[QA322.4.H34 1974] 515'.73 74-10673

All rights reserved.

No part of this book may be translated or reproduced in
any form without written permission from Springer-Verlag.

© 1967 by American Book Company and

1974 by Springer-Verlag New York Inc.

Softcover reprint of the hardcover 1st edition 1974

ISBN-13: 978-1-4615-9978-4

e-ISBN-13: 978-1-4615-9976-0

DOI: 10.1007/978-1-4615-9976-0

To J. U. M.

Preface

The only way to learn mathematics is to do mathematics. That tenet is the foundation of the do-it-yourself, Socratic, or Texas method, the method in which the teacher plays the role of an omniscient but largely uncommunicative referee between the learner and the facts. Although that method is usually and perhaps necessarily oral, this book tries to use the same method to give a written exposition of certain topics in Hilbert space theory.

The right way to read mathematics is first to read the definitions of the concepts and the statements of the theorems, and then, putting the book aside, to try to discover the appropriate proofs. If the theorems are not trivial, the attempt might fail, but it is likely to be instructive just the same. To the passive reader a routine computation and a miracle of ingenuity come with equal ease, and later, when he must depend on himself, he will find that they went as easily as they came. The active reader, who has found out what does not work, is in a much better position to understand the reason for the success of the author's method, and, later, to find answers that are not in books.

This book was written for the active reader. The first part consists of problems, frequently preceded by definitions and motivation, and sometimes followed by corollaries and historical remarks. Most of the problems are statements to be proved, but some are questions (is it?, what is?), and some are challenges (construct, determine). The second part, a very short one, consists of hints. A hint is a word, or a paragraph, usually intended to help the reader find a solution. The hint itself is not necessarily a condensed solution of the problem; it may just point to what I regard as the heart of the matter. Sometimes a problem contains a trap, and the hint may serve to chide the reader for rushing in too recklessly. The third part, the longest, consists of solutions: proofs, answers, or constructions, depending on the nature of the problem.

The problems are intended to be challenges to thought, not legal technicalities. A reader who offers solutions in the strict sense only (this is what was asked, and here is how it goes) will miss a lot of the point, and he will miss a lot of fun. Do not just answer the question, but try to think of related questions, of generalizations (what if the operator is not normal?), and of special cases (what happens in the finite-

dimensional case?). What makes an assertion true? What would make it false?

If you cannot solve a problem, and the hint did not help, the best thing to do at first is to go on to another problem. If the problem was a statement, do not hesitate to use it later; its use, or possible misuse, may throw valuable light on the solution. If, on the other hand, you solved a problem, look at the hint, and then the solution, anyway. You may find modifications, generalizations, and specializations that you did not think of. The solution may introduce some standard nomenclature, discuss some of the history of the subject, and mention some pertinent references.

The topics treated range from, fairly standard textbook material to the boundary of what is known. I made an attempt to exclude dull problems with routine answers; every problem in the book puzzled me once. I did not try to achieve maximal generality in all the directions that the problems have contact with. I tried to communicate ideas and techniques and to let the reader generalize for himself.

To get maximum profit from the book the reader should know the elementary techniques and results of general topology, measure theory, and real and complex analysis. I use, with no apology and no reference, such concepts as subbase for a topology, precompact metric spaces, Lindelöf spaces, connectedness, and the convergence of nets, and such results as the metrizability of compact spaces with a countable base, and the compactness of the Cartesian product of compact spaces. (Reference: Kelley [1955].) From measure theory, I use concepts such as σ -fields and L^p spaces, and results such as that L^p convergent sequences have almost everywhere convergent subsequences, and the Lebesgue dominated convergence theorem. (Reference: Halmos [1950 b].) From real analysis I need, at least, the facts about the derivatives of absolutely continuous functions, and the Weierstrass polynomial approximation theorem. (Reference: Hewitt-Stromberg [1965].) From complex analysis I need such things as Taylor and Laurent series, sub-uniform convergence, and the maximum modulus principle. (Reference: Ahlfors [1953].)

This is not an introduction to Hilbert space theory. Some knowledge of that subject is a prerequisite; at the very least, a study of the elements of Hilbert space theory should proceed concurrently with the reading of this book. Ideally the reader should know something like as the first two chapters of Halmos [1951].

I tried to indicate where I learned the problems and the solutions and

where further information about them is available, but in many cases I could find no reference. When I ascribe a result to someone without an accompanying bracketed date (the date is an indication that the details of the source are in the list of references), I am referring to an oral communication or an unpublished preprint. When I make no ascription, I am not claiming originality; more than likely the result is a folk theorem.

The notation and terminology are mostly standard and used with no explanation. As far as Hilbert space is concerned, I follow Halmos [1951], except in a few small details. Thus, for instance, I now use f and g for vectors, instead of x and y (the latter are too useful for points in measure spaces and such), and, in conformity with current fashion, I use “kernel” instead of “null-space”. (The triple use of the word, to denote (1) null-space, (2) the continuous analogue of a matrix, and (3) the reproducing function associated with a functional Hilbert space, is regrettable but unavoidable; it does not seem to lead to any confusion.) Incidentally “kernel” and “range” are abbreviated as \ker and ran , “dimension” is abbreviated as \dim , “trace” is abbreviated as tr , and real and imaginary parts are denoted, as usual, by Re and Im . The “signum” of a complex number z , i.e., $z/|z|$ or 0 according as $z \neq 0$ or $z = 0$, is denoted by $\text{sgn } z$. The *co-dimension* of a subspace of a Hilbert space is the dimension of its orthogonal complement (or, equivalently, the dimension of the quotient space it defines). The symbol \vee is used to denote span, so that $\mathbf{M} \vee \mathbf{N}$ is the smallest closed linear manifold that includes both \mathbf{M} and \mathbf{N} , and, similarly, $\bigvee_j \mathbf{M}_j$ is the smallest closed linear manifold that includes each \mathbf{M}_j . Subspace, by the way, means closed linear manifold, and operator means bounded linear transformation.

The arrow has two uses: $f_n \rightarrow f$ indicates that a sequence $\{f_n\}$ tends to the limit f , and $x \rightarrow x^2$ denotes the function φ defined by $\varphi(x) = x^2$. Since the inner product of two vectors f and g is always denoted by (f, g) , another symbol is needed for their ordered pair; I use $\langle f, g \rangle$. This leads to the systematic use of the angular bracket to enclose the coordinates of a vector, as in $\langle f_0, f_1, f_2, \dots \rangle$. In accordance with inconsistent but widely accepted practice, I use braces to denote both sets and sequences; thus $\{x\}$ is the set whose only element is x , and $\{x_n\}$ is the sequence whose n -th term is x_n , $n = 1, 2, 3, \dots$. This could lead to confusion, but in context it does not seem to do so. For the complex conjugate of a complex number z , I use z^* . This tends to make mathematicians nervous, but it is widely used by physicists, it is in harmony

with the standard notation for the adjoints of operators, and it has typographical advantages. (The image of a set M of complex numbers under the mapping $z \rightarrow z^*$ is M^* ; the symbol \bar{M} suggests topological closure.)

For many years I have battled for proper values, and against the one and a half times translated German-English hybrid that is often used to refer to them. I have now become convinced that the war is over, and eigenvalues have won it; in this book I use them.

Since I have been teaching Hilbert space by the problem method for many years, I owe thanks for their help to more friends among students and colleagues than I could possibly name here. I am truly grateful to them all just the same. Without them this book could not exist; it is not the sort of book that could have been written in isolation from the mathematical community. My special thanks are due to Ronald Douglas, Eric Nordgren, and Carl Pearcy; each of them read the whole manuscript (well, almost the whole manuscript) and stopped me from making many foolish mistakes.

P. R. H.

The University of Michigan

Contents

Chapter	Page no. for		
	PROBLEM	HINT	SOLUTION
Preface	vii		
I. VECTORS AND SPACES			
1. Limits of quadratic forms	3	143	167
2. Representation of linear functionals	3	143	167
3. Strict convexity	4	143	168
4. Continuous curves	5	143	169
5. Linear dimension	5	143	170
6. Infinite Vandermondes	6	143	171
7. Approximate bases	6	143	172
8. Vector sums	7	143	174
9. Lattice of subspaces	7	143	175
10. Local compactness and dimension	8	143	175
11. Separability and dimension	8	143	176
12. Measure in Hilbert space	8	143	176
2. WEAK TOPOLOGY			
13. Weak closure of subspaces	10	144	178
14. Weak continuity of norm and inner product	11	144	178
15. Weak separability	12	144	179
16. Uniform weak convergence	12	144	179
17. Weak compactness of the unit ball	12	144	180
18. Weak metrizability of the unit ball	12	144	181
19. Weak metrizability and separability	12	144	183
20. Uniform boundedness	12	144	183
21. Weak metrizability of Hilbert space	13	144	185
22. Linear functionals on l^2	14	145	185
23. Weak completeness	14	145	186
3. ANALYTIC FUNCTIONS			
24. Analytic Hilbert spaces	15	145	187
25. Basis for \mathbf{A}^2	15	145	189

<i>Chapter</i>	<i>Page no. for</i>		
	PROBLEM	HINT	SOLUTION
26. Real functions in \mathbf{H}^2	15	145	190
27. Products in \mathbf{H}^2	17	145	191
28. Analytic characterization of \mathbf{H}^2	17	145	192
29. Functional Hilbert spaces	18	145	193
30. Kernel functions	19	146	193
31. Continuity of extension	19	146	194
32. Radial limits	19	146	195
33. Bounded approximation	20	146	196
34. Multiplicativity of extension	20	146	198
35. Dirichlet problem	20	146	198
 4. INFINITE MATRICES			
36. Column-finite matrices	21	146	201
37. Schur test	22	146	202
38. Hilbert matrix	23	146	203
 5. BOUNDEDNESS AND INVERTIBILITY			
39. Boundedness on bases	24	146	204
40. Uniform boundedness of linear transformations	24	147	204
41. Invertible transformations	25	147	205
42. Preservation of dimension	27	147	205
43. Projections of equal rank	27	147	206
44. Closed graph theorem	28	147	206
45. Unbounded symmetric transformations	28	147	207
 6. MULTIPLICATION OPERATORS			
46. Diagonal operators	29	147	209
47. Multiplications on l^2	29	147	209
48. Spectrum of a diagonal operator	30	147	210
49. Norm of a multiplication	31	147	210
50. Boundedness of multipliers	31	147	212
51. Boundedness of multiplications	31	147	213
52. Spectrum of a multiplication	32	148	214
53. Multiplications on functional Hilbert spaces	32	148	214
54. Multipliers of functional Hilbert spaces	33	148	217

<i>Chapter</i>	<i>Page no. for</i>		
	PROBLEM	HINT	SOLUTION
7. OPERATOR MATRICES			
55. Commutative operator determinants	34	148	218
56. Operator determinants	35	148	219
57. Operator determinants with a finite entry	36	149	223
8. PROPERTIES OF SPECTRA			
58. Spectra and conjugation	37	149	225
59. Spectral mapping theorem	38	149	225
60. Similarity and spectrum	38	149	226
61. Spectrum of a product	39	149	227
62. Closure of approximate point spectrum	39	149	227
63. Boundary of spectrum	39	149	228
9. EXAMPLES OF SPECTRA			
64. Residual spectrum of a normal operator	40	149	229
65. Spectral parts of a diagonal operator	40	149	229
66. Spectral parts of a multiplication	40	149	229
67. Unilateral shift	40	149	230
68. Bilateral shift	41	150	232
69. Spectrum of a functional multiplication	42	150	232
70. Relative spectrum of shift	43	150	233
71. Closure of relative spectrum	43	150	234
10. SPECTRAL RADIUS			
72. Analyticity of resolvents	44	150	236
73. Non-emptiness of spectra	44	150	237
74. Spectral radius	45	150	237
75. Weighted shifts	46	150	238
76. Similarity of weighted shifts	47	151	239
77. Norm and spectral radius of a weighted shift	48	151	239
78. Eigenvalues of weighted shifts	48	151	240
79. Weighted sequence spaces	48	151	241
80. One-point spectrum	49	151	242
81. Spectrum of a direct sum	50	151	243
82. Reid's inequality	51	151	244

<i>Chapter</i>	<i>Page no. for</i>		
	PROBLEM	HINT	SOLUTION
11. NORM TOPOLOGY			
83. Metric space of operators	52	151	245
84. Continuity of inversion	52	151	245
85. Continuity of spectrum	53	151	246
86. Semicontinuity of spectrum	53	151	247
87. Continuity of spectral radius	54	152	248
12. STRONG AND WEAK TOPOLOGIES			
88. Topologies for operators	55	152	250
89. Continuity of norm	56	152	250
90. Continuity of adjoint	56	152	251
91. Continuity of multiplication	56	152	252
92. Separate continuity of multiplication	57	152	253
93. Sequential continuity of multiplication	57	152	254
94. Increasing sequences of Hermitian operators	57	152	254
95. Square roots	58	152	256
96. Infimum of two projections	58	152	257
13. PARTIAL ISOMETRIES			
97. Spectral mapping theorem for normal operators	60	152	259
98. Partial isometries	62	153	259
99. Maximal partial isometries	63	153	260
100. Closure and connectedness of partial isometries	64	153	260
101. Rank, co-rank, and nullity	65	153	261
102. Components of the space of partial isometries	66	153	261
103. Unitary equivalence for partial isometries	66	153	262
104. Spectrum of a partial isometry	67	153	263
105. Polar decomposition	68	153	263
106. Maximal polar representation	69	153	264
107. Extreme points	69	153	265
108. Quasinormal operators	69	154	266
109. Density of invertible operators	70	154	266
110. Connectedness of invertible operators	70	154	267
14. UNILATERAL SHIFT			
111. Reducing subspaces of normal operators	71	154	268
112. Products of symmetries	71	154	269

<i>Chapter</i>	<i>Page no. for</i>		
	PROBLEM	HINT	SOLUTION
113. Unilateral shift versus normal operators	71	154	270
114. Square root of shift	72	154	271
115. Commutant of the bilateral shift	72	154	271
116. Commutant of the unilateral shift	73	154	272
117. Commutant of the unilateral shift as limit	74	155	273
118. Characterization of isometries	74	155	274
119. Distance from shift to unitary operators	74	155	275
120. Reduction by the unitary part	74	155	275
121. Shifts as universal operators	75	155	276
122. Similarity to parts of shifts	76	155	278
123. Wandering subspaces	77	155	279
124. Special invariant subspaces of the shift	78	155	280
125. Invariant subspaces of the shift	78	155	281
126. Cyclic vectors	79	155	282
127. The F. and M. Riesz theorem	82	156	283
128. The F. and M. Riesz theorem generalized	82	156	283
129. Reducible weighted shifts	82	156	284
15. COMPACT OPERATORS			
130. Mixed continuity	84	156	286
131. Compact operators	84	156	287
132. Diagonal compact operators	86	156	287
133. Normal compact operators	86	156	288
134. Kernel of the identity	87	156	288
135. Hilbert-Schmidt operators	87	157	290
136. Compact versus Hilbert-Schmidt	89	157	290
137. Limits of operators of finite rank	89	157	291
138. Ideals of operators	90	157	292
139. Square root of a compact operator	90	157	292
140. Fredholm alternative	90	157	293
141. Range of a compact operator	91	157	294
142. Atkinson's theorem	91	157	294
143. Weyl's theorem	91	157	295
144. Perturbed spectrum	92	157	295
145. Shift modulo compact operators	92	157	295

<i>Chapter</i>	<i>Page no. for</i>		
	PROBLEM	HINT	SOLUTION
146. Bounded Volterra kernels	93	158	296
147. Unbounded Volterra kernels	94	158	297
148. The Volterra integration operator	94	158	300
149. Skew-symmetric Volterra operator	95	158	301
150. Norm 1, spectrum $\{1\}$	95	158	302
151. Donoghue lattice	95	158	303
 16. SUBNORMAL OPERATORS			
152. The Putnam-Fuglede theorem	98	158	306
153. Spectral measure of the unit disc	99	158	307
154. Subnormal operators	100	159	307
155. Minimal normal extensions	101	159	308
156. Similarity of subnormal operators	102	159	308
157. Spectral inclusion theorem	102	159	309
158. Filling in holes	102	159	310
159. Extensions of finite co-dimension	103	159	310
160. Hyponormal operators	103	159	311
161. Normal and subnormal partial isometries	105	159	312
162. Norm powers and power norms	105	159	313
163. Compact hyponormal operators	106	160	313
164. Powers of hyponormal operators	106	160	314
165. Contractions similar to unitary operators	107	160	316
 17. NUMERICAL RANGE			
166. The Toeplitz-Hausdorff theorem	108	160	317
167. Higher-dimensional numerical range	110	160	318
168. Closure of numerical range	111	160	319
169. Spectrum and numerical range	111	160	320
170. Quasinilpotence and numerical range	112	160	320
171. Normality and numerical range	112	160	321
172. Subnormality and numerical range	113	161	321

<i>Chapter</i>	<i>Page no. for</i>		
	PROBLEM	HINT	SOLUTION
173. Numerical radius	114	161	322
174. Normaloid, convexoid, and spectraloid operators	114	161	322
175. Continuity of numerical range	115	161	323
176. Power inequality	115	161	324
18. UNITARY DILATIONS			
177. Unitary dilations	118	161	326
178. Unitary power dilations	119	161	327
179. Ergodic theorem	121	162	329
180. Spectral sets	122	162	330
181. Dilations of positive definite sequences	124	162	331
19. COMMUTATORS OF OPERATORS			
182. Commutators	126	162	333
183. Limits of commutators	127	162	334
184. The Kleinecke-Shirokov theorem	128	162	334
185. Distance from a commutator to the identity	129	162	336
186. Operators with large kernels	129	162	336
187. Direct sums as commutators	131	162	338
188. Positive self-commutators	132	163	339
189. Projections as self-commutators	132	163	339
190. Multiplicative commutators	133	163	340
191. Unitary multiplicative com- mutators	134	163	341
192. Commutator subgroup	134	163	342
20. TOEPLITZ OPERATORS			
193. Laurent operators and matrices	135	163	345
194. Toeplitz operators and matrices	135	163	345
195. Toeplitz products	137	163	347
196. Spectral inclusion theorem for Toeplitz operators	138	163	348
197. Analytic Toeplitz operators	140	164	350
198. Eigenvalues of Hermitian Toeplitz operators	140	164	350
199. Spectrum of a Hermitian Toeplitz operator	140	164	350
REFERENCES			352
INDEX			359

Problems

Chapter 1. Vectors and spaces

1. Limits of quadratic forms. The objects of chief interest in the study of a Hilbert space are not the vectors in the space, but the operators on it. Most people who say they study the theory of Hilbert spaces in fact study operator theory. The reason is that the algebra and geometry of vectors, linear functionals, quadratic forms, subspaces and the like are easier than operator theory and are pretty well worked out. Some of these easy and known things are useful and some are amusing; perhaps some are both.

Recall to begin with that a bilinear functional on a complex vector space \mathbf{H} is sometimes defined as a complex-valued function on the Cartesian product of \mathbf{H} with itself that is linear in its first argument and conjugate linear in the second; cf. Halmos [1951, p. 12]. Some mathematicians, in this context and in other more general ones, use “semi-linear” instead of “conjugate linear”, and, incidentally, “form” instead of “functional”. Since “sesqui” means “one and a half” in Latin, it has been suggested that a bilinear functional is more accurately described as a sesquilinear form.

A quadratic form is defined in Halmos [1951, p. 12] as a function φ^- associated with a sesquilinear form φ via the equation $\varphi^-(f) = \varphi(f, f)$. (The symbol $\hat{\varphi}$ is used there instead of φ^- .) More honestly put, a quadratic form is a function ψ for which *there exists* a sesquilinear form φ such that $\psi(f) = \varphi(f, f)$. Such an existential definition makes it awkward to answer even the simplest algebraic questions, such as whether or not the sum of two quadratic forms is a quadratic form (yes), and whether or not the product of two quadratic forms is a quadratic form (no).

Problem 1. *Is the limit of a sequence of quadratic forms a quadratic form?*

2. Representation of linear functionals. The Riesz representation theorem says that to each bounded linear functional ξ on a Hilbert space \mathbf{H} there corresponds a vector g in \mathbf{H} such that $\xi(f) = (f, g)$ for all f .

The statement is “invariant” or “coordinate-free”, and therefore, according to current mathematical ethics, it is mandatory that the proof be such. The trouble is that most coordinate-free proofs (such as the one in Halmos [1951, p. 32]) are so elegant that they conceal what is really going on.

Problem 2. Find a coordinatized proof of the Riesz representation theorem.

3. Strict convexity. In a real vector space (and hence, in particular, in a complex vector space) the *segment* joining two vectors f and g is, by definition, the set of all vectors of the form $tf + (1 - t)g$, where $0 \leq t \leq 1$. A subset of a real vector space is *convex* if, for each pair of vectors that it contains, it contains all the vectors of the segment joining them. Convexity plays an increasingly important role in modern vector space theory. Hilbert space is so rich in other, more powerful, structure, that the role of convexity is sometimes not so clearly visible in it as in other vector spaces. An easy example of a convex set in a Hilbert space is the *unit ball*, which is, by definition, the set of all vectors f with $\|f\| \leq 1$. Another example is the *open unit ball*, the set of all vectors f with $\|f\| < 1$. (The adjective “closed” can be used to distinguish the unit ball from its open version, but is in fact used only when unusual emphasis is necessary.) These examples are of geometric interest even in the extreme case of a (complex) Hilbert space of dimension 1; they reduce then to the closed and the open unit disc, respectively, in the complex plane.

If $h = tf + (1 - t)g$ is a point of the segment joining f and g , and if $0 < t < 1$ (the emphasis is that $t \neq 0$ and $t \neq 1$), then h is called an *interior* point of that segment. If a point of a convex set does not belong to the interior of any segment in the set, then it is called an *extreme point* of the set. The extreme points of the closed unit disc in the complex plane are just the points on its perimeter (the unit circle). The open unit disc in the complex plane has no extreme points. The set of all those complex numbers z for which $|\operatorname{Re} z| + |\operatorname{Im} z| \leq 1$ is convex (it consists of the interior and boundary of the square whose vertices are $1, i, -1$, and $-i$); this convex set has just four extreme points (namely $1, i, -1$, and $-i$).

A closed convex set in a Hilbert space is called *strictly convex* if all its boundary points are extreme points. The expression “boundary point” is used here in its ordinary topological sense. Unlike convexity, the concept of strict convexity is not purely algebraic. It makes sense in many spaces other than Hilbert spaces, but in order for it to make sense the space must have a topology, preferably one that is properly related to the linear structure. The closed unit disc in the complex plane is strictly convex.

Problem 3. *The unit ball of every Hilbert space is strictly convex.*

The problem is stated here to call attention to a circle of ideas and to prepare the ground for some later work. No great intrinsic interest is claimed for it; it is very easy.

4. Continuous curves. An infinite-dimensional Hilbert space is even roomier than it looks; a striking way to demonstrate its spaciousness is to study continuous curves in it. A *continuous curve* in a Hilbert space \mathbf{H} is a continuous function from the closed unit interval into \mathbf{H} ; the curve is *simple* if the function is one-to-one. The *chord* of the curve f determined by the parameter interval $[a, b]$ is the vector $f(b) - f(a)$. Two chords, determined by the intervals $[a, b]$ and $[c, d]$ are *non-overlapping* if the intervals $[a, b]$ and $[c, d]$ have at most an end-point in common. If two non-overlapping chords are orthogonal, then the curve makes a right-angle turn during the passage between their farthest end-points. If a curve could do so for every pair of non-overlapping chords, then it would seem to be making a sudden right-angle turn at each point, and hence, in particular, it could not have a tangent at any point.

Problem 4. *Construct, for every infinite-dimensional Hilbert space, a simple continuous curve with the property that every two non-overlapping chords of it are orthogonal.*

5. Linear dimension. The concept of dimension can mean two different things for a Hilbert space \mathbf{H} . Since \mathbf{H} is a vector space, it has a *linear* dimension; since \mathbf{H} has, in addition, an inner product structure, it has an *orthogonal* dimension. A unified way to approach the two con-

cepts is first to prove that all bases of \mathbf{H} have the same cardinal number, and then to define the dimension of \mathbf{H} as the common cardinal number of all bases; the difference between the two concepts is in the definition of basis. A *Hamel basis* for \mathbf{H} (also called a *linear basis*) is a maximal linearly independent subset of \mathbf{H} . (Recall that an infinite set is called linearly independent if each finite subset of it is linearly independent. It is true, but for present purposes irrelevant, that every vector is a finite linear combination of the vectors in any Hamel basis.) An *orthonormal basis* for \mathbf{H} is a maximal orthonormal subset of \mathbf{H} . (The analogues of the finite expansions appropriate to the linear theory are the Fourier expansions always used in Hilbert space.)

Problem 5. *Does there exist a Hilbert space whose linear dimension is \aleph_0 ?*

6. Infinite Vandermondes. The Hilbert space l^2 consists, by definition, of all infinite sequences $\langle \xi_0, \xi_1, \xi_2, \dots \rangle$ of complex numbers such that $\sum_{n=0}^{\infty} |\xi_n|^2 < \infty$. The vector operations are coordinatewise and the inner product is defined by

$$(\langle \xi_0, \xi_1, \xi_2, \dots \rangle, \langle \eta_0, \eta_1, \eta_2, \dots \rangle) = \sum_{n=0}^{\infty} \xi_n \eta_n^*.$$

Problem 6. *If $0 < |\alpha| < 1$, and if*

$$f_k = \langle 1, \alpha^k, \alpha^{2k}, \alpha^{3k}, \dots \rangle, \quad k = 1, 2, 3, \dots,$$

determine the span of the set of all f_k 's in l^2 . Generalize (to other collections of vectors), and specialize (to finite-dimensional spaces).

7. Approximate bases.

Problem 7. *If $\{e_1, e_2, e_3, \dots\}$ is an orthonormal basis for a Hilbert space \mathbf{H} , and if $\{f_1, f_2, f_3, \dots\}$ is an orthonormal set in \mathbf{H} such that*

$$\sum_{j=1}^{\infty} \|e_j - f_j\|^2 < \infty,$$

then the vectors f_j span \mathbf{H} (and hence form an orthonormal basis for \mathbf{H}).

This is a hard one. There are many problems of this type; the first one is apparently due to Paley and Wiener. For a related exposition, and detailed references, see Riesz-Nagy [1952, No. 86]. The version above is discussed by Birkhoff-Rota [1960].

8. Vector sums. If \mathbf{M} and \mathbf{N} are orthogonal subspaces of a Hilbert space, then $\mathbf{M} + \mathbf{N}$ is closed (and therefore $\mathbf{M} + \mathbf{N} = \mathbf{M} \vee \mathbf{N}$). Orthogonality may be too strong an assumption, but it is sufficient to ensure the conclusion. It is known that something is necessary; if no additional assumptions are made, then $\mathbf{M} + \mathbf{N}$ need not be closed (see Halmos [1951, p. 28], and Problem 41 below). Here is the conclusion under another very strong but frequently usable additional assumption.

Problem 8. *If \mathbf{M} is a finite-dimensional linear manifold in a Hilbert space \mathbf{H} , and if \mathbf{N} is a subspace (a closed linear manifold) in \mathbf{H} , then the vector sum $\mathbf{M} + \mathbf{N}$ is necessarily closed (and is therefore equal to the span $\mathbf{M} \vee \mathbf{N}$).*

The result has the corollary (which it is also easy to prove directly) that every finite-dimensional linear manifold is closed; just put $\mathbf{N} = \{0\}$.

9. Lattice of subspaces. The collection of all subspaces of a Hilbert space is a *lattice*. This means that the collection is partially ordered (by inclusion), and that any two elements \mathbf{M} and \mathbf{N} of it have a least upper bound or supremum (namely the span $\mathbf{M} \vee \mathbf{N}$) and a greatest lower bound or infimum (namely the intersection $\mathbf{M} \cap \mathbf{N}$). A lattice is called *distributive* if (in the notation appropriate to subspaces)

$$\mathbf{L} \cap (\mathbf{M} \vee \mathbf{N}) = (\mathbf{L} \cap \mathbf{M}) \vee (\mathbf{L} \cap \mathbf{N})$$

identically in \mathbf{L} , \mathbf{M} , and \mathbf{N} .

There is a weakening of this distributivity condition, called modularity; a lattice is called *modular* if the distributive law, as written above, holds at least when $\mathbf{N} \subset \mathbf{L}$. In that case, of course, $\mathbf{L} \cap \mathbf{N} = \mathbf{N}$, and the identity becomes

$$\mathbf{L} \cap (\mathbf{M} \vee \mathbf{N}) = (\mathbf{L} \cap \mathbf{M}) \vee \mathbf{N}$$

(with the proviso $\mathbf{N} \subset \mathbf{L}$ still in force).

Since a Hilbert space is geometrically indistinguishable from any other Hilbert space of the same dimension, it is clear that the modularity or distributivity of its lattice of subspaces can depend on its dimension only.

Problem 9. *For which cardinal numbers m is the lattice of subspaces of a Hilbert space of dimension m modular? distributive?*

10. Local compactness and dimension. Many global topological questions are easy to answer for Hilbert space. The answers either are a simple yes or no, or depend on the dimension. Thus, for instance, every Hilbert space is connected, but a Hilbert space is compact if and only if it is the trivial space with dimension 0. The same sort of problem could be posed backwards: given some information about the dimension of a Hilbert space (e.g., that it is finite), find topological properties that distinguish such a space from Hilbert spaces of all other dimensions. Such problems sometimes have useful and elegant solutions.

Problem 10. *A Hilbert space is locally compact if and only if it is finite-dimensional.*

11. Separability and dimension.

Problem 11. *A Hilbert space H is separable if and only if $\dim H \leq \aleph_0$.*

12. Measure in Hilbert space. Infinite-dimensional Hilbert spaces are properly regarded as the most successful infinite-dimensional generalizations of finite-dimensional Euclidean spaces. Finite-dimensional Euclidean spaces have, in addition to their algebraic and topological structure, a measure; it might be useful to generalize that too to infinite dimensions. Various attempts have been made to do so (see Löwner [1939] and Segal [1965]). The unsophisticated approach is to seek a countably additive set function μ defined on (at least) the collection of all Borel sets (the σ -field generated by the open sets), so that $0 \leq \mu(M) \leq \infty$ for all Borel sets M . (Warning: the parenthetical definition of Borel sets in the preceding sentence is not the same as the

one in Halmos [1950 b].) In order that μ be suitably related to the other structure of the space, it makes sense to require that every open set have positive measure and that measure be invariant under translation. (The second condition means that $\mu(f + M) = \mu(M)$ for every vector f and for every Borel set M .) If, for now, the word “measure” is used to describe a set function satisfying just these conditions, then the following problem indicates that the unsophisticated approach is doomed to fail.

Problem 12. *For each measure in an infinite-dimensional Hilbert space, the measure of every non-empty ball is infinite.*

Chapter 2. Weak topology

13. Weak closure of subspaces. A Hilbert space is a metric space, and, as such, it is a topological space. The metric topology (or norm topology) of a Hilbert space is often called the *strong* topology. A base for the strong topology is the collection of open balls, i.e., sets of the form

$$\{f: \|f - f_0\| < \varepsilon\},$$

where f_0 (the center) is a vector and ε (the radius) is a positive number.

Another topology, called the *weak* topology, plays an important role in the theory of Hilbert spaces. A subbase (not a base) for the weak topology is the collection of all sets of the form

$$\{f: |(f - f_0, g_0)| < \varepsilon\}.$$

It follows that a base for the weak topology is the collection of all sets of the form

$$\{f: |(f - f_0, g_i)| < \varepsilon, \quad i = 1, \dots, k\},$$

where k is a positive integer, f_0, g_1, \dots, g_k are vectors, and ε is a positive number.

Facts about these topologies are described by the grammatically appropriate use of “weak” and “strong”. Thus, for instance, a function may be described as weakly continuous, or a sequence as strongly convergent; the meanings of such phrases should be obvious. The use of a topological word without a modifier always refers to the strong topology; this convention has already been observed in the preceding problems.

Whenever a set is endowed with a topology, many technical questions automatically demand attention. (Which separation axioms does the space satisfy? Is it compact? Is it connected?) If a large class of sets is in sight (for example, the class of all Hilbert spaces), then classification problems arise. (Which ones are locally compact? Which ones are

separable?) If the set (or sets) already had some structure, the connection between the old structure and the new topology should be investigated. (Is the closed unit ball compact? Are inner products continuous?) If, finally, more than one topology is considered, then the relations of the topologies to one another must be clarified. (Is a weakly compact set strongly closed?) Most such questions, though natural, and, in fact, unavoidable, are not likely to be inspiring; for that reason most such questions do not appear below. The questions that do appear justify their appearance by some (perhaps subjective) test, such as a surprising answer, a tricky proof, or an important application.

Problem 13. *Every weakly closed set is strongly closed, but the converse is not true. Nevertheless every subspace of a Hilbert space (i.e., every strongly closed linear manifold) is weakly closed.*

14. Weak continuity of norm and inner product. For each fixed vector g , the function $f \rightarrow (f, g)$ is weakly continuous; this is practically the definition of the weak topology. (A sequence, or a net, $\{f_n\}$ is weakly convergent to f if and only if $(f_n, g) \rightarrow (f, g)$ for each g .) This, together with the (Hermitian) symmetry of the inner product, implies that, for each fixed vector f , the function $g \rightarrow (f, g)$ is weakly continuous. These two assertions between them say that the mapping from ordered pairs $\langle f, g \rangle$ to their inner product (f, g) is separately weakly continuous in each of its two variables.

It is natural to ask whether the mapping is weakly continuous jointly in its two variables, but it is easy to see that the answer is no. A counterexample has already been seen, in Solution 13; it was used there for a slightly different purpose. If $\{e_1, e_2, e_3, \dots\}$ is an orthonormal sequence, then $e_n \rightarrow 0$ (weak), but $(e_n, e_n) = 1$ for all n . This example shows at the same time that the norm is not weakly continuous. It could, in fact, be said that the possible discontinuity of the norm is the only difference between weak convergence and strong convergence: a weakly convergent sequence (or net) on which the norm behaves itself is automatically strongly convergent.

Problem 14. *If $f_n \rightarrow f$ (weak) and $\|f_n\| \rightarrow \|f\|$, then $f_n \rightarrow f$ (strong).*

15. Weak separability. Since the strong closure of every set is included in its weak closure (see Solution 13), it follows that if a Hilbert space is separable (that is, strongly separable), then it is weakly separable. What about the converse?

Problem 15. *Is every weakly separable Hilbert space separable?*

16. Uniform weak convergence.

Problem 16. *Strong convergence is the same as weak convergence uniformly on the unit sphere. Precisely: $\|f_n - f\| \rightarrow 0$ if and only if $(f_n, g) \rightarrow (f, g)$ uniformly for $\|g\| = 1$.*

17. Weak compactness of the unit ball.

Problem 17. *The closed unit ball in a Hilbert space is weakly compact.*

The result is sometimes known as the Tychonoff-Alaoglu theorem. It is as hard as it is important. It is very important.

18. Weak metrizability of the unit ball. Compactness is good, but even compact sets are better if they are metric. Once the unit ball is known to be weakly compact, it is natural to ask if it is weakly metrizable also.

Problem 18. *Is the weak topology of the unit ball in a separable Hilbert space metrizable?*

19. Weak metrizability and separability.

Problem 19. *If the weak topology of the unit ball in a Hilbert space H is metrizable, must H be separable?*

20. Uniform boundedness. The celebrated “principle of uniform boundedness” (true for all Banach spaces) is the assertion that a point-wise bounded collection of bounded linear functionals is bounded. The assumption and the conclusion can be expressed in the terminology

appropriate to a Hilbert space \mathbf{H} , as follows. The assumption of pointwise boundedness for a subset \mathbf{T} of \mathbf{H} could also be called *weak boundedness*; it means that for each f in \mathbf{H} there exists a positive constant $\alpha(f)$ such that $|(f,g)| \leq \alpha(f)$ for all g in \mathbf{T} . The desired conclusion means that there exists a positive constant β such that $|(f,g)| \leq \beta \|f\|$ for all f in \mathbf{H} and all g in \mathbf{T} ; this conclusion is equivalent to $\|g\| \leq \beta$ for all g in \mathbf{T} . It is clear that every bounded subset of a Hilbert space is weakly bounded. The principle of uniform boundedness (for vectors in a Hilbert space) is the converse: every weakly bounded set is bounded. The proof of the general principle is a mildly involved category argument. A standard reference for a general treatment of the principle of uniform boundedness is Dunford-Schwartz [1958, p. 49].

Problem 20. *Find an elementary proof of the principle of uniform boundedness for Hilbert space.*

(In this context a proof is “elementary” if it does not use the Baire category theorem.)

A frequently used corollary of the principle of uniform boundedness is the assertion that a weakly convergent sequence must be bounded. The proof is completely elementary: since convergent sequences of numbers are bounded, it follows that a weakly convergent sequence of vectors is weakly bounded. Nothing like this is true for nets, of course. One easy generalization of the sequence result that is available is that every weakly compact set is bounded. Reason: for each f , the map $g \rightarrow (f,g)$ sends the g 's in a weakly compact set onto a compact and therefore bounded set of numbers, so that a weakly compact set is weakly bounded.

21. Weak metrizability of Hilbert space. Some of the preceding results, notably the weak compactness of the unit ball and the principle of uniform boundedness, show that for bounded sets the weak topology is well behaved. For unbounded sets it is not.

Problem 21. *The weak topology of an infinite-dimensional Hilbert space is not metrizable.*

The shortest proof of this is tricky.

22. Linear functionals on l^2 . If

$$\langle \alpha_1, \alpha_2, \alpha_3, \dots \rangle \in l^2 \quad \text{and} \quad \langle \beta_1, \beta_2, \beta_3, \dots \rangle \in l^2,$$

then

$$\langle \alpha_1\beta_1, \alpha_2\beta_2, \alpha_3\beta_3, \dots \rangle \in l^1.$$

The following assertion is a kind of converse; it says that l^2 sequences are the only ones whose product with every l^2 sequence is in l^1 .

Problem 22. If $\sum_n |\alpha_n\beta_n| < \infty$ whenever $\sum_n |\alpha_n|^2 < \infty$, then $\sum_n |\beta_n|^2 < \infty$.

23. Weak completeness. A sequence $\{g_n\}$ of vectors in a Hilbert space is a *weak Cauchy sequence* if (surely this definition is guessable) the numerical sequence $\{(f, g_n)\}$ is a Cauchy sequence for each f in the space. *Weak Cauchy nets* are defined exactly the same way: just replace “sequence” by “net” throughout. To say of a Hilbert space, or a subset of one, that it is *weakly complete* means that every weak Cauchy net has a weak limit (in the set under consideration). If the conclusion is known to hold for sequences only, the space is called *sequentially weakly complete*.

Problem 23. (a) No infinite-dimensional Hilbert space is weakly complete. (b) Which Hilbert spaces are sequentially weakly complete?

Chapter 3. Analytic functions

24. Analytic Hilbert spaces. Analytic functions enter Hilbert space theory in several ways; one of their roles is to provide illuminating examples. The typical way to construct these examples is to consider a region D (“region” means a non-empty open connected subset of the complex plane), let μ be planar Lebesgue measure in D , and let $\mathbf{A}^2(D)$ be the set of all complex-valued functions that are analytic throughout D and square-integrable with respect to μ . The most important special case is the one in which D is the open unit disc, $D = \{z: |z| < 1\}$; the corresponding function space will be denoted simply by \mathbf{A}^2 . No matter what D is, the set $\mathbf{A}^2(D)$ is a vector space with respect to pointwise addition and scalar multiplication. It is also an inner-product space with respect to the inner product defined by

$$(f, g) = \int_D f(z) \overline{g(z)} d\mu(z).$$

Problem 24. *Is the space $\mathbf{A}^2(D)$ of square-integrable analytic functions on a region D a Hilbert space, or does it have to be completed before it becomes one?*

25. Basis for \mathbf{A}^2 .

Problem 25. *If $e_n(z) = \sqrt{(n+1)/\pi} z^n$ for $|z| < 1$ and $n = 0, 1, 2, \dots$, then the e_n 's form an orthonormal basis for \mathbf{A}^2 . If $f \in \mathbf{A}^2$, with Taylor series $\sum_{n=0}^{\infty} \alpha_n z^n$, then $\alpha_n = \sqrt{(n+1)/\pi} (f, e_n)$ for $n = 0, 1, 2, \dots$.*

26. Real functions in \mathbf{H}^2 . Except for size (dimension) one Hilbert space is very like another. To make a Hilbert space more interesting than its neighbors, it is necessary to enrich it by the addition of some external structure. Thus, for instance, the spaces $\mathbf{A}^2(D)$ are of interest because of the analytic properties of their elements. Another important

Hilbert space, known as \mathbf{H}^2 (\mathbf{H} is for Hardy this time), endowed with some structure not usually found in a Hilbert space, is defined as follows.

Let C be the unit circle (that means circumference) in the complex plane, $C = \{z: |z| = 1\}$, and let μ be Lebesgue measure (the extension of arc length) on the Borel sets of C , normalized so that $\mu(C) = 1$ (instead of $\mu(C) = 2\pi$). If $e_n(z) = z^n$ for $|z| = 1$ ($n = 0, \pm 1, \pm 2, \dots$), then, by elementary calculus, the functions e_n form an orthonormal set in $L^2(\mu)$; it is an easy consequence of standard approximation theorems (e.g., the Weierstrass theorem on approximation by polynomials) that the e_n 's form an orthonormal basis for L^2 . (Finite linear combinations of the e_n 's are called trigonometric polynomials.) The space \mathbf{H}^2 is, by definition, the subspace of L^2 spanned by the e_n 's with $n \geq 0$; equivalently \mathbf{H}^2 is the orthogonal complement in L^2 of $\{e_{-1}, e_{-2}, e_{-3}, \dots\}$. A related space, playing a role dual to that of \mathbf{H}^2 , is the span of the e_n 's with $n \leq 0$; it will be denoted by \mathbf{H}^{2*} .

Fourier expansions with respect to the orthonormal basis $\{e_n: n = 0, \pm 1, \pm 2, \dots\}$ are formally similar to the Laurent expansions that occur in analytic function theory. The analogy motivates calling the functions in \mathbf{H}^2 the *analytic* elements of L^2 ; the elements of \mathbf{H}^{2*} are called *co-analytic*. A subset of \mathbf{H}^2 (a linear manifold but not a subspace) of considerable technical significance is the set \mathbf{H}^∞ of bounded functions in \mathbf{H}^2 ; equivalently, \mathbf{H}^∞ is the set of all those functions in L^∞ for which $\int f e_n^* d\mu = 0$ ($n = -1, -2, -3, \dots$). Similarly \mathbf{H}^1 is the set of all those elements f of L^1 for which these same equations hold. What gives \mathbf{H}^1 , \mathbf{H}^2 , and \mathbf{H}^∞ their special flavor is the structure of the semigroup of non-negative integers within the additive group of all integers.

It is customary to speak of the elements of spaces such as \mathbf{H}^1 , \mathbf{H}^2 , and \mathbf{H}^∞ as functions, and this custom was followed in the preceding paragraph. The custom is not likely to lead its user astray, as long as the qualification "almost everywhere" is kept in mind at all times. Thus "bounded" means "essentially bounded", and, similarly, all statements such as " $f = 0$ " or " f is real" or " $|f| = 1$ " are to be interpreted, when asserted, as holding almost everywhere.

Some authors define the Hardy spaces so as to make them honest function spaces (consisting of functions analytic on the unit disc). In that approach (see Problem 28) the almost everywhere difficulties are still present, but they are pushed elsewhere; they appear in questions

(which must be asked and answered) about the limiting behavior of the functions on the boundary.

Independently of the approach used to study them, the functions in \mathbf{H}^2 are anxious to behave like analytic functions. The following statement is evidence in that direction.

Problem 26. *If f is a real function in \mathbf{H}^2 , then f is a constant.*

27. Products in \mathbf{H}^2 . The deepest statements about the Hardy spaces have to do with their multiplicative structure. The following one is an easily accessible sample.

Problem 27. *The product of two functions in \mathbf{H}^2 is in \mathbf{H}^1 .*

A kind of converse of this statement is true: it says that every function in \mathbf{H}^1 is the product of two functions in \mathbf{H}^2 . (See Hoffman [1962, p. 52].) The direct statement is more useful in Hilbert space theory than the converse, and the techniques used in the proof of the direct statement are nearer to the ones appropriate to this book.

28. Analytic characterization of \mathbf{H}^2 . If $f \in \mathbf{H}^2$, with Fourier expansion $f = \sum_{n=0}^{\infty} \alpha_n e_n$, then $\sum_{n=0}^{\infty} |\alpha_n|^2 < \infty$, and therefore the radius of convergence of the power series $\sum_{n=0}^{\infty} \alpha_n z^n$ is greater than or equal to 1. It follows from the usual expression for the radius of convergence in terms of the coefficients that the power series $\sum_{n=0}^{\infty} \alpha_n z^n$ defines an analytic function \tilde{f} in the open unit disc D . The mapping $f \rightarrow \tilde{f}$ (obviously linear) establishes a one-to-one correspondence between \mathbf{H}^2 and the set $\tilde{\mathbf{H}}^2$ of those functions analytic in D whose series of Taylor coefficients is square-summable.

Problem 28. *If φ is an analytic function in the open unit disc, $\varphi(z) = \sum_{n=0}^{\infty} \alpha_n z^n$, and if $\varphi_r(z) = \varphi(rz)$ for $0 < r < 1$ and $|z| = 1$, then $\varphi_r \in \mathbf{H}^2$ for each r ; the series $\sum_{n=0}^{\infty} |\alpha_n|^2$ converges if and only if the norms $\|\varphi_r\|$ are bounded.*

Many authors define \mathbf{H}^2 to be $\tilde{\mathbf{H}}^2$; for them, that is, \mathbf{H}^2 consists of analytic functions in the unit disc with square-summable Taylor series,

or, equivalently, with bounded concentric L^2 norms. If φ and ψ are two such functions, with $\varphi(z) = \sum_{n=0}^{\infty} \alpha_n z^n$ and $\psi(z) = \sum_{n=0}^{\infty} \beta_n z^n$, then the inner product (φ, ψ) is defined to be $\sum_{n=0}^{\infty} \alpha_n \beta_n^*$. In view of the one-to-one correspondence $f \rightarrow \tilde{f}$ between H^2 and \tilde{H}^2 , it all comes to the same thing. If $f \in H^2$, its image \tilde{f} in \tilde{H}^2 may be spoken of as the *extension* of f into the interior (cf. Solution 32). Since H^∞ is included in H^2 , this concept makes sense for elements of H^∞ also; the set of all their extensions will be denoted by \tilde{H}^∞ .

29. Functional Hilbert spaces. Many of the popular examples of Hilbert spaces are called function spaces, but they are not. If a measure space has a non-empty set of measure zero (and this is usually the case), then the L^2 space over it consists not of functions, but of equivalence classes of functions modulo sets of measure zero, and there is no natural way to identify such equivalence classes with representative elements. There is, however, a class of examples of Hilbert spaces whose elements are bona fide functions; they will be called functional Hilbert spaces. A *functional Hilbert space* is a Hilbert space H of complex-valued functions on a (non-empty) set X ; the Hilbert space structure of H is related to X in two ways (the only two natural ways it could be). It is required that (1) if f and g are in H and if α and β are scalars, then $(\alpha f + \beta g)(x) = \alpha f(x) + \beta g(x)$ for each x in X , i.e., the evaluation functionals on H are linear, and (2) to each x in X there corresponds a positive constant γ_x , such that $|f(x)| \leq \gamma_x \|f\|$ for all f in H , i.e., the evaluation functionals on H are bounded. The usual sequence spaces are trivial examples of functional Hilbert spaces (whether the length of the sequences is finite or infinite); the role of X is played by the index set. More typical examples of functional Hilbert spaces are the spaces A^2 and \tilde{H}^2 of analytic functions.

There is a trivial way of representing every Hilbert space as a functional one. Given H , write $X = H$, and let \tilde{H} be the set of all those functions f on X ($= H$) that are bounded conjugate-linear functionals. There is a natural correspondence $f \rightarrow \tilde{f}$ from H to \tilde{H} , defined by $\tilde{f}(g) = (f, g)$ for all g in X . By the Riesz representation theorem the correspondence is one-to-one; since (f, g) depends linearly on f , the correspondence is linear. Write, by definition, $(\tilde{f}, \tilde{g}) = (f, g)$ (whence, in particular, $\|\tilde{f}\| = \|f\|$); it follows that \tilde{H} is a Hilbert space. Since

$|(\tilde{f}(g))| = |(f,g)| \leq \|f\| \cdot \|g\| = \|\tilde{f}\| \cdot \|g\|$, it follows that $\tilde{\mathbf{H}}$ is a functional Hilbert space. The correspondence $f \rightarrow \tilde{f}$ between \mathbf{H} and $\tilde{\mathbf{H}}$ is a Hilbert space isomorphism.

Problem 29. Give an example of a Hilbert space of functions such that the vector operations are pointwise, but not all the evaluation functionals are bounded.

An early and still useful reference for functional Hilbert spaces is Aronszajn [1950].

30. Kernel functions. If \mathbf{H} is a functional Hilbert space, over X say, then the linear functional $f \rightarrow f(y)$ on \mathbf{H} is bounded for each y in X , and, consequently, there exists, for each y in X , an element K_y of \mathbf{H} such that $f(y) = (f, K_y)$ for all f . The function K on $X \times X$, defined by $K(x, y) = K_y(x)$, is called the *kernel function* or the *reproducing kernel* of \mathbf{H} .

Problem 30. If $\{e_j\}$ is an orthonormal basis for a functional Hilbert space \mathbf{H} , then the kernel function K of \mathbf{H} is given by

$$K(x, y) = \sum_j e_j(x) e_j(y)^*.$$

What are the kernel functions of \mathbf{A}^2 and of $\tilde{\mathbf{H}}^2$?

The kernel functions of \mathbf{A}^2 and of $\tilde{\mathbf{H}}^2$ are known, respectively, as the *Bergman kernel* and the *Szegő kernel*.

31. Continuity of extension.

Problem 31. The extension mapping $f \rightarrow \tilde{f}$ (from \mathbf{H}^2 to $\tilde{\mathbf{H}}^2$) is continuous not only in the Hilbert space sense, but also in the sense appropriate to analytic functions. That is: if $f_n \rightarrow f$ in \mathbf{H}^2 , then $\tilde{f}_n(z) \rightarrow \tilde{f}(z)$ for $|z| < 1$, and, in fact, the convergence is uniform on each disc $\{z: |z| \leq r\}$, $0 < r < 1$.

32. Radial limits.

Problem 32. *If an element f of \mathbf{H}^2 is such that the corresponding analytic function \tilde{f} in $\tilde{\mathbf{H}}^2$ is bounded, then f is bounded, (i.e., $f \in \mathbf{H}^\infty$).*

33. Bounded approximation.

Problem 33. *If $f \in \mathbf{H}^\infty$, does it follow that \tilde{f} is bounded?*

34. Multiplicativity of extension.

Problem 34. *Is the mapping $f \rightarrow \tilde{f}$ multiplicative?*

35. Dirichlet problem.

Problem 35. *To each real function u in \mathbf{L}^2 there corresponds a unique real function v in \mathbf{L}^2 such that $(v, e_0) = 0$ and such that $u + iv \in \mathbf{H}^2$. Equivalently, to each u in \mathbf{L}^2 there corresponds a unique f in \mathbf{H}^2 such that (f, e_0) is real and such that $\operatorname{Re} f = u$.*

The relation between u and v is expressed by saying that they are *conjugate functions*; alternatively, v is the *Hilbert transform* of u .

Chapter 4. Infinite matrices

36. Column-finite matrices. Many problems about operators on finite-dimensional spaces can be solved with the aid of matrices; matrices reduce qualitative geometric statements to explicit algebraic computations. Not much of matrix theory carries over to infinite-dimensional spaces, and what does is not so useful, but it sometimes helps.

Suppose that $\{e_j\}$ is an orthonormal basis for a Hilbert space \mathbf{H} . If A is an operator on \mathbf{H} , then each Ae_j has a Fourier expansion,

$$Ae_j = \sum_i \alpha_{ij} e_i;$$

the entries of the matrix that arises this way are given by

$$\alpha_{ij} = (Ae_j, e_i).$$

The index set is arbitrary here; it does not necessarily consist of positive integers. Familiar words (such as row, column, diagonal) can nevertheless be used in their familiar senses. Note that if, as usual, the first index indicates rows and the second one columns, then the matrix is formed by writing the coefficients in the expansion of Ae_j as the j column.

The correspondence from operators to matrices (induced by a fixed basis) has the usual algebraic properties. The zero matrix and the unit matrix are what they ought to be, the linear operations on matrices are the obvious ones, adjoint corresponds to conjugate transpose, and operator multiplication corresponds to the matrix product defined by the familiar formula

$$\gamma_{ij} = \sum_k \alpha_{ik} \beta_{kj}.$$

There are several ways of showing that these sums do not run into convergence trouble; here is one. Since $\alpha_{ik} = (e_k, A^* e_i)$, it follows that for each fixed i the family $\{\alpha_{ik}\}$ is square-summable; since, similarly, $\beta_{kj} = (Be_j, e_k)$, it follows that for each fixed j the family $\{\beta_{kj}\}$ is square-

summable. Conclusion (via the Schwarz inequality): for fixed i and j the family $\{\alpha_{ik}\beta_{kj}\}$ is (absolutely) summable.

It follows from the preceding paragraph that each row and each column of the matrix of each operator is square-summable. These are necessary conditions on a matrix in order that it arise from an operator; they are not sufficient. (Example: the diagonal matrix whose n -th diagonal term is n .) A sufficient condition of the same kind is that the family of all entries be square-summable; if, that is, $\sum_i \sum_j |\alpha_{ij}|^2 < \infty$, then there exists an operator A such that $\alpha_{ij} = (Ae_j, e_i)$. (Proof: since $|\sum_j \alpha_{ij}(f, e_j)|^2 \leq \sum_j |\alpha_{ij}|^2 \cdot \|f\|^2$ for each i and each f , it follows that $\|\sum_i (\sum_j \alpha_{ij}(f, e_j))e_i\|^2 \leq \sum_i \sum_j |\alpha_{ij}|^2 \cdot \|f\|^2$.) This condition is not necessary. (Example: the unit matrix.) There are no elegant and usable necessary and sufficient conditions. It is perfectly possible, of course, to write down in matricial terms the condition that a linear transformation is everywhere defined and bounded, but the result is neither elegant nor usable. This is the first significant way in which infinite matrix theory differs from the finite version: every operator corresponds to a matrix, but not every matrix corresponds to an operator, and it is hard to say which ones do.

As long as there is a fixed basis in the background, the correspondence from operators to matrices is one-to-one; as soon as the basis is allowed to vary, one operator may be assigned many matrices. An enticing game is to choose the basis so as to make the matrix as simple as possible. Here is a sample theorem, striking but less useful than it looks.

Problem 36. *Every operator has a column-finite matrix. More precisely, if A is an operator on a Hilbert space \mathbf{H} , then there exists an orthonormal basis $\{e_j\}$ for \mathbf{H} such that, for each j , the matrix entry (Ae_j, e_i) vanishes for all but finitely many i 's.*

Reference: Toeplitz [1910].

37. Schur test. While the algebra of infinite matrices is more or less reasonable, the analysis is not. Questions about norms and spectra are likely to be recalcitrant. Each of the few answers that is known is considered a respectable mathematical accomplishment. The following result (due in substance to Schur) is an example.

Problem 37. If $\alpha_{ij} \geq 0$ ($i, j = 0, 1, 2, \dots$), if $p_i > 0$ ($i = 0, 1, 2, \dots$), and if β and γ are positive numbers such that

$$\sum_i \alpha_{ij} p_i \leq \beta p_j \quad (j = 0, 1, 2, \dots),$$

$$\sum_j \alpha_{ij} p_j \leq \gamma p_i \quad (i = 0, 1, 2, \dots),$$

then there exists an operator A (on a separable infinite-dimensional Hilbert space, of course) with $\|A\|^2 \leq \beta\gamma$ and with matrix $\langle \alpha_{ij} \rangle$ (with respect to a suitable orthonormal basis).

For a related result, and a pertinent reference, see Problem 135.

38. Hilbert matrix.

Problem 38. There exists an operator A (on a separable infinite-dimensional Hilbert space) with $\|A\| \leq \pi$ and with matrix $\langle 1/(i+j+1) \rangle$ ($i, j = 0, 1, 2, \dots$).

The matrix is named after Hilbert; the norm of the matrix is in fact equal to π (Hardy-Littlewood-Pólya [1934, p. 226]).

Chapter 5.

Boundedness and invertibility

39. Boundedness on bases. Boundedness is a useful and natural condition, but it is a very strong condition on a linear transformation. The condition has a profound effect throughout operator theory, from its mildest algebraic aspects to its most complicated topological ones. To avoid certain obvious mistakes, it is important to know that boundedness is more than just the conjunction of an infinite number of conditions, one for each element of a basis. If A is an operator on a Hilbert space \mathbf{H} with an orthonormal basis $\{e_1, e_2, e_3, \dots\}$, then the numbers $\|Ae_n\|$ are bounded; if, for instance, $\|A\| \leq 1$, then $\|Ae_n\| \leq 1$ for all n ; and, of course, if $A = 0$, then $Ae_n = 0$ for all n . The obvious mistakes just mentioned are based on the assumption that the converses of these assertions are true.

Problem 39. *Give an example of an unbounded linear transformation that is bounded on a basis; give examples of operators of arbitrarily large norms that are bounded by 1 on a basis; and give an example of an unbounded linear transformation that annihilates a basis.*

40. Uniform boundedness of linear transformations. Sometimes linear transformations between two Hilbert spaces play a role even when the center of the stage is occupied by operators on one Hilbert space. Much of the two-space theory is an easy adaptation of the one-space theory.

If \mathbf{H} and \mathbf{K} are Hilbert spaces, a linear transformation A from \mathbf{H} into \mathbf{K} is *bounded* if there exists a positive number α such that $\|Af\| \leq \alpha \|f\|$ for all f in \mathbf{H} ; the *norm* of A , in symbols $\|A\|$, is the infimum of all such values of α . Given a bounded linear transformation A , the inner product (Af, g) makes sense whenever f is in \mathbf{H} and g is in \mathbf{K} ; the inner product is formed in \mathbf{K} . For fixed g the inner product defines a bounded linear functional of f , and, consequently, it is identically equal to (f, \tilde{g}) for some \tilde{g} in \mathbf{H} . The mapping from g to \tilde{g} is the *adjoint* of A ;

it is a bounded linear transformation A^* from \mathbf{K} into \mathbf{H} . By definition

$$(Af, g) = (f, A^*g)$$

whenever $f \in \mathbf{H}$ and $g \in \mathbf{K}$; here the left inner product is formed in \mathbf{K} and the right one in \mathbf{H} . The algebraic properties of this kind of adjoint can be stated and proved the same way as for the classical kind. An especially important (but no less easily proved) connection between A and A^* is that the orthogonal complement of the range of A is equal to the kernel of A^* ; since $A^{**} = A$, this assertion remains true with A and A^* interchanged.

All these algebraic statements are trivialities; the generalization of the principle of uniform boundedness from linear functionals to linear transformations is somewhat subtler. The generalization can be formulated almost exactly the same way as the special case: a pointwise bounded collection of bounded linear transformations is uniformly bounded. The assumption of pointwise boundedness can be formulated in a "weak" manner and a "strong" one. A set \mathbf{Q} of linear transformations (from \mathbf{H} into \mathbf{K}) is *weakly bounded* if for each f in \mathbf{H} and each g in \mathbf{K} there exists a positive constant $\alpha(f, g)$ such that $|(Af, g)| \leq \alpha(f, g)$ for all A in \mathbf{Q} . The set \mathbf{Q} is *strongly bounded* if for each f in \mathbf{H} there exists a positive constant $\beta(f)$ such that $\|Af\| \leq \beta(f)$ for all A in \mathbf{Q} . It is clear that every bounded set is strongly bounded and every strongly bounded set is weakly bounded. The principle of uniform boundedness for linear transformations is the best possible converse.

Problem 40. *Every weakly bounded set of bounded linear transformations is bounded.*

41. Invertible transformations. A bounded linear transformation A from a Hilbert space \mathbf{H} to a Hilbert space \mathbf{K} is *invertible* if there exists a bounded linear transformation B (from \mathbf{K} into \mathbf{H}) such that $AB = 1$ (= the identity operator on \mathbf{K}) and $BA = 1$ (= the identity operator on \mathbf{H}). If A is invertible, then A is a one-to-one mapping of \mathbf{H} onto \mathbf{K} . In the sense of pure set theory the converse is true: if A maps \mathbf{H} one-to-one onto \mathbf{K} , then there exists a unique mapping A^{-1} from \mathbf{K} to \mathbf{H} such that $AA^{-1} = 1$ and $A^{-1}A = 1$; the mapping A^{-1} is linear.

It is not obvious, however, that the linear transformation A^{-1} must be bounded; it is conceivable that A could be invertible as a set-theoretic mapping but not invertible as an operator. To guarantee that A^{-1} is bounded it is customary to strengthen the condition that A be one-to-one. The proper strengthening is to require that A be bounded from below, i.e., that there exist a positive number δ such that $\|Af\| \geq \delta \|f\|$ for every f in \mathbf{H} . (It is trivial to verify that if A is bounded from below, then A is indeed one-to-one.) If that strengthened condition is satisfied, then the other usual condition (onto) can be weakened: the requirement that the range of A be equal to \mathbf{K} can be replaced by the requirement that the range of A be dense in \mathbf{K} . In sum: A is invertible if and only if it is bounded from below and has a dense range (see Halmos [1951, p. 38]). Observe that the linear transformations A and A^* are invertible together; if they are invertible, then each of A^{-1} and A^{*-1} is the adjoint of the other.

It is perhaps worth a short digression to discuss the possibility of the range of an operator not being closed, and its consequences. If, for instance, A is defined on l^2 by $A \langle \xi_1, \xi_2, \xi_3, \dots \rangle = \langle \xi_1, \frac{1}{2}\xi_2, \frac{1}{3}\xi_3, \dots \rangle$, then the range of A consists of all vectors

$$\langle \eta_1, \eta_2, \eta_3, \dots \rangle \quad \text{with} \quad \sum_n n^2 |\eta_n|^2 < \infty.$$

Since this range contains all finitely non-zero sequences, it is dense in l^2 ; since, however, it does not contain the sequence $\langle 1, \frac{1}{2}, \frac{1}{3}, \dots \rangle$, it is not closed. Another example: for f in $L^2(0,1)$, define $(Af)(x) = xf(x)$. These operators are, of course, not bounded from below; if they were, their ranges would be closed.

Operators with non-closed ranges can be used to give a very simple example of two subspaces whose vector sum is not closed; cf. Halmos [1951, p. 110]. Let A be an operator on a Hilbert space \mathbf{H} ; the construction itself takes place in the direct sum $\mathbf{H} \oplus \mathbf{H}$. Let \mathbf{M} be the "x-axis", i.e., the set of all vectors (in $\mathbf{H} \oplus \mathbf{H}$) of the form $\langle f, 0 \rangle$, and let \mathbf{N} be the "graph" of A , i.e., the set of all vectors of the form $\langle f, Af \rangle$. It is trivial to verify that both \mathbf{M} and \mathbf{N} are subspaces of $\mathbf{H} \oplus \mathbf{H}$. When does $\langle f, g \rangle$ belong to $\mathbf{M} + \mathbf{N}$? The answer is if and only if it has the form $\langle u, 0 \rangle + \langle v, Av \rangle = \langle u + v, Av \rangle$; since u and v are arbitrary, a vector in $\mathbf{H} \oplus \mathbf{H}$ has that form if and only if its second coordinate belongs to the range \mathbf{R} of the operator A . (In other words, $\mathbf{M} + \mathbf{N} =$

$\mathbf{H} \oplus \mathbf{R}$.) Is $\mathbf{M} + \mathbf{N}$ closed? This means: if $\langle f_n, g_n \rangle \rightarrow \langle f, g \rangle$, where $f_n \in \mathbf{H}$ and $g_n \in \mathbf{R}$, does it follow that $f \in \mathbf{H}$? (trivially yes), and does it follow that $g \in \mathbf{R}$? (possibly no). Conclusion: $\mathbf{M} + \mathbf{N}$ is closed in $\mathbf{H} \oplus \mathbf{H}$ if and only if \mathbf{R} is closed in \mathbf{H} . Since A can be chosen so that \mathbf{R} is not closed, the vector sum of two subspaces need not be closed either.

The theorems and the examples seem to indicate that set-theoretic invertibility and operatorial invertibility are indeed distinct; it is one of the pleasantest and most useful facts about operator theory that they are the same after all.

Problem 41. *If \mathbf{H} and \mathbf{K} are Hilbert spaces, and if A is a bounded linear transformation that maps \mathbf{H} one-to-one onto \mathbf{K} , then A is invertible.*

The corresponding statement about Banach spaces is usually proved by means of the Baire category theorem.

42. Preservation of dimension. An important question about operators is what do they do to the geometry of the underlying space. It is familiar from the study of finite-dimensional vector spaces that a linear transformation can lower dimension: the transformation 0, for an extreme example, collapses every space to a 0-dimensional one. If, however, a linear transformation on a finite-dimensional vector space is one-to-one (i.e., its kernel is $\{0\}$), then it cannot lower dimension; since the same can be said about the inverse transformation (from the range back to the domain), it follows that dimension is preserved. The following assertion is, in a sense, the generalization of this finite-dimensional result to arbitrary Hilbert spaces.

Problem 42. *If there exists a one-to-one bounded linear transformation from a Hilbert space \mathbf{H} into a Hilbert space \mathbf{K} , then $\dim \mathbf{H} \leq \dim \mathbf{K}$. If the image of \mathbf{H} is dense in \mathbf{K} , then equality holds.*

43. Projections of equal rank.

Problem 43. *If P and Q are projections such that $\|P - Q\| < 1$, then P and Q have the same rank.*

This is a special case of Problem 101.

44. Closed graph theorem. The *graph* of a linear transformation A (not necessarily bounded) between inner product spaces \mathbf{H} and \mathbf{K} (not necessarily complete) is the set of all those ordered pairs $\langle f, g \rangle$ (elements of $\mathbf{H} \oplus \mathbf{K}$) for which $Af = g$. (The terminology is standard. It is curious that it should be so, but it is. According to a widely adopted approach to the foundations of mathematics, a function, by definition, is a set of ordered pairs satisfying a certain univalence condition. According to that approach, the graph of A is A , and it is hard to see what is accomplished by giving it another name. Nevertheless most mathematicians cheerfully accept the unnecessary word; at the very least it serves as a warning that the same object is about to be viewed from a different angle.) A linear transformation is called *closed* if its graph is a closed set.

Problem 44. *A linear transformation from a Hilbert space into a Hilbert space is closed if and only if it is bounded.*

The assertion is known as the *closed graph theorem* for Hilbert spaces; its proof for Banach spaces is usually based on a category argument (Dunford-Schwartz [1958, p. 57]). The theorem does not make the subject of closed but unbounded linear transformations trivial. Such transformations occur frequently in the applications of functional analysis; what the closed graph theorem says is that they can occur on incomplete inner-product spaces only (or non-closed linear manifolds in Hilbert spaces).

45. Unbounded symmetric transformations. A linear transformation A (not necessarily bounded) on an inner-product space \mathbf{H} (not necessarily complete) is called *symmetric* if $(Af, g) = (f, Ag)$ for all f and g in \mathbf{H} . It is advisable to use this neutral term (rather than “Hermitian” or “self-adjoint”), because in the customary approach to Hermitian operators ($A = A^*$) boundedness is an assumption necessary for the very formulation of the definition. Is there really a distinction here?

Problem 45. (a) *Is a symmetric linear transformation on an inner-product space \mathbf{H} necessarily bounded?* (b) *What if \mathbf{H} is a Hilbert space?*

Chapter 6.

Multiplication operators

46. Diagonal operators. Operator theory, like every other part of mathematics, cannot be properly studied without a large stock of concrete examples. The purpose of several of the problems that follow is to build a stock of concrete operators, which can then be examined for the behavior of their norms, inverses, and spectra.

Suppose, for a modest first step, that \mathbf{H} is a Hilbert space and that $\{e_j\}$ is a family of vectors that constitute an orthonormal basis for \mathbf{H} . An operator A is called a *diagonal operator* if Ae_j is a scalar multiple of e_j , say $Ae_j = \alpha_j e_j$, for each j ; the family $\{\alpha_j\}$ may properly be called the *diagonal* of A .

The definition of a diagonal operator depends, of course, on the basis $\{e_j\}$, but in most discussions of diagonal operators a basis is (perhaps tacitly) fixed in advance, and then never mentioned again. Alternatively, diagonal operators can be characterized in invariant terms as normal operators whose eigenvectors span the space. (The proof of the characterization is an easy exercise.) Usually diagonal operators are associated with an orthonormal *sequence*; the emphasis is on both the cardinal number (\aleph_0) and the order (ω) of the underlying index set. That special case makes possible the use of some convenient language (e.g., “the first element of the diagonal”) and the use of some convenient techniques (e.g., constructions by induction).

Problem 46. *A necessary and sufficient condition that a family $\{\alpha_j\}$ be the diagonal of a diagonal operator is that it be bounded; if it is bounded, then the equations $Ae_j = \alpha_j e_j$ uniquely determine an operator A , and $\|A\| = \sup_j |\alpha_j|$.*

47. Multiplications on l^2 . Each sequence $\{\alpha_n\}$ of complex scalars induces a linear transformation A that maps l^2 into the vector space of all (not necessarily square-summable) sequences; by definition $A \langle \xi_1, \xi_2, \xi_3, \dots \rangle = \langle \alpha_1 \xi_1, \alpha_2 \xi_2, \alpha_3 \xi_3, \dots \rangle$. Half of Problem 46 implies

that if A is an operator (i.e., a bounded linear transformation of l^2 into itself), then the sequence $\{\alpha_n\}$ is bounded. What happens if the boundedness assumption on A is dropped?

Problem 47. *Can an unbounded sequence of scalars induce a (possibly unbounded) transformation of l^2 into itself?*

The emphasis is that all l^2 is in the domain of the transformation, i.e., that if $\langle \xi_1, \xi_2, \xi_3, \dots \rangle \in l^2$, then $\langle \alpha_1 \xi_1, \alpha_2 \xi_2, \alpha_3 \xi_3, \dots \rangle \in l^2$. The question should be compared with Problem 22. That problem considered sequences that multiply l^2 into l^1 (and concluded that they must belong to l^2); this one considers sequences that multiply l^2 into l^2 (and asks whether they must belong to l^∞). See Problem 51 for the generalization to L^2 .

48. Spectrum of a diagonal operator. The set of all bounded sequences $\{\alpha_n\}$ of complex numbers is an algebra (pointwise operations), with unit ($\alpha_n = 1$ for all n), with a conjugation ($\{\alpha_n\} \rightarrow \{\alpha_n^*\}$), and with a norm ($\|\{\alpha_n\}\| = \sup_n |\alpha_n|$). A bounded sequence $\{\alpha_n\}$ will be called *invertible* if it has an inverse in this algebra, i.e., if there exists a bounded sequence $\{\beta_n\}$ such that $\alpha_n \beta_n = 1$ for all n . A necessary and sufficient condition for this to happen is that $\{\alpha_n\}$ be bounded away from 0, i.e., that there exist a positive number δ such that $|\alpha_n| \geq \delta$ for all n .

If \mathbf{H} is a Hilbert space with an orthonormal basis $\{e_n\}$, then it is easy to verify that the correspondence $\{\alpha_n\} \rightarrow A$, where A is the operator on \mathbf{H} such that $Ae_n = \alpha_n e_n$ for all n , is an isomorphism (an embedding) of the sequence algebra into the algebra of operators on \mathbf{H} . The correspondence preserves not only the familiar algebraic operations, but also conjugation; that is, if $\{\alpha_n\} \rightarrow A$, then $\{\alpha_n^*\} \rightarrow A^*$. The correspondence preserves the norm also (see Problem 46).

Problem 48. *A diagonal operator with diagonal $\{\alpha_n\}$ is an invertible operator if and only if the sequence $\{\alpha_n\}$ is an invertible sequence. Consequence: the spectrum of a diagonal operator is the closure of the set of its diagonal terms.*

The result has the following useful corollary: every non-empty compact subset of the complex plane is the spectrum of some operator (and, in

fact, of some diagonal operator). Proof: find a sequence of complex numbers dense in the prescribed compact set, and form a diagonal operator with that sequence as its diagonal.

49. Norm of a multiplication. Diagonal operators are special cases of a general measure-theoretic construction. Suppose that X is a measure space with measure μ . If φ is a complex-valued bounded (i.e., essentially bounded) measurable function on X , then the *multiplication operator* (or just *multiplication*, for short) induced by φ is the operator A on $L^2(\mu)$ defined by

$$(Af)(x) = \varphi(x)f(x)$$

for all x in X . (Here, as elsewhere in measure theory, two functions are identified if they differ on a set of measure zero only. This applies to the bounded φ 's as well as to the square-integrable f 's.) If X is the set of all positive integers and μ is the counting measure (the measure of every set is the number of elements in it), then multiplication operators reduce to diagonal operators.

Problem 49. *What, in terms of the multiplier φ , is the norm of the multiplication induced by φ ?*

50. Boundedness of multipliers. Much of the theory of diagonal operators extends to multiplication operators on measure spaces, but the details become a little fussy at times. A sample is the generalization of the assertion that if a sequence is the diagonal of a diagonal operator, then it is bounded.

Problem 50. *If an operator A on L^2 (for a σ -finite measure) is such that $Af = \varphi \cdot f$ for all f in L^2 (for some function φ), then φ is measurable and bounded.*

51. Boundedness of multiplications. Each complex-valued measurable function φ induces a linear transformation A that maps L^2 into the vector space of all (not necessarily square-integrable) measurable functions; by definition $(Af)(x) = \varphi(x)f(x)$. Half of Problem 50 implies that if A is an operator (i.e., a bounded linear transformation of

L^2 into itself), then the function φ is bounded. What happens if the boundedness assumption on A is dropped?

Problem 51. *Can an unbounded function induce a (possibly unbounded) transformation of L^2 (for a σ -finite measure) into itself?*

This is the generalization to measures of Problem 47.

52. Spectrum of a multiplication. Some parts of the theory of diagonal operators extend to multiplication operators almost verbatim, as follows. The set of all bounded measurable functions (identified modulo sets of measure zero) is an algebra (pointwise operations), with unit ($\varphi(x) = 1$ for all x), with a conjugation ($\varphi \rightarrow \varphi^*$), and with a norm ($\|\varphi\|_\infty$). A bounded measurable function is *invertible* if it has an inverse in this algebra, i.e., if there exists a bounded measurable function ψ such that $\varphi(x)\psi(x) = 1$ for almost every x . A necessary and sufficient condition for this to happen is that φ be bounded away from 0 almost everywhere, i.e., that there exist a positive number δ such that $|\varphi(x)| \geq \delta$ for almost every x .

The correspondence $\varphi \rightarrow A$, where A is the multiplication operator defined by $(Af)(x) = \varphi(x)f(x)$, is an isomorphism (an embedding) of the function algebra into the algebra of operators on L^2 . The correspondence preserves not only the familiar algebraic operations, but also the conjugation; that is, if $\varphi \rightarrow A$, then $\varphi^* \rightarrow A^*$. If the measure is σ -finite, the correspondence preserves the norm also (see Solution 49).

The role played by the range of a sequence is played, in the general case, by the *essential range* of a function φ ; by definition, that is the set of all complex numbers λ such that for each neighborhood N of λ the set $\varphi^{-1}(N)$ has positive measure.

Problem 52. *The multiplication operator on L^2 (for a σ -finite measure) induced by φ is an invertible operator if and only if φ is an invertible function. Consequence: the spectrum of a multiplication is the essential range of the multiplier.*

53. Multiplications on functional Hilbert spaces. If a function φ multiplies L^2 into itself, then φ is necessarily bounded (Solution 51),

and therefore multiplication by φ is necessarily an operator on L^2 . Are the analogues of these assertions true for functional Hilbert spaces?

Problem 53. *Suppose that \mathbf{H} is a functional Hilbert space, over a set X say, and suppose that φ is a complex-valued function on X such that $\varphi \cdot f \in \mathbf{H}$ whenever $f \in \mathbf{H}$. (a) If $Af = \varphi \cdot f$, is the linear transformation A bounded? (b) If $Af = \varphi \cdot f$ and if A is bounded, is the function φ bounded?*

54. Multipliers of functional Hilbert spaces. Suppose that \mathbf{H} is a functional Hilbert space over a set X . A function φ on X is a *multiplier* of \mathbf{H} if $\varphi \cdot f \in \mathbf{H}$ for every f in \mathbf{H} . Solution 53 says that every multiplier is bounded. It is frequently interesting and important to determine all multipliers of a functional Hilbert space.

For l^2 , the easiest infinite-dimensional space, it is easy to prove that a necessary and sufficient condition that a function (i.e., a sequence) be a multiplier is that it be bounded. In a certain sense the space l^2 has too many multipliers: most of them do not belong to the space.

The space \mathbf{A}^2 behaves differently: for it a necessary and sufficient condition that a function be a multiplier is that it be bounded and belong to the space. In a certain sense the space has too few multipliers: most of the functions in the space are not among them.

If X is finite and if \mathbf{H} consists of all functions on X , then the set of multipliers of \mathbf{H} is neither too large nor too small: it consists exactly of the elements of \mathbf{H} . Can this happen for infinite-dimensional spaces?

Problem 54. *Construct an infinite-dimensional functional Hilbert space \mathbf{H} such that the multipliers of \mathbf{H} are exactly the elements of \mathbf{H} .*

To say that every element of \mathbf{H} is a multiplier is the same as to say that \mathbf{H} is closed under multiplication, i.e., that \mathbf{H} is an algebra. The constant function 1 is a multiplier of every \mathbf{H} ; hence, to say that every multiplier of \mathbf{H} belongs to \mathbf{H} is the same as to say that $1 \in \mathbf{H}$. If $1 \in \mathbf{H}$, then, of course, the algebra \mathbf{H} has a unit, but trivial examples show that the converse is not true. Thus, the construction of an infinite-dimensional functional Hilbert space that is an algebra with unit (under pointwise functional multiplication) is not quite, but almost, what the problem asks for.

Chapter 7. Operator matrices

55. Commutative operator determinants. An orthonormal basis serves to express a Hilbert space as the direct sum of one-dimensional subspaces. Some of the matrix theory associated with orthonormal bases deserves to be extended to more general direct sums. Suppose, to be specific, that $\mathbf{H} = \mathbf{H}_1 \oplus \mathbf{H}_2 \oplus \mathbf{H}_3 \oplus \cdots$. (Uncountable direct sums work just as well, and finite ones even better.) If the direct sum is viewed as an “internal” one, so that the \mathbf{H}_i ’s are subspaces of \mathbf{H} , then the elements f of \mathbf{H} are sums

$$f = f_1 + f_2 + f_3 + \cdots,$$

with f_i in \mathbf{H}_i . If A is an operator on \mathbf{H} , then

$$Af = Af_1 + Af_2 + Af_3 + \cdots.$$

Each Af_j , being an element of \mathbf{H} , has a decomposition:

$$Af_j = g_{1j} + g_{2j} + g_{3j} + \cdots,$$

with g_{ij} in \mathbf{H}_i . The g_{ij} ’s depend, of course, on f_j , and the dependence is linear and continuous. It follows that

$$g_{ij} = A_{ij}f_j,$$

where A_{ij} is a bounded linear transformation from \mathbf{H}_j to \mathbf{H}_i . The construction is finished: corresponding to each A on \mathbf{H} there is a matrix $\langle A_{ij} \rangle$, whose entry in row i and column j is the projection onto the i component of the restriction of A to \mathbf{H}_j .

The correspondence from operators to matrices (induced by a fixed direct decomposition) has all the right algebraic properties. If $A = 0$, then $A_{ij} = 0$ for all i and j ; if $A = 1$ (on \mathbf{H}), then $A_{ij} = 0$ when $i \neq j$ and $A_{ii} = 1$ (on \mathbf{H}_i). The linear operations on operator matrices are the obvious ones. The matrix of A^* is the adjoint transpose of the matrix

of A ; that is, the matrix of A^* has the entry A_{ji}^* in row i and column j . The multiplication of operators corresponds to the matrix product defined by $\sum_k A_{ik}B_{kj}$. There is no convergence trouble here, but there may be commutativity trouble; the order of the factors must be watched with care.

The theory of operator matrices does not become trivial even if the number of direct summands is small (say two) and even if all the direct summands are identical. The following situation is the one that occurs most frequently: a Hilbert space \mathbf{H} is given, the role of what was \mathbf{H} in the preceding paragraph is played now by the direct sum $\mathbf{H} \oplus \mathbf{H}$, and operators on that direct sum are expressed as two-by-two matrices whose entries are operators on \mathbf{H} .

Problem 55. *If A, B, C , and D are pairwise commutative operators on a Hilbert space, then a necessary and sufficient condition that the operator matrix*

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

be invertible is that the formal determinant $AD - BC$ be invertible.

56. Operator determinants. There are many situations in which the invertibility of an operator matrix

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

plays a central role but in which the entries are not commutative; any special case is worth knowing.

Problem 56. *If C and D commute, and if D is invertible, then a necessary and sufficient condition that*

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

be invertible is that $AD - BC$ be invertible. Construct examples to show that if the assumption that D is invertible is dropped, then the condition becomes unnecessary and insufficient.

For finite matrices more is known (cf. Schur [1917]): if C and D commute, then

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

and $AD - BC$ have the same determinant. The proof for the general case can be arranged so as to yield this strengthened version for the finite-dimensional case.

57. Operator determinants with a finite entry. If A , B , and D are operators on a Hilbert space \mathbf{H} , then the operator matrix

$$M = \begin{pmatrix} A & B \\ 0 & D \end{pmatrix}$$

induces (is) an operator on $\mathbf{H} \oplus \mathbf{H}$, and (cf. Problem 56) if both A and D are invertible, then M is invertible. The converse (if M is invertible, then A and D are) is not true (see Problem 56 again).

Operator matrices define operators on direct sums of Hilbert spaces whether the direct summands are identical or not. In at least one special case of interest the converse that was false in the preceding paragraph becomes true.

Problem 57. *If \mathbf{H} and \mathbf{K} are Hilbert spaces, with $\dim \mathbf{H} < \infty$, and if*

$$M = \begin{pmatrix} A & B \\ 0 & D \end{pmatrix}$$

is an invertible operator on $\mathbf{H} \oplus \mathbf{K}$, then both A and D are invertible. Consequence: the spectrum of M is the union of the spectra of A and D .

Note that A operates on \mathbf{H} , D operates on \mathbf{K} , and B maps \mathbf{K} into \mathbf{H} .

Chapter 8. Properties of spectra

58. Spectra and conjugation. It is often useful to ask of a point in the spectrum of an operator how it got there. To say that λ is in the spectrum of A means that $A - \lambda$ is not invertible. The question reduces therefore to this: why is a non-invertible operator not invertible? There are several possible ways of answering the question; they have led to several (confusingly overlapping) classifications of spectra.

Perhaps the simplest approach to the subject is to recall that if an operator is bounded from below and has a dense range, then it is invertible. Consequence: if $\Lambda(A)$ is the spectrum of A , if $\Pi(A)$ is the set of complex numbers λ such that $A - \lambda$ is not bounded from below, and if $\Gamma(A)$ is the set of complex numbers λ such that the closure of the range of $A - \lambda$ is a proper subspace of \mathbf{H} (i.e., distinct from \mathbf{H}), then

$$\Lambda(A) = \Pi(A) \cup \Gamma(A).$$

The set $\Pi(A)$ is called the *approximate point spectrum* of A ; a number λ belongs to $\Pi(A)$ if and only if there exists a sequence $\{f_n\}$ of unit vectors such that $\|(A - \lambda)f_n\| \rightarrow 0$. An important subset of the approximate point spectrum is the *point spectrum* $\Pi_0(A)$; a number λ belongs to it if and only if there exists a unit vector f such that $Af = \lambda f$ (i.e., $\Pi_0(A)$ is the set of all eigenvalues of A). The set $\Gamma(A)$ is called the *compression spectrum* of A . Schematically: think of the spectrum (Λ) as the union of two overlapping discs (Π and Γ), one of which (Π) is divided into two parts (Π_0 and $\Pi - \Pi_0$) by a diameter perpendicular to the overlap. The result is a partition of Λ into five parts, each one of which may be sometimes present and sometimes absent. The born taxonomist may amuse himself by trying to see which one of the 2^5 a priori possibilities is realizable, but he would be well advised to postpone the attempt until he has seen several more examples of operators than have appeared in this book so far.

This is a good opportunity to comment on a sometimes confusing aspect of the nomenclature of operator theory. There is something called the spectral theorem for normal operators (see Problem 97), and there

are things called spectra for all operators. The study of the latter might be called spectral theory, and sometimes it is. In the normal case the spectral theorem gives information about spectral theory, but, usually, that information can be bought cheaper elsewhere. Spectral theory in the present sense of the phrase is one of the easiest aspects of operator theory.

There is no consensus on which concepts and symbols are most convenient in this part of operator theory. Apparently every book introduces its own terminology, and the present one is no exception. A once popular approach was to divide the spectrum into three disjoint sets, namely the point spectrum Π_0 , the *residual spectrum* $\Gamma - \Pi_0$, and the *continuous spectrum* $\Pi - (\Gamma \cup \Pi_0)$. (The sets Π and Γ may overlap; examples will be easy to construct a little later.) As for symbols: the spectrum is often σ (or Σ) instead of Λ .

The best way to master these concepts is, of course, through illuminating special examples, but a few general facts should come first; they help in the study of the examples. The most useful things to know are the relations of spectra to the algebra and topology of the complex plane. Perhaps the easiest algebraic questions concern conjugation.

Problem 58. *What happens to the point spectrum, the compression spectrum, and the approximate point spectrum when an operator is replaced by its adjoint?*

59. Spectral mapping theorem. An assertion such as that if A is an operator and p is a polynomial, then $\Lambda(p(A)) = p(\Lambda(A))$ (see Halmos [1951, p. 53]) is called a *spectral mapping theorem*; other instances of it have to do with functions other than polynomials, such as inversion, conjugation, and wide classes of analytic functions (Dunford-Schwartz [1958, p. 569]).

Problem 59. *Is the spectral mapping theorem for polynomials true with Π_0 , or Π , or Γ in place of Λ ? What about the spectral mapping theorem for inversion ($p(z) = 1/z$ when $z \neq 0$), applied to invertible operators, with Π_0 , or Π , or Γ ?*

60. Similarity and spectrum. Two operators A and B are *similar* if there exists an invertible operator P such that $P^{-1}AP = B$.

Problem 60. *Similar operators have the same spectrum, the same point spectrum, the same approximate point spectrum, and the same compression spectrum.*

61. Spectrum of a product. If A and B are operators, and if at least one of them is invertible, then AB and BA are similar. (For the proof, apply $BA = A^{-1}(AB)A$ in case A is invertible or $AB = B^{-1}(BA)B$ in case B is.) This implies (Problem 60) that if at least one of A and B is invertible, then AB and BA have the same spectrum. In the finite-dimensional case more is known: with no invertibility assumptions, AB and BA always have the same characteristic polynomial. If neither A nor B is invertible, then, in the infinite-dimensional case, the two products need not have the same spectrum (many examples occur below), but their spectra cannot differ by much. Here is the precise assertion.

Problem 61. *The non-zero elements of $\Lambda(AB)$ and $\Lambda(BA)$ are the same.*

62. Closure of approximate point spectrum.

Problem 62. *Is the approximate point spectrum always closed?*

63. Boundary of spectrum.

Problem 63. *The boundary of the spectrum of an operator is included in the approximate point spectrum.*

Chapter 9. Examples of spectra

64. Residual spectrum of a normal operator. The time has come to consider special cases. The first result is that for normal operators, the most amenable large class known, the worst spectral pathology cannot occur.

Problem 64. *If A is normal, then $\Gamma(A) = \Pi_0(A)$ (and therefore $\Lambda(A) = \Pi(A)$). Alternative formulation: the residual spectrum of a normal operator is always empty.*

Recall that the residual spectrum of A is $\Gamma(A) - \Pi_0(A)$.

65. Spectral parts of a diagonal operator. The spectrum of a diagonal operator was determined (Problem 48) as the closure of its diagonal; the determination of the fine structure of the spectrum requires another look.

Problem 65. *For each diagonal operator, find its point spectrum, compression spectrum, and approximate point spectrum.*

66. Spectral parts of a multiplication.

Problem 66. *For each multiplication, find its point spectrum, compression spectrum, and approximate point spectrum.*

67. Unilateral shift. The most important single operator, which plays a vital role in all parts of Hilbert space theory, is called the *unilateral shift*. Perhaps the simplest way to define it is to consider the Hilbert space l^2 of square-summable sequences; the unilateral shift is the operator U on l^2 defined by

$$U \langle \xi_0, \xi_1, \xi_2, \dots \rangle = \langle 0, \xi_0, \xi_1, \xi_2, \dots \rangle.$$

(The unilateral shift has already occurred in this book, although it was not named until now; see Solution 56.) Linearity is obvious. As for

boundedness, it is true with room to spare. Norms are not only kept within reasonable bounds, but they are preserved exactly; the unilateral shift is an isometry. The range of U is not l^2 but a proper subspace of l^2 , the subspace of vectors with vanishing first coordinate. The existence of an isometry whose range is not the whole space is characteristic of infinite-dimensional spaces.

If e_n is the vector $\langle \xi_0, \xi_1, \xi_2, \dots \rangle$ for which $\xi_n = 1$ and $\xi_i = 0$ whenever $i \neq n$ ($n = 0, 1, 2, \dots$), then the e_n 's form an orthonormal basis for l^2 . The effect of U on this basis is described by

$$Ue_n = e_{n+1} \quad (n = 0, 1, 2, \dots).$$

These equations uniquely determine U , and in most of the study of U they may be taken as its definition.

A familiar space that comes equipped with an orthonormal basis indexed by non-negative integers is \mathbf{H}^2 (see Problem 26). Since, in that space, $e_n(z) = z^n$, the effect of shifting forward by one index is the same as the effect of multiplication by e_1 . In other words, the unilateral shift is the same as the multiplication operator on \mathbf{H}^2 defined by

$$(Uf)(z) = zf(z).$$

To say that it is the “same”, and, in fact, to speak of “the” unilateral shift is a slight abuse of language, a convenient one that will be maintained throughout the sequel. Properly speaking the unilateral shift is a unitary equivalence class of operators, but no confusion will result from regarding it as one operator with many different manifestations.

Problem 67. *What is the spectrum of the unilateral shift, and what are its parts (point spectrum, compression spectrum, and approximate point spectrum)? What are the answers to the same questions for the adjoint of the unilateral shift?*

68. Bilateral shift. A close relative of the unilateral shift is the *bilateral shift*. To define it, let \mathbf{H} be the Hilbert space of all two-way (bilateral) square-summable sequences. The elements of \mathbf{H} are most

conveniently written in the form

$$\langle \cdots, \xi_{-2}, \xi_{-1}, (\xi_0), \xi_1, \xi_2, \cdots \rangle;$$

the term in parentheses indicates the one corresponding to the index 0. The bilateral shift is the operator W on \mathbf{H} defined by

$$W \langle \cdots, \xi_{-2}, \xi_{-1}, (\xi_0), \xi_1, \xi_2, \cdots \rangle = \langle \cdots, \xi_{-3}, \xi_{-2}, (\xi_{-1}), \xi_0, \xi_1, \cdots \rangle.$$

Linearity is obvious, and boundedness is true with room to spare; the bilateral shift, like the unilateral one, is an isometry. Since the range of the bilateral shift is the entire space \mathbf{H} , it is even unitary.

If e_n is the vector $\langle \cdots, \xi_{-1}, (\xi_0), \xi_1, \cdots \rangle$ for which $\xi_n = 1$ and $\xi_i = 0$ whenever $i \neq n$ ($n = 0, \pm 1, \pm 2, \cdots$), then the e_n 's form an orthonormal basis for \mathbf{H} . The effect of W on this basis is described by

$$We_n = e_{n+1} \quad (n = 0, \pm 1, \pm 2, \cdots).$$

Problem 68. *What is the spectrum of the bilateral shift, and what are its parts (point spectrum, compression spectrum, and approximate point spectrum)? What are the answers to the same questions for the adjoint of the bilateral shift?*

69. Spectrum of a functional multiplication. Every operator studied so far has been a multiplication, either in the legitimate sense (on an L^2) or in the extended sense (on a functional Hilbert space). The latter kind is usually harder to study; it does, however, have the advantage of having a satisfactory characterization in terms of its spectrum.

Problem 69. *A necessary and sufficient condition that an operator A on a Hilbert space \mathbf{H} be representable as a multiplication on a functional Hilbert space is that the eigenvectors of A^* span \mathbf{H} .*

Caution: as the facts for multiplications on L^2 spaces show (cf. Solution 66) this characterization is applicable to functional Hilbert spaces only. The result seems to be due to P. R. Halmos and A. L. Shields.

70. Relative spectrum of shift. An operator A is *relatively invertible* if there exists an operator B such that $ABA = A$. This is a rather special concept, not particularly useful, but with some curious properties. Clearly every invertible operator is relatively invertible; in fact every operator that is either left invertible or right invertible is also relatively invertible. These remarks are obvious; it is much less obvious (but true) that every operator on a finite-dimensional space is relatively invertible. (Hint: write the operator as a direct sum of an invertible operator and a nilpotent one.) The concept belongs to general ring theory; the assertion about finite-dimensional spaces can be expressed by saying that a finite-dimensional full matrix algebra over the complex numbers is a *regular ring* (see von Neumann [1936]). The *relative spectrum* of an operator A (on a Hilbert space of any dimension) is the set of all those complex numbers λ for which $A - \lambda$ is not relatively invertible.

Problem 70. *What is the relative spectrum of the unilateral shift?*

The concept of relative spectrum was introduced and studied by Asplund [1958].

71. Closure of relative spectrum.

Problem 71. *Is the relative spectrum always closed?*

Chapter 10. Spectral radius

72. Analyticity of resolvents. Suppose that A is an operator on a Hilbert space \mathbf{H} . If λ does not belong to the spectrum of A , then the operator $A - \lambda$ is invertible; write $\rho(\lambda) = (A - \lambda)^{-1}$. (When it is necessary to indicate the dependence of the function ρ on the operator A , write $\rho = \rho_A$.) The function ρ is called the *resolvent* of A . The domain of ρ is the complement of the spectrum of A ; its values are operators on \mathbf{H} .

The definition of the resolvent is very explicit; this makes it seem plausible that the resolvent is a well-behaved function. To formulate this precisely, consider, quite generally, functions φ whose domains are open sets in the complex plane and whose values are operators on \mathbf{H} . Such a function φ will be called *analytic* if, for each f and g in \mathbf{H} , the numerical function $\lambda \rightarrow (\varphi(\lambda)f, g)$ (with the same domain as φ) is analytic in the usual sense. (To distinguish this concept from other closely related ones, it is sometimes called *weak* analyticity.) In case the function ψ defined by $\psi(\lambda) = \varphi(1/\lambda)$ can be assigned a value at the origin so that it becomes analytic there, then (just as for numerical functions) φ will be called analytic at ∞ , and φ is assigned at ∞ the value of ψ at 0.

Problem 72. *The resolvent of every operator is analytic at each point of its domain, and at ∞ ; its value at ∞ is (the operator) 0.*

For a detailed study of resolvents, see Dunford-Schwartz [1958, VII, 3].

73. Non-emptiness of spectra. Does every operator have a non-empty spectrum? The question was bound to arise sooner or later. Even the finite-dimensional case shows that the question is non-trivial. To say that every finite matrix has an eigenvalue is the same as to say that the characteristic polynomial of every finite matrix has at least one zero, and that is no more and no less general than to say that every polynomial equation (with complex coefficients) has at least one (com-

plex) zero. In other words, the finite-dimensional case of the general question about spectra is as deep as the fundamental theorem of algebra, whose proof is usually based on the theory of complex analytic functions. It should not be too surprising now that the theory of such functions enters the study of operators in every case (whether the dimension is finite or infinite).

Problem 73. *Every operator has a non-empty spectrum.*

74. Spectral radius. The *spectral radius* of an operator A , in symbols $r(A)$, is defined by

$$r(A) = \sup\{|\lambda| : \lambda \in \Lambda(A)\}.$$

Clearly $0 \leq r(A) \leq \|A\|$; the spectral mapping theorem implies also that $r(A^n) = (r(A))^n$ for every positive integer n . It frequently turns out that the spectral radius of an operator is easy to compute even when it is hard to find the spectrum; the tool that makes it easy is the following assertion.

Problem 74. *For each operator A ,*

$$r(A) = \lim_n \|A^n\|^{1/n},$$

in the sense that the indicated limit always exists and has the indicated value.

It is an easy consequence of this result that if A and B are commutative operators, then

$$r(AB) \leq r(A)r(B).$$

It is a somewhat less easy consequence, but still a matter of no more than a little fussy analysis with inequalities, that if A and B commute, then

$$r(A + B) \leq r(A) + r(B).$$

If no commutativity assumptions are made, then two-dimensional ex-

amples, such as

$$A = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},$$

show that neither the submultiplicative nor the subadditive property persists.

75. Weighted shifts. A *weighted shift* is the product of a shift (one-sided or two) and a compatible diagonal operator. More explicitly, suppose that $\{e_n\}$ is an orthonormal basis ($n = 0, 1, 2, \dots$, or else $n = 0, \pm 1, \pm 2, \dots$), and suppose that $\{\alpha_n\}$ is a bounded sequence of complex numbers (the set of n 's being the same as before). A weighted shift is an operator of the form SP , where S is a shift ($Se_n = e_{n+1}$) and P is a diagonal operator with diagonal $\{\alpha_n\}$ ($Pe_n = \alpha_n e_n$). Not everything about weighted shifts is known, but even the little that is makes them almost indispensable in the construction of examples and counterexamples.

Problem 75. If P and Q are diagonal operators, with diagonals $\{\alpha_n\}$ and $\{\beta_n\}$, and if $|\alpha_n| = |\beta_n|$ for all n , then the weighted shifts $A = SP$ and $B = SQ$ are unitarily equivalent.

A discussion of two weighted shifts should, by rights, refer to two orthonormal bases, but the generality gained that way is shallow. If $\{e_n\}$ and $\{f_n\}$ are orthonormal bases, then there exists a unitary operator U such that $Ue_n = f_n$ for all n , and U can be carried along gratis with any unitary equivalence proof.

The result about the unitary equivalence of weighted shifts has two useful consequences. First, the weighted shift with weights α_n is unitarily equivalent to the weighted shift with weights $|\alpha_n|$. Since unitarily equivalent operators are “abstractly identical”, there is never any loss of generality in restricting attention to weighted shifts whose weights are non-negative; this is what really justifies the use of the word “weight”. Second, if A is a weighted shift and if α is a complex number of modulus 1, then, since αA is a weighted shift, whose weights have the same moduli as the corresponding weights of A , it follows that A and αA are unitarily equivalent. In other words, to within unitary equivalence, a weighted shift is not altered by multiplication by a number of modulus 1. This

implies, for instance, that the spectrum of a weighted shift has circular symmetry: if λ is in the spectrum and if $|\alpha| = 1$, then $\alpha\lambda$ is in the spectrum.

76. Similarity of weighted shifts. Is the converse of Problem 75 true? Suppose, in other words, that A and B are weighted shifts, with weights $\{\alpha_n\}$ and $\{\beta_n\}$; if A and B are unitarily equivalent, does it follow that $|\alpha_n| = |\beta_n|$ for all n ? The answer can be quite elusive, but with the right approach it is easy. The answer is no; the reason is that, for bilateral shifts, a translation of the weights produces a unitarily equivalent shift. That is: if $Ae_n = \alpha_n e_{n+1}$ and $Be_n = \alpha_{n+1} e_{n+1}$ ($n = 0, \pm 1, \pm 2, \dots$), then A and B are unitarily equivalent. If, in fact, W is the bilateral shift ($We_n = e_{n+1}$, $n = 0, \pm 1, \pm 2, \dots$), then $W^*AW = B$; if, however, the sequence $\{|\alpha_n|\}$ is not constant, then there is at least one n such that $|\alpha_n| \neq |\alpha_{n+1}|$.

Unilateral shifts behave differently. If some of the weights are allowed to be zero, the situation is in part annoying and in part trivial. In the good case (no zero weights), the kernel of the adjoint A^* of a unilateral weighted shift is spanned by e_0 , the kernel of A^{*2} is spanned by e_0 and e_1 , and, in general, the kernel of A^{*n} is spanned by e_0, \dots, e_{n-1} ($n = 1, 2, 3, \dots$). If A and B are unitarily equivalent weighted shifts, then A^{*n} and B^{*n} are unitarily equivalent; if, say, $A = U^*BU$, then U must send $\ker A^{*n}$ onto $\ker B^{*n}$. This implies that the span of $\{e_0, \dots, e_{n-1}\}$ is invariant under U , and from this, in turn, it follows that U is a diagonal operator. Since the diagonal entries of a unitary diagonal matrix have modulus 1, it follows that, for each n , the effect of A on e_n can differ from that of B by a factor of modulus 1 only.

This settles the unitary equivalence theory for weighted shifts with non-zero weights; what about similarity?

Problem 76. *If A and B are unilateral weighted shifts, with non-zero weights $\{\alpha_n\}$ and $\{\beta_n\}$, then a necessary and sufficient condition that A and B be similar is that the sequence of quotients*

$$\left| \frac{\alpha_0 \cdots \alpha_n}{\beta_0 \cdots \beta_n} \right|$$

be bounded away from 0 and from ∞ .

Similarity is a less severe restriction than unitary equivalence; questions about similarity are usually easier to answer. By a modification of the argument for one-sided shifts, a modification whose difficulties are more notational than conceptual, it is possible to get a satisfactory condition, like that in Problem 76, for the similarity of two-sided shifts; this was done by R. L. Kelley.

77. Norm and spectral radius of a weighted shift.

Problem 77. *Express the norm and the spectral radius of a weighted shift in terms of its weights.*

78. Eigenvalues of weighted shifts. The exact determination of the spectrum and its parts for arbitrary weighted shifts is a non-trivial problem. Here is a useful fragment.

Problem 78. *Find all the eigenvalues of all unilateral weighted shifts (with non-zero weights) and of their adjoints.*

The possible presence of 0 among the weights is not a genuine difficulty but a nuisance. A unilateral weighted shift, one of whose weights vanishes, becomes thereby the direct sum of a finite-dimensional operator and another weighted shift. The presence of an infinite number of zero weights can cause some interesting trouble (cf. Problem 81), but the good problems about shifts have to do with non-zero weights.

79. Weighted sequence spaces. The expression “weighted shift” means one thing, but it could just as well have meant something else. What it does mean is to modify the ordinary shift on the ordinary sequence space l^2 by attaching weights to the transformation; what it could have meant is to modify by attaching weights to the space.

To get an explicit description of the alternative, let $p = \{p_0, p_1, p_2, \dots\}$ be a sequence of strictly positive numbers, and let $l^2(p)$ be the set of complex sequences $\langle \xi_0, \xi_1, \xi_2, \dots \rangle$ with $\sum_{n=0}^{\infty} p_n |\xi_n|^2 < \infty$. With respect to the coordinatewise linear operations and the inner product defined by

$$(\langle \xi_0, \xi_1, \xi_2, \dots \rangle, \langle \eta_0, \eta_1, \eta_2, \dots \rangle) = \sum_{n=0}^{\infty} p_n \xi_n \eta_n^*,$$

the set $l^2(p)$ is a Hilbert space; it may be called a weighted sequence space. (All this is unilateral; the bilateral case can be treated similarly.) When is the shift an operator on this space? When, in other words, is it true that if $f = \langle \xi_0, \xi_1, \xi_2, \dots \rangle \in l^2(p)$, then $Sf = \langle 0, \xi_0, \xi_1, \xi_2, \dots \rangle \in l^2(p)$, and, as f varies over $l^2(p)$, $\|Sf\|$ is bounded by a constant multiple of $\|f\|$? The answer is easy. An obviously necessary condition is that there exist a positive constant α such that $\|e_{n+1}\| \leq \alpha \|e_n\|$, where e_n , of course, is the vector whose coordinate with index n is 1 and all other coordinates are 0. Since $\|e_n\|^2 = p_n$, this condition says that the sequence $\{p_{n+1}/p_n\}$ is bounded. It is almost obvious that this necessary condition is also sufficient. If $p_{n+1}/p_n \leq \alpha^2$ for all n , then

$$\begin{aligned} \|Sf\|^2 &= \sum_{n=1}^{\infty} p_n |\xi_{n-1}|^2 = \sum_{n=1}^{\infty} \frac{p_n}{p_{n-1}} p_{n-1} |\xi_{n-1}|^2 \\ &\leq \alpha^2 \sum_{n=0}^{\infty} p_n |\xi_n|^2 = \alpha^2 \|f\|^2. \end{aligned}$$

Every question about weighted shifts on the ordinary sequence space can be re-asked about the ordinary shift on weighted sequence spaces; here is a sample.

Problem 79. If $p = \{p_n\}$ is a sequence of positive numbers such that $\{p_{n+1}/p_n\}$ is bounded, what, in terms of $\{p_n\}$, is the spectral radius of the shift on $l^2(p)$?

80. One-point spectrum. The proof in Problem 48 (every non-empty compact subset of the plane is the spectrum of some operator) is not sufficiently elastic to yield examples of all the different ways spectral parts can behave. That proof used diagonal operators, which always have eigenvalues; from that proof alone it is not possible to infer the existence of operators whose point spectrum is empty. Multiplication operators come to the rescue. If D is a bounded region, if $\varphi(z) = z$ for z in D , and if A is the multiplication operator induced by φ on the L^2 space of planar Lebesgue measure in D , then the spectrum of A is the closure \bar{D} , but the point spectrum of A is empty. Similar techniques show the existence of operators A with $\Pi_0(A) = \emptyset$ and $\Lambda(A) = [0, 1]$,

say; just use linear Lebesgue measure in $[0,1]$. Whenever a compact set M in the plane is the support of a measure (on the Borel sets) that gives zero weight to each single point, then M is the spectrum of an operator with no eigenvalues. (To say that M is the support of μ means that if N is an open set with $\mu(M \cap N) = 0$, then $M \cap N = \emptyset$.) It is a routine exercise in topological measure theory to prove that every non-empty, compact, perfect set (no isolated points) in the plane is the support of a measure (on the Borel sets) that gives zero weight to each single point. (The proof is of no relevance to Hilbert space theory.) It follows that every such set is the spectrum of an operator with no eigenvalues. What about sets that are not perfect?

A very satisfactory answer can be given in terms of the appropriate analytic generalization of the algebraic concept of nilpotence. An operator is nilpotent if some positive integral power of it is zero (and the least such power is the index of nilpotence); an operator A is *quasinilpotent* if $\lim_n \|A^n\|^{1/n} = 0$. It is obvious that nilpotence implies quasinilpotence. The spectral mapping theorem implies that if A is nilpotent, then $\Lambda(A) = \{0\}$. The expression for the spectral radius in terms of norms implies that if A is quasinilpotent, then $\Lambda(A) = \{0\}$, and that, moreover, the converse is true. A nilpotent operator always has a non-trivial kernel, and hence a non-empty point spectrum; for quasinilpotent operators that is not so.

Problem 80. *Construct a quasinilpotent operator whose point spectrum is empty.*

Observe that on finite-dimensional spaces such a construction is clearly impossible.

81. Spectrum of a direct sum. The spectrum of the direct sum of two operators is the union of their spectra, and the same is true of the point spectrum, the approximate point spectrum, and the compression spectrum. The extension of this result from two direct summands to any finite number is a trivial induction. What happens if the number of summands is infinite? A possible clue to the answer is the behavior of diagonal operators on infinite-dimensional spaces. Such an operator is an infinite direct sum, each summand of which is an operator on a

one-dimensional space, and its spectrum is the closure of the union of their spectra (Problem 48).

Problem 81. *Is the spectrum of a direct sum of operators always the closure of the union of their spectra?*

82. Reid's inequality. Algebraic properties of operators, such as being Hermitian or positive, have not played much of a role so far in this book. They occur in the next problem, but only incidentally; the main point of the problem in its present location is its reference to spectral radius.

Problem 82. *If A and B are operators such that A is positive and AB is Hermitian, then $|(ABf, f)| \leq r(B) \cdot (Af, f)$ for every vector f .*

A slightly weaker version of the result is due to Reid [1951]; it is weaker in that he has $\|B\|$ instead of $r(B)$.

Chapter 11. Norm topology

83. Metric space of operators. If the distance between two operators A and B is defined to be $\|A - B\|$, the set of all operators on a Hilbert space becomes a metric space. Some of the standard metric and topological questions about that space have more interesting answers than others. Thus, for instance, it is no more than minimum courtesy to ask whether or not the space is complete. The answer is yes. The proof is the kind of routine analysis every mathematician has to work through at least once in his life; it offers no surprises. The result, incidentally, has been tacitly used already. In Solution 72, the convergence of the series $\sum_{n=0}^{\infty} A^n$ was inferred from the assumption $\|A\| < 1$. The alert reader should have noted that the justification of this inference is in the completeness result just mentioned. (It takes less alertness to notice that the very concept of convergence refers to some topology.)

So much for completeness; what about separability? If the underlying Hilbert space is not separable, it is not to be expected that the operator space is, and, indeed, it is easy to prove that it is not. That leaves one more natural question along these lines.

Problem 83. *If a Hilbert space is separable, does it follow that the metric space of operators on it is separable?*

84. Continuity of inversion. Soon after the introduction of a topology on an algebraic structure, such as the space of operators on a Hilbert space, it is customary and necessary to ask about the continuity of the pertinent algebraic operations. In the present case it turns out that all the elementary algebraic operations (linear combination, conjugation, multiplication) are continuous in all their variables simultaneously, and the norm of an operator is also a continuous function of its argument. The proofs are boring.

The main algebraic operation not mentioned above is inversion. Since not every operator is invertible, the question of the continuity of inversion makes sense on only a subset of the space of operators.

Problem 84. *The set of invertible operators is open. Is the mapping $A \rightarrow A^{-1}$ of that set onto itself continuous?*

The statement that the set of invertible operators is open does not answer all questions about the geometry of that set. It does not say, for instance, whether or not invertible operators can completely surround a singular (= non-invertible) one. In more technical language: are there any isolated singular operators? The answer is no; the set of singular operators is (arcwise) connected. Reason: if A is singular, so is tA for all scalars t ; the mapping $t \rightarrow tA$ is a continuous curve that joins the operator 0 to the operator A . Is the open set of invertible operators connected also? That question is much harder; see Problem 110.

85. Continuity of spectrum. The spectrum (restricted for a moment to operators on just one fixed Hilbert space) is a function whose domain consists of operators and whose range consists of compact sets of complex numbers. It would be quite reasonable to try to define what it means for a function of this kind to be continuous. Is the spectrum continuous? The following example is designed to prove that however the question is interpreted, the answer is always no.

Problem 85. *If $k = 1, 2, 3, \dots$ and if $k = \infty$, let A_k be the two-sided weighted shift such that $A_k e_n$ is e_{n+1} or $(1/k)e_{n+1}$ according as $n \neq 0$ or $n = 0$. (Put $1/\infty = 0$.) What are the spectra of the operators A_k ($k = 1, 2, 3, \dots, \infty$)?*

86. Semicontinuity of spectrum. The example of Problem 85 shows that there exists an operator with a large spectrum in every neighborhood of which there are operators with relatively small spectra. Could it happen the other way? Is there a small spectrum with arbitrarily near large spectra? The answer turns out to be no. The precise assertion is that the spectrum is an upper semicontinuous function, in the following sense.

Problem 86. *To each operator A and to each open set Λ_0 that includes $\Lambda(A)$ there corresponds a positive number ϵ such that if $\|A - B\| < \epsilon$, then $\Lambda(B) \subset \Lambda_0$.*

This is a standard result. One standard reference is Hille-Phillips [1957, p. 167]; another is Rickart [1960, p. 35]. The semicontinuity of related functions is discussed in Halmos-Lumer [1954].

87. Continuity of spectral radius. Since the spectrum is upper semicontinuous (Problem 86), so is the spectral radius. That is: to each operator A and to each positive number δ there corresponds a positive number ϵ such that if $\|A - B\| < \epsilon$, then $r(B) < r(A) + \delta$. (The proof is immediate from Problem 86.) The spectrum is not continuous (Problem 85); what about the spectral radius?

Problem 87. *Is it true that to each operator A and to each positive number δ there corresponds a positive number ϵ such that if $\|A - B\| < \epsilon$, then $|r(A) - r(B)| < \delta$? Equivalently: if $A_n \rightarrow A$, does it follow that $r(A_n) \rightarrow r(A)$?*

This is hard. Note that the example in Problem 85 gives no information; in that case the spectral radius is equal to 1 for each term of the sequence and also for the limit.

Chapter 12.

Strong and weak topologies

88. Topologies for operators. A Hilbert space has two useful topologies (weak and strong); the space of operators on a Hilbert space has several. The metric topology induced by the norm is one of them; to distinguish it from the others, it is usually called the *norm topology* or the *uniform topology*. The next two are natural outgrowths for operators of the strong and weak topologies for vectors. A subbase for the *strong* operator topology is the collection of all sets of the form

$$\{A: \|(A - A_0)f\| < \varepsilon\};$$

correspondingly a base is the collection of all sets of the form

$$\{A: \|(A - A_0)f_i\| < \varepsilon, \quad i = 1, \dots, k\}.$$

Here k is a positive integer, f_1, \dots, f_k are vectors, and ε is a positive number. A subbase for the *weak* operator topology is the collection of all sets of the form

$$\{A: |((A - A_0)f, g)| < \varepsilon\},$$

where f and g are vectors and $\varepsilon > 0$; as above (as always) a base is the collection of all finite intersections of such sets. The corresponding concepts of convergence (for sequences and nets) are easy to describe: $A_n \rightarrow A$ strongly if and only if $A_n f \rightarrow A f$ strongly for each f (i.e., $\|(A_n - A)f\| \rightarrow 0$ for each f), and $A_n \rightarrow A$ weakly if and only if $A_n f \rightarrow A f$ weakly for each f (i.e., $(A_n f, g) \rightarrow (A f, g)$ for each f and g).

The easiest questions to settle are the ones about comparison. The weak topology is smaller (weaker) than the strong topology, and the strong topology is smaller than the norm topology. In other words, every weak open set is a strong open set, and every strong open set is norm open. In still other words: every weak neighborhood of each oper-

ator includes a strong neighborhood of that operator, and every strong neighborhood includes a metric neighborhood. Again: norm convergence implies strong convergence, and strong convergence implies weak convergence. These facts are immediate from the definitions. In the presence of uniformity on the unit sphere, the implications are reversible (cf. Problem 16).

Problem 88. *If $(A_n f, g) \rightarrow (A f, g)$ uniformly for $\|g\| = 1$, then $\|A_n f - A f\| \rightarrow 0$, and if $\|A_n f - A f\| \rightarrow 0$ uniformly for $\|f\| = 1$, then $\|A_n - A\| \rightarrow 0$.*

89. Continuity of norm. In the study of topological algebraic structures (such as the algebra of operators on a Hilbert space, endowed with one of the appropriate operator topologies) the proof that something is continuous is usually dull; the interesting problems arise in proving that something is not continuous. Thus, for instance, it is true that the linear operations on operators ($\alpha A + \beta B$) are continuous in all variables simultaneously, and the proof is a matter of routine. (Readers who have never been through this routine are urged to check it before proceeding.) Here is a related question that is easy but not quite so mechanical.

Problem 89. *Which of the three topologies (uniform, strong, weak) makes the norm (i.e., the function $A \rightarrow \|A\|$) continuous?*

90. Continuity of adjoint.

Problem 90. *Which of the three topologies (uniform, strong, weak) makes the adjoint (i.e., the mapping $A \rightarrow A^*$) continuous?*

91. Continuity of multiplication. The most useful, and most recalcitrant, questions concern products. Since a product (unlike the norm and the adjoint) is a function of two variables, a continuity statement about products has a “joint” and a “separate” interpretation. It is usual, when nothing is said to the contrary, to interpret such statements in the “joint” sense, i.e., to interpret them as referring to the mapping that sends an ordered pair $\langle A, B \rangle$ onto the product AB .

Problem 91. *Multiplication is continuous with respect to the uniform topology and discontinuous with respect to the strong and weak topologies.*

The proof is easy, but the counterexamples are hard; the quickest ones depend on unfair trickery.

92. Separate continuity of multiplication. Although multiplication is not jointly continuous with respect to either the strong topology or the weak, it is separately continuous in each of its arguments with respect to both topologies. A slightly more precise formulation runs as follows.

Problem 92. *Each of the mappings $A \rightarrow AB$ (for fixed B) and $B \rightarrow AB$ (for fixed A) is both strongly and weakly continuous.*

93. Sequential continuity of multiplication. Separate continuity (strong and weak) of multiplication is a feeble substitute for joint continuity; another feeble (but sometimes usable) substitute is joint continuity in the sequential sense.

Problem 93. (a) *If $\{A_n\}$ and $\{B_n\}$ are sequences of operators that strongly converge to A and B , respectively, then $A_n B_n \rightarrow AB$ strongly.*
 (b) *Does the assertion remain true if "strongly" is replaced by "weakly" in both hypothesis and conclusion?*

94. Increasing sequences of Hermitian operators. A bounded increasing sequence of Hermitian operators is weakly convergent (to a necessarily Hermitian operator). To see this, suppose that $\{A_n\}$ is an increasing sequence of Hermitian operators (i.e., $(A_n f, f) \leq (A_{n+1} f, f)$ for all n and all f), bounded by α (i.e., $(A_n f, f) \leq \alpha \|f\|^2$ for all n and all f). If $\psi_n(f) = (A_n f, f)$, then each ψ_n is a quadratic form. The assumptions imply that the sequence $\{\psi_n\}$ is convergent and hence (Solution 1) that the limit ψ is a quadratic form. It follows that $\psi(f) = (Af, f)$ for some (necessarily Hermitian) operator A ; polarization justifies the conclusion that $A_n \rightarrow A$ (weakly).

Does the same conclusion follow with respect to the strong and the uniform topologies?

Problem 94. *Is a bounded increasing sequence of Hermitian operators necessarily strongly convergent? uniformly convergent?*

95. Square roots. The assertion that a positive operator has a unique positive square root is an easy consequence of the spectral theorem. In some approaches to spectral theory, however, the existence of square roots is proved first, and the spectral theorem is based on that result. The following assertion shows how to get square roots without the spectral theorem.

Problem 95. *If A is an operator such that $0 \leq A \leq 1$, and if a sequence $\{B_n\}$ is defined recursively by the equations*

$$B_0 = 0 \quad \text{and} \quad B_{n+1} = \frac{1}{2}((1 - A) + B_n^2), \quad n = 0, 1, 2, \dots,$$

then the sequence $\{B_n\}$ is strongly convergent. If $\lim_n B_n = B$, then $(1 - B)^2 = A$.

96. Infimum of two projections. If E and F are projections with ranges \mathbf{M} and \mathbf{N} , then it is sometimes easy and sometimes hard to find, in terms of E and F , the projections onto various geometric constructs formed with \mathbf{M} and \mathbf{N} . Things are likely to be easy if E and F commute. Thus, for instance, if $\mathbf{M} \subset \mathbf{N}$, then it is easy to find the projection with range $\mathbf{N} \cap \mathbf{M}^\perp$, and if $\mathbf{M} \perp \mathbf{N}$, then it is easy to find the projection with range $\mathbf{M} \vee \mathbf{N}$. In the absence of such special assumptions, the problems become more interesting.

Problem 96. *If E and F are projections with ranges \mathbf{M} and \mathbf{N} , find the projection $E \wedge F$ with range $\mathbf{M} \cap \mathbf{N}$.*

The problem is to find an “expression” for the projection described. Although most mathematicians would read the statement of such a problem with sympathetic understanding, it must be admitted that rigorously speaking it does not really mean anything. The most obvious

way to make it precise is to describe certain classes of operators by the requirement that they be closed under some familiar algebraic and topological operations, and then try to prove that whenever E and F belong to such a class, then so does $E \wedge F$. The most famous and useful classes pertinent here are the von Neumann algebras (called “rings of operators” by von Neumann). A *von Neumann algebra* is an algebra of operators (i.e., a collection closed under addition and multiplication, and closed under multiplication by arbitrary scalars), self-adjoint (i.e., closed under adjunction), containing 1, and strongly closed (i.e., closed with respect to the strong operator topology). For von Neumann algebras, then, the problem is this: prove that if a von Neumann algebra contains two projections E and F , then it contains $E \wedge F$.

Reference: von Neumann [1950, vol. 2, p. 55].

Chapter 13. Partial isometries

97. Spectral mapping theorem for normal operators. Normal operators constitute the most important tractable class of operators known; the most important statement about them is the spectral theorem. Students of operator theory generally agree that the finite-dimensional version of the spectral theorem has to do with diagonal forms. (Every finite normal matrix is unitarily equivalent to a diagonal one.) The general version, applicable to infinite-dimensional spaces, does not have a universally accepted formulation. Sometimes bounded operator representations of function algebras play the central role, and sometimes Stieltjes integrals with unorthodox multiplicative properties. There is a short, simple, and powerful statement that does not attain maximal generality (it applies to only one operator at a time, not to algebras of operators), but that does have all classical formulations of the spectral theorem as easy corollaries, and that has the advantage of being a straightforward generalization of the familiar statement about diagonal forms. That statement will be called the spectral theorem in what follows; it says that *every normal operator is unitarily equivalent to a multiplication*. The statement can be proved by exactly the same techniques as are usually needed for the spectral theorem; see Halmos [1963], Dunford-Schwartz [1963, pp. 911–912].

The multiplication version of the spectral theorem has a technical drawback: the measures that it uses may fail to be σ -finite. This is not a tragedy, for two reasons. In the first place, the assumption of σ -finiteness in the treatment of multiplications is a matter of convenience, not of necessity (see Segal [1951]). In the second place, non- σ -finite measures need to be considered only when the underlying Hilbert space is not separable; the pathology of measures accompanies the pathology of operators. In the sequel when reference is made to the spectral theorem, the reader may choose one of two courses: treat the general case and proceed with the caution it requires, or restrict attention to the separable case and proceed with the ease that the loss of generality permits.

In some contexts some authors choose to avoid a proof that uses the spectral theorem even if the alternative is longer and more involved.

This sort of ritual circumlocution is common to many parts of mathematics; it is the fate of many big theorems to be more honored in evasion than in use. The reason is not just mathematical mischievousness. Often a long but "elementary" proof gives more insight, and leads to more fruitful generalizations, than a short proof whose brevity is made possible by a powerful but overly specialized tool.

This is not to say that use of the spectral theorem is to be avoided at all costs. Powerful general theorems exist to be used, and their willful avoidance can lose insight at least as often as gain it. Thus, for example, the spectral theorem yields an immediate and perspicuous proof that every positive operator has a positive square root (because every positive measurable function has one); the approximation trickery of Problem 95 is fun, and has its uses, but it is not nearly so transparent. For another example, consider the assertion that a Hermitian operator whose spectrum consists of the two numbers 0 and 1 is a projection. To prove it, let A be the operator, and write $B = A - A^2$. Clearly B is Hermitian, and, by the spectral mapping theorem, $\Lambda(B) = \{0\}$. This implies that $\|B\| = r(B) = 0$ and hence that $B = 0$. (It is true for all normal operators that the norm is equal to the spectral radius, but for Hermitian operators it is completely elementary; see Halmos [1951, p.55].) Compare this with the proof via the spectral theorem: if φ is a function whose range consists of the two numbers 0 and 1, then $\varphi^2 = \varphi$. For a final example, try to prove, without using the spectral theorem, that every normal operator with a real spectrum (i.e., with spectrum included in the real line) is Hermitian.

The spectral theorem makes possible a clear and efficient description of the so-called *functional calculus*. If A is a normal operator and if F is a bounded Borel measurable function on $\Lambda(A)$, then the functional calculus yields an operator $F(A)$. To define $F(A)$ represent A as a multiplication, with multiplier φ , say, on a measure space X ; the operator $F(A)$ is then the multiplication induced by the composite function $F \circ \varphi$. In order to be sure that this makes sense, it is necessary to know that φ maps almost every point of X into $\Lambda(A)$, i.e., that if the domain of φ is altered by, at worst, a set of measure zero, then the range of φ comes to be included in its essential range. The proof goes as follows. By definition, every point in the complement of $\Lambda(A)$ has a neighborhood whose inverse image under φ has measure zero. Since the plane is a

Lindelöf space, it follows that the complement of $\Lambda(A)$ is covered by a countable collection of neighborhoods with that property, and hence that the inverse image of the entire complement of $\Lambda(A)$ has measure zero.

The mapping $F \rightarrow F(A)$ has many pleasant properties. Its principal property is that it is an algebraic homomorphism that preserves conjugation also (i.e., $F^*(A) = (F(A))^*$); it follows, for instance, that if $F(\lambda) = |\lambda|^2$, then $F(A) = A^*A$. The functions F that occur in the applications of the functional calculus are not always continuous (e.g., characteristic functions of Borel sets are of importance), but continuous functions are sometimes easier to handle. The problem that follows is a spectral mapping theorem; it is very special in that it refers to normal operators only, but it is very general in that it allows all continuous functions.

Problem 97. *If A is a normal operator and if F is a continuous function on $\Lambda(A)$, then $\Lambda(F(A)) = F(\Lambda(A))$.*

Something like $F(A)$ might seem to make sense sometimes even for non-normal A 's, but the result is not likely to remain true. Suppose, for instance, that $F(\lambda) = \lambda^*\lambda (= |\lambda|^2)$, and define $F(A)$, for every operator A , as A^*A . There is no hope for the statement $F(\Lambda(A)) = \Lambda(F(A))$; for a counterexample, contemplate the unilateral shift.

98. Partial isometries. An *isometry* is a linear transformation U (from a Hilbert space into itself, or from one Hilbert space into another) such that $\|Uf\| = \|f\|$ for all f . An isometry is a distance-preserving transformation: $\|Uf - Ug\| = \|f - g\|$ for all f and g . A necessary and sufficient condition that a linear transformation U be an isometry is that $U^*U = 1$. Indeed: the conditions (1) $\|Uf\|^2 = \|f\|^2$, (2) $(U^*Uf, f) = (f, f)$, (3) $(U^*Uf, g) = (f, g)$, and (4) $U^*U = 1$ are mutually equivalent. (To pass from (2) to (3), polarize.) Caution: the conditions $U^*U = 1$ and $UU^* = 1$ are not equivalent. The latter condition is satisfied in case U^* is an isometry; in that case U is called a *co-isometry*.

It is sometimes convenient to consider linear transformations U that act isometrically on a subset (usually a linear manifold, but not neces-

sarily a subspace) of a Hilbert space; this just means that $\|Uf\| = \|f\|$ for all f in that subset. A *partial isometry* is a linear transformation that is isometric on the orthogonal complement of its kernel. There are two large classes of examples of partial isometries that are in a sense opposite extreme cases; they are the isometries (and, in particular, the unitary operators), and the projections. The definition of partial isometries is deceptively simple, and these examples continue the deception; the structure of partial isometries can be quite complicated. In any case, however, it is easy to verify that a partial isometry U is bounded; in fact if U is not 0, then $\|U\| = 1$.

The orthogonal complement of the kernel of a partial isometry is frequently called its *initial space*. The initial space of a partial isometry U turns out to be equal to the set of all those vectors f for which $\|Uf\| = \|f\|$. (What needs proof is that if $\|Uf\| = \|f\|$, then $f \perp \ker U$. Write $f = g + h$, with $g \in \ker U$ and $h \perp \ker U$; then $\|f\| = \|Uf\| = \|Ug + Uh\| = \|Uh\| = \|h\|$; since $\|f\|^2 = \|g\|^2 + \|h\|^2$, it follows that $g = 0$.) The range of a partial isometry is equal to the image of the initial space and is necessarily closed. (Since U is isometric on the initial space, the image is a complete metric space.) For partial isometries, the range is sometimes called the *final space*.

Problem 98. *A bounded linear transformation U is a partial isometry if and only if U^*U is a projection.*

Corollary 1. *If U is a partial isometry, then the initial space of U is the range of U^*U .*

Corollary 2. *The adjoint of a partial isometry is a partial isometry, with initial space and final space interchanged.*

Corollary 3. *A bounded linear transformation U is a partial isometry if and only if $U = UU^*U$.*

99. Maximal partial isometries. It is natural to define a (partial) order for partial isometries as follows: if U and V are partial isometries, write $U \leq V$ in case V agrees with U on the initial space of U . This implies that the initial space of U is included in the initial space of V .

(Cf. the characterization of initial spaces given in Problem 98.) It follows that if $U \leq V$ with respect to the present order, then $U^*U \leq V^*V$ with respect to the usual order for operators. (The “usual” order, usually considered for Hermitian operators only, is the one according to which $A \leq B$ if and only if $(Af, f) \leq (Bf, f)$ for all f . Note that $U^*U \leq V^*V$ in this sense is equivalent to $\|Uf\| \leq \|Vf\|$ for all f .) The converse is not true; if all that is known about the partial isometries U and V is that $U^*U \leq V^*V$, then, to be sure, the initial space of U is included in the initial space of V , but it cannot be concluded that U and V necessarily agree on the smaller initial space.

If $U^*U = 1$, i.e., if U is an isometry, then the only partial isometry that can dominate U is U itself: an isometry is a maximal partial isometry. Are there any other maximal partial isometries? One way to get the answer is to observe that if $U \leq V$, then the final space of U (i.e., the initial space of U^*) is included in the final space of V (the initial space of V^*), and, moreover, V^* agrees with U^* on the initial space of U^* . In other words, if $U \leq V$, then $U^* \leq V^*$, and hence, in particular, $UU^* \leq VV^*$. This implies that if $UU^* = 1$, i.e., if U is a co-isometry, then, again, U is maximal. If a partial isometry U is neither an isometry nor a co-isometry, then both U and U^* have non-zero kernels. In that case it is easy to enlarge U to a partial isometry that maps a prescribed unit vector in $\ker U$ onto a prescribed unit vector in $\ker U^*$ (and, of course, agrees with U on $(\ker U)^\perp$). Conclusion: a partial isometry is maximal if and only if either it or its adjoint is an isometry.

The easy way to be a maximal partial isometry is to be unitary. If U is unitary on \mathbf{H} and if \mathbf{M} is a subspace of \mathbf{H} , then a necessary and sufficient condition that \mathbf{M} reduce U is that $U\mathbf{M} = \mathbf{M}$. If U is merely a partial isometry, then it can happen that $U\mathbf{M} = \mathbf{M}$ but \mathbf{M} does not reduce U , and it can happen that \mathbf{M} reduces U but $U\mathbf{M} \neq \mathbf{M}$. What if U is a maximal partial isometry?

Problem 99. *Discover the implication relations between the statements “ $U\mathbf{M} = \mathbf{M}$ ” and “ \mathbf{M} reduces U ” when U is a maximal partial isometry.*

100. Closure and connectedness of partial isometries. Some statements about partial isometries are slightly awkward just because 0 must be counted as one of them. The operator 0 is an isolated point of

the set of partial isometries; it is the only partial isometry in the interior of the unit ball. For this reason, for instance, the set of all partial isometries is obviously not connected. What about the partial isometries on the boundary of the unit ball?

Problem 100. *The set of all non-zero partial isometries is closed but not connected (with respect to the norm topology of operators).*

101. Rank, co-rank, and nullity. If U is a partial isometry, write $\rho(U) = \dim \operatorname{ran} U$, $\rho'(U) = \dim (\operatorname{ran} U)^\perp$, and $\nu(U) = \dim \ker U$. (That U is a partial isometry is not really important in these definitions; similar definitions can be made for arbitrary operators.) These three cardinal numbers, called the *rank*, the *co-rank*, and the *nullity* of U , respectively, are not completely independent of one another; they are such that both $\rho + \rho'$ and $\rho + \nu$ are equal to the dimension of the underlying Hilbert space. (Caution: subtraction of infinite cardinal numbers is slippery; it does not follow that $\rho' = \nu$.) It is easy to see that if ρ , ρ' , and ν are any three cardinal numbers such that $\rho + \rho' = \rho + \nu$, then there exist partial isometries with rank ρ , co-rank ρ' , and nullity ν . (Symmetry demands the consideration of $\nu'(U) = \dim (\ker U)^\perp$, the *co-nullity* of U , but there is no point in it; since U is isometric on $(\ker U)^\perp$ it follows that $\nu' = \rho$.)

Recall that if U is a partial isometry, then so is U^* ; the initial space of U^* is the final space of U , and vice versa. It follows that $\nu(U^*) = \rho'(U)$ and $\rho'(U^*) = \nu(U)$.

One reason that the functions ρ , ρ' , and ν are useful is that they are continuous. To interpret this statement, use the norm topology for the space \mathbf{P} of partial isometries (on a fixed Hilbert space), and use the discrete topology for cardinal numbers. With this explanation the meaning of the continuity assertion becomes unambiguous: if U is sufficiently near to V , then U and V have the same rank, the same co-rank, and the same nullity. The following assertion is a precise quantitative formulation of the result.

Problem 101. *If U and V are partial isometries such that $\|U - V\| < 1$, then $\rho(U) = \rho(V)$, $\rho'(U) = \rho'(V)$, and $\nu(U) = \nu(V)$.*

For each fixed ρ , ρ' , and ν let $\mathbf{P}(\rho, \rho', \nu)$ be the set of partial isometries (on a fixed Hilbert space) with rank ρ , co-rank ρ' , and nullity ν . Clearly the sets of the form $\mathbf{P}(\rho, \rho', \nu)$ constitute a partition of the space \mathbf{P} of all partial isometries; it is a consequence of the statement of Problem 101 that each set $\mathbf{P}(\rho, \rho', \nu)$ is both open and closed. It follows that the set of all isometries ($\nu = 0$) is both open and closed, and so is the set of all unitary operators ($\rho' = \nu = 0$).

102. Components of the space of partial isometries. If φ is a measurable function on a measure space, such that $|\varphi| = 1$ almost everywhere, then there exists a measurable real-valued function θ on that space such that $\varphi = e^{i\theta}$ almost everywhere. This is easy to prove. What it essentially says is that a measurable function always has a measurable logarithm. The reason is that the exponential function has a Borel measurable inverse (in fact many of them) on the complement of the origin in the complex plane. (Choose a continuous logarithm on the complement of the negative real axis, and extend it by requiring one-sided continuity on, say, the upper half plane.)

In the language of the functional calculus, the result of the preceding paragraph can be expressed as follows: if U is a unitary operator, then there exists a Hermitian operator A such that $U = e^{iA}$. If $U_t = e^{itA}$, $0 \leq t \leq 1$, then $t \rightarrow U_t$ is a continuous curve of unitary operators joining $1 (= U_0)$ to $U (= U_1)$. Conclusion: the set of all unitary operators is arcwise connected. In the notation of Problem 101, the open-closed set $\mathbf{P}(\rho, 0, 0)$ (on a Hilbert space of dimension ρ) is connected; it is a component of the set \mathbf{P} of all partial isometries. Question: what are the other components? Answer: the sets of the form $\mathbf{P}(\rho, \rho', \nu)$.

Problem 102. *Each pair of partial isometries (on the same Hilbert space) with the same rank, co-rank, and nullity, can be joined by a continuous curve of partial isometries with the same rank, co-rank, and nullity.*

103. Unitary equivalence for partial isometries. If A is a contraction (that means $\|A\| \leq 1$), then $1 - AA^*$ is positive. It follows that there exists a unique positive operator whose square is $1 - AA^*$; call it A' .

Assertion: the operator matrix

$$M(A) = \begin{pmatrix} A & A^* \\ 0 & 0 \end{pmatrix}$$

is a partial isometry. Proof (via Problem 98): check that $MM^*M = M$.
Consequence: every contraction can be extended to a partial isometry.

Problem 103. *If A and B are unitarily equivalent contractions, then $M(A)$ and $M(B)$ are unitarily equivalent; if A and B are invertible contractions, then the converse is true.*

There are many ways that a possibly “bad” operator A can be used to manufacture a “good” one. Samples: $A + A^*$ and

$$\begin{pmatrix} 0 & A \\ A^* & 0 \end{pmatrix}.$$

None of these ways yields sufficiently many usable unitary invariants for A . It is usually easy to prove that if A and B are unitarily equivalent, then so are the various constructs in which they appear. It is, however, usually false that if the constructs are unitarily equivalent, then the original operators themselves are. The chief interest of the assertion of Problem 103 is that, for the special partial isometry construct it deals with, the converse happens to be true.

The result is that the unitary equivalence problem for an apparently very small class of operators (partial isometries) is equivalent to the problem for the much larger class of invertible contractions. The unitary equivalence problem for invertible contractions is, in turn, trivially equivalent to the unitary equivalence problem for arbitrary operators. The reason is that by a translation ($A \rightarrow A + \alpha$) and a change of scale ($A \rightarrow \beta A$) every operator becomes an invertible contraction, and translations and changes of scale do not affect unitary equivalence. The end product of all this is a reduction of the general unitary equivalence problem to the special case of partial isometries.

104. Spectrum of a partial isometry. What conditions must a set of complex numbers satisfy in order that it be the spectrum of some partial

isometry? Since a partial isometry is a contraction, its spectrum is necessarily a subset of the closed unit disc. If the spectrum of a partial isometry does not contain the origin, i.e., if a partial isometry is invertible, then it is unitary, and, therefore, its spectrum is a subset of the unit circle (perimeter). Since every non-empty compact subset of the unit circle is the spectrum of some unitary operator (cf. Problem 48), the problem of characterizing the spectra of invertible partial isometries is solved. What about the non-invertible ones?

Problem 104. *What conditions must a set of complex numbers satisfy in order that it be the spectrum of some non-unitary partial isometry?*

105. Polar decomposition. Every complex number is the product of a non-negative number and a number of modulus 1; except for the number 0, this polar decomposition is unique. The generalization to finite matrices says that every complex matrix is the product of a positive matrix and a unitary one. If the given matrix is invertible, and if the order of the factors is specified (UP or PU), then, once again, this polar decomposition is unique. It is possible to get a satisfactory uniqueness theorem for every matrix, but only at the expense of changing the kind of factors admitted; this is a point at which partial isometries can profitably enter the study of finite-dimensional vector spaces. In the infinite-dimensional case, partial isometries are unavoidable. It is not true that every operator on a Hilbert space is equal to a product UP , with U unitary and P positive, and it does not become true even if U is required to be merely isometric. (The construction of concrete counter-examples may not be obvious now, but it will soon be an easy by-product of the general theory.) The correct statements are just as easy for transformations between different spaces as for operators on one space.

Problem 105. *If A is a bounded linear transformation from a Hilbert space \mathbf{H} to a Hilbert space \mathbf{K} , then there exists a partial isometry U (from \mathbf{H} to \mathbf{K}) and there exists a positive operator P (on \mathbf{H}) such that $A = UP$. The transformations U and P can be found so that $\ker U = \ker P$, and this additional condition uniquely determines them.*

The representation of A as the product of the unique U and P satisfying the stated conditions is called the *polar decomposition* of A , or, more accurately, the right-handed polar decomposition of A . The corresponding left-handed theory ($A = PU$) follows by a systematic exploitation of adjoints.

Corollary 1. *If $A = UP$ is the polar decomposition of A , then $U^*A = P$.*

Corollary 2. *If $A = UP$ is the polar decomposition of A , then a necessary and sufficient condition that U be an isometry is that A be one-to-one, and a necessary and sufficient condition that U be a co-isometry is that the range of A be dense.*

106. Maximal polar representation.

Problem 106. *Every bounded linear transformation is the product of a maximal partial isometry and a positive operator.*

107. Extreme points. The closed unit ball in the space of operators is convex. For every interesting convex set, it is of interest to determine the extreme points.

Problem 107. *What are the extreme points of the closed unit ball in the space of operators on a Hilbert space?*

108. Quasinormal operators. The condition of normality can be weakened in various ways; the most elementary of these leads to the concept of quasinormality. An operator A is called *quasinormal* if A commutes with A^*A . It is clear that every normal operator is quasinormal. The converse is obviously false. If, for instance, A is an isometry, then $A^*A = 1$ and therefore A commutes with A^*A , but if A is not unitary, then A is not normal. (For a concrete example consider the unilateral shift.)

Problem 108. *An operator with polar decomposition UP is quasinormal if and only if $UP = PU$.*

Quasinormal operators (under another name) were first introduced and studied by Brown [1953].

109. Density of invertible operators. It sometimes happens that a theorem is easy to prove for invertible operators but elusive in the general case. This makes it useful to know that every finite (square) matrix is the limit of invertible matrices. In the infinite-dimensional case the approximation technique works, with no difficulty, for normal operators. (Invoke the spectral theorem to represent the given operator as a multiplication, and, by changing the small values of the multiplier, approximate it by operators that are bounded from below.) If, however, the space is infinite-dimensional and the operator is not normal, then there is trouble.

Problem 109. *The set of all operators that have either a left or a right inverse is dense, but the set of all operators that have both a left and a right inverse (i.e., the set of all invertible operators) is not.*

110. Connectedness of invertible operators.

Problem 110. *The set of all invertible operators is connected.*

Chapter 14. Unilateral shift

111. Reducing subspaces of normal operators. One of the principal achievements of the spectral theorem is to reduce the study of a normal operator to subspaces with various desirable properties. The following assertion is one way to say that the spectral theorem provides many reducing subspaces.

Problem 111. *If A is a normal operator on an infinite-dimensional Hilbert space \mathbf{H} , then \mathbf{H} is the direct sum of a countably infinite collection of subspaces that reduce A , all with the same infinite dimension.*

112. Products of symmetries. A *symmetry* is a unitary involution, i.e., an operator Q such that $Q^*Q = QQ^* = Q^2 = 1$. It may be pertinent to recall that if an operator possesses any two of the properties “unitary”, “involutory”, and “Hermitian”, then it possesses the third; the proof is completely elementary algebraic manipulation.

Problem 112. *Discuss the assertion: every unitary operator is the product of a finite number of symmetries.*

113. Unilateral shift versus normal operators. The main point of Problem 111 is to help solve Problem 112 (and, incidentally, to provide a non-trivial application of the spectral theorem). The main point of Problem 112 is to emphasize the role of certain shift operators. Shifts (including the simple unilateral and bilateral ones introduced before) are a basic tool in operator theory. The unilateral shift, in particular, has many curious properties, both algebraic and analytic. The techniques for discovering and proving these properties are frequently valuable even when the properties themselves have no visible immediate application. Here are three sample questions.

Problem 113. (a) *Is the unilateral shift the product of a finite number of normal operators?* (b) *What is the norm of the real part*

of the unilateral shift? (c) How far is the unilateral shift from the set of normal operators?

The last question takes seriously the informal question: "How far does the unilateral shift miss being normal?" The question can be asked for every operator and the answer is a unitary invariant that may occasionally be useful.

114. Square root of shift.

Problem 114. *Does the unilateral shift have a square root? In other words, if U is the unilateral shift, does there exist an operator V such that $V^2 = U$?*

115. Commutant of the bilateral shift. The *commutant* of an operator (or of a set of operators) is the set of all operators that commute with it (or with each operator in the set). The commutant is one of the most useful things to know about an operator. One of the most important purposes of the so-called multiplicity theory is to discuss the commutants of normal operators. In some special cases the determination of the commutant is accessible by relatively elementary methods; a case in point is the bilateral shift.

The bilateral shift W can be viewed as multiplication by e_1 on L^2 of the unit circle (cf. Problem 68). Here $e_n(z) = z^n$ ($n = 0, \pm 1, \pm 2, \dots$) whenever $|z| = 1$, and L^2 is formed with normalized Lebesgue measure.

Problem 115. *The commutant of the bilateral shift is the set of all multiplications.*

Corollary. *Each reducing subspace of the bilateral shift is determined by a Borel subset M of the circle as the set of all functions (in L^2) that vanish outside M .*

Both the main statement and the corollary have natural generalizations that can be bought at the same price. The generalizations are obtained by replacing the unit circle by an arbitrary bounded Borel set X in the complex plane and replacing Lebesgue measure by an arbitrary

finite Borel measure in X . The generalization of the bilateral shift is the multiplication induced by e_1 (where $e_1(z) = z$ for all z in X).

116. Commutant of the unilateral shift. The unilateral shift is the restriction of the bilateral shift to \mathbf{H}^2 . If the bilateral shift is regarded as a multiplication, then its commutant can be described as the set of all multiplications on the same \mathbf{L}^2 (Problem 115). The wording suggests a superficially plausible conjecture: perhaps the commutant of the unilateral shift consists of the restrictions to \mathbf{H}^2 of all multiplications. On second thought this is absurd: \mathbf{H}^2 need not be invariant under a multiplication, and, consequently, the restriction of a multiplication to \mathbf{H}^2 is not necessarily an operator on \mathbf{H}^2 . If, however, the multiplier itself is in \mathbf{H}^2 (and hence in \mathbf{H}^∞), then \mathbf{H}^2 is invariant under the induced multiplication (cf. Problem 27), and the conjecture makes sense.

Problem 116. *The commutant of the unilateral shift is the set of all restrictions to \mathbf{H}^2 of multiplications by multipliers in \mathbf{H}^∞ .*

Corollary. *The unilateral shift is irreducible, in the sense that its only reducing subspaces are $\{0\}$ and \mathbf{H}^2 .*

Just as for the bilateral shift, the main statement has a natural generalization. Replace the unit circle by an arbitrary bounded Borel subset of the complex plane, and replace Lebesgue measure by an arbitrary finite Borel measure μ in X . The generalization of \mathbf{H}^2 , sometimes denoted by $\mathbf{H}^2(\mu)$, is the span in $\mathbf{L}^2(\mu)$ of the functions e_n , $n = 0, 1, 2, \dots$, where $e_n(z) = z^n$ for all z in X . The generalization of the unilateral shift is the restriction to $\mathbf{H}^2(\mu)$ of the multiplication induced by e_1 .

The corollary does not generalize so smoothly as the main statement. The trouble is that the structure of $\mathbf{H}^2(\mu)$ within $\mathbf{L}^2(\mu)$ depends strongly on X and μ ; it can, for instance, happen that $\mathbf{H}^2(\mu) = \mathbf{L}^2(\mu)$.

The characterization of the commutant of the unilateral shift yields a curious alternative proof of, and corresponding insight into, the assertion that U has no square root (Solution 114). Indeed, if $V^2 = U$, then V commutes with U , and therefore V is the restriction to \mathbf{H}^2 of the multiplication induced by a function φ in \mathbf{H}^∞ . Apply V^2 to e_0 , apply U to e_0 ,

and infer that $(\varphi(z))^2 = z$ almost everywhere. This implies that $(\tilde{\varphi}(z))^2 = z$ in the unit disc (see Solution 34), i.e., that the function \tilde{e}_1 has an analytic square root; the contradiction has arrived.

117. Commutant of the unilateral shift as limit.

Problem 117. *Every operator that commutes with the unilateral shift is the limit (strong operator topology) of a sequence of polynomials in the unilateral shift.*

118. Characterization of isometries. What can an isometry look like? Some isometries are unitary, and some are not; an example of the latter kind is the unilateral shift. Since a direct sum (finite or infinite) of isometries is an isometry, a mixture of the two kinds is possible. More precisely, the direct sum of a unitary operator and a number of copies (finite or infinite) of the unilateral shift is an isometry. (There is no point in forming direct sums of unitary operators—they are no more unitary than the summands.) The useful theorem along these lines is that that is the only way to get isometries. It follows that the unilateral shift is more than just an example of an isometry, with interesting and peculiar properties; it is in fact one of the fundamental building blocks out of which all isometries are constructed.

Problem 118. *Every isometry is either unitary, or a direct sum of one or more copies of the unilateral shift, or a direct sum of a unitary operator and some copies of the unilateral shift.*

An isometry for which the unitary direct summand is absent is called *pure*.

119. Distance from shift to unitary operators.

Problem 119. *How far is the unilateral shift from the set of unitary operators?*

120. Reduction by the unitary part. Each isometry decomposes into a unitary part and a pure part (Problem 118). If the subspace on which

the unitary part acts is \mathbf{M} , then \mathbf{M} reduces the given isometry, and, therefore, \mathbf{M} reduces each polynomial in that isometry. Does \mathbf{M} reduce every operator that commutes with that isometry?

Problem 120. *Does the domain of the unitary part of an isometry reduce every operator that commutes with the isometry?*

121. Shifts as universal operators. If U is an isometry on a Hilbert space \mathbf{H} , and if there exists a unit vector e_0 in \mathbf{H} such that the vectors e_0, Ue_0, U^2e_0, \dots form an orthonormal basis for \mathbf{H} , then (obviously) U is unitarily equivalent to the unilateral shift, or, by a slight abuse of language, U is the unilateral shift. This characterization of the unilateral shift can be reformulated as follows: U is an isometry on a Hilbert space \mathbf{H} for which there exists a one-dimensional subspace \mathbf{N} such that the subspaces $\mathbf{N}, U\mathbf{N}, U^2\mathbf{N}, \dots$ are pairwise orthogonal and span \mathbf{H} . If there is such a subspace \mathbf{N} , then it must be equal to the *co-range* $(U\mathbf{H})^\perp$. In view of this comment another slight reformulation is possible: the unilateral shift is an isometry U of co-rank 1 on a Hilbert space \mathbf{H} such that the subspaces $(U\mathbf{H})^\perp, U(U\mathbf{H})^\perp, U^2(U\mathbf{H})^\perp, \dots$ span \mathbf{H} . (Since U is an isometry, it follows that they must be pairwise orthogonal.) Most of these remarks are implicit in Solution 118.

A generalization lies near at hand. Consider an isometry U on a Hilbert space \mathbf{H} such that the subspaces $(U\mathbf{H})^\perp, U(U\mathbf{H})^\perp, U^2(U\mathbf{H})^\perp, \dots$ are pairwise orthogonal and span \mathbf{H} , but make no demands on the value of the co-rank. Every such isometry may be called a *shift* (a unilateral shift). The co-rank of a shift (also called its *multiplicity*) constitutes a complete set of unitary invariants for it; the original unilateral shift is determined (to within unitary equivalence) as the shift of multiplicity 1 (the *simple* unilateral shift).

Unilateral shifts of higher multiplicities are just as important as the simple one. Problem 118 shows that they are exactly the pure isometries. They play a vital role in the study of all operators, not only isometries; they, or rather their adjoints, turn out to be universal operators.

A *part* of an operator is a restriction of it to an invariant subspace. Each part of an isometry is an isometry; the study of the parts of unilateral shifts does not promise anything new. What about parts of the adjoints of unilateral shifts? If U is a unilateral shift, then $\|U\| =$

$\|U^*\| = 1$, and it follows that if A is a part of U^* , then $\|A\| \leq 1$. Since, moreover, $U^{*n} \rightarrow 0$ in the strong topology (cf. Solution 90), it follows that $A^n \rightarrow 0$ (strong). The miraculous and useful fact is that these two obviously necessary conditions are also sufficient; cf. Foiaş [1963] and de Branges-Rovnyak [1964, 1965].

Problem 121. *Every contraction whose powers tend strongly to 0 is unitarily equivalent to a part of the adjoint of a unilateral shift.*

122. Similarity to parts of shifts. For many purposes similarity is just as good as unitary equivalence. When is an operator A similar to a part of the adjoint of a shift U ? Since similarity need not preserve norm, there is no obvious condition that $\|A\|$ must satisfy. There is, however, a measure of size that similarity does preserve, namely the spectral radius; since $r(U^*) = 1$, it follows that $r(A) \leq 1$. It is easy to see that this necessary condition is not sufficient. The reason is that one of the necessary conditions for unitary equivalence ($A^n \rightarrow 0$ strongly, cf. Problem 121) is necessary for similarity also. (That is: if $A^n \rightarrow 0$ strongly, and if $B = S^{-1}AS$, then $B^n \rightarrow 0$ strongly.) Since there are many operators A such that $r(A) \leq 1$ but A^n does not tend to 0 in any sense (example: 1), the condition on the spectral radius is obviously not sufficient. There is a condition on the spectral radius alone that is sufficient for similarity to a part of the adjoint of a shift, but it is quite a bit stronger than $r(A) \leq 1$; it is, in fact, $r(A) < 1$.

Problem 122. *Every operator whose spectrum is included in the interior of the unit disc is similar to a contraction whose powers tend strongly to 0.*

Corollary 1. *Every operator whose spectrum is included in the interior of the unit disc is similar to a part of the adjoint of a unilateral shift.*

Corollary 2. *Every operator whose spectrum is included in the interior of the unit disc is similar to a proper contraction.*

(A *proper* contraction is an operator A with $\|A\| < 1$.)

Corollary 3. *Every quasinilpotent operator is similar to operators with arbitrarily small norms.*

These simple but beautiful and general results are due to Rota [1960].

Corollary 4. *The spectral radius of every operator A is the infimum of the numbers $\|S^{-1}AS\|$ for all invertible operators S .*

123. Wandering subspaces. If A is an operator on a Hilbert space \mathbf{H} , a subspace \mathbf{N} of \mathbf{H} is called *wandering* for A if it is orthogonal to all its images under the positive powers of A . This concept is especially useful in the study of isometries. If U is an isometry and \mathbf{N} is a wandering subspace for U , then $U^m\mathbf{N} \perp U^n\mathbf{N}$ whenever m and n are distinct positive integers. In other words, if f and g are in \mathbf{N} , then $U^mf \perp U^ng$. (Proof: reduce to the case $m > n$, and note that $(U^mf, U^ng) = (U^{*n}U^mf, g) = (U^{m-n}f, g)$.) If U is unitary, even more is true: in that case $U^m\mathbf{N} \perp U^n\mathbf{N}$ whenever m and n are any two distinct integers, positive, negative, or zero. (Proof: find k so that $m+k$ and $n+k$ are positive and note that $(U^mf, U^ng) = (U^{m+k}f, U^{n+k}g)$.)

Wandering subspaces are important because they are connected with invariant subspaces, in this sense: if U is an isometry, then there is a natural one-to-one correspondence between all wandering subspaces \mathbf{N} and some invariant subspaces \mathbf{M} . The correspondence is given by setting $\mathbf{M} = \bigvee_{n=0}^{\infty} U^n\mathbf{N}$. (To prove that this correspondence is one-to-one, observe that $U\mathbf{M} = \bigvee_{n=1}^{\infty} U^n\mathbf{N}$, so that $\mathbf{N} = \mathbf{M} \cap (U\mathbf{M})^{\perp}$.) For at least one operator, namely the unilateral shift, the correspondence is invertible.

Problem 123. *If U is the (simple) unilateral shift and if \mathbf{M} is a non-zero subspace invariant under U , then there exists a (necessarily unique) one-dimensional wandering subspace \mathbf{N} such that $\mathbf{M} = \bigvee_{n=0}^{\infty} U^n\mathbf{N}$.*

The equation connecting \mathbf{M} and \mathbf{N} can be expressed by saying that every non-zero part of the simple unilateral shift is a shift. To add that $\dim \mathbf{N} = 1$ is perhaps an unsurprising sharpening, but a useful and non-trivial one. In view of these comments, the following concise state-

ment is just a reformulation of the problem: every non-zero part of the simple unilateral shift is (unitarily equivalent to) the simple unilateral shift. With almost no additional effort, and only the obviously appropriate changes in the statement, all these considerations extend to shifts of higher multiplicities.

124. Special invariant subspaces of the shift. One of the most recalcitrant unsolved problems of Hilbert space theory is whether or not every operator has a non-trivial invariant subspace. A promising, interesting, and profitable thing to do is to accumulate experimental evidence by examining concrete special cases and seeing what their invariant subspaces look like. A good concrete special case to look at is the unilateral shift.

There are two kinds of invariant subspaces: the kind whose orthogonal complement is also invariant (the reducing subspaces), and the other kind. The unilateral shift has no reducing subspaces (Problem 116); the question remains as to how many of the other kind it has and what they look like.

The easiest way to obtain an invariant subspace of the unilateral shift U is to fix a positive integer k , and consider the span \mathbf{M}_k of the e_n 's with $n \geq k$. After this elementary observation most students of the subject must stop and think; it is not at all obvious that any other invariant subspaces exist. A recollection of the spectral behavior of U is helpful here. Indeed, since each complex number λ of modulus less than 1 is a simple eigenvalue of U^* (Solution 67), with corresponding eigenvector $f_\lambda = \sum_{n=0}^{\infty} \lambda^n e_n$, it follows that the orthogonal complement of the singleton $\{f_\lambda\}$ is a non-trivial subspace invariant under U .

Problem 124. *If $\mathbf{M}_k(\lambda)$ is the orthogonal complement of $\{f_\lambda, \dots, U^{k-1}f_\lambda\}$ ($k = 1, 2, 3, \dots$), then $\mathbf{M}_k(\lambda)$ is invariant under U , $\dim \mathbf{M}_k^\perp(\lambda) = k$, and $\bigvee_{k=1}^{\infty} \mathbf{M}_k^\perp(\lambda) = \mathbf{H}^2$.*

Note that the spaces \mathbf{M}_k considered above are the same as the spaces $\mathbf{M}_k(0)$.

125. Invariant subspaces of the shift. What are the invariant subspaces of the unilateral shift? The spaces \mathbf{M}_k and their generalizations

$\mathbf{M}_k(\lambda)$ (see Problem 124) are examples. The lattice operations (intersection and span) applied to them yield some not particularly startling new examples, and then the well seems to run dry. New inspiration can be obtained by abandoning the sequential point of view and embracing the functional one; regard U as the restriction to \mathbf{H}^2 of the multiplication induced by e_1 .

Problem 125. *A non-zero subspace \mathbf{M} of \mathbf{H}^2 is invariant under U if and only if there exists a function φ in \mathbf{H}^∞ , of constant modulus 1 almost everywhere, such that \mathbf{M} is the range of the restriction to \mathbf{H}^2 of the multiplication induced by φ .*

This basic result is due to Beurling [1949]. It has received considerable attention since then; cf. Lax [1959], Halmos [1961], and Helson [1964].

In more informal language, \mathbf{M} can be described as the set of all *multiples* of φ (multiples by functions in \mathbf{H}^2 , that is). Correspondingly it is suggestive to write $\mathbf{M} = \varphi \cdot \mathbf{H}^2$. For no very compelling reason, the functions such as φ (functions in \mathbf{H}^∞ , of constant modulus 1) are called *inner functions*.

Corollary 1. *If φ and ψ are inner functions such that $\varphi \cdot \mathbf{H}^2 \subset \psi \cdot \mathbf{H}^2$, then φ is divisible by ψ , in the sense that there exists an inner function θ such that $\varphi = \psi \cdot \theta$. If $\varphi \cdot \mathbf{H}^2 = \psi \cdot \mathbf{H}^2$, then φ and ψ are constant multiples of one another, by constants of modulus 1.*

The characterization in terms of inner functions does not solve all problems about invariant subspaces of the shift, but it does solve some. Here is a sample.

Corollary 2. *If \mathbf{M} and \mathbf{N} are non-zero subspaces invariant under the unilateral shift, then $\mathbf{M} \cap \mathbf{N} \neq \{0\}$.*

Corollary 2 says that the lattice of invariant subspaces of the unilateral shift is about as far as can be from being complemented.

126. Cyclic vectors. An operator A on a Hilbert space \mathbf{H} has a *cyclic vector* f if the vectors f, Af, A^2f, \dots span \mathbf{H} . Equivalently, f is a

cyclic vector for A in case the set of all vectors of the form $p(A)f$, where p varies over all polynomials, is dense in \mathbf{H} . The simple unilateral shift has many cyclic vectors; a trivial example is e_0 .

On finite-dimensional spaces the existence of a cyclic vector indicates something like multiplicity 1. If, to be precise, A is a finite diagonal matrix, then a necessary and sufficient condition that A have a cyclic vector is that the diagonal entries be distinct (i.e., that the eigenvalues be simple). Indeed, if the diagonal entries are $\lambda_1, \dots, \lambda_n$, then $p(A) \langle \xi_1, \dots, \xi_n \rangle = \langle p(\lambda_1)\xi_1, \dots, p(\lambda_n)\xi_n \rangle$ for every polynomial p . For $f = \langle \xi_1, \dots, \xi_n \rangle$ to be cyclic, it is clearly necessary that $\xi_i \neq 0$ for each i ; otherwise the i coordinate of $p(A)f$ is 0 for all p . If the λ 's are not distinct, nothing is sufficient to make f cyclic. If, for instance, $\lambda_1 = \lambda_2$, then $\langle \xi_2^*, -\xi_1^*, 0, \dots, 0 \rangle$ is orthogonal to $p(A)f$ for all p . If, on the other hand, the λ 's are distinct, then $p(\lambda_1), \dots, p(\lambda_n)$ can be prescribed arbitrarily, so that if none of the ξ 's vanishes, then the $p(A)f$'s exhaust the whole space.

Some trace of the relation between the existence of cyclic vectors and multiplicity 1 is visible even for non-diagonal matrices. Thus, for instance, if A is a finite matrix, then the direct sum $A \oplus A$ cannot have a cyclic vector. Reason: by virtue of the Hamilton-Cayley equation, at most n of the matrices $1, A, A^2, \dots$ are linearly independent (where n is the size of A), and consequently, no matter what f and g are, at most n of the vectors $A^j f \oplus A^j g$ are linearly independent; it follows that their span can never be $2n$ -dimensional.

If A has multiplicity 1 in any sense, it is reasonable to expect that A^* also has; this motivates the conjecture that if A has a cyclic vector, then so does A^* . For finite matrices this is true. For a proof, note that, surely, if a matrix has a cyclic vector, then so does its complex conjugate, and recall that every matrix is similar to its transpose.

The methods of the preceding paragraphs are very parochially finite-dimensional; that indicates that the theory of cyclic vectors in infinite-dimensional spaces is likely to be refractory, and it is. There is, to begin with, a trivial difficulty with cardinal numbers. If there is a cyclic vector, then a countable set spans the space, and therefore the space is separable; in other words, in non-separable spaces there are no cyclic vectors. This difficulty can be got around; that is one of the achievements

of the multiplicity theory of normal operators (Halmos [1951, III]). For normal operators, the close connection between multiplicity 1 and the existence of cyclic vectors persists in infinite-dimensional spaces, and, suitably reinterpreted, even in spaces of uncountable dimension.

For non-normal operators, things are peculiar. It is possible for a direct sum $A \oplus A$ to have a cyclic vector, and it is possible for A to have a cyclic vector when A^* does not. These facts were first noticed by D. E. Sarason.

Problem 126. *If U is a unilateral shift of multiplicity not greater than \aleph_0 , then U^* has a cyclic vector.*

It is obvious that the simple unilateral shift has a cyclic vector, but it is not at all obvious that its adjoint has one. It does, but that by itself does not imply anything shocking. The first strange consequence of the present assertion is that if U is the simple unilateral shift, then $U^* \oplus U^*$ (which is the adjoint of a unilateral shift of multiplicity 2) has a cyclic vector. The promised strange behavior becomes completely exposed with the remark that $U \oplus U$ cannot have a cyclic vector (and, all the more, the same is true for direct sums with more direct summands). To prove the negative assertion, consider a candidate

$$\langle \langle \xi_0, \xi_1, \xi_2, \dots \rangle, \langle \eta_0, \eta_1, \eta_2, \dots \rangle \rangle$$

for a cyclic vector of $U \oplus U$. If $\langle \alpha, \beta \rangle$ is an arbitrary non-zero vector orthogonal to $\langle \xi_0, \eta_0 \rangle$ in the usual two-dimensional complex inner product space, then the vector

$$\langle \langle \alpha, 0, 0, \dots \rangle, \langle \beta, 0, 0, \dots \rangle \rangle$$

is orthogonal to

$$(U \oplus U)^n \langle \langle \xi_0, \xi_1, \xi_2, \dots \rangle, \langle \eta_0, \eta_1, \eta_2, \dots \rangle \rangle$$

for all $n (= 0, 1, 2, \dots)$, and that proves that the cyclic candidate fails.

127. The F. and M. Riesz theorem. It is always a pleasure to see a piece of current (soft) mathematics reach into the past to illuminate and simplify some of the work of the founding fathers on (hard) analysis; the characterization of the invariant subspaces of the unilateral shift does that. The elements of \mathbf{H}^2 are related to certain analytic functions on the unit disc (Problem 28), and, although they themselves are defined on the unit circle only, and only almost everywhere at that, they tend to imitate the behavior of analytic functions. A crucial property of an analytic function is that it cannot vanish very often without vanishing everywhere. An important theorem of F. and M. Riesz asserts that the elements of \mathbf{H}^2 exhibit the same kind of behavior; here is one possible formulation.

Problem 127. *A function in \mathbf{H}^2 vanishes either almost everywhere or almost nowhere.*

Corollary. *If f and g are in \mathbf{H}^2 and if $fg = 0$ almost everywhere, then $f = 0$ almost everywhere or $g = 0$ almost everywhere.*

Concisely: there are no zero-divisors in \mathbf{H}^2 .

For a more general discussion of the F. and M. Riesz theorem, see Hoffman [1962, p. 47].

128. The F. and M. Riesz theorem generalized. The F. and M. Riesz theorem says that if $f \in \mathbf{H}^2$ and if f vanishes on a set of positive measure, then $f = 0$ almost everywhere. To say " $f \in \mathbf{H}^2$ " is to say that the Fourier coefficients of f with negative index are zero. This implies that if $f = \sum_n \alpha_n e_n$, then $\alpha_n \alpha_{-n} = 0$ for all non-zero n 's. Is this condition sufficient to yield the conclusion of the F. and M. Riesz theorem?

Problem 128. *If $f \in L^2$ with Fourier expansion $f = \sum_n \alpha_n e_n$, if $\alpha_n \alpha_{-n} = 0$ for all non-zero n 's, and if f vanishes on a set of positive measure, does it follow that $f = 0$ almost everywhere?*

129. Reducible weighted shifts. Very little of the theory of reducing and invariant subspaces of the bilateral and the unilateral shift is known

for weighted shifts. There is, however, one striking fact that deserves mention; it has to do with the reducibility of two-sided weighted shifts. It is due to R. L. Kelley.

Problem 129. *If A is a bilateral weighted shift with strictly positive weights α_n , $n = 0, \pm 1, \pm 2, \dots$, then a necessary and sufficient condition that A be reducible is that the sequence $\{\alpha_n\}$ be periodic.*

Chapter 15. Compact operators

130. Mixed continuity. Corresponding to the strong (s) and weak (w) topologies for a Hilbert space \mathbf{H} , there are four possible interpretations of continuity for a transformation from \mathbf{H} into \mathbf{H} : they are the ones suggested by the symbols $(s \rightarrow s)$, $(w \rightarrow w)$, $(s \rightarrow w)$, and $(w \rightarrow s)$. Thus, to say that A is continuous $(s \rightarrow w)$ means that the inverse image under A of each w-open set is s-open; equivalently it means that the direct image under A of a net s-convergent to f is a net w-convergent to Af . Four different kinds of continuity would be too much of a good thing; it is fortunate that three of them collapse into one.

Problem 130. *For a linear transformation A the three kinds of continuity $(s \rightarrow s)$, $(w \rightarrow w)$, and $(s \rightarrow w)$ are equivalent (and hence each is equivalent to boundedness), and continuity $(w \rightarrow s)$ implies that A has finite rank.*

Corollary. *The image of the closed unit ball under an operator on a Hilbert space is always strongly closed.*

It is perhaps worth observing that for linear transformations of finite rank all four kinds of continuity are equivalent; this is a trivial finite-dimensional assertion.

131. Compact operators. A linear transformation on a Hilbert space is called *compact* (also *completely continuous*) if its restriction to the unit ball is $(w \rightarrow s)$ continuous (see Problem 130). Equivalently, a linear transformation is compact if it maps each bounded weakly convergent net onto a strongly convergent net. Since weakly convergent sequences are bounded, it follows that a compact linear transformation maps every weakly convergent sequence onto a strongly convergent one.

The image of the closed unit ball under a compact linear transformation is strongly compact. (Proof: the closed unit ball is weakly compact.) This implies that the image of each bounded set is precompact

(i.e., has a strongly compact closure). (Proof: a bounded set is included in some closed ball.) The converse implication is also true: if a linear transformation maps bounded sets onto precompact sets, then it maps the closed unit ball onto a compact set. To prove this, observe first that compact (and precompact) sets are bounded, and that therefore a linear transformation that maps bounded sets onto precompact sets is necessarily bounded itself. (This implies, incidentally, that every compact linear transformation is bounded.) It follows from the corollary to Problem 130 that the image of the closed unit ball is strongly closed; this, together with the assumption that that image is precompact, implies that that image is actually compact. (The converse just proved is not universally true for Banach spaces.) The compactness conditions, here treated as consequences of the continuity conditions used above to define compact linear transformations, can in fact be shown to be equivalent to those continuity conditions and are frequently used to define compact linear transformations. (See Dunford-Schwartz [1958, p. 484].)

An occasionally useful property of compact operators is that they “attain their norm”. Precisely said: if A is compact, then there exists a unit vector f such that $\|Af\| = \|A\|$. The reason is that the mapping $f \rightarrow Af$ is (w \rightarrow s) continuous on the unit ball, and the mapping $g \rightarrow \|g\|$ is strongly continuous; it follows that $f \rightarrow \|Af\|$ is weakly continuous on the unit ball. Since the unit ball is weakly compact, this function attains its maximum, so that $\|Af\| = \|A\|$ for some f with $\|f\| \leq 1$. If $A = 0$, then f can be chosen to have norm 1; if $A \neq 0$, then f necessarily has norm 1. Reason: since $f \neq 0$ and $1/\|f\| \geq 1$, it follows that

$$\|A\| \leq \frac{\|A\|}{\|f\|} = \frac{\|Af\|}{\|f\|} \leq \|A\|.$$

Problem 131. *The set \mathbf{C} of all compact operators on a Hilbert space is a closed self-adjoint (two-sided) ideal.*

Here “closed” refers to the norm topology, “self-adjoint” means that if $A \in \mathbf{C}$, then $A^* \in \mathbf{C}$, and “ideal” means that linear combinations of operators in \mathbf{C} are in \mathbf{C} and that products with at least one factor in \mathbf{C} are in \mathbf{C} .

132. Diagonal compact operators. Is the identity operator compact? Since in finite-dimensional spaces the strong and the weak topologies coincide, the answer is yes for them. For infinite-dimensional spaces, the answer is no; the reason is that the image of the unit ball is the unit ball, and in an infinite-dimensional space the unit ball cannot be strongly compact (Problem 10).

The indistinguishability of the strong and the weak topologies in finite-dimensional spaces yields a large class of examples of compact operators, namely all operators of finite rank. Examples of a slightly more complicated structure can be obtained by exploiting the fact that the set of compact operators is closed.

Problem 132. *A diagonal operator with diagonal $\{\alpha_n\}$ is compact if and only if $\alpha_n \rightarrow 0$ as $n \rightarrow \infty$.*

Corollary. *A weighted shift with weights $\{\alpha_n: n = 0, 1, 2, \dots\}$ is compact if and only if $\alpha_n \rightarrow 0$ as $n \rightarrow \infty$.*

133. Normal compact operators. It is easy to see that if a normal operator has the property that every non-zero element in its spectrum is isolated (i.e., is not a cluster point of the spectrum), then it is a diagonal operator. (For each non-zero eigenvalue λ of A , choose an orthonormal basis for the subspace $\{f: Af = \lambda f\}$; the union of all these little bases, together with a basis for the kernel of A , is a basis for the whole space.) If, moreover, each non-zero eigenvalue has finite multiplicity, then the operator is compact. (Compare Problem 132; note that under the assumed conditions the set of eigenvalues is necessarily countable.) The remarkable and useful fact along these lines goes in the converse direction.

Problem 133. *The spectrum of a compact normal operator is countable; all its non-zero elements are eigenvalues of finite multiplicity.*

Corollary. *Every compact normal operator is the direct sum of the operator 0 (on a space that can be anything from absent to non-separable) and a diagonal operator (on a separable space).*

A less sharp but shorter formulation of the corollary is this: every compact normal operator is diagonal.

134. Kernel of the identity. Matrices have valuable “continuous” generalizations. The idea is to replace sums by integrals, and it works—up to a point. To see where it goes wrong, consider a measure space X with measure μ (σ -finite as usual), and consider a measurable function K on the product space $X \times X$. A function of two variables, such as K , is what a generalized matrix can be expected to be. Suppose that A is an operator on $L^2(\mu)$ whose relation to K is similar to the usual relation of an operator to its matrix. In precise terms this means that if $f \in L^2(\mu)$, then

$$(Af)(x) = \int K(x,y)f(y)d\mu(y)$$

for almost every x . Under these conditions A is called an *integral operator* and K is called its *kernel*.

In the study of a Hilbert space H , to say “select an orthonormal basis” is a special case of saying “select a particular way of representing H as L^2 ”. Many phenomena in L^2 spaces are the natural “continuous” generalizations of more familiar phenomena in sequence spaces. One simple fact about sequence spaces is that every operator on them has a matrix, and this is true whether the sequences (families) that enter are finite or infinite. (It is the reverse procedure that goes wrong in the infinite case. From operators to matrices all is well; it is from matrices to operators that there is trouble.) On this evidence it is reasonable to guess that every operator on L^2 has a kernel, i.e., that every operator is an integral operator. This guess is wrong, hopelessly wrong. The trouble is not with wild operators, and it is not with wild measures; it arises already if the operator is the identity and if the measure is Lebesgue measure (in the line or in any interval).

Problem 134. *If μ is Lebesgue measure, then the identity is not an integral operator on $L^2(\mu)$.*

135. Hilbert–Schmidt operators. Under what conditions does a kernel induce an operator? Since the question includes the corresponding

question for matrices, it is not reasonable to look for necessary and sufficient conditions. A somewhat special sufficient condition, which is nevertheless both natural and useful, is that the kernel be square integrable.

Suppose, to be quite precise, that X is a measure space with σ -finite measure μ , and suppose that K is a complex-valued measurable function on $X \times X$ such that $|K|^2$ is integrable with respect to the product measure $\mu \times \mu$. It follows that, for almost every x , the function $y \rightarrow K(x, y)$ is in $L^2(\mu)$, and hence that the product function $y \rightarrow K(x, y)f(y)$ is integrable whenever $f \in L^2(\mu)$. Since, moreover,

$$\begin{aligned} \int \left| \int K(x, y)f(y) d\mu(y) \right|^2 d\mu(x) \\ \leq \int \left(\int |K(x, y)|^2 d\mu(y) \cdot \int |f(y)|^2 d\mu(y) \right) d\mu(x) = \|K\|^2 \cdot \|f\|^2 \end{aligned}$$

(where $\|K\|$ is the norm of K in $L^2(\mu \times \mu)$), it follows that the equation

$$(Af)(x) = \int K(x, y)f(y) d\mu(y)$$

defines an operator (with kernel K) on $L^2(\mu)$. The inequality implies also that

$$\|A\| \leq \|K\|.$$

Integral operators with kernels of this type (i.e., kernels in $L^2(\mu \times \mu)$) are called *Hilbert–Schmidt operators*. A good reference for their properties is Schatten [1960].

The correspondence $K \rightarrow A$ is a one-to-one linear mapping from $L^2(\mu \times \mu)$ to operators on $L^2(\mu)$. If A has a kernel K (in $L^2(\mu \times \mu)$), then A^* has the kernel \tilde{K} defined by

$$\tilde{K}(x, y) = (K(y, x))^*.$$

If A and B have kernels H and K (in $L^2(\mu \times \mu)$), then AB has the

kernel HK defined by

$$(HK)(x,y) = \int H(x,z)K(z,y)d\mu(z).$$

The proofs of all these algebraic assertions are straightforward computations with integrals.

On the analytic side, the situation is just as pleasant. If $\{K_n\}$ is a sequence of kernels in $L^2(\mu \times \mu)$ such that $K_n \rightarrow K$ (in the norm of $L^2(\mu \times \mu)$), and if the corresponding operators are A_n (for K_n) and A (for K), then $\|A_n - A\| \rightarrow 0$. The proof is immediate from the inequality between the norm of an integral operator and the norm of its kernel.

Problem 135. *Every Hilbert-Schmidt operator is compact.*

These considerations apply, in particular, when the space is the set of positive integers with the counting measure. It follows that if the entries of a matrix are square-summable, then it is bounded (in the sense that it defines an operator) and compact (in view of the assertion of Problem 135). It should also be remarked that the Schur test (Problem 37) for the boundedness of a matrix has a straightforward generalization to a theorem about kernels; see Brown-Halmos-Shields [1965].

136. Compact versus Hilbert-Schmidt.

Problem 136. *Is every compact operator a Hilbert-Schmidt operator?*

137. Limits of operators of finite rank. Every example of a compact operator seen so far (diagonal operators, weighted shifts, integral operators) was proved to be compact by showing it to be a limit of operators of finite rank. That is no accident.

Problem 137. *Every compact operator is the limit (in the norm) of operators of finite rank.*

The generalization of the assertion to arbitrary Banach spaces is an unsolved problem.

138. Ideals of operators. An ideal of operators is *proper* if it does not contain every operator. An easy example of an ideal of operators on a Hilbert space is the set of all operators of finite rank on that space; if the space is infinite-dimensional, that ideal is proper. Another example is the set of all compact operators; again, if the space is infinite-dimensional, that ideal is proper. The second of these examples is closed; in the infinite-dimensional case the first one is not.

Problem 138. *If H is a separable Hilbert space, then the collection of compact operators is the only non-zero closed proper ideal of operators on H .*

Similar results hold for non-separable spaces, but the formulations and proofs are fussier and much less interesting.

139. Square root of a compact operator. It is easy to construct non-compact operators whose square is compact; in fact, it is easy to construct non-compact operators that are nilpotent of index 2. (Cf. Problem 80.) What about the normal case?

Problem 139. *Do there exist non-compact normal operators whose square is compact?*

140. Fredholm alternative. The principal spectral fact about a compact operator (normal or no) on a Hilbert space is that a non-zero number can get into the spectrum via the point spectrum only. More precisely: if C is compact, and if λ is a non-zero complex number, then either λ is an eigenvalue of C or $C - \lambda$ is invertible. Division by λ shows that it is sufficient to treat $\lambda = 1$. The statement has thus been transformed into the following one.

Problem 140. *If C is compact and if $\ker (1 - C) = \{0\}$, then $1 - C$ is invertible.*

The statement is frequently referred to as the *Fredholm alternative*. The Fredholm alternative has a facetious but not too inaccurate reformulation in terms of the equation $(1 - C)f = g$, in which g is regarded as given and f as unknown; according to that formulation, if the solution is unique, then it exists.

Corollary. *A compact operator whose point spectrum is empty is quasilinear.*

141. Range of a compact operator.

Problem 141. *Every (closed) subspace included in the range of a compact operator is finite-dimensional.*

Corollary. *Every eigenvalue of a compact operator has finite multiplicity.*

142. Atkinson's theorem. An operator A is called a *Fredholm operator* if (1) $\text{ran } A$ is closed and both $\ker A$ and $(\text{ran } A)^\perp$ are finite-dimensional. (The last two conditions can be expressed by saying that the nullity and the co-rank of A are finite.) An operator A is *invertible modulo the ideal of operators of finite rank* if (2) there exists an operator B such that both $1 - AB$ and $1 - BA$ have finite rank. An operator A is *invertible modulo the ideal of compact operators* if (3) there exists an operator B such that both $1 - AB$ and $1 - BA$ are compact.

Problem 142. *An operator A is (1) a Fredholm operator if and only if it is (2) invertible modulo the ideal of operators of finite rank, or, alternatively, if and only if it is (3) invertible modulo the ideal of compact operators.*

The result is due to Atkinson [1951].

143. Weyl's theorem. The process of adding a compact operator to a given one is sometimes known as *perturbation*. The accepted attitude toward perturbation is that compact operators are "small"; the addition of a compact operator cannot (or should not) make for radical changes.

Problem 143. *If the difference between two operators is compact, then their spectra are the same except for eigenvalues. More explicitly: if $A - B$ is compact, and if $\lambda \in \Lambda(A) - \Pi_0(A)$, then $\lambda \in \Lambda(B)$.*

Note that for $B = 0$ the statement follows from Problem 140.

144. Perturbed spectrum. The spectrum of an operator changes, of course, when a compact operator is added to it, but in some sense not very much. Eigenvalues may come and go, but otherwise the spectrum remains invariant. In another sense, however, the spectrum can be profoundly affected by the addition of a compact operator.

Problem 144. *There exists a unitary operator U and there exists a compact operator C such that the spectrum of $U + C$ is the entire unit disc.*

145. Shift modulo compact operators. Weyl's theorem (Problem 143) implies that if U is the unilateral shift and if C is compact, then the spectrum of $U + C$ includes the unit disc. (Here is a small curiosity. The reason the spectrum of $U + C$ includes the unit disc is that U has no eigenvalues. The adjoint U^* has many eigenvalues, so that this reasoning does not apply to it, but the conclusion does. Reason: the spectrum of $U^* + C$ is obtained from the spectrum of $U + C^*$ by reflection through the real axis, and C^* is just as compact as C .) More is true: Stampfli [1965] proved that every point of the open unit disc is an eigenvalue of $(U + C)^*$.

It follows from the preceding paragraph that $U + C$ can never be invertible (the spectrum cannot avoid 0), and it follows also that $U + C$ can never be quasinilpotent (the spectrum cannot consist of 0 alone). Briefly: if invertibility and quasinilpotence are regarded as good properties, then not only is U bad, but it cannot be improved by a perturbation. Perhaps the best property an operator can have (and U does not have) is normality; can a perturbation improve U in this respect?

Problem 145. *If U is the unilateral shift, does there exist a compact operator C such that $U + C$ is normal?*

Freeman [1965] has a result that is pertinent to this circle of ideas; he proves that, for a large class of compact operators C , the perturbed shift $U + C$ is similar to the unperturbed shift U .

146. Bounded Volterra kernels. Integral operators are generalized matrices. Experience with matrices shows that the more zeros they have, the easier they are to compute with; triangular matrices, in particular, are usually quite tractable. Which integral operators are the right generalizations of triangular matrices? For the answer it is convenient to specialize drastically the measure spaces considered; in what follows the only X will be the unit interval, and the only μ will be Lebesgue measure. (The theory can be treated somewhat more generally; see Ringrose [1962].)

A *Volterra kernel* is a kernel K in $L^2(\mu \times \mu)$ such that $K(x, y) = 0$ when $x < y$. Equivalently: a Volterra kernel is a Hilbert-Schmidt kernel that is triangular in the sense that it vanishes above the diagonal ($x = y$) of the unit square. In view of this definition, the effect of the integral operator A (*Volterra operator*) induced by a Volterra kernel K can be described by the equation

$$(Af)(x) = \int_0^x K(x, y)f(y)dy.$$

If the diagonal terms of a finite triangular matrix vanish, then the matrix is nilpotent. Since the diagonal of the unit square has measure 0, and since from the point of view of Hilbert space sets of measure 0 are negligible, the condition of vanishing on the diagonal does not have an obvious continuous analogue. It turns out nevertheless that the zero values of a Volterra kernel above the diagonal win out over the non-zero values below.

Problem 146. *A Volterra operator with a bounded kernel is quasi-nilpotent.*

Caution: "bounded" here refers to the kernel, not to the operator; the assumption is that the kernel is bounded almost everywhere in the unit square.

147. Unbounded Volterra kernels. How important is the boundedness assumption in Problem 146?

Problem 147. *Is every Volterra operator quasinilpotent?*

148. The Volterra integration operator. The simplest non-trivial Volterra operator is the one whose kernel is the characteristic function of the triangle $\{ \langle x, y \rangle : 0 \leq y \leq x \leq 1 \}$. Explicitly this is the Volterra operator V defined on $L^2(0,1)$ by

$$(Vf)(x) = \int_0^x f(y) dy.$$

In still other words, V is indefinite integration, with the constant of integration adjusted so that every function in the range of V vanishes at 0. (Note that every function in the range of V is continuous. Better: every vector in the range of V , considered as an equivalence class of functions modulo sets of measure 0, contains a unique continuous function.)

Since V^* is the integral operator whose kernel is the “conjugate transpose” of the kernel of V , so that the kernel of V^* is the characteristic function of the triangle $\{ \langle x, y \rangle : 0 \leq x \leq y \leq 1 \}$, it follows that $V + V^*$ is the integral operator whose kernel is equal to the constant function 1 almost everywhere. (The operators V^* and $V + V^*$ are of course not Volterra operators.) This is a pleasantly simple integral operator; a moment’s reflection should serve to show that it is the projection whose range is the (one-dimensional) space of constants. It follows that $\operatorname{Re} V$ has rank 1; since $V = \operatorname{Re} V + i \operatorname{Im} V$, it follows that V is a perturbation (by an operator of rank 1 at that) of a skew Hermitian operator.

The theory of Hilbert-Schmidt operators in general and Volterra operators in particular answers many questions about V . Thus, for instance, V is compact (because it is a Hilbert-Schmidt operator), and it is quasinilpotent (because it is a Volterra operator). There are many other natural questions about V ; some are easy to answer and some are not. Here is an easy one: does V annihilate any non-zero vectors? (Equivalently: “does V have a nontrivial kernel?”, but that way terminological confusion lies.) The answer is no. If $\int_0^x f(y) dy = 0$ for

almost every x , then, by continuity, the equation holds for every x . Since the functions in the range of V are not only continuous but, in fact, differentiable almost everywhere, the equation can be differentiated; the result is that $f(x) = 0$ for almost every x . As for natural questions that are not so easily disposed of, here is a simple sample.

Problem 148. *What is the norm of V ?*

149. Skew-symmetric Volterra operator. There is an operator V_0 on $L^2(-1, +1)$ (Lebesgue measure) that bears a faint formal resemblance to the operator V on $L^2(0, 1)$; by definition

$$(V_0 f)(x) = \int_{-x}^{+x} f(y) dy.$$

Note that V_0 is the integral operator induced by the kernel that is the characteristic function of the butterfly $\{ \langle x, y \rangle : 0 \leq |y| \leq |x| \leq 1 \}$.

Problem 149. *Find the spectrum and the norm of the skew-symmetric Volterra operator V_0 .*

150. Norm 1, spectrum $\{1\}$. Every finite matrix is unitarily equivalent to a triangular matrix. If a triangular matrix has only 1's on the main diagonal, then its norm is at least 1; the norm can be equal to 1 only in case the matrix is the identity. The conclusion is that on a finite-dimensional Hilbert space the identity is the only contraction with spectrum $\{1\}$. The reasoning that led to this conclusion was very finite-dimensional; can it be patched up to yield the same conclusion for infinite-dimensional spaces?

Problem 150. *Is there an operator A , other than 1, such that $\Lambda(A) = \{1\}$ and $\|A\| = 1$?*

151. Donoghue lattice. One of the most important, most difficult, and most exasperating unsolved problems of operator theory is the problem of invariant subspaces. The question is simple to state: does every operator on an infinite-dimensional Hilbert space have a non-trivial invariant subspace? "Non-trivial" means different from both $\{0\}$

and the whole space; “invariant” means that the operator maps it into itself. For finite-dimensional spaces there is, of course, no problem; as long as the complex field is used, the fundamental theorem of algebra implies the existence of eigenvectors.

According to a dictum of Pólya’s, for each unanswered question there is an easier unanswered question, and the scholar’s first task is to find the latter. Even that dictum is hard to apply here; many weakenings of the invariant subspace problem are either trivial or as difficult as the full-strength problem. If, for instance, in an attempt to get a positive result, “subspace” is replaced by “linear manifold” (not necessarily closed), then the answer is yes, and easy. (For an elegant discussion, see Schaefer [1963].) If, on the other hand, in an attempt to get a counterexample, “Hilbert space” is replaced by “Banach space”, nothing happens; no one has succeeded in finding a counterexample in any space.

Positive results are known for some special classes of operators. The cheapest way to get one is to invoke the spectral theorem and to conclude that normal operators always have non-trivial invariant subspaces. The earliest non-trivial result along these lines is the assertion that compact operators always have non-trivial invariant subspaces (Aronszajn-Smith [1954]). That result has been generalized (Bernstein-Robinson [1966], Halmos [1966]), but the generalization is still closely tied to compactness. Non-compact results are few; here is a sample. If A is a contraction such that neither of the sequences $\{A^n\}$ and $\{A^{*n}\}$ tends strongly to 0, then A has a non-trivial invariant subspace (Nagy-Foiaş [1964]). A relatively recent bird’s-eye view of the subject was given by Helson [1964]; a more extensive bibliography is in Dunford-Schwartz [1963].

It is helpful to approach the subject from a different direction: instead of searching for counterexamples, study the structure of some non-counterexamples. One way to do this is to fix attention on a particular operator and to characterize all its invariant subspaces; the first significant step in this direction is the work of Beurling [1949] (Problem 125).

Nothing along these lines is easy. The second operator whose invariant subspaces have received detailed study is the Volterra integration operator (Brodskii [1957], Donoghue [1957 b], Kalisch [1957], Sakhnovich [1957]). The results for it are easier to describe than for the shift, but harder to prove. If $(Vf)(x) = \int_0^x f(y)dy$ for f in $L^2(0,1)$, and if, for each α in $[0,1]$, \mathbf{M}_α is the subspace of those functions that vanish almost

everywhere on $[0, \alpha]$, then \mathbf{M}_α is invariant under V ; the principal result is that every invariant subspace of V is one of the \mathbf{M}_α 's. An elegant way of obtaining these results is to reduce the study of the Volterra integration operator (as far as invariant subspaces are concerned) to that of the unilateral shift; this was done by Sarason [1965].

The collection of all subspaces invariant under some particular operator is a lattice (closed under the formation of intersections and spans). One way to state the result about V is to say that its lattice of invariant subspaces is anti-isomorphic to the closed unit interval. ("Anti-" because as α grows \mathbf{M}_α shrinks.) The lattice of invariant subspaces of V^* is in an obvious way isomorphic to the closed unit interval.

Is there an operator whose lattice of invariant subspaces is isomorphic to the positive integers? The question must be formulated with a little more care: every invariant subspace lattice has a largest element. The exact formulation is easy: is there an operator for which there is a one-to-one and order-preserving correspondence $n \rightarrow \mathbf{M}_n$, $n = 0, 1, 2, 3, \dots, \infty$, between the indicated integers (including ∞) and all invariant subspaces? The answer is yes. The first such operator was discovered by Donoghue [1957 b]; a wider class of them is described by Nikolskii [1965].

Suppose that $\{\alpha_n\}$ is a monotone sequence ($\alpha_n \geq \alpha_{n+1}$, $n = 0, 1, 2, \dots$) of positive numbers ($\alpha_n > 0$) such that $\sum_{n=0}^{\infty} \alpha_n^2 < \infty$. The unilateral weighted shift with the weight sequence $\{\alpha_n\}$ will be called a *monotone l^2 shift*. The span of the basis vectors $e_n, e_{n+1}, e_{n+2}, \dots$ is invariant under such a shift, $n = 0, 1, 2, \dots$. The orthogonal complement, i.e., the span \mathbf{M}_n of e_0, \dots, e_{n-1} , is invariant under the adjoint, $n = 1, 2, 3, \dots$; the principal result is that every invariant subspace of that adjoint is one of these orthogonal complements.

Problem 151. *If A is the adjoint of a monotone l^2 shift, and if \mathbf{M} is a non-trivial subspace invariant under A , then there exists an integer n ($= 1, 2, 3, \dots$) such that $\mathbf{M} = \mathbf{M}_n$.*

Chapter 16. Subnormal operators

152. The Putnam-Fuglede theorem. Some of the natural questions about normal operators have the same answers for finite-dimensional spaces as for infinite-dimensional ones, and the techniques used to prove the answers are the same. Some questions, on the other hand, are properly infinite-dimensional, in the sense that for finite-dimensional spaces they are either meaningless or trivial; questions about shifts, or, more generally, questions about subnormal operators are likely to belong to this category (see Problem 154). Between these two extremes there are the questions for which the answers are invariant under change of dimension, but the techniques are not. Sometimes, to be sure, either the question or the answer must be reformulated in order to bring the finite and the infinite into harmony. As for the technique, experience shows that an infinite-dimensional proof can usually be adapted to the finite-dimensional case; to say that the techniques are different means that the natural finite-dimensional techniques are not generalizable to infinite-dimensional spaces. It should be added, however, that sometimes the finite and the infinite proofs are intrinsically different, so that neither can be adapted to yield the result of the other; a case in point is the statement that any two bases have the same cardinal number. A familiar and typical example of a theorem whose statement is easily generalizable from the finite to the infinite, but whose proof is not, is the spectral theorem. A more striking example is the Fuglede commutativity theorem. It is more striking because it was for many years an unsolved problem. For finite-dimensional spaces the statement was known to be true and trivial; for infinite-dimensional spaces it was unknown.

The Fuglede theorem (cf. Solution 115) can be formulated in several ways. The algebraically simplest formulation is that if A is a normal operator and if B is an operator that commutes with A , then B commutes with A^* also. Equivalently: if A^* commutes with A , and A commutes with B , then A^* commutes with B . In the latter form the assertion is that in a certain special situation commutativity is transitive. (In general it is not.)

The operator A plays a double role in the Fuglede theorem; the modified assertion, obtained by splitting the two roles of A between two normal operators, is true and useful. Here is a precise formulation.

Problem 152. *If A_1 and A_2 are normal operators and if B is an operator such that $A_1B = BA_2$, then $A_1^*B = BA_2^*$.*

Observe that the Fuglede theorem is trivial in case B is Hermitian (even if A is not necessarily normal); just take the adjoint of the assumed equation $AB = BA$. The Putnam generalization (i.e., Problem 152) is, however, not obvious even if B is Hermitian; the adjoint of $A_1B = BA_2$ is, in that case, $BA_1^* = A_2^*B$, which is not what is wanted.

Corollary. *If two normal operators are similar, then they are unitarily equivalent.*

Is the product of two commutative normal operators normal? The answer is yes, and the proof is the same for spaces of all dimensions; the proof seems to need the Fuglede theorem. In this connection it should be mentioned that the product of not necessarily commutative normal operators is very reluctant to be normal. A pertinent positive result was obtained by Wiegmann [1948]; it says that if \mathbf{H} is a finite-dimensional Hilbert space, and if A and B are normal operators on \mathbf{H} such that AB is normal, then BA also is normal. Away from finite-dimensional spaces even this result becomes recalcitrant. It remains true for compact operators (Wiegmann [1949]), but it is false in the general case (Kaplansky [1953]).

153. Spectral measure of the unit disc. One of the techniques that can be used to prove the Fuglede theorem is to characterize in terms of the geometry of Hilbert space the spectral subspaces associated with a normal operator. That technique is useful in other contexts too.

A necessary and sufficient condition that a complex number have modulus less than or equal to 1 is that all its powers have the same property. This trivial observation extends to complex-valued functions: $\{x: |\varphi(x)| \leq 1\} = \{x: |\varphi(x)|^n \leq 1, n = 1, 2, 3, \dots\}$. There is a close

connection between complex-valued functions and normal operators. The operatorial analogue of the preceding numerical observations should be something like this: if A is a normal operator on a Hilbert space \mathbf{H} , then the set \mathbf{E} of those vectors f in \mathbf{H} for which $\|A^n f\| \leq \|f\|$, $n = 1, 2, 3, \dots$, should be, in some sense, the part of \mathbf{H} on which A is below 1. This is true; the precise formulation is that \mathbf{E} is a subspace of \mathbf{H} and the projection on \mathbf{E} is the value of the spectral measure associated with A on the closed unit disc in the complex plane. The same result can be formulated in a more elementary manner in the language of multiplication operators.

Problem 153. *If A is the multiplication operator induced by a bounded measurable function φ on a measure space, and if $D = \{z: |z| \leq 1\}$, then a necessary and sufficient condition that an element f in \mathbf{L}^2 be such that $\chi_{\varphi^{-1}(D)} f = f$ is that $\|A^n f\| \leq \|f\|$ for every positive integer n .*

Here, as usual, χ denotes the characteristic function of the set indicated by its subscript.

By translations and changes of scale the spectral subspaces associated with all discs can be characterized similarly; in particular, a necessary and sufficient condition that a vector f be invariant under multiplication by the characteristic function of $\{x: |\varphi(x)| \leq \varepsilon\}$ ($\varepsilon > 0$) is that $\|A^n f\| \leq \varepsilon^n \|f\|$ for all n . One way this result can sometimes be put to good use is this: if, for some positive number ε , there are no f 's in \mathbf{L}^2 (other than 0) such that $\|A^n f\| \leq \varepsilon^n \|f\|$ for all n , then the subspace of f 's that vanish on the complement of the set $\{x: |\varphi(x)| \leq \varepsilon\}$ is $\{0\}$, and therefore the set $\{x: |\varphi(x)| \leq \varepsilon\}$ is (almost) empty. Conclusion: under these circumstances $|\varphi(x)| > \varepsilon$ almost everywhere, and consequently the operator A is invertible.

154. Subnormal operators. The theory of normal operators is so successful that much of the theory of non-normal operators is modeled after it. A natural way to extend a successful theory is to weaken some of its hypotheses slightly and hope that the results are weakened only slightly. One weakening of normality is quasinormality (see Problem 108). Subnormal operators constitute a considerably more useful and

deeper generalization, which goes in an altogether different direction. An operator is *subnormal* if it has a normal extension. More precisely, an operator A on a Hilbert space \mathbf{H} is subnormal if there exists a normal operator B on a Hilbert space \mathbf{K} such that \mathbf{H} is a subspace of \mathbf{K} , the subspace \mathbf{H} is invariant under the operator B , and the restriction of B to \mathbf{H} coincides with A .

Every normal operator is trivially subnormal. On finite-dimensional spaces every subnormal operator is normal, but that takes a little proving; cf. Solution 159 or Problem 160. A more interesting and typical example of a subnormal operator is the unilateral shift; the bilateral shift is a normal extension.

Problem 154. *Every quasinormal operator is subnormal.*

Normality implies quasinormality, but not conversely (witness the unilateral shift). The present assertion is that quasinormality implies subnormality, but, again, the converse is false. To get a counterexample, add a non-zero scalar to the unilateral shift. The result is just as subnormal as the unilateral shift, but a straightforward computation shows that if it were also quasinormal, then the unilateral shift would be normal.

155. Minimal normal extensions. A normal extension B (on \mathbf{K}) of a subnormal operator A (on \mathbf{H}) is *minimal* if there is no reducing subspace of B between \mathbf{H} and \mathbf{K} . In other words, B is minimal over A if whenever \mathbf{M} reduces B and $\mathbf{H} \subset \mathbf{M}$, it follows that $\mathbf{M} = \mathbf{K}$. What is the right article for minimal normal extensions: “a” or “the”?

Problem 155. *If B_1 and B_2 (on \mathbf{K}_1 and \mathbf{K}_2) are minimal normal extensions of the subnormal operator A on \mathbf{H} , then there exists an isometry U from \mathbf{K}_1 onto \mathbf{K}_2 that carries B_1 onto B_2 (i.e., $UB_1 = B_2U$) and is equal to the identity on \mathbf{H} .*

In view of this result, it is permissible to speak of “the” minimal normal extension of a subnormal operator, and everyone does. Typical example: the minimal normal extension of the unilateral shift is the bilateral shift.

156. Similarity of subnormal operators. For normal operators similarity implies unitary equivalence (Problem 152). Subnormal operators are designed to imitate the properties of normal ones; is this one of the respects in which they succeed?

Problem 156. *Are two similar subnormal operators necessarily unitarily equivalent?*

157. Spectral inclusion theorem. If an operator A is a restriction of an operator B to an invariant subspace \mathbf{H} of B , and if f is an eigenvector of A (i.e., $f \in \mathbf{H}$ and $Af = \lambda f$ for some scalar λ), then f is an eigenvector of B . Differently expressed: if $A \subset B$, then $\Pi_0(A) \subset \Pi_0(B)$, or, as an operator grows, its point spectrum grows. An equally easy verification shows that as an operator grows, its approximate point spectrum grows. In view of these very natural observations, it is tempting to conjecture that as an operator grows, its spectrum grows, and hence that, in particular, if A is subnormal and B is its minimal normal extension, then $\Lambda(A) \subset \Lambda(B)$. The first non-trivial example of a subnormal operator shows that this conjecture is false: if A is the unilateral shift and B is the bilateral shift, then $\Lambda(A)$ is the unit disc, whereas $\Lambda(B)$ is only the perimeter of the unit disc. It turns out that this counterexample illustrates the general case better than do the plausibility arguments based on eigenvalues, exact or approximate.

Problem 157. *If A is subnormal and if B is its minimal normal extension, then $\Lambda(B) \subset \Lambda(A)$.*

Reference: Halmos [1952 a].

158. Filling in holes. The spectral inclusion theorem (Problem 157) for subnormal operators can be sharpened in an interesting and surprising manner. The result is that the spectrum of a subnormal operator is always obtained from the spectrum of its minimal normal extension by “filling in some of the holes”. This informal expression can be given a precise technical meaning. A *hole* in a compact subset of the complex plane is a bounded component of its complement.

Problem 158. *If A is subnormal, if B is its minimal normal extension, and if Δ is a hole of $\Lambda(B)$, then Δ is either included in or disjoint from $\Lambda(A)$.*

159. Extensions of finite co-dimension.

Problem 159. *Can a subnormal but non-normal operator on a Hilbert space \mathbf{H} have a normal extension to a Hilbert space \mathbf{K} when $\dim(\mathbf{K} \cap \mathbf{H}^\perp)$ is finite?*

160. Hyponormal operators. If A (on \mathbf{H}) is subnormal, with normal extension B (on \mathbf{K}), what is the relation between A^* and B^* ? The answer is best expressed in terms of the projection P from \mathbf{K} onto \mathbf{H} . If f and g are in \mathbf{H} , then

$$(A^*f, g) = (f, Ag) = (f, Bg) = (B^*f, g) = (B^*f, Pg) = (PB^*f, g).$$

Since the operator PB^* on \mathbf{K} leaves \mathbf{H} invariant, its restriction to \mathbf{H} is an operator on \mathbf{H} , and, according to the preceding chain of equations, that restriction is equal to A^* . That is the answer:

$$A^*f = PB^*f$$

for every f in \mathbf{H} .

This relation between A^* and B^* has a curious consequence. If $f \in \mathbf{H}$, then

$$\|A^*f\| = \|PB^*f\| \leq \|B^*f\| = \|Bf\| \text{ (by normality)} = \|Af\|.$$

The result ($\|A^*f\| \leq \|Af\|$) can be reformulated in another useful way; it is equivalent to the operator inequality

$$AA^* \leq A^*A.$$

Indeed: $\|A^*f\|^2 = (AA^*f, f)$ and $\|Af\|^2 = (A^*Af, f)$.

The curious inequality that subnormal operators always satisfy can also be obtained from an illuminating matrix calculation. Corresponding to the decomposition $\mathbf{K} = \mathbf{H} \oplus \mathbf{H}^\perp$, every operator on \mathbf{K} can be expressed

as an operator matrix, and, in particular, that is true for B . It is easy to express the relation ($A \subset B$) between A and B in terms of the matrix of B ; a necessary and sufficient condition for it is that (1) the principal (northwest) entry is A , and (2) the one below it (southwest) is 0. The condition (2) says that \mathbf{H} is invariant under B , and (1) says that the restriction of B to \mathbf{H} is A . Thus

$$B = \begin{pmatrix} A & R \\ 0 & S \end{pmatrix},$$

so that

$$B^* = \begin{pmatrix} A^* & 0 \\ R^* & S^* \end{pmatrix}.$$

Since B is normal, it follows that the matrix

$$B^*B - BB^* = \begin{pmatrix} A^*A & A^*R \\ R^*A & R^*R + S^*S \end{pmatrix} - \begin{pmatrix} AA^* + RR^* & RS^* \\ SR^* & SS^* \end{pmatrix}$$

must vanish. This implies that $A^*A - AA^* = RR^*$, and hence that $A^*A - AA^* \geq 0$.

There is a curious lack of symmetry here: why should A^*A play a role so significantly different from that of AA^* ? A little meditation on the unilateral shift may help. If $A = U$, the unilateral shift, then A is subnormal, and $A^*A = 1$, whereas AA^* is a non-trivial projection; clearly $A^*A \geq AA^*$. If $A = U^*$, then A is not subnormal. (Reason: if it were, then it would satisfy the inequality $A^*A \geq AA^*$, i.e., $UU^* \geq U^*U$, and then U would be normal.) If it were deemed absolutely essential, symmetry could be restored to the universe by the introduction of the dual concept of co-subnormality. (Proposed definition: the adjoint is subnormal.) If A is co-subnormal in this sense, then $AA^* \geq A^*A$. An operator A such that $A^*A \geq AA^*$ has been called *hyponormal*. (The dual kind might be called co-hyponormal. Note that "hypon" is in Greek what "sub" is in Latin. The nomenclature is not especially suggestive, but this is how it grew, and it seems to be here to stay.) The result of the preceding paragraphs is that every subnormal operator is hyponormal. The dull dual result is, of course, that every co-subnormal operator is co-hyponormal.

On a finite-dimensional space every hyponormal operator is normal. The most efficient proof of this assertion is a trace argument, as follows. Since $\text{tr}(AB)$ is always equal to $\text{tr}(BA)$, it follows that $\text{tr}(A^*A - AA^*)$ is always 0; if $A^*A \geq AA^*$, then $A^*A - AA^*$ is a positive operator with trace 0, and therefore $A^*A - AA^* = 0$. What was thus proved is a generalization of the statement that on a finite-dimensional space every subnormal operator is normal (cf. Problem 154).

Problem 160. *Give an example of a hyponormal operator that is not subnormal.*

This is not easy. The techniques used are almost sufficient to yield an intrinsic characterization of subnormality obtained by Halmos [1950 a] and sharpened by Bram [1955]. “Intrinsic” means that the characterization is expressed in terms of the action of the operator on the vectors in its domain, and not in terms of the existence of something outside that domain. The characterization is of “finite character”, in the sense that it depends on the behavior of the operator on all possible finite sets of vectors. With still more work of the same kind an elegant topological characterization of subnormality can be obtained; this was first done by Bishop [1957]. Bishop’s result is easy to state: the set of all subnormal operators is exactly the strong closure of the set of all normal operators.

161. Normal and subnormal partial isometries.

Problem 161. *A partial isometry is normal if and only if it is the direct sum of a unitary operator and zero; it is subnormal if and only if it is the direct sum of an isometry and zero.*

In both cases, one or the other of the direct summands may be absent.

162. Norm powers and power norms. The set \mathbf{T} of those operators A such that $\|A^n\| = \|A\|^n$ for every positive integer n has, at the very least, a certain curiosity value. If $A \in \mathbf{T}$, then $\|A^n\|^{1/n} = \|A\|$, and therefore $r(A) = \|A\|$; if, conversely, $r(A) = \|A\|$, then

$\|A^n\| \leq \|A\|^n = (r(A))^n = r(A^n) \leq \|A^n\|$, so that equality holds all the way through. Conclusion: $A \in \mathbf{T}$ if and only if $r(A) = \|A\|$.

The definition of \mathbf{T} implies that every normal operator belongs to \mathbf{T} (and so does the conclusion of the preceding paragraph). For two-by-two matrices an unpleasant computation proves a strong converse: if $\|A^2\| = \|A\|^2$, then A is normal. Since neither the assertion nor its proof have any merit, the latter is omitted. As soon as the dimension becomes greater than 2, the converse becomes false. If, for example,

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix},$$

then $\|A^n\| = 1$ for all n , but A is certainly not normal.

The quickest (but not the most elementary) proof of the direct assertion (if A is normal, then $A \in \mathbf{T}$) is to refer to the spectral theorem. Since for subnormal and hyponormal operators that theorem is not available, a natural question remains unanswered. The answer turns out to be affirmative.

Problem 162. *If A is hyponormal, then $\|A^n\| = \|A\|^n$ for every positive integer n .*

Corollary. *The only hyponormal quasinilpotent operator is 0.*

163. Compact hyponormal operators. It follows from the discussion of hyponormal operators on finite-dimensional spaces (Problem 160) that a hyponormal operator of finite rank (on a possibly infinite-dimensional space) is always normal. What about limits of operators of finite rank?

Problem 163. *Every compact hyponormal operator is normal.*

Reference: Andô [1963], Berberian [1962], Stampfli [1962].

164. Powers of hyponormal operators. Every power of a normal operator is normal. This trivial observation has as an almost equally

trivial consequence the statement that every power of a subnormal operator is subnormal. For hyponormal operators the facts are different.

Problem 164. *Give an example of a hyponormal operator whose square is not hyponormal.*

This is not easy. It is, in fact, bound to be at least as difficult as the construction of a hyponormal operator that is not subnormal (Problem 160), since any solution of Problem 164 is automatically a solution of Problem 160. The converse is not true; the hyponormal operator used in Solution 160 has the property that all its powers are hyponormal also.

165. Contractions similar to unitary operators.

Problem 165. *Is a contraction similar to a unitary operator necessarily unitary?*

Chapter 17. Numerical range

166. The Toeplitz-Hausdorff theorem. In early studies of Hilbert space (by Hilbert, Hellinger, Toeplitz, and others) the objects of chief interest were quadratic forms. Nowadays they play a secondary role. First comes an operator A on a Hilbert space \mathbf{H} , and then, apparently as an afterthought, comes the numerical-valued function $f \rightarrow (Af, f)$ on \mathbf{H} . This is not to say that the quadratic point of view is dead; it still suggests questions that are interesting with answers that can be useful.

Most quadratic questions about an operator are questions about its numerical range, sometimes called its field of values. The *numerical range* of an operator A is the set $W(A)$ of all complex numbers of the form (Af, f) , where f varies over all vectors on the unit sphere. (Important: $\|f\| = 1$, not $\|f\| \leq 1$.) The numerical range of A is the range of the restriction to the unit sphere of the quadratic form associated with A . One reason for the emphasis on the image of the unit sphere is that the image of the unit ball, and also the entire range, are easily described in terms of it, but not vice versa. (The image of the unit ball is the union of all the closed segments that join the origin to points of the numerical range; the entire range is the union of all the closed rays from the origin through points of the numerical range.)

The determination of the numerical range of an operator is sometimes easy. Here are some sample results. If

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix},$$

then $W(A)$ is the closed unit interval (easy); if

$$A = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix},$$

then $W(A)$ is the closed disc with center 0 and radius $\frac{1}{2}$ (easy, but more

interesting); if

$$A = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix},$$

then $W(A)$ is the closed elliptical disc with foci at 0 and 1, minor axis 1 and major axis $\sqrt{2}$ (analytic geometry at its worst). There is a theorem that covers all these cases. If A is a two-by-two matrix with distinct eigenvalues α and β , and corresponding eigenvectors f and g , so normalized that $\|f\| = \|g\| = 1$, then $W(A)$ is a closed elliptical disc with foci at α and β ; if $\gamma = |(f, g)|$ and $\delta = \sqrt{1 - \gamma^2}$, then the minor axis is $\gamma \|\alpha - \beta\|/\delta$ and the major axis is $\|\alpha - \beta\|/\delta$. If A has only one eigenvalue α , then $W(A)$ is the (circular) disc with center α and radius $\frac{1}{2} \|A - \alpha\|$.

A couple of three-dimensional examples will demonstrate that the two-dimensional case is not typical. If

$$A = \begin{pmatrix} 0 & 0 & \lambda \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix},$$

where λ is a complex number of modulus 1, then $W(A)$ is the equilateral triangle (interior and boundary) whose vertices are the three cube roots of λ . (Cf. Problem 171.) If

$$A = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

then $W(A)$ is the union of all the closed segments that join the point 1 to points of the closed disc with center 0 and radius $\frac{1}{2}$. (Cf. Problem 171.)

The higher the dimension, the stranger the numerical range can be. If A is the Volterra integration operator (see Problem 148), then $W(A)$

is the set lying between the curves

$$t \rightarrow \frac{1 - \cos t}{t^2} \pm i \frac{t - \sin t}{t^2}, \quad 0 \leq t \leq 2\pi$$

(where the value at 0 is taken to be the limit from the right).

The following assertion describes the most important common property of all these examples.

Problem 166. *The numerical range of an operator is always convex.*

The result is known as the *Toeplitz-Hausdorff theorem*. Every known proof is computational. The computations can be arranged well, or they can be arranged badly, but any way they are arranged they are still computations. A conceptual proof would be desirable even (or especially?) if the concepts it uses are less elementary than the ones in a computational proof.

One more comment is pertinent. Consideration of real and imaginary parts shows that the Toeplitz-Hausdorff theorem is a special case ($n = 2$) of the following general assertion: if A_1, \dots, A_n are Hermitian operators, then the set of all n -tuples of the form $\langle (A_1 f, f), \dots, (A_n f, f) \rangle$, where $\|f\| = 1$, is a convex subset of n -dimensional real Euclidean space. True or false, the assertion seems to be a natural generalization of the Toeplitz-Hausdorff theorem; it is a pity that it is so very false. It is false for $n = 3$ in dimension 2; counterexamples are easy to come by.

The first paper on the subject was by Toeplitz [1918], who proved that the boundary of $W(A)$ is a convex curve, but left open the possibility that it had interior holes. Hausdorff [1919] proved that it did not. Donoghue [1957 a] re-examined the facts and presented some pertinent computations. The result about the Volterra integration operator is due to A. Brown.

167. Higher-dimensional numerical range. The numerical range can be regarded as the one-dimensional case of a multi-dimensional concept. To see how that goes, recall the expression of a projection P of rank 1 in terms of a unit vector f in its range:

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

for all g . If A is an arbitrary operator, then PAP is an operator of rank 1, and therefore a finite-dimensional concept such as trace makes sense for it. The trace of PAP can be computed by finding the (one-by-one) matrix of the restriction of PAP to the range of P , with respect to the (one-element) basis $\{f\}$; since $Pf = f$, the value of that trace is

$$(PAPf, f) = (APf, Pf) = (Af, f).$$

These remarks can be summarized as follows: $W(A)$ is equal to the set of all complex numbers of the form $\text{tr } PAP$, where P varies over all projections of rank 1. Replace 1 by an arbitrary positive integer k , and obtain the k -numerical range of A , in symbols $W_k(A)$: it is the set of all complex numbers of the form $\text{tr } PAP$, where P varies over all projections of rank k . The ordinary numerical range is the k -numerical range with $k = 1$.

Problem 167. *Is the k -numerical range of an operator always convex?*

168. Closure of numerical range.

Problem 168. *Give examples of operators whose numerical range is not closed.*

Observe that in the finite-dimensional case the numerical range of an operator is a continuous image of a compact set, and hence necessarily compact.

169. Spectrum and numerical range.

Problem 169. *The closure of the numerical range includes the spectrum.*

The trivial corollary that asserts that if $A = B + iC$, with B and C Hermitian, then $\Lambda(A) \subset \overline{W(B)} + i\overline{W(C)}$ is the *Bendixson-Hirsch theorem*.

170. Quasinilpotence and numerical range. If A is a quasinilpotent operator, then, by Problem 169, $0 \in \overline{W(A)}$. By Solution 168, the set $W(A)$ may fail to be closed, so that from $0 \in \overline{W(A)}$ it does not follow that $0 \in W(A)$. Is it true just the same?

Problem 170. Give an example of a quasinilpotent operator A such that $0 \notin W(A)$.

Observe that any such example is a solution of Problem 168.

171. Normality and numerical range. Can the closure of the numerical range be very much larger than the spectrum? The answer is yes. A discouraging example is

$$\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix};$$

the spectrum is small ($\{0\}$), but the numerical range is large ($\{z: |z| \leq \frac{1}{2}\}$). Among normal operators such extreme examples do not exist; for them the closure of the numerical range is as small as the universal properties of spectra and numerical ranges permit.

To formulate the result precisely, it is necessary to introduce the concept of the *convex hull* of a set M , in symbols $\text{conv } M$. By definition, $\text{conv } M$ is the smallest convex set that includes M ; in other words, $\text{conv } M$ is the intersection of all the convex sets that include M . It is a non-trivial fact of finite-dimensional Euclidean geometry that the convex hull of a compact set is closed. Perhaps the most useful formulation of this fact for the plane goes as follows: the convex hull of a compact set is the intersection of all the closed half planes that include it. A useful reference for all this is Valentine [1964].

So much for making convex sets out of closed sets. The reverse process of making closed sets out of convex sets is much simpler to deal with; it is true and easy to prove that the closure of a convex set is convex.

Problem 171. The closure of the numerical range of a normal operator is the convex hull of its spectrum.

As an application consider the matrix

$$\begin{pmatrix} 0 & 0 & \lambda \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix},$$

where $|\lambda| = 1$. Since this matrix is unitary, and therefore normal, the result just proved implies that its numerical range is the convex hull of its eigenvalues. The eigenvalues are easy to compute (they are the cube roots of λ), and this proves the assertion (made in passing in Problem 166) that the numerical range of this particular matrix is the triangle whose vertices are the cube roots of λ .

The general result includes the special assertion that the numerical range of every finite diagonal matrix is the convex hull of its diagonal entries. A different generalization of this special assertion is that the numerical range of a direct sum is the convex hull of the numerical ranges of its summands. The proof of the generalization is straightforward. For an example, consider the direct sum of

$$\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

and (1), and recapture the assertion (made in passing in Problem 166) about the domed-cone shape of the numerical range of

$$\begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

172. Subnormality and numerical range.

Problem 172. *Does the conclusion of Problem 171 remain true if in the hypothesis “normal” is replaced by “subnormal”?*

173. Numerical radius. The numerical range, like the spectrum, associates a set with each operator; it is a set-valued function of operators. There is a closely related numerical function w , called the *numerical radius*, defined by

$$w(A) = \sup \{ |\lambda| : \lambda \in W(A) \}.$$

(Cf. the definition of spectral radius, Problem 74.) Some of the properties of the numerical radius lie near the surface; others are quite deep.

It is easy to prove that w is a norm. That is: $w(A) \geq 0$, and $w(A) = 0$ if and only if $A = 0$; $w(\alpha A) = |\alpha| \cdot w(A)$ for each scalar α ; and $w(A + B) \leq w(A) + w(B)$. This norm is equivalent to the ordinary operator norm, in the sense that each is bounded by a constant multiple of the other:

$$\frac{1}{2} \|A\| \leq w(A) \leq \|A\|.$$

(See Halmos [1951, p. 33].) The norm w has many other pleasant properties; thus, for instance, $w(A^*) = w(A)$, $w(A^*A) = \|A\|^2$, and w is unitarily invariant, in the sense that $w(U^*AU) = w(A)$ whenever U is unitary.

Since $\Lambda(A) \subset \overline{W(A)}$ (Problem 169), there is an easy inequality between the spectral radius and the numerical radius:

$$r(A) \leq w(A).$$

The existence of quasinilpotent (or, for that matter, nilpotent) operators shows that nothing like the reverse inequality could be true.

Problem 173. (a) If $w(1 - A) < 1$, then A is invertible. (b) If $w(A) = \|A\|$, then $r(A) = \|A\|$.

174. Normaloid, convexoid, and spectraloid operators. If A is normal, then $w(A) = \|A\|$. Wintner called an operator A with $w(A) = \|A\|$ *normaloid*. Another useful (but nameless) property of a normal operator A (Problem 171) is that $\overline{W(A)}$ is the convex hull of $\Lambda(A)$. To have a temporary label for (not necessarily normal) operators with this property, call them *convexoid*. Still another (nameless) property

of a normal operator A is that $r(A) = w(A)$; call an operator with this property *spectraloid*. It is a consequence of Problem 173 that every normaloid operator is spectraloid. It is also true that every convexoid operator is spectraloid. Indeed, since the closed disc with center 0 and radius $r(A)$ includes $\Lambda(A)$ and is convex, it follows that if A is convexoid, then that disc includes $W(A)$. This implies that $w(A) \leq r(A)$, and hence that A is spectraloid.

Problem 174. *Discuss the implication relations between the properties of being convexoid and normaloid.*

175. Continuity of numerical range. In what sense is the numerical range a continuous function of its argument? (Cf. Problems 85 and 86.) The best way to ask the question is in terms of the *Hausdorff metric* for compact subsets of the plane. To define that metric, write

$$M + (\epsilon) = \{z + \alpha : z \in M, |\alpha| < \epsilon\}$$

for each set M of complex numbers and each positive number ϵ . In this notation, if M and N are compact sets, the Hausdorff distance $d(M, N)$ between them is the infimum of all positive numbers ϵ such that both $M \subset N + (\epsilon)$ and $N \subset M + (\epsilon)$.

Since the Hausdorff metric is defined for compact sets, the appropriate function to discuss is \bar{W} , not W . As for the continuity question, it still has as many interpretations as there are topologies for operators. Is \bar{W} weakly continuous? strongly? uniformly? And what about w ? The only thing that is immediately obvious is that if \bar{W} is continuous with respect to any topology, then so is w , and, consequently, if w is discontinuous, then so is \bar{W} .

Problem 175. *Discuss the continuity of \bar{W} and w in the weak, strong, and uniform operator topologies.*

176. Power inequality. The good properties of the numerical range and the numerical radius have to do with convexity and linearity; the relations between the numerical range and the multiplicative properties of operators are less smooth. Thus, for instance, w is certainly not multi-

plicative, i.e., $w(AB)$ is not always equal to $w(A)w(B)$. (Example with commutative normal operators: if

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix},$$

then $w(A) = w(B) = 1$ and $w(AB) = 0$.) The next best thing would be for w to be submultiplicative ($w(AB) \leq w(A)w(B)$), but that is false too. (Example: if

$$A = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},$$

then $w(A) = w(B) = \frac{1}{2}$ and $w(AB) = 1$.) Since $w(AB) \leq \|AB\| \leq \|A\| \cdot \|B\|$, it follows that for normal operators w is submultiplicative (because if A and B are normal, then $\|A\| = w(A)$ and $\|B\| = w(B)$), and for operators in general $w(AB) \leq 4w(A)w(B)$ (because $\|A\| \leq 2w(A)$ and $\|B\| \leq 2w(B)$). The example used to show that w is not submultiplicative shows also that the constant 4 is best possible here.

Commutativity sometimes helps; here it does not. Examples of commutative operators A and B for which $w(AB) > w(A)w(B)$ are a little harder to come by, but they exist. Here is one:

$$A = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

and $B = A^2$. It is easy to see that $w(A^2) = w(A^3) = \frac{1}{2}$. The value of $w(A)$ is slightly harder to compute, but it is not needed; the almost obvious relation $w(A) < 1$ will do. Indeed: $w(AB) = w(A^3) = \frac{1}{2} > w(A) \cdot \frac{1}{2} = w(A)w(B)$.

The only shred of multiplicative behavior that has not yet been ruled out is the *power inequality*

$$w(A^n) \leq (w(A))^n.$$

This turns out to be true, but remarkably tricky. Even for two-by-two matrices there is no simple computation that yields the result. If not the dimension but the exponent is specialized, if, say, $n = 2$, then relatively easy proofs exist, but even they require surprisingly delicate handling. The general case requires either brute force or ingenuity.

Problem 176. *If A is an operator such that $w(A) \leq 1$, then $w(A^n) \leq 1$ for every positive integer n .*

The statement is obviously a consequence of the power inequality. To show that it also implies the power inequality, reason as follows. If $w(A) = 0$, then $A = 0$, and everything is trivial. If $w(A) \neq 0$, then write $B = A/w(A)$, note that $w(B) \leq 1$, use the statement of Problem 176 to infer that $w(B^n) \leq 1$, and conclude that $w(A^n) \leq (w(A))^n$.

Generalizations of the theorem are known. Here is a nice one: if p is a polynomial such that $p(0) = 0$ and $|p(z)| \leq 1$ whenever $|z| \leq 1$, and if A is an operator such that $w(A) \leq 1$, then $w(p(A)) \leq 1$. With a little care, polynomials can be replaced by analytic functions, and, with a lot of care, the unit disc (which enters by the emphasis on the inequality $|z| \leq 1$) can be replaced by other compact convex sets.

The first proof of the power inequality is due to C. A. Berger; the first generalizations along the lines mentioned in the preceding paragraph were derived by J. G. Stampfli. The first published version, in a quite general form, was given by Kato [1965]. An interesting generalization along completely different lines appears in Nagy-Foiaş [1966].

Chapter 18. Unitary dilations

177. Unitary dilations. Suppose that \mathbf{H} is a subspace of a Hilbert space \mathbf{K} , and let P be the (orthogonal) projection from \mathbf{K} onto \mathbf{H} . Each operator B on \mathbf{K} induces in a natural way an operator A on \mathbf{H} defined for each f in \mathbf{H} by

$$Af = PBf.$$

The relation between A and B can also be expressed by

$$AP = PBP.$$

Under these conditions the operator A is called the *compression* of B to \mathbf{H} and B is called a *dilation* of A to \mathbf{K} . This geometric definition of compression and dilation is to be contrasted with the customary concepts of restriction and extension: if it happens that \mathbf{H} is invariant under B , then it is not necessary to project Bf back into \mathbf{H} (it is already there), and, in that case, A is the restriction of B to \mathbf{H} and B is an extension of A to \mathbf{K} . Restriction-extension is a special case of compression-dilation, the special case in which the operator on the larger space leaves the smaller space invariant.

There are algebraic roads that lead to compressions and dilations, as well as geometric ones. One such road goes via quadratic forms. It makes sense to consider the quadratic form associated with B and to consider it for vectors of \mathbf{H} only (i.e., to restrict it to \mathbf{H}). This restriction is a quadratic form on \mathbf{H} , and, therefore, it is induced by an operator on \mathbf{H} ; that operator is the compression A . In other words, compression and dilation for operators are not only analogous to (and generalizations of) restriction and extension, but, in the framework of quadratic forms, they *are* restriction and extension: the quadratic form of A is the restriction of the quadratic form of B to \mathbf{H} , and the quadratic form of B is an extension of the quadratic form of A to \mathbf{K} .

Still another manifestation of compressions and dilations in Hilbert

space theory is in connection with operator matrices. If \mathbf{K} is decomposed into \mathbf{H} and \mathbf{H}^\perp , and, correspondingly, operators on \mathbf{K} are written in terms of matrices (whose entries are operators on \mathbf{H} and \mathbf{H}^\perp and linear transformations between \mathbf{H} and \mathbf{H}^\perp), then a necessary and sufficient condition that B be a dilation of A is that the matrix of B have the form

$$\begin{pmatrix} A & X \\ Y & Z \end{pmatrix}.$$

- Problem 177.** (a) *If $\|A\| \leq 1$, then A has a unitary dilation.*
 (b) *If $0 \leq A \leq 1$, then A has a dilation that is a projection.*

Note that in both cases the assumptions are clearly necessary. If A has a dilation B that is a contraction, then $\|Af\| = \|PBf\| \leq \|Bf\| \leq \|f\|$ for all f in \mathbf{H} , and if A has a positive dilation B , then $(Af, f) = (Bf, f) \geq 0$ for all f in \mathbf{H} .

Corollary. *Every operator has a normal dilation.*

178. Unitary power dilations. The least unitary looking contraction is 0, but even it has a unitary dilation. The construction of Solution 177 exhibits it as

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

The construction is canonical, in a sense, but it does not have many useful algebraic properties. It is not necessarily true, for instance, that the square of a dilation is a dilation of the square; indeed, the square of the dilation of 0 exhibited above is

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

which is not a dilation of the square of 0. Is there a unitary dilation of 0 that is fair to squares? The answer is yes:

$$\begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

is an example. The square of this dilation is

$$\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix},$$

which is a dilation of the square of 0. Unfortunately, however, this dilation is not perfect either; its cube is

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

which is not a dilation of the cube of 0. The cube injustice can be remedied by passage to

$$\begin{pmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix},$$

but then fourth powers fail. There is no end to inductive greed; the clearly suggested final demand is for a unitary dilation of 0 with the property that all its powers are dilations of 0. In matrix language the

demand is for a unitary matrix with the property that one of its diagonal entries is 0 and that, moreover, the corresponding entry in all its powers is also 0. Brief meditation on the preceding finite examples, or just inspired guessing, might suggest the answer; the bilateral shift will work, with the $\langle 0,0 \rangle$ entry playing the distinguished role. (Caution: the unilateral shift is not unitary.) The general definition suggested by the preceding considerations is this: an operator B is a *power dilation* (sometimes called a *strong dilation*) of an operator A if B^n is a dilation of A^n for $n = 1, 2, 3, \dots$.

Problem 178. *Every contraction has a unitary power dilation.*

In all fairness to dilations, it should be mentioned that they all have at least one useful algebraic property: if B is a dilation of A , then B^* is a dilation of A^* . The quickest proof is via quadratic forms: if $\langle Af, f \rangle = \langle Bf, f \rangle$ for each f in the domain of A , then, for the same f 's, $\langle A^*f, f \rangle = \langle f, Af \rangle = \langle Af, f \rangle^* = \langle Bf, f \rangle^* = \langle f, Bf \rangle = \langle B^*f, f \rangle$. One consequence of this is that if B is a power dilation of A , then B^* is a power dilation of A^* .

The power dilation theorem was first proved by Nagy [1953]. The subject has received quite a lot of attention since then; good summaries of results are in Nagy [1955] and Mlak [1965]. An especially interesting aspect of the theory concerns minimal unitary power dilations. Their definition is similar to that of minimal normal extensions (Problem 155), and they too are uniquely determined by the given operator (to within unitary equivalence). The curious fact is that knowledge of the minimal unitary power dilation of an operator is not so helpful as one might think. Schreiber [1956] proved that all proper contractions (see Problem 122) on separable Hilbert spaces have the same minimal unitary power dilation, namely a bilateral shift; Nagy [1957] extended the result to non-separable spaces.

179. Ergodic theorem. If u is a complex number of modulus 1, then the averages

$$\frac{1}{n} \sum_{j=0}^{n-1} u^j$$

form a convergent sequence. This is an amusing and simple piece of

classical analysis, whose generalizations are widely applicable. To prove the statement, consider separately the cases $u = 1$ and $u \neq 1$. If $u = 1$, then each average is equal to 1, and the limit is 1. If $u \neq 1$, then

$$\left| \frac{1}{n} \sum_{j=0}^{n-1} u^j \right| = \left| \frac{1 - u^n}{n(1 - u)} \right| \leq \frac{2}{n |1 - u|},$$

and the limit is 0.

The most plausible operatorial generalization of the result of the preceding paragraph is known as the *mean ergodic theorem* for unitary operators; it asserts that if U is a unitary operator on a Hilbert space, then the averages

$$\frac{1}{n} \sum_{j=0}^{n-1} U^j$$

form a strongly convergent sequence. A more informative statement of the ergodic theorem might go on to describe the limit; it is, in fact, the projection whose range is the subspace $\{f: Uf = f\}$, i.e., the subspace of fixed points of U .

It is less obvious that a similar ergodic theorem is true not only for unitary operators but for all contractions.

Problem 179. *If A is a contraction on a Hilbert space \mathbf{H} , then*

$$\left\{ \frac{1}{n} \sum_{j=0}^{n-1} A^j \right\}$$

is a strongly convergent sequence of operators on \mathbf{H} .

180. Spectral sets. If F is a bounded complex-valued function defined on a set M , write

$$\|F\|_M = \sup \{ |F(\lambda)| : \lambda \in M \}.$$

If A is a normal operator with spectrum Λ , and if F is a bounded Borel measurable function on Λ , then $\|F(A)\| \leq \|F\|_\Lambda$. (Equality does not hold in general; F may take a few large values that have no measure-

theoretically detectable influence on $F(A)$.) It is not obvious how this inequality can be generalized to non-normal operators. There are two obstacles: in general, $F(A)$ does not make sense, and, when it does, the result can be false. There is an easy way around both obstacles: consider only such functions F for which $F(A)$ does make sense, and consider only such sets, in the role of Λ , for which the inequality does hold. A viable theory can be built on these special considerations.

If the only functions considered are polynomials, then they can be applied to every operator. If, however, the spectrum of the operator is too small, the inequality between norms will fail. If, for instance, A is quasinilpotent and $p(z) = z$, then $\|p(A)\| = \|A\|$ and $\|p\|_{\Lambda(A)} = 0$; the inequality $\|p(A)\| \leq \|p\|_{\Lambda(A)}$ holds only if $A = 0$. The earliest positive result, which is still the most incisive and informative statement along these lines, is known as the *von Neumann-Heinz theorem*. (Reference: von Neumann [1951], Heinz [1952].)

Problem 180. *If $\|A\| \leq 1$ and if D is the closed unit disc, then*

$$\|p(A)\| \leq \|p\|_D$$

for every polynomial p .

The general context to which the theorem belongs is the theory of spectral sets. That theory is concerned with rational functions instead of just polynomials. Roughly speaking, a spectral set for an operator is a set such that the appropriate norm inequality holds for all rational functions on the set. Precisely, a *spectral set* for A is a set M such that $\Lambda(A) \subset M$ and such that if F is a bounded rational function on M (i.e., a rational function with no poles in the closure of M), then $\|F(A)\| \leq \|F\|_M$. (Note that the condition on the poles of the admissible F 's implies that $F(A)$ makes sense for each such F .) It turns out that the theory loses no generality if the definition of spectral set demands that the set be closed, or even compact, and that is usually done. To demand the norm inequality for polynomials only does, however, seriously change the definition. A moderately sophisticated complex function argument (cf. Lebow [1963]) can be used to show that the polynomial definition and the rational function definition are the same in case the

set in question is sufficiently simple. (For this purpose a set is sufficiently simple if it is compact and its complement is connected.) In view of the last remark, the von Neumann-Heinz theorem is frequently stated as follows: the closed unit disc is a spectral set for every contraction.

181. Dilations of positive definite sequences. The theorem on unitary power dilations (Problem 178) says that certain sequences of operators can be obtained by compressing the sequence of powers of one unitary operator. More precisely: if A is a contraction on \mathbf{H} , and if $A_n = A^n$ for $n \geq 0$ and $A_n = A^{*n}$ for $n \leq 0$, then there exists a unitary operator U , on a space that includes \mathbf{H} , such that the compression of U^n to \mathbf{H} is A_n , $n = 0, \pm 1, \pm 2, \dots$. What other sequences $\{A_n\}$ can be obtained in this way? Is there a usable intrinsic characterization of such sequences? The answers to these questions are best formulated in terms of positive definite sequences of operators. The sequence of powers of a unitary operator turns out to be positive definite in a strong sense of that phrase; the compression of a positive definite sequence is itself positive definite; and, for suitably normalized sequences, positive definiteness is, in fact, a necessary and sufficient condition for the possession of a unitary dilation. The detailed explanations and definitions follow.

A sequence $\{A_n: n = 0, \pm 1, \pm 2, \dots\}$ is *positive definite* if

$$\sum_i \sum_j (A_{i-j} f_i, f_j) \geq 0$$

for every finitely non-zero sequence $\{f_n\}$ of vectors. A standard elementary argument via polarization shows that a positive definite sequence is Hermitian symmetric in the sense that $A_n^* = A_{-n}$ for all n .

If $A_n = U^n$ with U unitary, then

$$(A_{i-j} f_i, f_j) = (U^{i-j} f_i, f_j) = (U^j f_i, U^j f_j),$$

so that

$$\sum_i \sum_j (A_{i-j} f_i, f_j) = \sum_i \sum_j (U^j f_i, U^j f_j) = \left\| \sum_i U^j f_i \right\|^2 \geq 0;$$

the sequence of powers of U is positive definite.

Suppose now that U is unitary on \mathbf{K} , that \mathbf{H} is a subspace of \mathbf{K} , and that P is the projection from \mathbf{K} to \mathbf{H} . If $A_n f = P U^n f$ for each f in \mathbf{H} (i.e., if A_n is the compression of U^n to \mathbf{H}), and if $\{f_n\}$ is a finitely non-zero sequence of vectors in \mathbf{H} , then

$$(A_{i-j} f_i, f_j) = (P U^{i-j} f_i, f_j) = (U^{i-j} f_i, f_j);$$

it follows that $\{A_n\}$ is positive definite. The sequence $\{A_n\}$ is also normalized in the sense that $A_0 = 1$; the reason is that $U^0 = 1$.

The preceding paragraphs have introduced positive definiteness and have shown it to be a necessary condition for the possession of a unitary dilation. (To speak of a unitary dilation of a sequence means, as it has meant above, that each term of the sequence is the compression of the corresponding power of a fixed unitary operator.) The main task is to prove that the condition is sufficient. (Reference: Nagy [1955].)

Problem 181. *If $\{A_n\}$ is a positive definite sequence of operators on a Hilbert space \mathbf{H} , such that $A_0 = 1$, then there exists a unitary operator U on a Hilbert space that includes \mathbf{H} such that the compression of U^n to \mathbf{H} is A_n .*

The theorem is non-trivial even if the given Hilbert space \mathbf{H} is one-dimensional. It says in that case that if $\{\alpha_n\}$ is a positive definite sequence of complex numbers with $\alpha_0 = 1$, then there exists a Hilbert space \mathbf{K} , there exists a unit vector f in \mathbf{K} , and there exists a unitary operator U on \mathbf{K} such that $\alpha_n = (U^n f, f)$ for every integer n . If U is represented as a multiplication induced by a multiplier φ on some $L^2(\mu)$, the result is that $\alpha_n = \int \varphi^n |f|^2 d\mu$. By a standard change of variables, this can be reformulated as follows: there exists a normalized measure ν on the Borel sets of the unit circle such that $\alpha_n = \int \lambda^n d\nu(\lambda)$ for all n . In this form the statement is sometimes called *Herglotz's theorem*; it has been extensively generalized to groups other than the additive group of integers. The derivation of Herglotz's theorem from dilation theory and spectral theory is probably not the most economical way of getting at it, but it is a way. In any case it is good to be aware of still another connection between Hilbert space theory and classical analysis.

Chapter 19.

Commutators of operators

182. Commutators. A mathematical formulation of the famous Heisenberg uncertainty principle is that a certain pair of linear transformations P and Q satisfies, after suitable normalizations, the equation $PQ - QP = 1$. It is easy enough to produce a concrete example of this behavior; consider $L^2(-\infty, +\infty)$ and let P and Q be the differentiation transformation and the position transformation, respectively (that is, $(Pf)(x) = f'(x)$ and $(Qf)(x) = xf(x)$). These are not bounded linear transformations, of course, their domains are far from being the whole space, and they misbehave in many other ways. Can this misbehavior be avoided?

To phrase the question precisely, define a *commutator* as an operator of the form $PQ - QP$, where P and Q are operators on a Hilbert space. More general uses of the word can be found in the literature (e.g., commutators on Banach spaces), and most of them do not conflict with the present definition; the main thing that it is intended to exclude is the unbounded case. The question of the preceding paragraph can be phrased this way: "Is 1 a commutator?" The answer is no.

Problem 182. *The only scalar commutator is 0.*

The finite-dimensional case is easy to settle. The reason is that in that case the concept of trace is available. Trace is linear, and the trace of a product of two factors is independent of their order. It follows that the trace of a commutator is always zero; the only scalar with trace 0 is 0 itself. That settles the negative statement. More is known: in fact a finite square matrix is a commutator if and only if it has trace 0 (Shoda [1936], Albert-Muckenhoupt [1957]).

For the general (not necessarily finite-dimensional) case, two beautiful proofs are known, quite different from one another; they are due to Wintner [1947] and Wielandt [1949]. Both apply, with no change, to arbitrary complex normed algebras with unit. A *normed algebra* is a

normed vector space that is at the same time an algebra such that

$$\|fg\| \leq \|f\| \cdot \|g\|$$

for all f and g . A *unit* in a normed algebra is, of course, an element e such that $ef = fe = f$ for all f ; it is customary to require, moreover, that $\|e\| = 1$. The algebraic character of the Wintner and Wielandt proofs can be used to get more information about commutators, as follows.

The identity is a projection; it is the unique projection with nullity 0. (Recall that the nullity of an operator is the dimension of its kernel.) What about a projection (on an infinite-dimensional Hilbert space) with nullity 1; can it be a commutator? Intuition cries out for a negative answer, and, for once, intuition is right (Halmos [1963 a]). Consider the normed algebra of all operators and in it the ideal of compact operators. The quotient algebra is a normed algebra. In that algebra the unit element is not a commutator (by Wintner and Wielandt); translated back to operators, this means that the identity cannot be equal to the sum of a commutator and a compact operator. Since a projection with nullity 1 is a very special example of such a sum, the proof is complete. The following statement summarizes what the proof proves.

Corollary. *The sum of a compact operator and a non-zero scalar is not a commutator.*

The corollary gives a sufficient condition that an operator be a non-commutator; the most surprising fact in this subject is that on separable spaces the condition is necessary also (Brown-Pearcy [1965]). In other words: on a separable space every operator that is not the sum of a non-zero scalar and a compact operator is a commutator. The proof is not short.

183. Limits of commutators. Granted that the identity is not a commutator, is it at least a limit of commutators? Do there, in other words, exist sequences $\{P_n\}$ and $\{Q_n\}$ of operators such that $\|1 - (P_n Q_n - Q_n P_n)\| \rightarrow 0$ as $n \rightarrow \infty$? The Brown-Pearcy characterization of commutators (see Problem 182) implies that the answer is yes. (See also Problem 187.) A more modest result is more easily accessible.

Problem 183. *If $\{P_n\}$ and $\{Q_n\}$ are bounded sequences of operators (i.e., if there exists a positive number α such that $\|P_n\| \leq \alpha$ and $\|Q_n\| \leq \alpha$ for all n), and if the sequence $\{P_n Q_n - Q_n P_n\}$ converges in the norm to an operator C , then $C \neq 1$.*

In other words: the identity cannot be the limit of commutators formed from bounded sequences. Reference: Brown-Halmos-Pearcy [1965].

184. The Kleinecke-Shirokov theorem. The result of Problem 182 says that if $C = PQ - QP$ and if C is a scalar, then $C = 0$. How does the proof use the assumption that C is a scalar? An examination of Wielandt's proof suggests at least part of the answer: it is important that C commutes with P . Commutators with this sort of commutativity property have received some attention; the original question ($PQ - QP = 1$?) fits into the context of their theory. An easy way for $PQ - QP$ to commute with P is for it to be equal to P . Example:

$$P = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad Q = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

If that happens, then an easy inductive argument proves that $P^n Q - Q P^n = n P^n$, and this implies that

$$n \|P^n\| \leq 2 \|P^n\| \cdot \|Q\|$$

for every positive integer n . Since it is impossible that $n \leq 2 \|Q\|$ for all n , it follows that $P^n = 0$ for some n , i.e., that $P (= PQ - QP)$ is nilpotent.

The first general theorem of this sort is due to Jacobson [1935], who proved, under suitable finiteness assumptions, that if $C = PQ - QP$ and C commutes with P , then C is nilpotent. This is a not unreasonable generalization of the theorem about scalars; after all the only nilpotent scalar is 0. In infinite-dimensional Hilbert spaces finiteness conditions are not likely to be satisfied. Kaplansky conjectured that if nilpotence is replaced by its appropriate generalization, quas-nilpotence, then the Jacobson theorem will extend to operators, and he turned out to be

right. The proof was discovered, independently, by Kleinecke [1957] and Shirokov [1956].

Problem 184. *If P and Q are operators, if $C = PQ - QP$, and if C commutes with P , then C is quasinilpotent.*

185. Distance from a commutator to the identity. By Wintner and Wielandt, commutators cannot be equal to 1; by Brown-Pearcy, commutators can come arbitrarily near to 1. Usually, however, a commutator is anxious to stay far from 1.

Problem 185. (a) *If $C = PQ - QP$ and if P is hyponormal (hence, in particular, if P is an isometry, or if P is normal), then $\|1 - C\| \geq 1$.* (b) *If C commutes with P , then $\|1 - C\| \geq 1$.*

If the underlying Hilbert space is finite-dimensional, then it is an easy exercise in linear algebra to prove that $\|1 - C\| \geq 1$ for all commutators C .

186. Operators with large kernels. As far as the construction of commutators is concerned, all the results of the preceding problems are negative; they all say that something is not a commutator.

To get a positive result, suppose that \mathbf{H} is an infinite-dimensional Hilbert space and consider the infinite direct sum $\mathbf{H} \oplus \mathbf{H} \oplus \mathbf{H} \oplus \cdots$. Operators on this large space can be represented as infinite matrices whose entries are operators on \mathbf{H} . If, in particular, A is an arbitrary operator on \mathbf{H} (it could even be the identity), then the matrix

$$P = \begin{pmatrix} 0 & A & 0 & 0 & & \\ 0 & 0 & A & 0 & & \\ 0 & 0 & 0 & A & & \\ 0 & 0 & 0 & 0 & & \\ & & & & \ddots & \\ & & & & & \ddots & \\ & & & & & & \ddots \end{pmatrix}$$

defines an operator; if

$$Q = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ & & & \ddots \\ & & & & \ddots \end{pmatrix},$$

then it can be painlessly verified that

$$PQ - QP = \begin{pmatrix} A & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ & & & \ddots \\ & & & & \ddots \end{pmatrix}.$$

Since the direct sum of infinitely many copies of \mathbf{H} is the direct sum of the first copy and the others, and since the direct sum of the others is isomorphic (unitarily equivalent) to \mathbf{H} , it follows that every two-by-two operator matrix of the form

$$\begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix}$$

is a commutator (Halmos [1952 b], [1954]).

It is worth while reformulating the result without matrices. Call a subspace \mathbf{M} of a Hilbert space \mathbf{H} *large* if $\dim \mathbf{M} = \dim \mathbf{H}$. (The idea has appeared before, even if the word has not; cf. Problem 111.) In this language, if \mathbf{H} is infinite-dimensional, then \mathbf{H} (regarded as one of the

axes of the direct sum $\mathbf{H} \oplus \mathbf{H}$) is a large subspace of $\mathbf{H} \oplus \mathbf{H}$. If the matrix of an operator on $\mathbf{H} \oplus \mathbf{H}$ is

$$\begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix},$$

then that operator has a large kernel, and, moreover, that kernel reduces A . If, conversely, an operator on an infinite-dimensional Hilbert space has a large reducing kernel, then that operator can be represented by a matrix of the form

$$\begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix}.$$

(Represent the space as the direct sum of the kernel and its orthogonal complement. If the dimension of that orthogonal complement is too small, enlarge it by adjoining “half” the kernel.) In view of these remarks the matrix result of the preceding paragraph can be formulated as follows: every operator with a large reducing kernel is a commutator. This result can be improved (Percy [1965]).

Problem 186. *Every operator with a large kernel is a commutator.*

Corollary 1. *On an infinite-dimensional Hilbert space commutators are strongly dense.*

Corollary 2. *Every operator on an infinite-dimensional Hilbert space is the sum of two commutators.*

Corollary 2 shows that nothing like a trace can exist on the algebra of all operators on an infinite-dimensional Hilbert space. The reason is that a linear functional that deserves the name “trace” must vanish on all commutators, and hence, by Corollary 2, identically.

187. Direct sums as commutators.

Problem 187. *If an operator A on a separable Hilbert space is not a scalar, then the infinite direct sum $A \oplus A \oplus A \oplus \dots$ is a commutator.*

Even though this result is far from a complete characterization of commutators, it answers many of the obvious questions about them. Thus, for instance, it is an immediate corollary that the spectrum of a commutator is quite arbitrary; more precisely, each non-empty compact subset of the plane (i.e., any set that can be a spectrum at all) is the spectrum of some commutator. Another immediate corollary is that the identity is the limit (in the norm) of commutators; compare Problems 183 and 185.

The techniques needed for the proof contain the germ (a very rudimentary germ, to be sure) of what is needed for the general characterization of commutators (Brown-Pearcy [1965]).

188. Positive self-commutators. The *self-commutator* of an operator A is the operator $A^*A - AA^*$. The theory of self-commutators has some interest. It is known that a finite square matrix is a self-commutator if and only if it is Hermitian and has trace 0 (Thompson [1958]). An obvious place where self-commutators could enter is in the theory of hyponormal operators; a necessary and sufficient condition that A be hyponormal is that the self-commutator of A be positive. That self-commutators can be non-trivially positive is a relatively rare phenomenon (which, by the way, is strictly infinite-dimensional). It is natural to ask just how positive a self-commutator can be, and the answer is not very.

Problem 188. *A positive self-commutator cannot be invertible.*

Reference: Putnam [1951 a].

189. Projections as self-commutators. If a self-commutator $C = A^*A - AA^*$ is positive, then, by Problem 188, C is not invertible. The easiest way for C to be not invertible is to have a non-trivial kernel. Among the positive operators with non-trivial kernels, the most familiar ones are the projections. Can C be a projection, and, if so, how?

The most obvious way for C to be a projection is for A to be normal; in that case $C = 0$. Whatever other ways there might be, they can always be combined with a normal operator (direct sum) to yield still another way, which, however, is only trivially different. The interesting question here concerns what may be called *abnormal* operators, i.e., operators that have no normal direct summands. Otherwise said, A is abnormal if no non-zero subspace of the kernel of $A^*A - AA^*$ reduces A .

It is not difficult to produce an example of an abnormal operator whose self-commutator is a projection: any non-normal isometry will do. If A is a non-normal isometry (i.e., the direct sum of a unilateral shift of non-zero multiplicity and a unitary operator—see Problem 118), then $\|A\| = 1$ and $C = A^*A - AA^* = 1 - AA^*$ is the projection onto the kernel of A^* . What is interesting is that in the presence of the norm condition ($\|A\| = 1$) this is the only way to produce examples.

Problem 189. (a) *If A is an abnormal operator of norm 1, such that $A^*A - AA^*$ is a projection, then A is an isometry.* (b) *Does the statement remain true if the norm condition is not assumed?*

190. Multiplicative commutators. The word “commutator” occurs in two distinct mathematical contexts. In ring theory it means $PQ - QP$ (additive commutators); in group theory it means $PQP^{-1}Q^{-1}$ (multiplicative commutators). A little judicious guessing about trace versus determinant, and, more generally, about logarithm versus exponential, is likely to lead to the formulation of multiplicative analogues of the results about additive commutators. Some of those analogues are true. What about the analogue of the additive theorem according to which the only scalar that is an additive commutator is 0?

Problem 190. *If H is an infinite-dimensional Hilbert space, then a necessary and sufficient condition that a scalar α acting on H be a multiplicative commutator is that $|\alpha| = 1$.*

For finite-dimensional spaces determinants can be brought into play. The determinant of a multiplicative commutator is 1, and the only scalars whose determinants are 1 are the roots of unity of order equal to the dimension of the space. This proves that on an n -dimensional

space a necessary condition for a scalar α to be a multiplicative commutator is that $\alpha^n = 1$; a modification of the argument that works for infinite-dimensional spaces shows that the condition is sufficient as well.

It turns out that the necessity proof is algebraic, just as in the additive theory, in the sense that it yields the same necessary condition for an arbitrary complex normed algebra with unit. From this, in turn, it follows, just as in the additive theory, that if a commutator is congruent to a scalar modulo the ideal of compact operators, then that scalar must have modulus 1.

191. Unitary multiplicative commutators. The positive assertion of Problem 190 can be greatly strengthened. One of the biggest steps toward the strengthened theory is the following assertion.

Problem 191. *On an infinite-dimensional Hilbert space every unitary operator is a multiplicative commutator.*

192. Commutator subgroup. The *commutator subgroup* of a group is the smallest subgroup that contains all elements of the form $PQP^{-1}Q^{-1}$; in other words it is the subgroup generated by all commutators (multiplicative ones, of course). The set of all invertible operators on a Hilbert space is a multiplicative group; in analogy with standard finite-dimensional terminology, it may be called the *full linear group* of the space.

Problem 192. *What is the commutator subgroup of the full linear group of an infinite-dimensional Hilbert space?*

Chapter 20. Toeplitz operators

193. Laurent operators and matrices. Multiplications are the prototypes of normal operators, and most of the obvious questions about them (e.g., those about numerical range, norm, and spectrum) have obvious answers. (This is not to say that every question about them has been answered.) Multiplications are, moreover, not too sensitive to a change of space; aside from the slightly fussy combinatorics of atoms, and aside from the pathology of the uncountable, what happens on the unit interval or the unit circle is typical of what can happen anywhere.

If φ is a bounded measurable function on the unit circle, then the multiplication induced by φ on L^2 (with respect to normalized Lebesgue measure μ) is sometimes called the *Laurent operator* induced by φ , in symbols L_φ . The matrix of L_φ with respect to the familiar standard orthonormal basis in L^2 ($e_n(z) = z^n, n = 0, \pm 1, \pm 2, \dots$) has a simple form, elegantly related to φ . To describe the relation, define a *Laurent matrix* as a (bilaterally) infinite matrix $\langle \lambda_{ij} \rangle$ such that

$$\lambda_{i+1,j+1} = \lambda_{ij}$$

for all i and j ($= 0, \pm 1, \pm 2, \dots$). In words: a Laurent matrix is one all of whose diagonals (parallel to the main diagonal) are constants.

Problem 193. *A necessary and sufficient condition that an operator on L^2 be a Laurent operator L_φ is that its matrix $\langle \lambda_{ij} \rangle$ with respect to the basis $\{e_n: n = 0, \pm 1, \pm 2, \dots\}$ be a Laurent matrix; if that condition is satisfied, then $\lambda_{ij} = \alpha_{i-j}$, where $\varphi = \sum_n \alpha_n e_n$ is the Fourier expansion of φ .*

194. Toeplitz operators and matrices. Laurent operators (multiplications) are distinguished operators on L^2 (of the unit circle), and H^2 is a distinguished subspace of L^2 ; something interesting is bound to happen if Laurent operators are compressed to H^2 . The description of what happens is called the theory of Toeplitz operators. Explicitly: if P is the

projection from L^2 onto H^2 , and if φ is a bounded measurable function, then the *Toeplitz operator* T_φ induced by φ is defined by

$$T_\varphi f = P(\varphi \cdot f)$$

for all f in H^2 . The simplest non-trivial example of a Laurent operator is the bilateral shift $W (= L_{e_1})$; correspondingly, the simplest non-trivial example of a Toeplitz operator is the unilateral shift $U (= T_{e_1})$.

There is a natural basis in L^2 ; the matrix of a Laurent operator with respect to that basis has an especially simple form. The corresponding statements are true about H^2 and Toeplitz operators. To state them, define a *Toeplitz matrix* as a (unilaterally) infinite matrix $\langle \lambda_{ij} \rangle$ such that

$$\lambda_{i+1, j+1} = \lambda_{ij}$$

for all i and j ($= 0, 1, 2, \dots$). In words: a Toeplitz matrix is one all of whose diagonals (parallel to the main diagonal) are constants. The structural differences between the Laurent theory and the Toeplitz theory are profound, but the difference between the two kinds of matrices is superficial and easy to describe; for Laurent matrices both indices go both ways from 0, but for Toeplitz matrices they go forward only.

Problem 194. *A necessary and sufficient condition that an operator on H^2 be a Toeplitz operator T_φ is that its matrix $\langle \lambda_{ij} \rangle$ with respect to the basis $\{e_n: n = 0, 1, 2, \dots\}$ be a Toeplitz matrix; if that condition is satisfied, then $\lambda_{ij} = \alpha_{i-j}$, where $\varphi = \sum_n \alpha_n e_n$ is the Fourier expansion of φ .*

The necessity of the condition should not be surprising: in terms of an undefined but self-explanatory phrase, it is just that the compressed operator has the compressed matrix.

The unilateral shift U does for Toeplitz operators what the bilateral shift W does for Laurent operators—but does it differently.

Corollary 1. *A necessary and sufficient condition that an operator A on H^2 be a Toeplitz operator is that $U^*AU = A$.*

Since W is unitary, there is no difference between $W^*AW = A$ and $AW = WA$. The corresponding equations for U say quite different things. The first, $U^*AU = A$, characterizes Toeplitz operators. The second, $AU = UA$, characterizes analytic Toeplitz operators (see Problem 116). The Toeplitz operator T_φ induced by φ is called *analytic* in case φ is analytic (see Problem 26), i.e., in case φ is not only in L^∞ but in H^∞ . (To justify the definition, note that the statement of Problem 194 implies that the correspondence $\varphi \rightarrow T_\varphi$ is one-to-one.) Observe that an analytic Toeplitz operator is subnormal; it is not only a compression but a restriction of the corresponding Laurent operator.

Corollary 2. *The only compact Toeplitz operator is 0.*

195. Toeplitz products. The algebraic structure of the set of all Laurent operators holds no surprises: everything is true and everything is easy. The mapping $\varphi \rightarrow L_\varphi$ from bounded measurable functions to operators is an algebraic homomorphism (it preserves unit, linear operations, multiplication, and conjugation), and an isometry (supremum norm to operator norm); the spectrum of L_φ is the essential range of φ . Since the Laurent operators constitute the commutant of W (Problem 115), and since the product $W^{-1}AW$ is weakly continuous in its middle factor, it follows that the set of all Laurent operators is weakly (and hence strongly) closed.

Some of the corresponding Toeplitz statements are true and easy, but some are hard, or false, or unknown. The easiest statements concern unit, linear operations, and conjugation: since both the mappings $\varphi \rightarrow L_\varphi$ and $L_\varphi \rightarrow (PL_\varphi)|_{H^2}$ ($=$ the restriction of PL_φ to $H^2 = T_\varphi$) preserve the algebraic structures named, the same is true of their composite, which is $\varphi \rightarrow T_\varphi$. (The preservation of adjunction is true for compressions in general; see Problem 178.) The argument that proved that the set of all Laurent operators is weakly closed works for Toeplitz operators too; just replace $W^{-1}AW$ by U^*AU (cf. Corollary 1 of Problem 194).

It is a trivial consequence of the preceding paragraph that a Toeplitz operator T_φ is Hermitian if and only if φ is real; indeed $T_\varphi = T_\varphi^*$ if and only if $\varphi = \varphi^*$. It is also true that T_φ is positive if and only if φ is positive. Indeed, since $(T_\varphi f, f) = (L_\varphi f, f)$ whenever $f \in H^2$, it follows that T_φ is positive if and only if $(L_\varphi f, f) \geq 0$ for all f in H^2 . The latter condition

is equivalent to this one: $(W^n L_\varphi f, W^n f) \geq 0$ whenever $f \in \mathbf{H}^2$ (and n is an arbitrary integer). Since W commutes with L_φ , the condition can also be expressed in this form: $(L_\varphi W^n f, W^n f) \geq 0$ whenever $f \in \mathbf{H}^2$. Since the set of all $W^n f$'s, with f in \mathbf{H}^2 , is dense in \mathbf{L}^2 , the condition is equivalent to $L_\varphi \geq 0$, and hence to $\varphi \geq 0$.

The easiest statements about the multiplicative properties of Toeplitz operators are negative: the set of all Toeplitz operators is certainly not commutative and certainly not closed under multiplication. A counterexample for both assertions is given by the unilateral shift and its adjoint. Both U and U^* are Toeplitz operators, but the product U^*U (which is equal to the Toeplitz operator 1) is not the same as the product UU^* (which is not a Toeplitz operator). One way to prove that UU^* is not a Toeplitz operator is to use Corollary 1 of Problem 194: since $U^*(UU^*)U = (U^*U)(U^*U) = 1 (\neq UU^*)$, everything is settled. Alternatively, this negative result could have been obtained via Problem 194 by a direct look at the matrix of UU^* .

When is the product of two Toeplitz operators a Toeplitz operator? The answer is: rarely. Reference: Brown-Halmos [1963].

Problem 195. *A necessary and sufficient condition that the product $T_\varphi T_\psi$ of two Toeplitz operators be a Toeplitz operator is that either φ^* or ψ be analytic; if the condition is satisfied, then $T_\varphi T_\psi = T_{\varphi\psi}$.*

The Toeplitz operator T_φ induced by φ is called *co-analytic* in case φ is co-analytic (see Problem 26). In this language, Problem 195 says that the product of two Toeplitz operators is a Toeplitz operator if and only if the first factor is co-analytic or the second one is analytic.

Corollary. *A necessary and sufficient condition that the product of two Toeplitz operators be zero is that at least one factor be zero.*

Concisely: among the Toeplitz operators there are no zero-divisors.

196. Spectral inclusion theorem for Toeplitz operators. Questions about the norms and the spectra of Toeplitz operators are considerably more difficult than those for Laurent operators. As for the norm of T_φ , for instance, all that is obvious at first glance is that $\|T_\varphi\| \leq \|L_\varphi\|$

($= \|\varphi\|_\infty$); that much is obvious because T is a compression of L . About the spectrum of T nothing is obvious, but there is a relatively easy inequality (due to Hartman and Wintner [1950]) that answers some of the natural questions.

Problem 196. *If L and T are the Laurent and the Toeplitz operators induced by a bounded measurable function, then $\Pi(L) \subset \Pi(T)$.*

This is a spectral inclusion theorem, formally similar to Problem 157; here, too, the “larger” operator has the smaller spectrum. The result raises a hope that it is necessary to nip in the bud. If T_φ is bounded from below, so that $0 \notin \Pi(T_\varphi)$, then, by Problem 196, $0 \notin \Pi(L_\varphi)$. This is equivalent to L_φ being bounded from below and hence to φ being bounded away from 0. If the converse were true, then the spectral structure of T_φ would be much more easily predictable from φ than in fact it is; unfortunately the converse is false. If, indeed, $\varphi = e_{-1}$, then φ is bounded away from 0, but $T_\varphi e_0 = P e_{-1} = 0$, so that T_φ has a non-trivial kernel.

Although the spectral behavior of Toeplitz operators is relatively bad, Problem 196 can be used to show that in some respects Toeplitz operators behave as if they were normal. Here are some samples.

Corollary 1. *If φ is a bounded measurable function, then $r(T_\varphi) = \|T_\varphi\| = \|\varphi\|_\infty$.*

Corollary 1 says, among other things, that the correspondence $\varphi \rightarrow T_\varphi$ is norm-preserving; this recaptures the result (cf. Problem 194) that that correspondence is one-to-one.

Corollary 2. *There are no quasinilpotent Toeplitz operators other than 0.*

Corollary 3. *Every Toeplitz operator with a real spectrum is Hermitian.*

Corollary 4. *The closure of the numerical range of a Toeplitz operator is the convex hull of its spectrum.*

197. Analytic Toeplitz operators. The easiest Toeplitz operators are the analytic ones, but even for them much more care is needed than for multiplications. The operative word is “analytic”. Recall that associated with each φ in \mathbf{H}^∞ there is a function $\tilde{\varphi}$ analytic in the open unit disc D (see Problem 28). The spectral behavior of T_φ is influenced by the complex analytic behavior of $\tilde{\varphi}$ rather than by the merely set-theoretic behavior of φ . Reference: Wintner [1929].

Problem 197. *If $\varphi \in \mathbf{H}^\infty$, then the spectrum of T_φ is the closure of the image of the open unit disc D under the corresponding element $\tilde{\varphi}$ of $\tilde{\mathbf{H}}^\infty$; in other words $\Lambda(T_\varphi) = \overline{\tilde{\varphi}(D)}$.*

Here is still another way to express the result. If $\varphi \in \mathbf{L}^\infty$, then the spectrum of L_φ is the essential range of φ ; if $\varphi \in \mathbf{H}^\infty$, then the spectrum of T_φ is what may be called the essential range of $\tilde{\varphi}$.

198. Eigenvalues of Hermitian Toeplitz operators. Can an analytic Toeplitz operator have an eigenvalue? Except in the trivial case of scalar operators, the answer is no. The reason is that if φ is analytic and $\varphi \cdot f = \lambda f$ for some f in \mathbf{H}^2 , then the F. and M. Riesz theorem (Problem 127) implies that either $\varphi = \lambda$ or $f = 0$. Roughly speaking, the reason is that an analytic function cannot take a constant value on a set of positive measure without being a constant. For Hermitian Toeplitz operators this reasoning does not apply: there is nothing to stop a non-constant real-valued function from being constant on a set of positive measure.

Problem 198. *Given a real-valued function φ in \mathbf{L}^∞ , determine the point spectrum of the Hermitian Toeplitz operator T_φ .*

199. Spectrum of a Hermitian Toeplitz operator.

Problem 199. *Given a real-valued function φ in \mathbf{L}^∞ , determine the spectrum of the Hermitian Toeplitz operator T_φ .*

For more recent and more general studies of the spectra of Toeplitz operators, see Widom [1960], [1964].

Hints

Chapter 1. Vectors and spaces

Problem 1. Polarize.

Problem 2. Use uniqueness: if $f = \sum_j \alpha_j e_j$, then $\xi(f) = \sum_j \alpha_j \xi(e_j)$.

Problem 3. Use inner products to reduce the problem to the strict convexity of the unit disc.

Problem 4. Consider characteristic functions in $L^2(0,1)$. An alternative hint, for those who know about spectral measures, is to contemplate spectral measures.

Problem 5. A countably infinite set has an uncountable collection of infinite subsets such that the intersection of two distinct ones among them is always finite.

Problem 6. Determine the orthogonal complement of the span.

Problem 7. If $f_0 \perp f_i$ for all i and if $\sum_{i>n} \|e_i - f_i\|^2 < 1$, then f_0, f_1, \dots, f_n are linearly dependent.

Problem 8. Prove that $\mathbf{M} + \mathbf{N}$ is complete. There is no loss of generality in assuming that $\dim \mathbf{M} = 1$.

Problem 9. In an infinite-dimensional space there always exist two subspaces whose vector sum is different from their span.

Problem 10. How many basis elements can an open ball of diameter $\sqrt{2}$ contain?

Problem 11. Given a countable basis, use rational coefficients. Given a countable dense set, approximate each element of a basis close enough to exclude all other basis elements.

Problem 12. Fit infinitely many balls of the same radius inside any given ball of positive radius.

Chapter 2. Weak topology

Problem 13. Consider orthonormal sets. Caution: is weak closure the same as weak sequential closure?

Problem 14. Expand $\|f_n - f\|^2$.

Problem 15. The span of a weakly dense set is the whole space.

Problem 16. Assume that $|(f_n, g)| < \varepsilon$ for all unit vectors g , and replace g by $f_n/\|f_n\|$.

Problem 17. Consider the set of all complex-valued functions ξ on \mathbf{H} such that $|\xi(f)| \leq \|f\|$ for all f , endowed with the product topology, and show that the linear functionals of norm less than or equal to 1 form a closed subset.

Problem 18. Given a countable dense set, define all possible basic weak neighborhoods of each of its elements, using finite subsets of itself for the vector parameters and reciprocals of positive integers for the numerical parameters of the neighborhoods; show that the resulting collection of neighborhoods is a base for the weak topology. Alternatively, given an orthonormal basis $\{e_1, e_2, e_3, \dots\}$, define a metric by

$$d(f, g) = \sum_j \frac{1}{2^j} |(f - g, e_j)|.$$

Problem 19. If the unit ball is weakly metrizable, then it is weakly separable.

Problem 20. If the conclusion is false, then construct, inductively, an orthonormal sequence such that the inner product of each term with a suitable element of the given weakly bounded set is very large; then form a suitable (infinite) linear combination of the terms of that orthonormal sequence.

Problem 21. Construct a sequence that has a weak cluster point but whose norms tend to ∞ .

Problem 22. Consider partial sums and use the principle of uniform boundedness.

Problem 23. (a) Given an unbounded linear functional ξ , use a Hamel basis to construct a Cauchy net $\{g_J\}$ such that $(f, g_J) \rightarrow \xi(f)$ for each f . (b) If $\{g_n\}$ is a weak Cauchy sequence, then $\xi(f) = \lim_n (f, g_n)$ defines a bounded linear functional.

Chapter 3. Analytic functions

Problem 24. The value of an analytic function at the center of a disc is equal to its average over the disc. This implies that evaluation at a point of D is a bounded linear functional on $A^2(D)$, and hence that Cauchy sequences in the norm are Cauchy sequences in the sense of uniform convergence on compact sets.

Problem 25. What is the connection between the concepts of convergence appropriate to power series and Fourier series?

Problem 26. Is conjugation continuous?

Problem 27. Is the Fourier series of a product the same as the formal product of the Fourier series?

Problem 28. A necessary and sufficient condition that

$$\sum_{n=0}^{\infty} |\alpha_n|^2 < \infty$$

is that the numbers

$$\sum_{n=0}^{\infty} |\alpha_n| r^{2n} \quad (0 < r < 1)$$

be bounded. Use continuity of the partial sums at $r = 1$.

Problem 29. Start with a well-behaved functional Hilbert space and adjoin a point to its domain.

Problem 30. To evaluate the Bergman and the Szegő kernels, use the general expression of a kernel function as a series.

Problem 31. Use the kernel function of $\tilde{\mathbf{H}}^2$.

Problem 32. Approximate f , in the norm, by the values of \tilde{f} on expanding concentric circles.

Problem 33. Use the maximum modulus principle and Fejér's theorem about the Cesaro convergence of Fourier series.

Problem 34. Assume that one factor is bounded and use Problem 27.

Problem 35. Given the Fourier expansion of an element of \mathbf{H}^2 , first find the Fourier expansion of its real part, and then try to invert the process.

Chapter 4. Infinite matrices

Problem 36. Treat the case of dimension \aleph_0 only. Construct the desired orthonormal set inductively; ensure that it is a basis by choosing every other element of it so that the span of that element and its predecessors includes the successive terms of a prescribed basis.

Problem 37. Write $\sum_j \alpha_{ij} \xi_j$ as

$$\sum_j (\sqrt{\alpha_{ij}} \sqrt{p_j}) \left(\frac{\sqrt{\alpha_{ij}} \xi_j}{\sqrt{p_j}} \right)$$

and apply the Schwarz inequality.

Problem 38. Apply Problem 37 with $p_i = 1/\sqrt{i + \frac{1}{2}}$.

Chapter 5. Boundedness and invertibility

Problem 39. To get the unbounded examples, extend an orthonormal basis to a Hamel basis; for the bounded ones use a matrix with a large but finite first row.

Problem 40. Apply the principle of uniform boundedness for linear functionals twice.

Problem 41. Prove that A^* is bounded from below by proving that the inverse image under A^* of the unit sphere in \mathbf{H} is bounded.

Problem 42. Write the given linear transformation as a matrix (with respect to orthonormal bases of \mathbf{H} and \mathbf{K}); if $\aleph_0 \leq \dim \mathbf{K} < \dim \mathbf{H}$, then there must be a row consisting of nothing but 0's.

Problem 43. Use Problem 42.

Problem 44. Apply Problem 41 to the mapping that projects the graph onto the domain.

Problem 45. For a counterexample, look at unbounded diagonal matrices. For a proof, apply the closed graph theorem.

Chapter 6. Multiplication operators

Problem 46. $|\alpha_j| = \|Ae_j\|$ and

$$\sum_j |\alpha_j \xi_j|^2 \leq (\sup_j |\alpha_j|)^2 \cdot \sum_j |\xi_j|^2.$$

Problem 47. If $|\alpha_n| \geq n$, then the sequence $\{1/\alpha_n\}$ belongs to l^2 .

Problem 48. The inverse operator must send e_n onto $(1/\alpha_n)e_n$.

Problem 49. If $\varepsilon > 0$ and if f is the characteristic function of a set of positive finite measure on which $|\varphi(x)| > \|\varphi\|_\infty - \varepsilon$, then $\|Af\| \geq (\|\varphi\|_\infty - \varepsilon) \cdot \|f\|$.

Problem 50. If $\|A\| = 1$, then $\|\varphi^n \cdot f\| \leq \|f\|$ for every positive integer n and for every f in L^2 ; this implies that $|\varphi(x)| \leq 1$ whenever $f(x) \neq 0$.

Problem 51. Imitate the discrete case (Solution 47), or prove that a multiplication is necessarily closed and apply the closed graph theorem.

Problem 52. Imitate the discrete case (Solution 48).

Problem 53. For the boundedness of the multiplication, use the closed graph theorem. For the boundedness of the multiplier, assume that if $x \in X$, then there exists an f in \mathbf{H} such that $f(x) \neq 0$; imitate the “slick” proof in Solution 50.

Problem 54. Consider the set of all those absolutely continuous functions on $[0,1]$ whose derivatives belong to L^2 .

Chapter 7. Operator matrices

Problem 55. If $AD - BC$ is invertible, then the formal inverse of

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

can be formed, in analogy with two-by-two numerical matrices. If

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

is invertible, then $AD - BC$ is bounded from below; imitate the elementary process of solving two linear equations in two unknowns.

Problem 56. Multiply on the right by

$$\begin{pmatrix} 1 & 0 \\ T & 1 \end{pmatrix},$$

with T chosen so as to annihilate the lower left entry of the product. Look for counterexamples formed out of the operator on l^2 defined by

$$\langle \xi_0, \xi_1, \xi_2, \dots \rangle \rightarrow \langle 0, \xi_0, \xi_1, \xi_2, \dots \rangle,$$

and its adjoint.

Problem 57. If a finite-dimensional subspace is invariant under an invertible operator, then it is invariant under the inverse.

Chapter 8. Properties of spectra

Problem 58. The kernel of an operator is the orthogonal complement of the range of its adjoint.

Problem 59. To prove $\Pi_0(p(A)) \subset p(\Pi_0(A))$, given α in $\Pi_0(p(A))$, factor $p(\lambda) - \alpha$. Use the same technique for Π , and, for Γ , apply the result with A^* in place of A .

Problem 60. For Π : if $\|f_n\| = 1$, then the numbers $\|P^{-1}f_n\|$ are bounded from below by $1/\|P\|$. For Γ : the range of $P^{-1}AP$ is included in the image under P^{-1} of the range of A .

Problem 61. Pretend that it is legitimate to expand $(1 - AB)^{-1}$ into a geometric series.

Problem 62. Prove that the complement is open.

Problem 63. Suppose that $\lambda_n \notin \Lambda(A)$, $\lambda \in \Lambda(A)$, and $\lambda_n \rightarrow \lambda$. If $f \neq 0$ and $f \perp \text{ran}(A - \lambda)$, then

$$\frac{(A - \lambda)(A - \lambda_n)^{-1}f}{\|(A - \lambda_n)^{-1}f\|} \rightarrow 0.$$

Chapter 9. Examples of spectra

Problem 64. If A is normal, then $\Pi_0(A) = (\Pi_0(A^*))^*$.

Problem 65. Use Problem 64.

Problem 66. If $\varphi \cdot f = \lambda f$ almost everywhere, then $\varphi = \lambda$ whenever $f \neq 0$.

Problem 67. Verify that $U^* \langle \xi_0, \xi_1, \xi_2, \dots \rangle = \langle \xi_1, \xi_2, \xi_3, \dots \rangle$. Compute that $\Pi_0(U)$ is empty and $\Pi_0(U^*)$ is the open unit disc. If $|\lambda| < 1$, then $U - \lambda$ is bounded from below.

Problem 68. Represent W as a multiplication.

Problem 69. Use a spanning set of eigenvectors of A^* for the domain; for each f in that domain, define the multiplier as the conjugate of the corresponding eigenvalue.

Problem 70. For an operator with a trivial kernel, relative invertibility is the same as left invertibility; for all operators, left invertibility is the same as boundedness from below.

Problem 71. Consider the operator matrix

$$\begin{pmatrix} 1 & 0 \\ 1 & U \end{pmatrix},$$

where U is the unilateral shift.

Chapter 10. Spectral radius

Problem 72. If λ_0 is not in the spectrum of A and if $|\lambda - \lambda_0|$ is sufficiently small, then

$$\rho_A(\lambda) = (A - \lambda_0)^{-1} \sum_{n=0}^{\infty} ((A - \lambda_0)^{-1}(\lambda - \lambda_0))^n.$$

Problem 73. Apply Liouville's theorem on bounded entire functions to the resolvent.

Problem 74. Write

$$\tau(\lambda) = \left(A - \frac{1}{\lambda} \right)^{-1}.$$

Use the analyticity of the resolvent to conclude that τ is analytic for $|\lambda| < 1/r(A)$, and then use the principle of uniform boundedness.

Problem 75. Look for a diagonal operator D such that $AD = DB$.

Problem 76. If $A = S^{-1}BS$, then the matrix of S must be lower triangular; find the matrix entries in row $n + 1$, column n , $n = 0, 1, 2, \dots$.

Problem 77. For the norm: S is an isometry, and therefore $\|S\| = \|SP\|$. For the spectral radius: use Problem 74.

Problem 78. Imitate the coordinate technique used for the unweighted unilateral shift.

Problem 79. If $f = \langle \xi_0, \xi_1, \xi_2, \dots \rangle \in \ell^2(p)$, write

$$Uf = \langle \sqrt{p_0}\xi_0, \sqrt{p_1}\xi_1, \sqrt{p_2}\xi_2, \dots \rangle,$$

and prove that U is an isometry from $\ell^2(p)$ onto ℓ^2 that transforms the shift on $\ell^2(p)$ onto a weighted shift on ℓ^2 .

Problem 80. Try unilateral weighted shifts; apply Solution 77.

Problem 81. Try unilateral weighted shifts with infinitely many zero weights; apply Solution 77.

Problem 82. Use induction to prove the inequality $|(ABf, f)|^{2^n} \leq (AB^{2^n}f, f) \cdot (Af, f)^{2^n-1}$.

Chapter 11. Norm topology

Problem 83. Think of projections on $L^2(0,1)$.

Problem 84. If A_0 is invertible, then $1 - AA_0^{-1} = (A_0 - A)A_0^{-1}$; use the geometric series trick to prove that A is invertible and to obtain a bound on $\|A^{-1}\|$.

Problem 85. Find the spectral radius of both A_k and A_k^{-1} .

Problem 86. The distance from $A - \lambda$ to the set of singular operators is positive on the complement of $\Lambda(A)$. Alternatively, the norm of the resolvent is bounded on the complement of Λ_0 ; the reciprocal of a bound is a suitable ϵ .

Problem 87. Approximate a weighted unilateral shift with positive spectral radius by weighted shifts with enough zero weights to make them nilpotent.

Chapter 12. Strong and weak topologies

Problem 88. Use the results and the methods of Problem 16.

Problem 89. For a counterexample with respect to the strong topology, consider the projections onto a decreasing sequence of subspaces.

Problem 90. For a counterexample with respect to the strong topology, consider the powers of the adjoint of the unilateral shift.

Problem 91. The set of all nilpotent operators of index 2 is strongly dense.

Problem 92. Use nets.

Problem 93. (a) Use the principle of uniform boundedness.
(b) Look at powers of the unilateral shift.

Problem 94. If $\{A_n\}$ is increasing and converges to A weakly, then the positive square root of $A - A_n$ converges to 0 strongly. For a counterexample with respect to the uniform topology, consider sequences of projections.

Problem 95. The B_n 's form a bounded increasing sequence.

Problem 96. Study the sequence of powers of EFE .

Chapter 13. Partial isometries

Problem 97. If N is a neighborhood of $F(\lambda)$, then $F^{-1}(N)$ is a neighborhood of λ . If $\lambda \notin F(\Lambda(A))$, then some neighborhood of λ is disjoint from $F(\Lambda(A))$.

Problem 98. If U is a partial isometry with initial space \mathbf{M} , evaluate (U^*Uf, f) when $f \in \mathbf{M}$ and when $f \perp \mathbf{M}$; if U^*U is a projection with range \mathbf{M} , do the same thing.

Problem 99. The only troublesome part is to find a co-isometry U and a non-reducing subspace \mathbf{M} such that $U\mathbf{M} = \mathbf{M}$; for this let U be the adjoint of the unilateral shift and let \mathbf{M} be the (one-dimensional) subspace of eigenvectors belonging to a non-zero eigenvalue.

Problem 100. For closure: A is a partial isometry if and only if $A = AA^*A$. For connectedness: if U is a partial isometry, if V is an isometry, and if $\|U - V\| < 1$, then U is an isometry.

Problem 101. For rank: the restriction of U to the initial space of V is one-to-one. For nullity: if $f \in \ker V$ and $f \perp \ker U$, then

$$\|Uf - Vf\| = \|f\|.$$

Problem 102. Find a unitary operator that matches up initial spaces, and another that matches up final spaces, and find continuous curves that join each of them to the identity.

Problem 103. If A and B are invertible contractions, and if a unitary operator transforms $M(A)$ onto $M(B)$, then it maps the subspace of all vectors of the form $\langle f, 0 \rangle$ onto itself.

Problem 104. If a compact subset Λ of the closed unit disc contains 0, find a contraction A with spectrum Λ , and extend A to a partial isometry.

Problem 105. Put $P^2 = A^*A$, and define U by $UPf = Af$ on $\text{ran } P$ and by $Uf = 0$ on $\ker P$.

Problem 106. Every partial isometry has a maximal enlargement.

Problem 107. To prove that maximal partial isometries are extreme points, use Problem 3. To prove the converse, show that every

contraction is the average of two maximal partial isometries; use Problem 106.

Problem 108. If UP commutes with P^2 , then it commutes with P , so that $UP - PU$ annihilates $\text{ran } P$.

Problem 109. For the positive result, apply Problem 106. For the negative one: a non-invertible operator that has a one-sided inverse cannot be the limit of invertible operators.

Problem 110. Consider polar decompositions UP and join both U and P to 1.

Chapter 14. Unilateral shift

Problem 111. Assume that \mathbf{H} is separable, and argue that it is enough to prove the existence of two orthogonal reducing subspaces of infinite dimension. Prove it by the consideration of spectral measures.

Problem 112. Apply Problem 111, and factor the given unitary operator into two operators, one of which shifts the resulting two-way sequence of subspaces forward and the other backward.

Problem 113. (a) If a normal operator has a one-sided inverse, then it is invertible. (b) Since 1 is an approximate eigenvalue of the unilateral shift, the same is true of the real part. (c) There is no invertible operator within 1 of the unilateral shift.

Problem 114. If $V^2 = U^*$, then $\dim \ker V \leq 1$ and V maps the underlying Hilbert space onto itself.

Problem 115. If W commutes with an operator A , and if ψ is a bounded measurable function on the circle, then, by the Fuglede commutativity theorem, $\psi(W)$ commutes with A . Put $Ae_0 = \varphi$, prove that $A\psi = \varphi \cdot \psi$, and use the technique of Solution 50.

Problem 116. Begin as for Solution 115; use Solution 50; imitate Solution 51.

Problem 117. Every function in H^∞ is the limit almost everywhere of a bounded sequence of polynomials; cf. Solution 33.

Problem 118. If V is an isometry on H , and if N is the orthogonal complement of the range of V , then $\bigcap_{n=0}^{\infty} V^n H = \bigcap_{n=0}^{\infty} (V^n N)^\perp$.

Problem 119. Use Problem 118, and recall that -1 belongs to the spectrum of the unilateral shift.

Problem 120. Consider the direct sum of the unilateral shift and infinitely many copies of the bilateral shift.

Problem 121. If $\|A\| \leq 1$ and $A^n \rightarrow 0$ strongly, write $T = \sqrt{1 - A^*A}$ and assign to each vector f the sequence

$$\langle Tf, T Af, T A^2 f, \dots \rangle.$$

Problem 122. If $r(A) < 1$, then $\sum_{n=0}^{\infty} \|A^n\|_{2Z^n}^2$ converges at $z = 1$, and consequently an equivalent norm is defined by $\|f\|_0^2 = \sum_{n=0}^{\infty} \|A^n f\|^2$.

Problem 123. Write $N = M \cap (UM)^\perp$ and apply the results of Solution 118. To prove $\dim N = 1$, assume the existence of two orthogonal unit vectors f and g in N and use Parseval's equation to compute $\|f\|^2 + \|g\|^2$. It is helpful to regard U as the restriction of the bilateral shift.

Problem 124. Prove that $M_k^\perp(\lambda)$ is invariant under U^* .

Problem 125. Use Problem 123 to express M in terms of a wandering subspace N , and examine the Fourier expansion of a unit vector in N .

Problem 126. For the simple shift, consider a vector $\langle \xi_0, \xi_1, \xi_2, \dots \rangle$ such that

$$\lim_k \frac{1}{|\xi_k|^2} \sum_{n=1}^{\infty} |\xi_{n+k}|^2 = 0.$$

For shifts of higher multiplicity, form vectors whose components are subsequences of this sequence $\{\xi_n\}$.

Problem 127. Given f in H^2 , let M be the least subspace of H^2 that contains f and is invariant under U , and apply Problem 125 to M .

Problem 128. Let g be a non-zero element of L^2 that vanishes on a set of positive measure, and write $f(z) = zg(z^3)$.

Problem 129. Necessity: consider a Hermitian operator that commutes with A (and hence with A^* and with A^*A), and examine its matrix. Sufficiency: assume $\{\alpha_n\}$ periodic of period p ; let M_j be the span of the e_j 's with $n \equiv j \pmod{p}$; observe that each vector has a unique representation in the form $f_0 + \cdots + f_p$, with f_j in M_j ; for each measurable subset E of the circle, consider the set of all those f 's for which $f_j(z) = 0$ for all j and for all z in the complement of E .

Chapter 15. Compact operators

Problem 130. Use nets. In the discussion of $(w \rightarrow s)$ continuity recall that a basic weak neighborhood depends on a finite set of vectors, and consider the orthogonal complement of their span.

Problem 131. To prove self-adjointness, use the polar decomposition.

Problem 132. Approximate by diagonal operators of finite rank.

Problem 133. If the restriction of a compact operator to an invariant subspace is invertible, then the subspace is finite-dimensional. Infer, via the spectral theorem, that the part of the spectrum of a normal compact operator that lies outside a closed disc with center at the origin consists of a finite number of eigenvalues with finite multiplicities.

Problem 134. If the identity has kernel K , then

$$\mu(F \cap G) = \iint_{F \times G} K(x, y) d\mu(x) d\mu(y)$$

whenever F and G are measurable sets; it follows that the indefinite integral of K is concentrated on the diagonal.

Problem 135. Approximate by simple functions.

Problem 136. If A is a Hilbert-Schmidt operator, then the sum of the eigenvalues of A^*A is finite.

Problem 137. Use the polar decomposition and Problem 133.

Problem 138. Every operator of rank 1 belongs to every non-zero ideal. Every non-compact Hermitian operator is bounded from below on some infinite-dimensional invariant subspace; its restriction to such a subspace is invertible.

Problem 139. Use the spectral theorem.

Problem 140. (1) If $\text{ran}(1 - C) = \mathbf{H}$, then $\ker(1 - C) = \{0\}$. (The sequence $\{\ker A, \ker A^2, \ker A^3, \dots\}$ is strictly increasing.) (2) $1 - C$ is bounded from below on $(\ker(1 - C))^\perp$. After proving (1) and (2), apply them not only to C but also to C^* .

Problem 141. If \mathbf{M} is a subspace included in $\text{ran } A$, the restriction of A to the inverse image of \mathbf{M} is invertible.

Problem 142. From (1) to (2): the restriction of A to $(\ker A)^\perp$ is invertible. From (3) to (1): if $1 - BA$ is compact, apply Solution 140 to $1 - BA$.

Problem 143. Assume $\lambda = 0$; note that if B is invertible, then $A = B(1 + B^{-1}(A - B))$.

Problem 144. Perturb the bilateral shift by an operator of rank 1.

Problem 145. If C is compact and $U + C$ is normal, then the spectrum of $U + C$ is large; but the spectrum of $(U + C)^*(U + C)$ is small.

Problem 146. If A is a Volterra operator with kernel bounded by c , then A^n is a Volterra operator with kernel bounded by $c^n/(n-1)!$.

Problem 147. Prove first that if A is a Volterra operator, and if ϵ is a positive number, then there exist Volterra operators B and C and there exists a positive integer k such that (1) $A = B + C$, (2) $\|B\| < \epsilon$, and (3) every product of B 's and C 's in which k or more factors are equal to C is equal to 0. To get the kernel of B , redefine the kernel of A to be 0 on a thin strip parallel to the diagonal.

Problem 148. Express V^*V as an integral operator. By differentiation convert the equation $V^*Vf = \lambda f$ into a differential equation, and solve it.

Problem 149. Identify $L^2(-1, +1)$ with $L^2(0, 1) \oplus L^2(0, 1)$, and determine the two-by-two operator matrix corresponding to such an identification. Caution: there is more than one interesting way of making the identification.

Problem 150. Put $A = (1 + V)^{-1}$, where V is the Volterra integration operator.

Problem 151. Reduce to the case where \mathbf{M} contains a vector f with infinitely many non-zero Fourier coefficients; in that case prove that there exist scalars λ_n such that $\lambda_n A^n f \rightarrow e_0$, so that \mathbf{M} contains e_0 ; use induction to conclude that \mathbf{M} contains e_k for every positive integer k .

Chapter 16. Subnormal operators

Problem 152. Apply Fuglede's theorem to two-by-two operator matrices made out of A_1 , A_2 , and B .

Problem 153. If $\|A^n f\| \leq \|f\|$ for all n , and if

$$M_r = \{x: |\varphi(x)| \geq r > 1\},$$

then $\|f\|^2 \geq \int_M r^{2n} |f|^2 d\mu$.

Problem 154. Show that $\ker A$ reduces A and throw it away. Once $\ker A = \{0\}$, consider the polar decomposition of A , extend the isometric factor to a unitary two-by-two matrix, extend the positive factor to a positive two-by-two matrix, and do all this so that the two extensions commute.

Problem 155. The desired isometry U must be such that if $\{f_1, \dots, f_n\}$ is a finite subset of \mathbf{H} , then $U(\sum_j B_1^* f_j) = \sum_j B_2^* f_j$.

Problem 156. Consider the measure space consisting of the unit circle together with its center, with measure defined so as to be normalized Lebesgue measure in the circle and a unit mass at the center. Form a subnormal operator by restricting a suitable multiplication on L^2 to the closure of the set of all polynomials.

Problem 157. It is sufficient to prove that if A is invertible, then so is B . Use Problem 153.

Problem 158. Both $\Delta - \Lambda(A)$ and $\Delta \cap \Lambda(A)$ are open. Use Problem 63.

Problem 159. Every finite-dimensional subspace invariant under a normal operator B reduces B .

Problem 160. If A (on \mathbf{H}) is subnormal, and if f_0, \dots, f_n are vectors in \mathbf{H} , then the matrix $\langle \langle A^i f_j, A^j f_i \rangle \rangle$ is positive definite. A weighted shift with weights $\{\alpha_0, \alpha_1, \alpha_2, \dots\}$ is hyponormal if and only if $|\alpha_n|^2 \leq |\alpha_{n+1}|^2$ for all n .

Problem 161. Use Problem 118.

Problem 162. If A is hyponormal, then

$$\|A^n f\|^2 \leq \|A^{n+1}\| \cdot \|A^{n-1}\| \cdot \|f\|^2$$

for every vector f .

Problem 163. If A is hyponormal, then the span of the eigenvectors of A reduces A . If A is compact also, then consider the restriction of A to the orthogonal complement of that span, and apply Problem 140 and Problem 162.

Problem 164. Let \mathbf{H} be the (infinite, bilateral) direct sum of copies of a two-dimensional Hilbert space, and consider an operator-weighted shift on \mathbf{H} .

Problem 165. If $C = P^{-1}UP$, where P is positive and U is unitary, then a necessary and sufficient condition that C be a contraction is that UP be hyponormal.

Chapter 17. Numerical range

Problem 166. It is sufficient to prove that if f and g are unit vectors such that $(Af, f) = 1$, $(Ag, g) = 0$, and (Af, g) is real, then $W(A)$ includes the whole unit interval. Consider $tf + (1 - t)g$, $0 \leq t \leq 1$, and argue by continuity.

Problem 167. If \mathbf{M} and \mathbf{N} are k -dimensional Hilbert spaces and if T is a linear transformation from \mathbf{M} to \mathbf{N} , then there exist orthonormal bases $\{f_1, \dots, f_k\}$ for \mathbf{M} , and $\{g_1, \dots, g_k\}$ for \mathbf{N} , and there exist positive scalars $\alpha_1, \dots, \alpha_k$ such that $Tf_i = \alpha_i g_i$, $i = 1, \dots, k$. If P and Q are projections of rank k , apply this statement to the restriction of QP to the range of P , and apply the Toeplitz-Hausdorff theorem k times.

Problem 168. Try a diagonal operator. Try the unilateral shift.

Problem 169. The closure of the numerical range includes both the compression spectrum (the complex conjugate of the point spectrum of the adjoint) and the approximate point spectrum.

Problem 170. Let V be the Volterra integration operator and consider $1 - (1 + V)^{-1}$.

Problem 171. Use the spectral theorem; reduce the thing to be proved to the statement that if the values of a function are in the right

half plane, then so is the value of its integral with respect to a positive measure.

Problem 172. Use Problems 157, 169, and 171.

Problem 173. (a) Prove the contrapositive. (b) If $\|A\| = 1$ and $(Af_n, f_n) \rightarrow 1$, then $Af_n - f_n \rightarrow 0$.

Problem 174. Write

$$M = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix},$$

let N be a normal operator whose spectrum is the closed disc with center 0 and radius $\frac{1}{2}$, and consider

$$\begin{pmatrix} M & 0 \\ 0 & N \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} M & 0 \\ 0 & 1 \end{pmatrix}.$$

Problem 175. If $\|A - B\| < \epsilon$ and $\|f\| = 1$, then $(Af, f) \in W(B) + (\epsilon)$. Let U be the unilateral shift and consider U^{*n} , $n = 1, 2, 3, \dots$.

Problem 176. A necessary and sufficient condition that $w(A) \leq 1$ is that $\operatorname{Re}(1 - zA)^{-1} \geq 0$ for every z in the open unit disc. Write down the partial fraction expansion of $1/(1 - z^n)$ and replace z by zA .

Chapter 18. Unitary dilations

Problem 177. (a) Suppose that the given Hilbert space is one-dimensional real Euclidean space and the dilation space is a plane. Examine the meaning of the assertion in this case, use analytic geometry to prove it, and let the resulting formulas suggest the solution in the general case. (b) Imitate (a).

Problem 178. Look for a bilaterally infinite matrix that does the job; use the techniques and results of Solution 177.

Problem 179. Use the spectral theorem to prove the assertion for unitary operators, and then use the existence of unitary power dilations to infer it for all contractions.

Problem 180. Find a unitary power dilation of A .

Problem 181. If $\{f_n\}$ and $\{g_n\}$ are finitely non-zero bilateral sequences of vectors in \mathbf{H} , and if $u_j = \sum_i A_{i-j} f_i$ and $v_j = \sum_i A_{i-j} g_i$, write $[u, v] = \sum_j (u_j, g_j)$. Use this as a definition of inner product for the dilation space \mathbf{K} .

Chapter 19. Commutators of operators

Problem 182. Wintner: assume that P is invertible and examine the spectral implications of $PQ = QP + 1$. Wielandt: assume $PQ - QP = 1$, evaluate $P^n Q - Q P^n$, and use that evaluation to estimate its norm.

Problem 183. Consider the Banach space of all bounded sequences of vectors, modulo null sequences, and observe that each bounded sequence of operators induces an operator on that space.

Problem 184. Fix P and consider $\Delta Q = PQ - QP$ as a function of Q ; determine $\Delta^n Q^n$.

Problem 185. (a) Generalize the formula for the “derivative” of a power to the non-commutative case, and imitate Wielandt’s proof. (b) Use the Kleinecke-Shirokov theorem.

Problem 186. Represent the space as an infinite direct sum in such a way that all summands after the first are in the kernel. Examine the corresponding matrix representation of the given operator, and try to represent it as $PQ - QP$, where P is the pertinent unilateral shift.

Problem 187. Find an invertible operator T such that $A + T^{-1}AT$ has a non-zero kernel; apply Problem 186 to the direct sum of $A + T^{-1}AT$

with itself countably many times. Prove and use the lemma that if $B + C$ is a commutator, then so is $B \oplus C$.

Problem 188. If $C = A^*A - AA^* \geq 0$, choose λ in $\Pi(A)$, find $\{f_n\}$ so that $\|f_n\| = 1$ and $(A - \lambda)f_n \rightarrow 0$, and prove that $Cf_n \rightarrow 0$.

Problem 189. (a) Prove that (1) A is quasinormal, (2) $\ker(1 - A^*A)$ reduces A , and (3) $\ker(1 - A^*A)^\perp \subset \ker(A^*A - AA^*)$. (b) Consider a weighted bilateral shift, with all the weights equal to either 1 or $\sqrt{2}$.

Problem 190. For sufficiency, try a (bilateral) diagonal operator and a bilateral shift; for necessity, adapt the Wintner argument from the additive theory.

Problem 191. Use Problem 111, and then try a diagonal operator matrix and a bilateral shift, in an operator matrix imitation of the technique that worked in Problem 190.

Problem 192. Use Problem 111, together with a multiplicative adaptation of the introduction to Problem 186, to prove that every invertible normal operator is the product of two commutators.

Chapter 20. Toeplitz operators

Problem 193. For necessity: compute. For sufficiency: use Problem 115.

Problem 194. For necessity: compute. For sufficiency: write $A_n f = W^{*n} A P W^n f$ for all f in \mathbf{L}^2 , $n = 0, 1, 2, \dots$, and prove that the sequence $\{A_n\}$ is weakly convergent.

Problem 195. If $\langle \gamma_{ij} \rangle$ is the matrix of $T_\varphi T_\psi$, then $\gamma_{i+1, j+1} = \gamma_{ij} + \alpha_{i+1} \beta_{-j-1}$, where $\varphi = \sum_i \alpha_i e_i$ and $\psi = \sum_j \beta_j e_j$.

Problem 196. Prove that $W^{*n} T P W^n \rightarrow L$ strongly, and use that to prove that if $0 \in \Pi(L)$, then $0 \in \Pi(T)$.

Problem 197. Let K be the kernel function of $\tilde{\mathbf{H}}^2$, and, for a fixed y in D and a fixed \tilde{f} in $\tilde{\mathbf{H}}^2$, write $\tilde{g}(z) = (\tilde{\varphi}(z) - \tilde{\varphi}(y))\tilde{f}(z)$. Since $\tilde{g}(y) = 0$, it follows that $\tilde{g} \perp K_y$ and hence that $\tilde{\varphi}(y)$ is in the (compression) spectrum of T_φ .

Problem 198. If φ is real and $T_\varphi f = 0$, then $\varphi \cdot f^* \cdot f$ is real and belongs to \mathbf{H}^1 .

Problem 199. If φ is real and T_φ is invertible, then $\varphi \cdot f^* \in \mathbf{H}^2$, and this implies that $\text{sgn } \varphi$ is constant.

Solutions

Chapter 1. Vectors and spaces

Solution 1. *The limit of a sequence of quadratic forms is a quadratic form.*

Proof. Associated with each function φ of two variables there is a function φ^- of one variable, defined by $\varphi^-(f) = \varphi(f, f)$; associated with each function ψ of one variable there is a function ψ^+ of two variables, defined by

$$\begin{aligned}\psi^+(f, g) &= \psi\left(\frac{1}{2}(f + g)\right) - \psi\left(\frac{1}{2}(f - g)\right) \\ &\quad + i\psi\left(\frac{1}{2}(f + ig)\right) - i\psi\left(\frac{1}{2}(f - ig)\right).\end{aligned}$$

If φ is a sesquilinear form, then $\varphi = \varphi^{-+}$; if ψ is a quadratic form, then $\psi = \psi^{+-}$. If $\{\psi_n\}$ is a sequence of quadratic forms and if $\psi_n \rightarrow \psi$ (that is, $\psi_n(f) \rightarrow \psi(f)$ for each vector f), then $\psi_n^+ \rightarrow \psi^+$ and $\psi_n^{+-} \rightarrow \psi^{+-}$. Since each ψ_n is a quadratic form, it follows that each ψ_n^+ is a sesquilinear form and hence that ψ^+ is one too. Since, moreover, $\psi_n = \psi_n^{+-}$, it follows that $\psi = \psi^{+-}$, and hence that ψ is a quadratic form.

The index set for sequences (i.e., the set of natural numbers) has nothing to do with the facts here; the proof is just as valid for ordered sequences of arbitrary length, and, more generally, for nets of arbitrary structure.

Solution 2. To motivate the approach, assume for a moment that it is already known that $\xi(f) = (f, g)$ for some g . Choose an arbitrary but fixed orthonormal basis $\{e_j\}$ and expand g accordingly: $g = \sum_j \beta_j e_j$. Since

$$\beta_j = (g, e_j) = (e_j, g)^* = \xi(e_j)^*,$$

the vector g could be captured by writing

$$g = \sum_j \xi(e_j)^* e_j.$$

If the existence of the Riesz representation is known, this reasoning proves uniqueness and exhibits the coordinates of the representing vector. The main problem, from the point of view of the present approach to the existence proof, is to prove the convergence of the series $\sum_j \beta_j e_j$, where $\beta_j = \xi(e_j)^*$.

For each finite set J of indices, write $g_J = \sum_{j \in J} \beta_j e_j$. Then

$$\xi(g_J) = \sum_{j \in J} |\beta_j|^2,$$

and therefore

$$\sum_{j \in J} |\beta_j|^2 \leq \|\xi\| \cdot \|g_J\| = \|\xi\| \cdot \sqrt{\sum_{j \in J} |\beta_j|^2}.$$

This implies that

$$\sqrt{\sum_{j \in J} |\beta_j|^2} \leq \|\xi\|,$$

and hence that

$$\sum_j |\beta_j|^2 < \infty.$$

This result justifies writing $g = \sum_j \beta_j e_j$. If $f = \sum_j \alpha_j e_j$, then

$$\xi(f) = \sum_j \alpha_j \xi(e_j) = \sum_j \alpha_j \beta_j^* = (f, g),$$

and the proof is complete.

Solution 3. The boundary points of the closed unit ball are the vectors on the unit sphere (that is, the unit vectors, the vectors f with $\|f\| = 1$). The thing to prove therefore is that if $f = tg + (1 - t)h$, where $0 \leq t \leq 1$, $\|f\| = 1$, $\|g\| \leq 1$, and $\|h\| \leq 1$, then $f = g = h$. Begin by observing that

$$1 = (f, f) = (f, tg + (1 - t)h) = t(f, g) + (1 - t)(f, h).$$

Since $|(f, g)| \leq 1$ and $|(f, h)| \leq 1$, it follows that $(f, g) = (f, h) = 1$; this step uses the strict convexity of the closed unit disc. The result says that the Schwarz inequality degenerates, both for f and g and for f and h , and this implies that both g and h are multiples of f . Write

$g = \alpha f$ and $h = \beta f$. Since $1 = (f, g) = (f, \alpha f) = \alpha^*$, and, similarly, $1 = \beta^*$, the proof is complete.

Solution 4. Since every infinite-dimensional Hilbert space has a subspace isomorphic to $L^2(0,1)$, it is sufficient to describe the construction for that special space. The description is easy. If $0 \leq t \leq 1$, let $f(t)$ be the characteristic function of the interval $[0, t]$; in other words, $(f(t))(s) = 1$ or 0 according as $0 \leq s \leq t$ or $t < s \leq 1$. If $0 \leq a \leq b \leq 1$, then

$$\begin{aligned} \|f(b) - f(a)\|^2 &= \int |(f(b))(s) - (f(a))(s)|^2 ds \\ &= \int_a^b ds = b - a; \end{aligned}$$

this implies that f is continuous. The verifications of simplicity and of the orthogonality conditions are obvious.

As for the existence of tangents: it is easy to see that the difference quotients do not tend to a limit at any point. Indeed,

$$\left\| \frac{f(t+h) - f(t)}{h} \right\|^2 = \left| \frac{h}{h^2} \right| = \left| \frac{1}{h} \right|,$$

which shows quite explicitly that f is not differentiable anywhere.

Although there is nothing mathematically unique about this construction, it is a curious empirical fact that the example is psychologically unique; everyone who tries it seems to come up with the same answer.

Infinite-dimensionality was explicitly used in the particular proof given above, but that does not imply that it is unavoidable. Is it? An examination of the finite-dimensional situation is quite instructive.

Constructions similar to the one given above are familiar in the theory of spectral measures (cf. Halmos [1951, p. 58]). If E is the spectral measure on the Borel sets of $[0,1]$ such that $E(M)$ is, for each Borel set M , multiplication by the characteristic function of M , and if e is the function constantly equal to 1, then the curve f above is given by

$$f(t) = E([0, t])e.$$

This remark shows how to construct many examples of suddenly turning continuous curves: use different spectral measures and apply them to different vectors. It is not absolutely necessary to consider only continuous spectral measures whose support is the entire interval, but it is wise; those assumptions guarantee that every non-zero vector will work in the role of e .

Solution 5. *If the orthogonal dimension of a Hilbert space is infinite, then its linear dimension is greater than or equal to 2^{\aleph_0} .*

(Recall that if either the linear dimension or the orthogonal dimension of a Hilbert space is finite, then so is the other, and the two are equal.)

Proof. The main tool is the following curious piece of set theory, which has several applications: there exists a collection $\{J_i\}$, of cardinal number 2^{\aleph_0} , consisting of infinite sets of positive integers, such that $J_s \cap J_t$ is finite whenever $s \neq t$. Here is a quick outline of a possible construction. Since there is a one-to-one correspondence between the positive integers and the rational numbers, it is sufficient to prove the existence of sets of rational numbers with the stated property. For each real number t , let J_t be an infinite set of rational numbers that has t as its only cluster point.

Suppose now that $\{e_1, e_2, e_3, \dots\}$ is a countably infinite orthonormal set in a Hilbert space \mathbf{H} , and let $f = \sum_n \xi_n e_n$ (Fourier expansion) be an arbitrary vector such that $\xi_n \neq 0$ for all n . If $\{J_i\}$ is a collection of sets of positive integers of the kind described above, write $f_i = \sum_{n \in J_i} \xi_n e_n$. Assertion: the collection $\{f_i\}$ of vectors is linearly independent. Suppose, indeed, that a finite linear combination of the f 's vanishes, say $\sum_{i=1}^k \alpha_i f_{t_i} = 0$. Since, for each $i \neq 1$, the set J_{t_1} contains infinitely many integers that do not belong to J_{t_i} , it follows that J_{t_1} contains at least one integer, say n , that does not belong to any J_{t_i} ($i \neq 1$). It follows that $\alpha_1 \xi_n = 0$, and hence, since $\xi_n \neq 0$, that $\alpha_1 = 0$. The same argument proves, of course, that $\alpha_i = 0$ for each $i = 1, \dots, k$.

This result is the main reason why the concept of linear dimension is of no interest in Hilbert space theory. In a Hilbert space context "dimension" always means "orthogonal dimension".

There are shorter solutions of the problem, but the preceding argument has the virtue of being elementary in a sense in which they are not. Thus, for instance, every infinite-dimensional Hilbert space may be assumed to include $L^2(0,1)$, and the vectors $f(t)$, $0 < t \leq 1$, exhibited in Solution 4, constitute a linearly independent set with cardinal number 2^{\aleph_0} . Alternatively, every infinite-dimensional Hilbert space may be assumed to include l^2 , and the vectors

$$g(t) = \langle 1, t, t^2, \dots \rangle, \quad 0 < t < 1,$$

constitute a linearly independent set with cardinal number 2^{\aleph_0} .

Solution 6. If $0 < |\alpha| < 1$ and $f_k = \langle 1, \alpha^k, \alpha^{2k}, \alpha^{3k}, \dots \rangle$ for $k = 1, 2, 3, \dots$, then the f_k 's span l^2 .

Proof. Perhaps the quickest approach is to look for a vector f orthogonal to all the f_k 's. If $f = \langle \xi_0, \xi_1, \xi_2, \dots \rangle$, then

$$0 = (f, f_k) = \sum_{n=0}^{\infty} \xi_n \alpha^{*nk}.$$

In other words, the power series $\sum_{n=0}^{\infty} \xi_n z^n$ vanishes for $z = \alpha^{*k}$ ($k = 1, 2, 3, \dots$), and consequently it vanishes identically. Conclusion: $\xi_n = 0$ for all n , and therefore $f = 0$.

The phrasing of the problem is deceptive. The solution has nothing to do with the arithmetic structure of the powers α^k ; the same method applies if the powers α^k are replaced by arbitrary numbers α_k (and, correspondingly, α^{nk} is replaced by α_k^n), provided only that the numbers α_k cluster somewhere in the interior of the unit disc. (Note that if $\sum_{n=0}^{\infty} |\xi_n|^2 < \infty$, then the power series $\sum_{n=0}^{\infty} \xi_n z^n$ has radius of convergence greater than or equal to 1.)

The result is a shallow generalization of the well known facts about Vandermonde matrices, and the proof suggested above is adaptable to the finite-dimensional case. If l_m^2 is the m -dimensional Hilbert space of all sequences $\langle \xi_0, \dots, \xi_{m-1} \rangle$ of length m ($= 1, 2, 3, \dots$), and if the vectors f_k ($k = 1, \dots, m$) are defined by $f_k = \langle 1, \alpha_k, \dots, \alpha_k^{m-1} \rangle$ (where

$0 \leq |\alpha_k| < 1$ and the α_k 's are distinct), then the span of $\{f_1, \dots, f_m\}$ is l_m^2 . Indeed, if $f = \langle \xi_0, \dots, \xi_{m-1} \rangle$ is orthogonal to each f_k , then $\sum_{n=0}^{m-1} \xi_n \alpha_k^{*n} = 0$, i.e., the polynomial $\sum_{n=0}^{m-1} \xi_n z^n$ of degree $m-1$ (at most) vanishes at m distinct points, and hence identically.

Solution 7. It is to be proved that if $f_0 \perp f_i$ ($i = 1, 2, 3, \dots$), then $f_0 = 0$. If $f_0 \neq 0$, then $\{f_0, f_1, f_2, \dots\}$ is an orthogonal set of non-zero vectors, and therefore linearly independent; the purpose of the argument that follows is to show that this cannot happen. The argument is essentially the same as the one used by Birkhoff-Rota, but somewhat simpler; it was discovered by J. T. Rosenbaum.

Begin by choosing a positive integer n so that

$$\sum_{j>n} \|e_j - f_j\|^2 < 1;$$

it will turn out that for this n the vectors f_0, f_1, \dots, f_n are linearly dependent. Write

$$g_k = \sum_{j=1}^n (f_k, e_j) e_j, \quad k = 0, 1, \dots, n.$$

Since each g_k belongs to the (n -dimensional) span of e_1, \dots, e_n , and since the number of g_k 's is $n+1$, it follows that the g_k 's are linearly dependent; say

$$\sum_{k=0}^n \alpha_k g_k = 0.$$

This implies that

$$0 = \sum_{k=0}^n \alpha_k \sum_{j=1}^n (f_k, e_j) e_j = \sum_{j=1}^n \left(\sum_{k=0}^n \alpha_k f_k, e_j \right) e_j.$$

Since the e_j 's are linearly independent, the coefficients in the last sum must vanish; in other words, if

$$h = \sum_{k=0}^n \alpha_k f_k,$$

then $h \perp e_1, \dots, e_n$. Compute the norm of h from its Fourier expansion in terms of the e_j 's:

$$\|h\|^2 = \sum_{j>n} |(h, e_j)|^2$$

(because $h \perp e_j$ for $j \leq n$)

$$= \sum_{j>n} |(h, e_j) - (h, f_j)|^2$$

(because, by definition, h belongs to the span of f_0, f_1, \dots, f_n , so that $h \perp f_j$ for $j > n$)

$$\leq \sum_{j>n} \|h\|^2 \cdot \|e_j - f_j\|^2.$$

The definition of n implies that, unless $h = 0$, the last term is strictly less than $\|h\|^2$. It follows that $h = 0$, i.e., that f_0, f_1, \dots, f_n are indeed linearly dependent.

There is an alternative way to look at this proof; it is a shade less elementary, but it is less computational, and perhaps more transparent. Find n as above, and define a linear transformation A , first on the linear combinations of the e_j 's only, by writing

$$Ae_j = e_j \quad \text{if } j \leq n,$$

and

$$Ae_j = f_j \quad \text{if } j > n.$$

If $f = \sum_j \xi_j e_j$ (finite sum), then

$$\begin{aligned} \|f - Af\|^2 &= \left\| \sum_{j>n} \xi_j (e_j - f_j) \right\|^2 \\ &\leq \sum_{j>n} |\xi_j|^2 \cdot \sum_{j>n} \|e_j - f_j\|^2 \\ &\leq \|f\|^2 \cdot \sum_{j>n} \|e_j - f_j\|^2. \end{aligned}$$

It follows that $1 - A$ is bounded (as far as it is defined) by

$$\sum_{j>n} \|e_j - f_j\|^2,$$

which is strictly less than 1. This implies (Halmos [1951, p. 52]) that A has a (unique) extension to an invertible operator on \mathbf{H} (which may as well be denoted by A again). The invertibility of A implies that the vectors $e_1, \dots, e_n, f_{n+1}, f_{n+2}, \dots$ (the images under A of $e_1, \dots, e_n, e_{n+1}, e_{n+2}, \dots$) span \mathbf{H} . It follows that if \mathbf{M} is the span of f_{n+1}, f_{n+2}, \dots , then $\dim \mathbf{M}^\perp = n$. Conclusion: the vectors $f_1, \dots, f_n, f_{n+1}, f_{n+2}, \dots$ span \mathbf{H} .

Solution 8. It is sufficient to prove that if $\dim \mathbf{M} = 1$, then $\mathbf{M} + \mathbf{N}$ is closed; the general case is obtained by induction on the dimension. Suppose, therefore, that \mathbf{M} is spanned by a single vector f_0 , so that $\mathbf{M} + \mathbf{N}$ consists of all the vectors of the form $\alpha f_0 + g$, where α is a scalar and $g \in \mathbf{N}$. If $f_0 \in \mathbf{N}$, then $\mathbf{M} + \mathbf{N} = \mathbf{N}$; in this case there is nothing to prove. If $f_0 \notin \mathbf{N}$, let g_0 be the projection of f_0 in \mathbf{N} ; that is, g_0 is the unique vector in \mathbf{N} for which $f_0 - g_0 \perp \mathbf{N}$.

Observe now that if g is a vector in \mathbf{N} , then

$$\begin{aligned} \|\alpha f_0 + g\|^2 &= \|\alpha(f_0 - g_0) + (\alpha g_0 + g)\|^2 \\ &\geq |\alpha|^2 \|f_0 - g_0\|^2 \end{aligned}$$

(since $f_0 - g_0 \perp \alpha g_0 + g$), or

$$|\alpha| \leq \frac{\|\alpha f_0 + g\|}{\|f_0 - g_0\|},$$

and therefore

$$\begin{aligned} \|g\| &= \|(\alpha f_0 + g) - \alpha f_0\| \\ &\leq \|\alpha f_0 + g\| + \frac{\|\alpha f_0 + g\|}{\|f_0 - g_0\|} \cdot \|f_0\|. \end{aligned}$$

These inequalities imply that $\mathbf{M} + \mathbf{N}$ (the set of all $\alpha f_0 + g$'s) is closed. Indeed, if $\alpha_n f_0 + g_n \rightarrow h$, so that $\{\alpha_n f_0 + g_n\}$ is a Cauchy sequence, then the inequalities imply that both $\{\alpha_n\}$ and $\{g_n\}$ are Cauchy sequences. It follows that $\alpha_n \rightarrow \alpha$ and $g_n \rightarrow g$, say, with g in \mathbf{N} of course, and consequently $h = \lim_n \alpha_n f_0 + g_n = \alpha f_0 + g$.

Solution 9. *The lattice of subspaces of a Hilbert space \mathbf{H} is modular if and only if $\dim \mathbf{H} < \aleph_0$ (i.e., \mathbf{H} is finite-dimensional); it is distributive if and only if $\dim \mathbf{H} \leq 1$.*

Proof. If \mathbf{H} is infinite-dimensional, then it has subspaces \mathbf{M} and \mathbf{N} such that $\mathbf{M} \cap \mathbf{N} = \{0\}$ and $\mathbf{M} + \mathbf{N} \neq \mathbf{M} \vee \mathbf{N}$ (cf. Problem 41). Given \mathbf{M} and \mathbf{N} , find a vector f_0 in $\mathbf{M} \vee \mathbf{N}$ that does not belong to $\mathbf{M} + \mathbf{N}$, and let \mathbf{L} be the span of \mathbf{N} and f_0 . By Problem 8, \mathbf{L} is equal to the vector sum of \mathbf{N} and the one-dimensional space spanned by f_0 , i.e., every vector in \mathbf{L} is of the form $\alpha f_0 + g$, where α is a scalar and g is in \mathbf{N} .

Both \mathbf{L} and $\mathbf{M} \vee \mathbf{N}$ contain f_0 , and, therefore, so does their intersection. On the other hand, $\mathbf{L} \cap \mathbf{M} = \{0\}$. Reason: if $\alpha f_0 + g \in \mathbf{M}$ (with g in \mathbf{N}), then $\alpha f_0 \in \mathbf{M} + \mathbf{N}$; this implies that $\alpha = 0$ and hence that $g = 0$. Conclusion: $(\mathbf{L} \cap \mathbf{M}) \vee \mathbf{N} = \mathbf{N}$, which does not contain f_0 .

The preceding argument is the only part of the proof in which infinite-dimensionality plays any role. All the remaining parts depend on easy finite-dimensional geometry only. They should be supplied by the reader, who is urged to be sure he can do so before he abandons the subject.

Solution 10. In a Hilbert space of dimension n ($< \aleph_0$) the (closed) unit ball is a closed and bounded subset of $2n$ -dimensional real Euclidean space, and therefore the closed unit ball is compact. It follows, since translations and changes of scale are homeomorphisms, that every closed ball is compact; since the open balls constitute a base for the topology, it follows that the space is locally compact.

Suppose, conversely, that \mathbf{H} is a locally compact Hilbert space. The argument in the preceding paragraph reverses to this extent: the assumption of local compactness implies that each closed ball is compact, and, in particular, so is the closed unit ball. To infer finite-dimensionality,

recall that the distance between two orthogonal unit vectors is $\sqrt{2}$, so that each open ball of diameter $\sqrt{2}$ (or less) can contain at most one element of each orthonormal basis. The collection of all open balls of diameter $\sqrt{2}$ is an open cover of the closed unit ball; the compactness of the latter implies that every orthonormal basis is finite, and hence that \mathbf{H} is finite-dimensional.

Solution 11. If $\dim \mathbf{H} \leq \aleph_0$, then \mathbf{H} has a countable orthonormal basis. Since every vector in \mathbf{H} is the limit of finite linear combinations of basis vectors, it follows that every vector in \mathbf{H} is the limit of such linear combinations with coefficients whose real and imaginary parts are rational. The set of all such rational linear combinations is countable, and consequently \mathbf{H} is separable.

Suppose, conversely, that $\{f_1, f_2, f_3, \dots\}$ is a countable set dense in \mathbf{H} . If $\{g_j\}$ is an orthonormal basis for \mathbf{H} , then for each index j there exists an index n_j such that $\|f_{n_j} - g_j\| < \sqrt{2}/2$. Since two open balls of radius $\sqrt{2}/2$ whose centers are distinct g_j 's are disjoint, the mapping $j \rightarrow n_j$ is one-to-one; this implies that the cardinal number of the set of indices j is not greater than \aleph_0 .

The Gram-Schmidt process yields an alternative approach to the converse. Since that process is frequently described for linearly independent sequences only, begin by discarding from the sequence $\{f_n\}$ all terms that are linear combinations of earlier ones. Once that is done, apply Gram-Schmidt to orthonormalize. The resulting orthonormal set is surely countable; since its span is the same as that of the original sequence $\{f_n\}$, it is a basis.

Solution 12. Since a measure is, by definition, invariant under translation, there is no loss of generality in considering balls with center at 0 only. If \mathbf{B} is such a ball, with radius r (> 0), and if $\{e_1, e_2, e_3, \dots\}$ is an infinite orthonormal set in the space, consider the open balls \mathbf{B}_n with center at $(r/2)e_n$ and radius $r/4$; that is, $\mathbf{B}_n = \{f: \|f - (r/2)e_n\| < r/4\}$. If $f \in \mathbf{B}_n$, then

$$\|f\| \leq \left\| f - \frac{r}{2} e_n \right\| + \left\| \frac{r}{2} e_n \right\| < r,$$

so that $\mathbf{B}_n \subset \mathbf{B}$. If $f \in \mathbf{B}_n$ and $g \in \mathbf{B}_m$, then

$$\left\| \frac{r}{2} e_n - \frac{r}{2} e_m \right\| \leq \left\| \frac{r}{2} e_n - f \right\| + \|f - g\| + \left\| g - \frac{r}{2} e_m \right\|.$$

This implies that if $n \neq m$, then

$$\|f - g\| \geq \frac{r\sqrt{2}}{2} - \frac{r}{4} - \frac{r}{4} > 0,$$

and hence that if $n \neq m$, then \mathbf{B}_n and \mathbf{B}_m are disjoint. Since, by invariance, all the \mathbf{B}_n 's have the same measure, it follows that \mathbf{B} includes infinitely many disjoint Borel sets of the same positive measure, and hence that the measure of \mathbf{B} must be infinite.

Chapter 2. Weak topology

Solution 13. If \mathbf{S} is a weakly closed set in \mathbf{H} and if $\{f_n\}$ is a sequence of vectors in \mathbf{S} with $f_n \rightarrow f$ (strong), then

$$|(f_n, g) - (f, g)| \leq \|f_n - f\| \cdot \|g\| \rightarrow 0,$$

so that $f_n \rightarrow f$ (weak), and therefore $f \in \mathbf{S}$. This proves that weakly closed sets are strongly closed; in fact, the proof shows that the strong closure of each set is included in its weak closure. The falsity of the converse (i.e., that a strongly closed set need not be weakly closed) can be deduced from the curious observation that if $\{e_1, e_2, e_3, \dots\}$ is an orthonormal sequence, then $e_n \rightarrow 0$ (weak). Reason: for each vector f , the inner products (f, e_n) are the Fourier coefficients of f , and, therefore, they are the terms of an absolutely square-convergent series. It follows that the set of all e_n 's is not closed in the weak topology; in the strong topology it is discrete and therefore closed. Another way of settling the converse is to exhibit a strongly open set that is not weakly open; one such set is the open unit ball. To prove what needs proof, observe that in an infinite-dimensional space weakly open sets are unbounded.

It remains to prove that subspaces are weakly closed. If $\{f_n\}$ is a sequence in a subspace \mathbf{M} , and if $f_n \rightarrow f$ (weak), then, by definition, $(f_n, g) \rightarrow (f, g)$ for every g . Since each f_n is orthogonal to \mathbf{M}^\perp , it follows that $f \perp \mathbf{M}^\perp$ and hence that $f \in \mathbf{M}$. This argument shows that \mathbf{M} contains the limits of all weakly convergent sequences in \mathbf{M} , but that does not yet justify the conclusion that \mathbf{M} is weakly closed. At this point in this book the weak topology is not known to be metrizable; sequential closure may not be the same as closure. The remedy, however, is easy; just observe that the sequential argument works without the change of a single symbol if the word "sequence" is replaced by "net", and net closure is always the same as closure.

Solution 14. The proof depends on a familiar trivial computation:

$$\|f_n - f\|^2 = (f_n - f, f_n - f) = \|f_n\|^2 - (f, f_n) - (f_n, f) + \|f\|^2.$$

Since $f_n \rightarrow f$ (weak), the terms with minus signs tend to $\|f\|^2$, and, by assumption, so does the first term. Conclusion: $\|f_n - f\|^2 \rightarrow 0$, as asserted.

Solution 15. *Every weakly separable Hilbert space is separable.*

Proof. The span of a countable set is always a (strongly) separable subspace; it is therefore sufficient to prove that if a countable set \mathbf{S} is weakly dense in a Hilbert space \mathbf{H} , then the span of \mathbf{S} is equal to \mathbf{H} . Looked at from the right point of view this is obvious. The span of \mathbf{S} is, by definition, a (strongly closed) subspace, and hence, by Problem 13, it is weakly closed; being at the same time weakly dense in \mathbf{H} , it must be equal to \mathbf{H} .

Caution: it is not only more elegant but it is also safer to argue without sequences. It is not a priori obvious that if f is in the weak closure of \mathbf{S} , then f is the limit of a sequence in \mathbf{S} .

Solution 16. It is sufficient to treat the case $f = 0$. If $\|f_n\| \rightarrow 0$, then, since $|(f_n, g)| \leq \|f_n\| \cdot \|g\| = \|f_n\|$ whenever $\|g\| = 1$, it follows that $(f_n, g) \rightarrow 0$ uniformly, as stated.

Suppose, conversely, that, for each positive number ε , if n is sufficiently large, then

$$|(f_n, g)| < \varepsilon \quad \text{whenever } \|g\| = 1;$$

uniformity manifests itself in that the size of the n that is needed does not depend on g . It follows that if n is sufficiently large, then

$$\left| \left(f_n, \frac{g}{\|g\|} \right) \right| < \varepsilon \quad \text{whenever } g \neq 0,$$

and hence that

$$|(f_n, g)| \leq \varepsilon \|g\| \quad \text{for all } g.$$

Hence, in particular, if n is sufficiently large, then (put $g = f_n$)

$$\|f_n\|^2 \leq \varepsilon \|f_n\|$$

or

$$\|f_n\| \leq \varepsilon.$$

Note that the argument is perfectly general; it applies to all nets, and not to sequences only.

Solution 17. Given the Hilbert space \mathbf{H} , for each f in \mathbf{H} let D_f be the closed disc $\{z: |z| \leq \|f\|\}$ in the complex plane, and let \mathbf{D} be the Cartesian product of all the D_f 's, with the customary product topology. For each g in the unit ball, the mapping $f \rightarrow (f, g)$ is a point, say $\delta(g)$, in \mathbf{D} . The mapping δ thus defined is a homeomorphism from the unit ball (with the weak topology) into \mathbf{D} (with the product topology). Indeed, if $\delta(g_1) = \delta(g_2)$, that is, if $(f, g_1) = (f, g_2)$ for all f , then clearly $g_1 = g_2$, so that δ is one-to-one. As for continuity:

$$g_j \rightarrow g \text{ (weak) if and only if } (f, g_j) \rightarrow (f, g)$$

for each f in \mathbf{H} , and that, in turn, happens if and only if $\delta(g_j) \rightarrow \delta(g)$ in \mathbf{D} . The Riesz theorem on the representation of linear functionals on \mathbf{H} implies that the range of δ consists exactly of those elements ξ of \mathbf{D} (complex-valued functions on \mathbf{H}) that are in fact linear functionals of norm less than or equal to 1 on \mathbf{H} .

The argument so far succeeded in constructing a homeomorphism δ from the unit ball into the compact Hausdorff space \mathbf{D} , and it succeeded in identifying the range of δ . The remainder of the argument will show that that range is closed (and therefore compact) in \mathbf{D} ; as soon as that is done, the weak compactness of the unit ball will follow.

The property of being a linear functional is a property of "finite character". That is: ξ is a linear functional if and only if it satisfies equations (infinitely many of them) each of which involves only a finite number of elements of \mathbf{H} ; this implies that the set of all linear functionals is closed in \mathbf{D} . In more detail, consider fixed pairs of scalars α_1 and α_2 and vectors f_1 and f_2 , and form the subset $\mathbf{E}(\alpha_1, \alpha_2, f_1, f_2)$ of \mathbf{D} defined by

$$\mathbf{E}(\alpha_1, \alpha_2, f_1, f_2) = \{\xi \in \mathbf{D}: \xi(\alpha_1 f_1 + \alpha_2 f_2) = \alpha_1 \xi(f_1) + \alpha_2 \xi(f_2)\}.$$

The assertion about properties of finite character amounts to this: the set of all linear functionals in \mathbf{D} (the range of δ) is the intersection of all the sets of the form $\mathbf{E}(\alpha_1, \alpha_2, f_1, f_2)$. Since the definition of product topology implies that each of the functions $\xi \rightarrow \xi(f_1)$, $\xi \rightarrow \xi(f_2)$, and

$\xi \rightarrow \xi(\alpha_1 f_1 + \alpha_2 f_2)$ is continuous on \mathbf{D} , it follows that each set $\mathbf{E}(\alpha_1, \alpha_2, f_1, f_2)$ is closed, and hence that the range of δ is compact.

The proof above differs from the proof of a more general Tychonoff-Alaoglu theorem (the unit ball of the conjugate space of a Banach space is weak * compact) in notation only.

Solution 18. *In a separable Hilbert space the weak topology of the unit ball is metrizable.*

Proof 1. Since the unit ball \mathbf{B} is weakly compact (Problem 17), it is sufficient to prove the existence of a countable base for the weak topology of \mathbf{B} . For this purpose, let $\{h_j: j = 1, 2, 3, \dots\}$ be a countable set dense in the space, and consider the basic weak neighborhoods (in \mathbf{B}) defined by

$$\mathbf{U}(p, q, r) = \left\{ f \in \mathbf{B}: |(f - h_p, h_j)| < \frac{1}{q}, j = 1, \dots, r \right\},$$

where $p, q, r = 1, 2, 3, \dots$. To prove: if $f_0 \in \mathbf{B}$, k is a positive integer, g_1, \dots, g_k are arbitrary vectors, and ϵ is a positive number, and if

$$\mathbf{U} = \{f \in \mathbf{B}: |(f - f_0, g_i)| < \epsilon, i = 1, \dots, k\},$$

then there exist integers p, q , and r such that

$$f_0 \in \mathbf{U}(p, q, r) \subset \mathbf{U}.$$

The proof is based on the usual inequality device:

$$\begin{aligned} |(f - f_0, g_i)| &\leq |(f - h_p, h_j)| + |(h_p - f_0, h_j)| + |(f - f_0, g_i - h_j)| \\ &\leq |(f - h_p, h_j)| + \|h_p - f_0\| \cdot \|h_j\| + \|f - f_0\| \cdot \|g_i - h_j\|. \end{aligned}$$

Argue as follows: for each i ($= 1, \dots, k$) choose j_i so that $\|g_i - h_{j_i}\|$ is small, and choose p so that $\|h_p - f_0\|$ is very small. Specifically: choose q so that $1/q < \epsilon/3$, choose j_i so that $\|g_i - h_{j_i}\| < 1/2q$, choose r so that $j_i \leq r$ for $i = 1, \dots, k$, and, finally, choose p so that

$\|h_p - f_0\| < 1/qm$, where $m = \max\{\|h_j\| : j = 1, \dots, r\}$. If $j = 1, \dots, r$, then

$$|(f_0 - h_p, h_j)| \leq \|f_0 - h_p\| \cdot \|h_j\| < \frac{1}{qm} \cdot m = \frac{1}{q},$$

so that $f_0 \in \mathbf{U}(p, q, r)$. If $f \in \mathbf{U}(p, q, r)$ and $i = 1, \dots, k$, then

$$|(f - h_p, h_{j_i})| < \frac{1}{q} < \frac{\varepsilon}{3},$$

$$\|h_p - f_0\| \cdot \|h_{j_i}\| < \frac{1}{qm} \cdot m < \frac{\varepsilon}{3},$$

and

$$\|f - f_0\| \cdot \|g_i - h_{j_i}\| < 2 \cdot \frac{1}{2q} < \frac{\varepsilon}{3},$$

(recall that $\|f\| \leq 1$ and $\|f_0\| \leq 1$). It follows that $f \in \mathbf{U}$, and the proof is complete.

Proof 2. There is an alternative procedure that sheds some light on the problem and has the merit, if merit it be, that it exhibits a concrete metric for the weak topology of \mathbf{B} . Let $\{e_1, e_2, e_3, \dots\}$ be an orthonormal basis for \mathbf{H} . (There is no loss of generality in assuming that the basis is infinite; in the finite-dimensional case all these topological questions become trivial.) For each vector f write

$$|f| = \sum_j \frac{1}{2^j} |(f, e_j)|;$$

since $|(f, e_j)| \leq \|f\|$, the series converges and defines a norm. If $d(f, g) = |f - g|$ whenever f and g are in \mathbf{B} , then d is a metric for \mathbf{B} . To show that d metrizes the weak topology of \mathbf{B} , it is sufficient to prove that $f_n \rightarrow 0$ (weak) if and only if $|f_n| \rightarrow 0$. (Caution: the metric d is defined for all \mathbf{H} but its relation to the weak topology of \mathbf{H} is not the same as its relation to the weak topology of \mathbf{B} . The uniform boundedness of the elements of \mathbf{B} is what is needed in the argument below.)

Assume that $f_n \rightarrow 0$ (weak), so that, in particular, $(f_n, e_j) \rightarrow 0$ as $n \rightarrow \infty$ for each j . The tail of the series for $|f_n|$ is uniformly small for

all n (in fact, the tail of the series for $|f|$ is uniformly small for all f in \mathbf{B}). In the present case the assumed weak convergence implies that each particular partial sum of the series for $|f_n|$ becomes small as n becomes large, and it follows that $|f_n| \rightarrow 0$.

Assume that $|f_n| \rightarrow 0$. Since the sum of the series for $|f_n|$ dominates each term, it follows that $(f_n, e_j) \rightarrow 0$ as $n \rightarrow \infty$ for each j . This implies that if g is a finite linear combination of e_j 's, then $(f_n, g) \rightarrow 0$. Such linear combinations are dense. If $h \in \mathbf{H}$, then

$$|(f_n, h)| \leq |(f_n, h - g)| + |(f_n, g)|.$$

Choose g so as to make $\|h - g\|$ small (and therefore $|(f_n, h - g)|$ will be just as small), and then choose n so as to make $|(f_n, g)|$ small. (This is a standard argument that is sometimes isolated as a lemma: a bounded sequence that satisfies the condition for weak convergence on a dense set is weakly convergent.) Conclusion: $f_n \rightarrow 0$ (weak).

Solution 19. *If the weak topology of the unit ball in a Hilbert space \mathbf{H} is metrizable, then \mathbf{H} is separable.*

Proof. If \mathbf{B} , the unit ball, is weakly metrizable, then it is weakly separable (since it is weakly compact). Let $\{f_n: n = 1, 2, 3, \dots\}$ be a countable set weakly dense in \mathbf{B} . The set of all vectors of the form mf_n , $m, n = 1, 2, 3, \dots$, is weakly dense in \mathbf{H} . (Reason: for fixed m , the mf_n 's are weakly dense in $m\mathbf{B}$, and $\bigcup_m m\mathbf{B} = \mathbf{H}$.) The proof is completed by recalling (Solution 15) that weakly separable Hilbert spaces are separable.

Solution 20. Suppose that \mathbf{T} is a weakly bounded set in \mathbf{H} and that, specifically, $|(f, g)| \leq \alpha(f)$ for all g in \mathbf{T} . If \mathbf{H} is finite-dimensional, the proof is easy. Indeed, if $\{e_1, \dots, e_n\}$ is an orthonormal basis for \mathbf{H} , then

$$\begin{aligned} |(f, g)| &= |(\sum_{i=1}^n (f, e_i) e_i, g)| = |\sum_{i=1}^n (f, e_i) (e_i, g)| \\ &\leq \sqrt{\sum_{i=1}^n |(f, e_i)|^2} \cdot \sqrt{\sum_{i=1}^n |(e_i, g)|^2} \\ &\leq \|f\| \cdot n \cdot \max\{\alpha(e_1), \dots, \alpha(e_n)\}, \end{aligned}$$

and all is well.

Assume now that \mathbf{H} is infinite-dimensional, and assume that the conclusion is false. A consequence of this assumption is the existence of an element g_1 of \mathbf{T} and a unit vector e_1 such that $|(e_1, g_1)| \geq 1$. Are the linear functionals induced by \mathbf{T} (i.e., the mappings $f \rightarrow (f, g)$ for g in \mathbf{T}) bounded on the orthogonal complement of the at most two-dimensional space spanned by e_1 and g_1 ? If so, then they are bounded on \mathbf{H} , contrary to the present assumption. A consequence of this argument is the existence of an element g_2 of \mathbf{T} and a unit vector orthogonal to e_1 and g_1 , such that $|(e_2, g_2)| \geq 2(\alpha(e_1) + 2)$. Continue in the same vein. Argue, as before, that the linear functionals induced by \mathbf{T} cannot be bounded on the orthogonal complement of the at most four-dimensional space spanned by e_1, e_2, g_1, g_2 ; arrive, as before, to the existence of an element g_3 of \mathbf{T} and a unit vector e_3 orthogonal to e_1, e_2 and g_1, g_2 , and such that

$$|(e_3, g_3)| \geq 3(\alpha(e_1) + \frac{1}{2}\alpha(e_2) + 3).$$

Induction yields, after n steps, an element g_{n+1} of \mathbf{T} and a unit vector e_{n+1} orthogonal to e_1, \dots, e_n and g_1, \dots, g_n , such that

$$|(e_{n+1}, g_{n+1})| \geq (n+1) \left(\sum_{i=1}^n \frac{1}{i} \alpha(e_i) + n+1 \right).$$

Now put $f = \sum_{i=1}^{\infty} (1/i) e_i$. Since

$$\begin{aligned} |(f, g_{n+1})| &= \left| \sum_{i=1}^n \frac{1}{i} (e_i, g_{n+1}) + \frac{1}{n+1} (e_{n+1}, g_{n+1}) \right| \\ &\geq - \sum_{i=1}^n \frac{1}{i} \alpha(e_i) + \frac{1}{n+1} (n+1) \cdot \left(\sum_{i=1}^n \frac{1}{i} \alpha(e_i) + n+1 \right) \\ &= n+1, \end{aligned}$$

it follows that if \mathbf{T} is not bounded, then it cannot be weakly bounded either.

This proof is due to D. E. Sarason. Special cases of it occur in von Neumann [1929, footnote 32] and Stone [1932, p. 59]; almost the general case is in Akhiezer-Glazman [1961, p. 45].

Solution 21. Let $\{e_1, e_2, e_3, \dots\}$ be an infinite orthonormal set in \mathbf{H} and let \mathbf{E} be the set of all vectors of the form $\sqrt{n} \cdot e_n$, $n = 1, 2, 3, \dots$. Assertion: the origin belongs to the weak closure of \mathbf{E} . Suppose indeed that

$$\{f: |(f, g_i)| < \varepsilon, \quad i = 1, \dots, k\}$$

is a basic weak neighborhood of 0. Since $\sum_{n=1}^{\infty} |(g_i, e_n)|^2 < \infty$ for each i , it follows that $\sum_{n=1}^{\infty} (\sum_{i=1}^k |(g_i, e_n)|)^2 < \infty$. (The sum of a finite number of square-summable sequences is square-summable.) It follows that there is at least one value of n for which $\sum_{i=1}^k |(g_i, e_n)| < \varepsilon / \sqrt{n}$; (otherwise square both sides and contemplate the harmonic series). If n is chosen so that this inequality is satisfied, then, in particular, $|(g_i, e_n)| < \varepsilon / \sqrt{n}$ for each i , and therefore $|(\sqrt{n} \cdot e_n, g_i)| < \varepsilon$ for each i ($= 1, \dots, k$).

The weak non-metrizability of \mathbf{H} can be established by proving that no sequence in \mathbf{E} converges weakly to 0. Since no infinite subset of \mathbf{E} is bounded, the desired result is an immediate consequence of the principle of uniform boundedness.

The first construction of this kind is due to von Neumann [1929, p. 380]. The one above is simpler; it was discovered by A. L. Shields.

Solution 22. Write $g_k = \{\beta_1^*, \dots, \beta_k^*, 0, 0, 0, \dots\}$, so that clearly $g_k \in l^2$, $k = 1, 2, 3, \dots$. If $f = \{\alpha_1, \alpha_2, \alpha_3, \dots\}$ is in l^2 , then $(f, g_k) = \sum_{j=1}^k \alpha_j \beta_j \rightarrow \sum_{j=1}^{\infty} \alpha_j \beta_j$. It follows that, for each f in l^2 , the sequence $\{(f, g_k)\}$ is bounded, i.e., that the sequence $\{g_k\}$ of vectors in l^2 is weakly bounded. Conclusion (from the principle of uniform boundedness): there exists a positive constant β such that $\|g_k\|^2 \leq \beta$ for all k , and, therefore, $\sum_{j=1}^{\infty} |\beta_j|^2 \leq \beta$.

The method generalizes to many measure spaces, including all σ -finite ones. Suppose that X is a measure space with σ -finite measure μ , and suppose that g is a measurable function on X with the property that its product with every function in $L^2(\mu)$ belongs to $L^1(\mu)$; the conclusion is that g belongs to $L^2(\mu)$.

Let $\{E_k\}$ be an increasing sequence of sets of finite measure such that $\bigcup_k E_k = X$ and such that g is bounded on each E_k . (Here is where σ -finiteness comes in.) Write $g_k = \chi_{E_k} g$ (where χ_{E_k} is the characteristic function of E_k), $k = 1, 2, 3, \dots$. The rest of the proof is the obvious modification of the preceding discrete proof; just replace sums by integrals.

For those who know about the closed graph theorem, it provides an alternative approach; apply it to the linear transformation $f \rightarrow fg^*$ from \mathbf{L}^2 into \mathbf{L}^1 . For a discussion of an almost, but not quite, sufficiently general version of the closed graph theorem, see Problem 44.

Solution 23. (a) The idea is that a sufficiently “large” Cauchy net can turn out to be anxious to converge to an “unbounded vector”, i.e., to something not in the space. To make this precise, let ξ be an unbounded linear functional, fixed throughout what follows; on an infinite-dimensional Hilbert space such things always exist. (Use a Hamel basis to make one.) Then let $\{e_j\}$ be a Hamel basis, and, corresponding to each finite subset J of the index set, let \mathbf{M}_J be the (finite-dimensional) subspace spanned by the e_j 's with j in J . Consider the linear functional ξ_J that is equal to ξ on \mathbf{M}_J and equal to 0 on \mathbf{M}_J^\perp . Since the ξ_J 's are bounded (finite-dimensionality), there exists a net $J \rightarrow g_J$ of vectors such that $\xi_J(f) = (f, g_J)$ for each f and for each J . (The finite sets J are ordered by inclusion, of course.) Given f_0 , let J_0 be a finite set such that $f_0 \in \mathbf{M}_{J_0}$. If both J and K include J_0 , then $(f, g_J) - (f, g_K) = 0$; it follows that $\{g_J\}$ is a weak Cauchy net. This Cauchy net cannot possibly converge weakly to anything. Suppose indeed that $g_J \rightarrow g$ weakly, so that $\xi_J(f_0) \rightarrow (f_0, g)$ for each fixed f_0 . As soon as J_0 is so large that $f_0 \in \mathbf{M}_{J_0}$, then $\xi_{J_0}(f_0) = \xi(f_0)$; it follows that $\xi(f_0) = (f_0, g)$ for each f_0 . Since ξ is unbounded, that is impossible.

(b) *Every Hilbert space is sequentially weakly complete.*

Proof. If $\{g_n\}$ is a weak Cauchy sequence in \mathbf{H} , then $\{(f, g_n)\}$ is Cauchy, and therefore bounded, for each f in \mathbf{H} , so that $\{g_n\}$ is weakly bounded. It follows from the principle of uniform boundedness that $\{g_n\}$ is bounded. Since $\lim_n (f, g_n) = \xi(f)$ exists for each f in \mathbf{H} , and since the boundedness of $\{g_n\}$ implies that the linear functional ξ is bounded, it follows that there exists a vector g in \mathbf{H} such that $\lim_n (f, g_n) = (f, g)$ for all f . This means that $g_n \rightarrow g$ (weak), so that $\{g_n\}$ does indeed have a weak limit.

Chapter 3. Analytic functions

Solution 24. For each region D , the inner-product space $\mathbf{A}^2(D)$ is complete.

Proof. It is convenient to present the proof in three steps.

(1) If D is an open disc with center λ and radius r , and if $f \in \mathbf{A}^2(D)$, then

$$f(\lambda) = \frac{1}{\pi r^2} \int_D f(z) d\mu(z).$$

There is no loss of generality in restricting attention to the unit disc D_1 in the role of D , $D_1 = \{z: |z| < 1\}$; the general case reduces to this special case by an appropriate translation and change of scale. Suppose, accordingly, that $f \in \mathbf{A}^2 (= \mathbf{A}^2(D_1))$ with Taylor series $\sum_{n=0}^{\infty} \alpha_n z^n$, and let D_r be the disc $\{z: |z| < r\}$, $0 < r < 1$. In each D_r , $0 < r < 1$, the Taylor series of f converges uniformly, and, consequently, it is term-by-term integrable. This implies that

$$\begin{aligned} \int_{D_r} f(z) d\mu(z) &= \sum_{n=0}^{\infty} \alpha_n \int_{D_r} z^n d\mu(z) \\ &= \alpha_0 \cdot \pi r^2. \end{aligned}$$

Since $|f|$ is integrable over D_1 , it follows that $\int_{D_r} f d\mu \rightarrow \int_{D_1} f d\mu$ as $r \rightarrow 1$; since $\alpha_0 = f(0)$, the proof of (1) is complete.

Return now to the case of a general region D .

(2) If $v_\lambda(f) = f(\lambda)$ whenever $\lambda \in D$ and $f \in \mathbf{A}^2(D)$, then, for each fixed λ , the functional v_λ is linear. If $r = r(\lambda)$ is the radius of the largest open disc with center λ that is entirely included in D , then

$$|v_\lambda(f)| \leq \frac{1}{\sqrt{\pi}r} \|f\|.$$

Let D_0 be the largest open disc with center λ that is entirely included in D . Since

$$\begin{aligned} \|f\|^2 &= \int_D |f(z)|^2 d\mu(z) \geq \int_{D_0} |f(z)|^2 d\mu(z) \\ &\geq \frac{1}{\pi r^2} \left| \int_{D_0} f(z) d\mu(z) \right|^2 \quad (\text{by the Schwarz inequality}) \\ &= \pi r^2 \left| \frac{1}{\pi r^2} \int_{D_0} f(z) d\mu(z) \right|^2 \\ &= \pi r^2 |f(\lambda)|^2 \quad (\text{by (1)}), \end{aligned}$$

the proof of (2) is complete.

(3) The proof of the main assertion is now within reach. Suppose that $\{f_n\}$ is a Cauchy sequence in $\mathbf{A}^2(D)$. It follows from (2) that

$$|f_n(\lambda) - f_m(\lambda)| \leq \frac{1}{\sqrt{\pi} \cdot r(\lambda)} \|f_n - f_m\|$$

for every λ in D ; here, as before, $r(\lambda)$ is the radius of the largest open disc with center at λ that is entirely included in D . It follows that if K is a compact subset of D , so that $r(\lambda)$ is bounded away from 0 when $\lambda \in K$, then the sequence $\{f_n\}$ of functions is uniformly convergent on K . This implies that there exists an analytic function f on D such that $f_n(\lambda) \rightarrow f(\lambda)$ for all λ in D . At the same time the completeness of the Hilbert space $\mathbf{L}^2(\mu)$ implies the existence of a complex-valued, square-integrable, but not necessarily analytic function g on D such that $f_n \rightarrow g$ in the mean of order 2. It follows that a subsequence of $\{f_n\}$ converges to g almost everywhere, and hence that $f = g$ almost everywhere. This implies that f is square-integrable, i.e., that $f \in \mathbf{A}^2(D)$, and hence that $\mathbf{A}^2(D)$ is complete.

These facts were first discussed by Bergman [1947, p. 24]; the proof above is due to Halmos-Lumer-Schäffer [1953]. The latter makes explicit use of the Riesz-Fischer theorem (the completeness of \mathbf{L}^2), instead of proving it in the particular case at hand, and consequently, from the point of view of the standard theory of Hilbert spaces, it is simpler than the analytic argument given by Bergman.

Solution 25. The evaluation of the inner products (e_n, e_m) is routine calculus. If, in fact, $D_r = \{z: |z| < r\}$, then

$$\begin{aligned} \int_{D_r} z^n \bar{z}^m d\mu(z) &= \int_0^{2\pi} \int_0^r e^{i(n-m)\theta} \rho^{n+m} \rho d\rho d\theta \\ &= 2\pi \delta_{nm} \frac{r^{n+m+2}}{n+m+2}. \end{aligned}$$

It follows (put $r = 1$) that if $n \neq m$, then $(e_n, e_m) = 0$, and it follows also (put $n = m$) that $\|e_n\|^2 = 1$. This proves orthonormality.

To prove that the e_n 's form a *complete* orthonormal set, it is tempting to argue as follows. If $f \in \mathbf{A}^2$, with Taylor series $\sum_{n=0}^{\infty} \alpha_n z^n$, then $f(z) = \sum_{n=0}^{\infty} \alpha_n \sqrt{\pi/(n+1)} \cdot e_n(z)$; this shows that each f in \mathbf{A}^2 is a linear combination of the e_n 's, q.e.d. The argument is almost right. The trouble is that the kind of convergence it talks about is wrong. Although $\sum_{n=0}^{\infty} \alpha_n z^n$ converges to $f(z)$ at each z , and even uniformly in each compact subset of the disc, these facts by themselves do not imply that the series converges in the metric (norm) of \mathbf{A}^2 .

There is a simple way around the difficulty: prove something else. Specifically, it is sufficient to prove that if $f \in \mathbf{A}^2$ and $f \perp e_n$ for $n = 0, 1, 2, \dots$, then $f = 0$; and this is an immediate consequence of the second statement in Problem 25 (the statement about the relation between the Taylor and Fourier coefficients). That statement is a straightforward generalization of (1) in Solution 24 (which is concerned with e_0 only). The proof of the special case can be adapted to the general case, as follows. In each D_r , $0 < r < 1$, the series

$$f(z) \bar{z}^m = \sum_{n=0}^{\infty} \alpha_n z^n \bar{z}^m$$

converges uniformly, and, consequently, it is term-by-term integrable. This implies that

$$\begin{aligned} \int_{D_r} f(z) \bar{z}^m d\mu(z) &= \sum_{n=0}^{\infty} \alpha_n \cdot 2\pi \delta_{nm} \frac{r^{n+m+2}}{n+m+2} \\ &= \alpha_m \frac{\pi \cdot r^{2m+2}}{m+1}. \end{aligned}$$

Since $|f \cdot e_m^*|$ is integrable over D_1 , it follows that $\int_{D_1} f \cdot e_m^* d\mu \rightarrow \int_{D_1} f \cdot e_m^* d\mu = (f, e_m)$ as $r \rightarrow 1$, and the proof is complete.

Note that the above argument makes tacit use of the completeness of \mathbf{A}^2 . The argument proves that the orthonormal set $\{e_0, e_1, e_2, \dots\}$ is maximal; a maximal orthonormal set deserves to be called a basis only if the space is complete. The point is that in the absence of completeness the convergence of Fourier expansions cannot be guaranteed.

An alternative proof that the e_n 's form a basis, which uses completeness in a less underhanded manner, is this. If $f \in \mathbf{A}^2$, with Taylor series $\sum_{n=0}^{\infty} \alpha_n z^n$, then $(f, e_n) = \sqrt{\pi/(n+1)} \cdot \alpha_n$. This implies, via the Bessel inequality, that

$$\sum_{n=0}^{\infty} \frac{\pi |\alpha_n|^2}{n+1}$$

converges. It follows that the series whose n -th term is

$$\sqrt{\pi/(n+1)} \cdot \alpha_n e_n(z)$$

converges in the mean of order 2; this conclusion squarely meets and overcomes the obstacle that stops the naive argument via power series expansions.

The result establishes a natural isomorphism between \mathbf{A}^2 and the Hilbert space of all those sequences $\langle \alpha_0, \alpha_1, \alpha_2, \dots \rangle$ for which

$$\sum_{n=0}^{\infty} \frac{|\alpha_n|^2}{n+1} < \infty,$$

with the inner product of $\langle \alpha_0, \alpha_1, \alpha_2, \dots \rangle$ and $\langle \beta_0, \beta_1, \beta_2, \dots \rangle$ given by

$$\sum_{n=0}^{\infty} \frac{\pi \alpha_n \beta_n^*}{n+1}.$$

Solution 26. Formally the assertion is almost obvious. For any f in \mathbf{L}^2 (not only \mathbf{H}^2), with Fourier expansion $f = \sum_n \alpha_n e_n$, complex conjugation yields

$$f^* = \sum_n \alpha_n^* e_n^* = \sum_n \alpha_n^* e_{-n} = \sum_n \alpha_{-n}^* e_n;$$

it follows that if $f = f^*$, then $\alpha_n = \alpha_{-n}^*$ for all n . If, moreover, $f \in \mathbf{H}^2$,

so that $\alpha_n = 0$ whenever $n < 0$, then it follows that $\alpha_n = 0$ whenever $n \neq 0$, and hence that $f = \alpha_0$.

The trouble with this argument is its assumption that complex conjugation distributes over Fourier expansion; that assumption must be justified or avoided. It can be justified this way: the finite subsums of $\sum_n \alpha_n e_n$ converge to f in the sense of the norm of \mathbf{L}^2 , i.e., in the mean of order 2; it follows that a subsequence of them converges to f almost everywhere, and the desired result follows from the continuity of conjugation. The assumption can be avoided this way: since $\alpha_n = \int f e_n^* d\mu$, it follows that $\alpha_{-n}^* = (\int f e_{-n}^* d\mu)^* = \int f^* e_n^* d\mu$, so that if $f = f^*$, then, indeed, $\alpha_n = \alpha_{-n}^*$. It is sometimes useful to know that this last argument applies to \mathbf{L}^1 as well as to \mathbf{L}^2 ; it follows that the constants are the only real functions in \mathbf{H}^1 .

Solution 27. Like the assertion (Problem 26) about real functions in \mathbf{H}^2 , the assertion is formally obvious. If f and g are in \mathbf{L}^2 , with Fourier expansions

$$f = \sum_n \alpha_n e_n, \quad g = \sum_m \beta_m e_m,$$

then

$$fg = \sum_n \sum_m \alpha_n \beta_m e_n e_m = \sum_k \left(\sum_n \alpha_n \beta_{k-n} \right) e_k.$$

If, moreover, f and g are in \mathbf{H}^2 , so that $\alpha_n = \beta_n = 0$ whenever $n < 0$, then $\sum_n \alpha_n \beta_{k-n} = 0$ whenever $k < 0$. Reason: for each term $\alpha_n \beta_{k-n}$, either $n < 0$, in which case $\alpha_n = 0$, or $n \geq 0$, in which case $k - n < 0$ and therefore $\beta_{k-n} = 0$.

The trouble with this argument is the assumption that the Fourier series of a product is equal to the formal product of the Fourier series of the factors. This assumption can be justified by appeal to the subsequence technique used in Solution 26. Alternatively the assumption can be avoided, as follows. The inner product (f, g^*) is equal to $\sum \alpha_n \beta_{-n}$ (by Parseval, and by the results of Solution 26 on the Fourier coefficients of complex conjugates); in other words, the 0-th Fourier coefficient of fg is given by

$$\int fg d\mu = \sum_n \alpha_n \beta_{-n}.$$

Apply this result with g replaced by the product ge_k^* . Since the Fourier coefficients γ_n of ge_k^* are given by

$$\gamma_n = \int ge_k^* e_n^* d\mu = \int ge_{k+n}^* d\mu = \beta_{k+n},$$

it follows that

$$\int fge_k^* d\mu = \sum_n \alpha_n \gamma_{-n} = \sum_n \alpha_n \beta_{k-n}.$$

It is an immediate corollary of this result that if $f \in \mathbf{H}^\infty$ and $g \in \mathbf{H}^2$, then $fg \in \mathbf{H}^2$. It is true also that if $f \in \mathbf{H}^\infty$ and $g \in \mathbf{H}^1$, then $fg \in \mathbf{H}^1$, but the proof requires one additional bit of analytic complication.

Solution 28. If $\varphi(z) = \sum_{n=0}^{\infty} \alpha_n z^n$ for $|z| < 1$, then $\varphi_r(z) = \sum_{n=0}^{\infty} \alpha_n r^n z^n$ for $0 < r < 1$ and $|z| = 1$. Since, for each fixed r , the latter series converges uniformly on the unit circle, it converges in every other useful sense; it follows, in particular, that the Fourier series expansion of φ_r in \mathbf{L}^2 is $\sum_{n=0}^{\infty} \alpha_n r^n e_n$, and hence that $\varphi_r \in \mathbf{H}^2$.

Since $\|\varphi_r\|^2 = \sum_{n=0}^{\infty} |\alpha_n|^2 r^{2n}$, the second (and principal) assertion reduces to this: if $\beta = \sum_{n=0}^{\infty} |\alpha_n|^2$ and $\beta_r = \sum_{n=0}^{\infty} |\alpha_n|^2 r^{2n}$, then a necessary and sufficient condition that $\beta < \infty$ is that the β_r 's ($0 < r < 1$) be bounded. In one direction the result is trivial; since $\beta_r \leq \beta$ for all r , it follows that if $\beta < \infty$, then the β_r 's are bounded. Suppose now, conversely, that $\beta_r \leq \gamma$ for all r . It follows that for each positive integer k ,

$$\begin{aligned} \sum_{n=0}^k |\alpha_n|^2 &= \left(\sum_{n=0}^k |\alpha_n|^2 - \sum_{n=0}^k |\alpha_n|^2 r^{2n} \right) + \sum_{n=0}^k |\alpha_n|^2 r^{2n} \\ &\leq \sum_{n=0}^k |\alpha_n|^2 (1 - r^{2n}) + \beta_r \\ &\leq \sum_{n=0}^k |\alpha_n|^2 (1 - r^{2n}) + \gamma. \end{aligned}$$

For k fixed, choose r so that $\sum_{n=0}^k |\alpha_n|^2 (1 - r^{2n}) < 1$; this can be done because the finite sum is a polynomial in r (and hence continuous) that

vanishes when $r = 1$. Conclusion: $\sum_{n=0}^k |\alpha_n|^2 \leq 1 + \gamma$ for all k , and this implies that $\sum_{n=0}^{\infty} |\alpha_n|^2 \leq 1 + \gamma$.

Solution 29. Start with an arbitrary infinite-dimensional functional Hilbert space \mathbf{H} , over a set X say, and adjoin a point that acts like an unbounded linear functional. To be specific: let φ be an unbounded linear functional on \mathbf{H} (such a thing exists because \mathbf{H} is infinite-dimensional), and write $X^+ = X \cup \{\varphi\}$. Let \mathbf{H}^+ be the set (pointwise vector space) of all those functions f^+ on X^+ whose restriction to X , say f , is in \mathbf{H} , and whose value at φ is equal to $\varphi(f)$. (Equivalently: extend each f in \mathbf{H} to a function f^+ on X^+ by writing $f^+(\varphi) = \varphi(f)$.) If f^+ and g^+ are in \mathbf{H}^+ , with restrictions f and g , write $(f^+, g^+) = (f, g)$. (Equivalently: define the inner product of the extensions of f and g to be equal to the inner product of f and g .) The vector space \mathbf{H}^+ with this inner product is isomorphic to \mathbf{H} with its original inner product (e.g., via the restriction mapping), and, consequently, \mathbf{H}^+ is a Hilbert space of functions. Since $\varphi \in X^+$ and $f^+(\varphi) = \varphi(f)$ for all f in \mathbf{H} , and since φ is not bounded, it follows that $|\varphi(f)|$ can be large for unit vectors f , and therefore that $|f^+(\varphi)|$ can be large for unit vectors f^+ .

Solution 30. If \mathbf{H} is a functional Hilbert space, over a set X say, with orthonormal basis $\{e_j\}$ and kernel function K , write $K_y(x) = K(x, y)$, and, for each y in X , consider the Fourier expansion of K_y :

$$K_y = \sum_j (K_y, e_j) e_j = \sum_j e_j(y) {}^* e_j.$$

Parseval's identity implies that

$$K(x, y) (K_y, K_z) = \sum_j e_j(x) e_j(y) {}^* e_j(z).$$

In \mathbf{A}^2 the functions e_n defined by

$$e_n(z) = \sqrt{(n+1)/\pi} \cdot z^n \quad \text{for } |z| < 1 \quad (n = 0, 1, 2, \dots)$$

form an orthonormal basis (see Problem 25); it follows (by the result

just obtained) that the kernel function K of \mathbf{A}^2 is given by

$$K(x, y) = \frac{1}{\pi} \sum_{n=0}^{\infty} (n+1) x^n y^{*n}.$$

(Note that x and y here are complex numbers in the open unit disc.) Since $\sum_{n=0}^{\infty} (n+1) z^n = 1/(1-z)^2$ when $|z| < 1$ (discover this by integrating the left side, or verify it by expanding the right), it follows that

$$K(x, y) = \frac{1}{\pi} \frac{1}{(1 - xy^*)^2}.$$

As for $\tilde{\mathbf{H}}^2$: by definition it consists of the functions \tilde{f} on the unit disc that correspond to the elements f of \mathbf{H}^2 . If $f = \sum_{n=0}^{\infty} \alpha_n e_n$ and if $|y| < 1$, then $\tilde{f}(y) = \sum_{n=0}^{\infty} \alpha_n y^n$ and consequently $\tilde{f}(y) = (f, K_y)$, where $K_y = \sum_{n=0}^{\infty} y^{*n} e_n$. This proves two things at once: it proves that $\tilde{f} \rightarrow \tilde{f}(y)$ is a bounded linear functional (so that $\tilde{\mathbf{H}}^2$ is a functional Hilbert space), and it proves that the kernel function of $\tilde{\mathbf{H}}^2$ is given by

$$K(x, y) = \sum_{n=0}^{\infty} x^n y^{*n}, \quad |x| < 1, |y| < 1.$$

In closed form: $K(x, y) = 1/(1 - xy^*)$.

Solution 31. Let K be the kernel function of $\tilde{\mathbf{H}}^2$ (see Solution 30). If $f_n \rightarrow f$ in \mathbf{H}^2 , and if $|y| < 1$, then

$$|\tilde{f}_n(y) - \tilde{f}(y)| = |(f_n - f, K_y)| \leq \|f_n - f\| \cdot \|K_y\|.$$

Since

$$\|K_y\|^2 = \sum_{n=0}^{\infty} |y|^{2n} = \frac{1}{1 - |y|^2},$$

it follows that if $|y| \leq r$, then

$$|\tilde{f}_n(y) - \tilde{f}(y)| \leq \|f_n - f\| \cdot \frac{1}{1 - r^2}.$$

Solution 32. The function \tilde{f} determines f — but how? Taylor and Fourier expansions do not reveal much about such structural properties as boundedness. The most useful way to approach the problem is to prove that the values of f (on the unit circle) are, in some sense, limits of the values of \tilde{f} (on the unit disc). For this purpose, write

$$f_r(z) = \tilde{f}(rz), \quad 0 < r < 1, \quad |z| = 1.$$

The functions f_r are in \mathbf{H}^2 (see Problem 28); the assertion is that $f_r \rightarrow f$ (as $r \rightarrow 1$) in the sense of convergence in the norm of \mathbf{H}^2 . (The boundedness of \tilde{f} is not relevant yet.)

To prove the assertion, recall that if $f = \sum_{n=0}^{\infty} \alpha_n e_n$, then $f_r = \sum_{n=0}^{\infty} \alpha_n r^n e_n$, and that, consequently,

$$\|f - f_r\|^2 = \sum_{n=0}^{\infty} |\alpha_n|^2 (1 - r^n)^2.$$

It follows that for each positive integer k

$$\|f - f_r\|^2 \leq \sum_{n=0}^k |\alpha_n|^2 (1 - r^n)^2 + \sum_{n=k+1}^{\infty} |\alpha_n|^2.$$

The desired result ($\|f - f_r\|$ is small when r is near to 1) is now easy: choose k large enough to make the second summand small (this is independent of r), and then choose r near enough to 1 to make the first summand small.

Since convergence in \mathbf{L}^2 implies the existence of subsequences converging almost everywhere, it follows that $f_{r_n} \rightarrow f$ almost everywhere for a suitable subsequence $\{r_n\}$, $r_n \rightarrow 1$; the assertion about the boundedness of f is an immediate consequence.

The assertion $f_r \rightarrow f$ is true in a sense different from (better than ?) convergence in the mean of order 2; in fact, it is true that $f_r \rightarrow f$ almost everywhere. This says, in other words, that if a point z in the disc tends radially to a boundary point z_0 , then the function value $\tilde{f}(z)$ tends to $f(z_0)$, for almost every z_0 . The result can be strengthened; radial convergence, for instance, can be replaced by non-tangential convergence. These analytic delicacies are at the center of the stage for some parts

of mathematics; in the context of Hilbert space, norm convergence is enough.

Solution 33. If $f \in \mathbf{H}^\infty$, then \tilde{f} is bounded, and, in fact, $\|\tilde{f}\|_\infty = \|f\|_\infty$.

(The norms are the supremum of \tilde{f} on the disc and the essential supremum of f on the boundary.)

Proof. Consider the following two assertions. (1) If $f \in \mathbf{L}^\infty$, and, say, $|f| \leq 1$, then there exists a sequence $\{f_n\}$ of trigonometric polynomials converging to f in the norm of \mathbf{L}^2 , such that $|f_n| \leq 1$ for all n ; if, moreover, f is in \mathbf{H}^∞ , then so are the f_n 's. (2) If p is a polynomial and if $|p(z)| \leq 1$ whenever $|z| = 1$, then $|p(z)| \leq 1$ whenever $|z| < 1$. Both these assertions are known parts of analysis: (1) is a consequence of Fejér's theorem about the Cesaro convergence of Fourier series, and (2) is the maximum modulus principle for polynomials. Of the two assertions, (2) seems to be far better known. In any case, (2) will be used below without any further apology; (1) will be used also, but after that it will be buttressed by the outline of a proof.

It is easy to derive the boundedness conclusion about \tilde{f} from the two assertions of the preceding paragraph. Given f in \mathbf{H}^∞ , assume (this is just a matter of normalization) that $|f| \leq 1$, and, using (1), find trigonometric polynomials f_n such that $|f_n| \leq 1$ and such that $f_n \rightarrow f$ in the norm of \mathbf{L}^2 . Since, according to (1), the f_n 's themselves are (can be chosen to be) in \mathbf{H}^∞ , it follows that their extensions \tilde{f}_n into the interior are polynomials. Since $f_n \rightarrow f$ in the norm of \mathbf{H}^2 , it follows from Problem 31 that $\tilde{f}_n(z) \rightarrow \tilde{f}(z)$ whenever $|z| < 1$. By (2), $|\tilde{f}_n(z)| \leq 1$ for all n and all z . Conclusion: $|\tilde{f}(z)| \leq 1$ for all z .

The inequality $\|\tilde{f}\|_\infty \leq \|f\|_\infty$ is implicit in the proof above. To get the reverse inequality, use Solution 32 ($f_r \rightarrow f$ as $r \rightarrow 1$).

It remains to look at a proof of (1). If $f = \sum_n \alpha_n e_n$, write $s_k = \sum_{j=-k}^{+k} \alpha_j e_j$ ($k = 0, 1, 2, \dots$). Clearly $s_k \rightarrow f$ in \mathbf{L}^2 , but this is not good enough; it does not yield the necessary boundedness results. (If $|f| \leq 1$, it does not follow that $|s_k| \leq 1$.) The remedy is to consider the averages

$$t_n = \frac{1}{n} \sum_{k=0}^{n-1} s_k \quad (n = 1, 2, 3, \dots).$$

(Note that if $f \in \mathbf{H}^2$, then so are the t_n 's.) Clearly $t_n \rightarrow f$ in \mathbf{L}^2 . (In fact it is known that $t_n \rightarrow f$ almost everywhere, but the proof is non-trivial, and, fortunately, the fact is not needed here.) This turns out to be good enough: if $|f| \leq 1$, then it does follow that $|t_n| \leq 1$.

For the proof, write $D_k = \sum_{j=-k}^{+k} e_j$ ($k = 0, 1, 2, \dots$), and $K_n = (1/n) \sum_{k=0}^{n-1} D_k$ ($n = 1, 2, 3, \dots$); the sequences of functions D_k and K_n are known as the Dirichlet and the Fejér kernels, respectively. Since $\int D_k d\mu = \int e_0 d\mu = 1$, it follows that $\int K_n d\mu = 1$. The principal property of the K_n 's is that their values are real, and in fact positive. This is proved by computation. For $z = 1$, it is obvious; for $z \neq 1$ (but, of course, $|z| = 1$) write $D_k(z) = 1 + 2 \operatorname{Re} \sum_{j=1}^k z^j$ ($k = 1, 2, 3, \dots$), and apply the formula for the sum of a geometric series to get

$$D_k(z) = 2 \operatorname{Re} \left(\frac{z^k - z^{k+1}}{|1 - z|^2} \right).$$

(Computational trick: note that if $|z| = 1$, then $|1 - z|^2 = 2 \operatorname{Re}(1 - z)$.) Substitute into the expression for K_n , observe that the sum telescopes, and get

$$K_n(z) = \frac{2}{n} \operatorname{Re} \frac{1 - z^n}{|1 - z|^2}.$$

This makes it obvious that K_n is real. Since, moreover, $\operatorname{Re} z^n \leq 1$ (recall that $|z^n| = |z|^n = 1$), i.e., $1 - \operatorname{Re} z^n \geq 0$, it follows that $K_n(z) \geq 0$, as asserted.

To apply these results to f , note that

$$\begin{aligned} s_k(z) &= \sum_{j=-k}^{+k} \int f(y) y^{*j} z^j d\mu(y) \\ &= \int D_k(y^* z) f(y) d\mu(y) \\ &= \int D_k(y) f(y^* z) d\mu(y), \end{aligned}$$

and hence that

$$t_n(z) = \int K_n(y) f(y^*z) d\mu(y);$$

this implies that if $|f| \leq 1$, then

$$|t_n(z)| \leq \int K_n(y) |f(y^*z)| d\mu(y) \leq \int K_n(y) d\mu(y) = 1.$$

The proof is over. Here is one more technical remark that is sometimes useful: under the assumptions of (1) it makes no difference whether the convergence in the conclusion is in the norm or almost everywhere. Reason: if it is in the norm, then a subsequence converges almost everywhere; if it is almost everywhere, then, by the Lebesgue bounded convergence theorem, it is also in the norm.

Solution 34. If $f \in \mathbf{H}^\infty$, $g \in \mathbf{H}^2$, and $h = fg$, then $\tilde{h} = \tilde{f}\tilde{g}$.

Proof. The trouble with the question as phrased is that it is easier to answer than to interpret. If f and g are in \mathbf{H}^2 and $h = fg$, then h is not necessarily in \mathbf{H}^2 , and hence the definition given in Problem 28 does not apply to h ; no such thing as \tilde{h} is defined. The simplest way out is to assume that one factor (say f) is bounded; in that case Problem 27 shows that $h \in \mathbf{H}^2$, and the question makes sense. (There is another way out, namely to note that $h \in \mathbf{H}^1$, by Problem 27, and to extend the process of passage into the interior to \mathbf{H}^1 . This way leads to some not overwhelming but extraneous analytic difficulties.) Once the question makes sense, the answer is automatic from Solution 27; the result there is that the Fourier coefficients of h are expressed in terms of those of f and g in exactly the same way as the Taylor coefficients of $\tilde{f} \cdot \tilde{g}$ are expressed in terms of those of \tilde{f} and \tilde{g} . In other words: formal multiplication applies to both Fourier and Taylor series, and, consequently, the mapping from one to the other is multiplicative.

Solution 35. In order to motivate the construction of, say, f from u , it is a good idea to turn the process around and to study the way u is obtained from f . Suppose therefore that $f \in \mathbf{H}^2$ with Fourier expansion $f = \sum_{n=0}^{\infty} \alpha_n e_n$, and write $u = \operatorname{Re} f$. Since $|u| \leq |f|$, the function u

is in L^2 . If the Fourier expansion of u is $u = \sum_n \xi_n e_n$, then (see Solution 26)

$$\begin{aligned} u &= \frac{1}{2}(f + f^*) = \frac{1}{2}\left(\sum_{n \geq 0} \alpha_n e_n + \sum_{n \geq 0} \alpha_n^* e_{-n}\right) \\ &= \operatorname{Re} \alpha_0 + \sum_{n > 0} \frac{1}{2} \alpha_n e_n + \sum_{n < 0} \frac{1}{2} \alpha_{-n}^* e_n, \end{aligned}$$

and therefore

$$\xi_0 = \operatorname{Re} \alpha_0 \quad \text{and} \quad \xi_n = \begin{cases} \frac{1}{2} \alpha_n & (n > 0), \\ \frac{1}{2} \alpha_{-n}^* & (n < 0). \end{cases}$$

It is now clear how to go in the other direction. Given $u = \sum_n \xi_n e_n$, with $\xi_n = \xi_{-n}^*$, and, in particular, with ξ_0 real, write

$$\alpha_0 = \xi_0 \quad \text{and} \quad \alpha_n = 2\xi_n = 2\xi_{-n}^* = \xi_n + \xi_{-n}^* \quad (n > 0).$$

Since the sequence of α 's is square-summable, an element f of \mathbf{H}^2 is defined by $f = \sum_{n \geq 0} \alpha_n e_n$. Write

$$f = Du$$

(D for Dirichlet); then $\operatorname{Re} Du = u$ for every real u in L^2 . It is not quite true that $D \operatorname{Re} f = f$ for every f in \mathbf{H}^2 , but it is almost true; the difference $f - D \operatorname{Re} f$ is a purely imaginary constant that can be prescribed arbitrarily.

As for the formulation in terms of v : given u , put $v = \operatorname{Im} Du$. Since $\operatorname{Im} Du = -\operatorname{Re}(iDu)$, it is easy to get explicit expressions for the Fourier coefficients of v . If, as above, $u = \sum_n \xi_n e_n$ and $f = Du = \sum_{n \geq 0} \alpha_n e_n$, then

$$v = \operatorname{Im} f = \frac{i}{2}(f^* - f) = \frac{i}{2}\left(\sum_{n \geq 0} \alpha_n^* e_{-n} - \sum_{n \geq 0} \alpha_n e_n\right).$$

If $v = \sum \eta_n e_n$, then

$$\eta_0 = \operatorname{Im} \xi_0 \quad \text{and} \quad \eta_n = \begin{cases} -\frac{i}{2} \cdot 2\xi_n = -i\xi_n & (n > 0), \\ \frac{i}{2} \cdot 2\xi_n = i\xi_n & (n < 0). \end{cases}$$

If $\text{Im } \xi_0 = 0$, the result can be concisely expressed:

$$\eta_n = (-i \operatorname{sgn} n) \xi_n$$

for all n .

As far as L^2 functions on the unit circle are concerned, these algebraic trivialities are all there is to the Dirichlet problem on the unit disc. The formal expression for v in terms of u makes sense even when u is not necessarily real, and the terminology (conjugate function, Hilbert transform) remains the same. It is important to note that the Hilbert transform of a bounded function need not be bounded, or, in other words (consider extensions to the interior) that unbounded analytic functions can have bounded real parts. Standard example: $f(z) = i \log(1 - z)$.

Chapter 4. Infinite matrices

Solution 36. Since \mathbf{H} is the direct sum of separable subspaces that reduce A , there is no loss of generality in assuming that \mathbf{H} is separable in the first place. This comment, while only feebly used in the proof, eliminates the discomfort of having to worry about the pathology of the uncountable.

There is a tempting attack on the proof that is doomed to failure but is illuminating just the same. Let e_1 be an arbitrary unit vector. Since e_1 and Ae_1 span a subspace of dimension at most 2, it follows that, unless $\dim \mathbf{H} = 1$, there exists a unit vector e_2 orthogonal to e_1 such that $Ae_1 \in \mathbf{V}\{e_1, e_2\}$. Since e_1 , e_2 , and Ae_2 span a subspace of dimension at most 3, it follows that, unless $\dim \mathbf{H} = 2$, there exists a unit vector e_3 orthogonal to e_1 and e_2 such that $Ae_2 \in \mathbf{V}\{e_1, e_2, e_3\}$. An inductive repetition of this argument yields an orthonormal sequence $\{e_1, e_2, e_3, \dots\}$ (which is finite only in case $\dim \mathbf{H} < \infty$) such that $Ae_n \in \mathbf{V}\{e_1, \dots, e_n, e_{n+1}\}$. The finite-dimensional case is transparent and, from the present point of view, uninteresting. In the infinite-dimensional case $(Ae_j, e_i) = 0$ when $i > j + 1$, and everything seems to be settled. There is a difficulty, however: there is no reason to suppose that the e_n 's form a basis. If they do not, then the process of embedding them into an orthonormal basis may ruin the column-finiteness of the matrix. That is, it could happen that for some e orthogonal to all the e_n 's infinitely many of the Fourier coefficients (Ae, e_i) are different from 0. If A happens to be Hermitian, then no such troubles can arise. The span of the e_n 's is, in any case, invariant under A , and hence, for Hermitian A , reduces A ; it follows that when the e_n 's are embedded into an orthonormal basis, the new matrix elements do not interfere with the old columns. This proof, in the Hermitian case, shows more than was promised: it shows that every Hermitian operator has a Jacobi matrix. (A Jacobi matrix is a Hermitian matrix all whose non-zero entries are on either the main diagonal or its two neighboring diagonals. Some authors require also that the matrix be irreducible, i.e., that none of the elements on the diagonals next to the main one vanish.) Indeed: if $(Ae_j, e_i) = 0$ when $i > j + 1$, then $(e_j, Ae_i) = 0$ when $i > j + 1$; the argument is com-

pleted by inductively enlarging the e_n 's to an orthonormal basis selected the same way as the e_n 's were.

In the non-Hermitian case the argument has to be refined (and the conclusion weakened to the form originally given) as follows. Let $\{f_1, f_2, f_3, \dots\}$ be an orthonormal basis for \mathbf{H} . Put $e_1 = f_1$. Find a unit vector e_2 orthogonal to e_1 such that $Ae_1 \in \mathbf{V}\{e_1, e_2\}$. (Once again restrict attention to the infinite-dimensional case.) Next find a unit vector e_3 orthogonal to e_1 and e_2 such that $f_2 \in \mathbf{V}\{e_1, e_2, e_3\}$, and then find a unit vector e_4 orthogonal to e_1, e_2 , and e_3 such that $Ae_2 \in \mathbf{V}\{e_1, e_2, e_3, e_4\}$. Continue in this way, catching alternately one of the f_n 's and the next as yet uncaught Ae_n . The selection of the needed new e is always possible. The general lemma is this: for each finite-dimensional subspace \mathbf{M} and for each vector g , there exists a unit vector e orthogonal to \mathbf{M} such that $g \in \mathbf{M} \vee \{e\}$. Conclusion: the sequence $\{e_1, e_2, e_3, \dots\}$ is orthonormal by construction; it forms a basis because its span contains each f_n ; and it has the property that for each n there is an i_n (calculable in case of need) such that $Ae_n \in \mathbf{V}\{e_1, \dots, e_{i_n}\}$. This last condition implies that $(Ae_j, e_i) = 0$ whenever $i > i_j$, and the proof is complete.

Solution 37. If $\langle \xi_0, \xi_1, \xi_2, \dots \rangle$ is a finitely non-zero sequence of complex numbers (i.e., $\xi_n = 0$ for n sufficiently large), then

$$\begin{aligned}
 \sum_i \left| \sum_j \alpha_{ij} \xi_j \right|^2 &= \sum_i \left| \sum_j (\sqrt{\alpha_{ij}} \sqrt{p_j}) \left(\frac{\sqrt{\alpha_{ij}} \xi_j}{\sqrt{p_j}} \right) \right|^2 \\
 &\leq \sum_i \left(\sum_j \alpha_{ij} p_j \right) \left(\sum_j \frac{\alpha_{ij} |\xi_j|^2}{p_j} \right) \\
 &\leq \sum_i \gamma p_i \sum_j \frac{\alpha_{ij} |\xi_j|^2}{p_j} \\
 &= \gamma \sum_j \frac{|\xi_j|^2}{p_j} \sum_i \alpha_{ij} p_i \\
 &\leq \gamma \sum_j \frac{|\xi_j|^2}{p_j} \cdot \beta p_j = \beta \cdot \gamma \sum_j |\xi_j|^2.
 \end{aligned}$$

These inequalities imply that the operator A on l^2 defined by

$$A \langle \xi_0, \xi_1, \xi_2, \dots \rangle = \langle \sum_j \alpha_{0j} \xi_j, \sum_j \alpha_{1j} \xi_j, \sum_j \alpha_{2j} \xi_j, \dots \rangle$$

satisfies the conditions.

Solution 38. The result is a corollary of Problem 37. For the proof, apply Problem 37 with $p_i = 1/\sqrt{i + \frac{1}{2}}$. Since the Hilbert matrix is symmetric, it is sufficient to verify one of the two inequalities (with $\beta = \gamma = \pi$). The verification depends on elementary calculus, as follows:

$$\begin{aligned} \sum_i \alpha_{ij} p_i &= \sum_i \frac{1}{(i + \frac{1}{2} + j + \frac{1}{2}) \sqrt{i + \frac{1}{2}}} \\ &< \int_0^\infty \frac{dx}{(x + j + \frac{1}{2}) \sqrt{x}} \\ &= 2 \int_0^\infty \frac{du}{u^2 + j + \frac{1}{2}} \\ &= \frac{2}{\sqrt{j + \frac{1}{2}}} \int_0^\infty \frac{du}{u^2 + 1} = \frac{\pi}{\sqrt{j + \frac{1}{2}}}. \end{aligned}$$

Chapter 5.

Boundedness and invertibility

Solution 39. Let $\{e_1, e_2, e_3, \dots\}$ be an orthonormal basis for a Hilbert space \mathbf{H} , and find a Hamel basis for \mathbf{H} that contains each e_n . Let f_0 be an arbitrary but fixed element of that Hamel basis distinct from each e_n (see Solution 5). A unique linear transformation A is defined on \mathbf{H} by the requirement that $Af_0 = f_0$ and $Af = 0$ for all other elements of the selected basis; in particular, $Ae_n = 0$ for all n . If A were bounded, then its vanishing on each e_n would imply that $A = 0$. This solves the first and the third parts of the problem.

For the second part, choose an arbitrary but fixed positive integer k and define an operator A (depending on k) by

$$Af = (f, e_1 + \dots + e_k)e_1.$$

It follows that $Ae_n = e_1$ or 0 according as $n \leq k$ or $n > k$, and hence that $\|Ae_n\| \leq 1$ for all n . Since (easy computation) $A^*f = (f, e_1)(e_1 + \dots + e_k)$ for all f , so that, in particular, $A^*e_1 = e_1 + \dots + e_k$, it follows that

$$\|A\| = \|A^*\| \geq \|A^*e_1\| = \|e_1 + \dots + e_k\| = \sqrt{k}.$$

A simple alternative way to say all this is to describe the matrix of A with respect to the basis $\{e_1, e_2, e_3, \dots\}$: all the entries are 0 except the first k entries in the first row, which are 1 's.

Solution 40. The conclusion can be obtained from two successive applications of the principle of uniform boundedness for vectors (Problem 20). Suppose that \mathbf{Q} is a weakly bounded set of bounded linear transformations from \mathbf{H} to \mathbf{K} , and that, specifically, $|(Af, g)| \leq \alpha(f, g)$ for all A in \mathbf{Q} . Fix an arbitrary vector g_0 and write $\mathbf{T}_0 = \{A^*g_0: A \in \mathbf{Q}\}$. Since

$$|(f, A^*g_0)| = |(Af, g_0)| \leq \alpha(f, g_0),$$

the set \mathbf{T}_0 is weakly bounded in \mathbf{H} , and therefore there exists a constant $\beta(g_0)$ such that $\|A^*g_0\| \leq \beta(g_0)$ for all A in \mathbf{Q} .

Next, write $\mathbf{T} = \{Af: A \in \mathbf{Q}, f \in \mathbf{B}\}$, where \mathbf{B} is the unit ball of \mathbf{H} . Since

$$|(g, Af)| = |(A^*g, f)| \leq \beta(g) \cdot \|f\| \leq \beta(g),$$

the set \mathbf{T} is weakly bounded in \mathbf{K} , and therefore there exists a constant γ such that

$$\|Af\| \leq \gamma$$

whenever $A \in \mathbf{Q}$ and $f \in \mathbf{B}$. This implies that

$$\|A\| \leq \gamma,$$

and the proof is complete.

Solution 41. It is sufficient to prove that A^* is invertible. The range of A^* is dense in \mathbf{H} (because the kernel of A is trivial), and, consequently, it is sufficient to prove that A^* is bounded from below. This means that $\|A^*g\| \geq \delta \|g\|$ for some δ (and all g in \mathbf{K}). To prove it, it is sufficient to prove that if $\|A^*g\| = 1$, then $\|g\| \leq 1/\delta$ for some δ . Caution: the last reduction uses the assumption that the kernel of A^* is trivial, which is true because the range of A is dense in \mathbf{K} . (The full force of the assumption that A maps \mathbf{H} onto \mathbf{K} will be used in a moment.) To see the difficulty, consider the transformation 0 in the role of A^* : for it the implication from $\|A^*g\| = 1$ to $\|g\| \leq 1/\delta$ is vacuously valid. Summary: it is sufficient to prove that if $\mathbf{S} = \{h: \|A^*h\| = 1\}$, then the set \mathbf{S} is bounded, and that can be done by proving that it is weakly bounded. To do that, take g in \mathbf{K} , find f in \mathbf{H} so that $Af = g$, and observe that

$$|(g, h)| = |(Af, h)| = |(f, A^*h)| \leq \|f\|$$

for all h in \mathbf{S} . The proof is complete.

Solution 42. If $\dim \mathbf{K} < \dim \mathbf{H}$, then there is no loss of generality in assuming that $\mathbf{K} \subset \mathbf{H}$. Suppose, accordingly, that A is an operator on \mathbf{H} with range included in \mathbf{K} ; it is to be proved that $\ker A$ is not trivial.

Assume that $\dim \mathbf{K}$ is infinite; this assumption excludes trivial cases only. Let $\{f_i\}$ and $\{g_j\}$ be orthonormal bases of \mathbf{H} and \mathbf{K} , respectively. Each A^*g_j can be expanded in terms of countably many f 's; the assumed inequality between the dimensions of \mathbf{H} and \mathbf{K} implies the existence of an i such that $(f_i, A^*g_j) = 0$ for all j . Since $(f_i, A^*g_j) = (Af_i, g_j)$, it follows that Af_i is orthogonal to each g_j and therefore to \mathbf{K} . Since, however, the range of A is included in \mathbf{K} , it follows that $Af_i = 0$.

Consider next the statement about equality. If $\dim \mathbf{H}$ is finite, all is trivial. If $\dim \mathbf{H}$ is infinite, then a set of cardinal number $\dim \mathbf{H}$ is dense in \mathbf{H} . (Use rational linear combinations; cf. Solution 11.) It follows that a set of cardinal number $\dim \mathbf{H}$ is dense in \mathbf{K} , and this implies that $\dim \mathbf{K} \leq \dim \mathbf{H}$.

The proof above is elementary, but, for a statement that is completely natural, it is not at all completely obvious. (It is due, incidentally, to G. L. Weiss; cf. Halmos-Lumer [1954].) There is a quick proof, which, however, is based on a non-trivial theory (polar decomposition). It goes as follows. If A is a one-to-one linear transformation from \mathbf{H} into \mathbf{K} , with polar decomposition UP (see Problem 105), then, since $\ker A$ is $\{0\}$, it follows that U is an isometry. As for the case of equality: if $\text{ran } A$ is dense in \mathbf{K} , then $\text{ran } U$ is equal to \mathbf{K} .

Solution 43. Observe first that no non-zero vector in the range of P is annihilated by Q . Indeed, if $Pf = f$ and $Qf = 0$, then $\|Pf - Qf\| = \|f\|$, and therefore $f \neq 0$ would imply $\|P - Q\| \geq 1$. From this it follows that the restriction of Q to the range of P is a one-to-one bounded linear transformation from that range into the range of Q , and therefore that the rank of P is less than or equal to the rank of Q (Problem 42). The conclusion follows by symmetry.

Solution 44. Suppose that A is a linear transformation from \mathbf{H} to \mathbf{K} , and suppose, first, that A is bounded. Let $\{\langle f_n, Af_n \rangle\}$ be a sequence of vectors in the graph of A converging to something, say $\langle f_n, Af_n \rangle \rightarrow \langle f, g \rangle$. Since $f_n \rightarrow f$ and A is continuous, it follows that $Af_n \rightarrow Af$; since at the same time $Af_n \rightarrow g$, it follows that $g = Af$, and hence that $\langle f, g \rangle$ is in the graph of A .

The proof of the converse is less trivial; it is a trick based on Problem 41. Let \mathbf{G} be the graph of A , and consider the linear transformation B

from \mathbf{G} to \mathbf{H} defined by $B\langle f, Af \rangle = f$. Clearly B is a one-to-one mapping from \mathbf{G} onto \mathbf{H} ; since

$$\|B\langle f, Af \rangle\|^2 = \|f\|^2 \leq \|f\|^2 + \|Af\|^2 = \|\langle f, Af \rangle\|^2,$$

it follows that B is bounded. Since \mathbf{G} is a closed subset of the complete space $\mathbf{H} \oplus \mathbf{K}$, it is complete, and all is ready for an application of Problem 41; the conclusion is that B is invertible. Equivalently the conclusion says that the mapping B^{-1} from \mathbf{H} into \mathbf{G} , defined by $B^{-1}f = \langle f, Af \rangle$ is a bounded linear transformation. This means, by definition, that

$$\|f\|^2 + \|Af\|^2 \leq \alpha \|f\|^2$$

for some α (and all f in \mathbf{H}); the boundedness of A is an immediate consequence.

It is worth remarking that the derivation of the result from Problem 41 is reversible; the assertion there is a special case of the closed graph theorem. This, of course, is not an especially helpful comment for someone who wants to know how to prove the closed graph theorem, and not just how to bounce back and forth between it and a reformulation.

Solution 45. (a) *On incomplete inner-product spaces unbounded symmetric transformations do exist.* (b) *On a Hilbert space, every symmetric linear transformation is bounded.*

Proof. (a) Let \mathbf{H} be the complex vector space of all finitely non-zero infinite sequences. That is, an element of \mathbf{H} is a sequence $\langle \xi_1, \xi_2, \xi_3, \dots \rangle$, with $\xi_n = 0$ for all sufficiently large n ; the “sufficiently large” may vary with the sequence. Define inner product in \mathbf{H} the natural way: if $f = \langle \xi_1, \xi_2, \xi_3, \dots \rangle$ and $g = \langle \eta_1, \eta_2, \eta_3, \dots \rangle$, write $(f, g) = \sum_{n=1}^{\infty} \xi_n \eta_n^*$. Let A be the linear transformation that maps each sequence $\langle \xi_1, \xi_2, \xi_3, \dots \rangle$ onto $\langle \xi_1, 2\xi_2, 3\xi_3, \dots \rangle$; in an obvious manner A is determined by the diagonal matrix whose sequence of diagonal terms is $\langle 1, 2, 3, \dots \rangle$. The linear transformation A is symmetric; indeed both (Af, g) and (f, Ag) are equal to $\sum_{n=1}^{\infty} n \xi_n \eta_n^*$. The linear transformation A is not bounded; indeed if $\{f_n\}$ is the sequence whose n -th term is 1 and all other terms are 0, then $\|f_n\| = 1$ and $\|Af_n\| = n$.

(b) This is an easy consequence of the closed graph theorem. Indeed, if A is symmetric, and if $f_n \rightarrow f$ and $Af_n \rightarrow f'$, then, for all g ,

$$(f', g) = \lim_n (Af_n, g) = \lim_n (f_n, Ag) = (f, Ag) = (Af, g),$$

and therefore $f' = Af$; this proves that A is closed, and hence that A is bounded.

Chapter 6.

Multiplication operators

Solution 46. If A is a diagonal operator, with $Ae_j = \alpha_j e_j$, then

$$|\alpha_j| = \| \alpha_j e_j \| = \| Ae_j \| \leq \| A \| \cdot \| e_j \| = \| A \|,$$

so that $\{\alpha_j\}$ is bounded and $\sup_j |\alpha_j| \leq \| A \|$. The reverse inequality follows from the relations

$$\begin{aligned} \| A \sum_j \xi_j e_j \|^2 &= \| \sum_j \alpha_j \xi_j e_j \|^2 = \sum_j |\alpha_j \xi_j|^2 \\ &\leq (\sup_j |\alpha_j|)^2 \cdot \sum_j |\xi_j|^2 = (\sup_j |\alpha_j|)^2 \cdot \| \sum_j \xi_j e_j \|^2. \end{aligned}$$

Given a bounded family $\{\alpha_j\}$, define A by $A \sum_j \xi_j e_j = \sum_j \alpha_j \xi_j e_j$; the preceding computations imply that A is an operator. Clearly A is a diagonal operator, and the diagonal of A is exactly the sequence $\{\alpha_j\}$. The proof of uniqueness is implicit in the construction: via Fourier expansions the behavior of an operator on a basis determines its behavior everywhere.

Solution 47. If $\{\alpha_n\}$ is a sequence of complex scalars, such that $\sum_n |\alpha_n \xi_n|^2 < \infty$ whenever $\sum_n |\xi_n|^2 < \infty$, then $\{\alpha_n\}$ is bounded.

Proof. Expressed contrapositively, the assertion is this: if $\{\alpha_n\}$ is not bounded, then there exists a sequence $\{\xi_n\}$ such that $\sum_n |\xi_n|^2 < \infty$ but $\sum_n |\alpha_n \xi_n|^2 = \infty$. The construction is reasonably straightforward. If $\{\alpha_n\}$ is not bounded, then $|\alpha_n|$ takes arbitrarily large values. There is no loss of generality in assuming that $|\alpha_n| \geq n$; all it takes is a slight change of notation, and, possibly, the omission of some α 's. If, in that

case, $\xi_n = 1/\alpha_n$, $n = 1, 2, 3, \dots$, then

$$\sum_n |\xi_n|^2 \leq \sum_n \frac{1}{n^2} < \infty,$$

but $\sum_n |\alpha_n \xi_n|^2$ diverges.

Solution 48. The assertion is that if A is a diagonal operator with diagonal $\{\alpha_n\}$, then A and $\{\alpha_n\}$ are invertible together. Indeed, if $\{\beta_n\}$ is a bounded sequence such that $\alpha_n \beta_n = 1$ for all n , then the diagonal operator B with diagonal $\{\beta_n\}$ acts as the inverse of A . Conversely: if A is invertible, then $A^{-1}(\alpha_n e_n) = e_n$, so that

$$A^{-1}e_n = \frac{1}{\alpha_n} e_n;$$

since $\|A^{-1}e_n\| \leq \|A^{-1}\|$, this implies that the sequence $\{1/\alpha_n\}$ is bounded, and hence that the sequence $\{\alpha_n\}$ is invertible.

As for the spectrum: the assertion here is that $A - \lambda$ is invertible if and only if λ does not belong to the closure of the diagonal $\{\alpha_n\}$. (The purist has a small right to object. The diagonal is a *sequence* of complex numbers, and, therefore, not just a *set* of complex numbers; “the closure of the diagonal” does not make rigorous sense. The usage is an instance of a deservedly popular kind of abuse of language, unambiguous and concise; it would be a pity to let the purist have his way.) The assertion is equivalent to this: $\{\alpha_n - \lambda\}$ is bounded away from 0 if and only if λ is not in the closure of $\{\alpha_n\}$. Contrapositively: the sequence $\{\alpha_n - \lambda\}$ has 0 as a limit point if and only if the set $\{\alpha_n\}$ has λ as a cluster point. Since this is obvious, the proof is complete.

Solution 49. *If A is the multiplication induced by a bounded measurable function φ on a σ -finite measure space, then $\|A\| = \|\varphi\|_\infty$ (= the essential supremum of $|\varphi|$).*

Proof. Let μ be the underlying measure. It is instructive to see how far the proof can get without the assumption that μ is σ -finite; until

further notice that assumption will not be used. Since

$$\|Af\|^2 = \int |\varphi \cdot f|^2 d\mu \leq \|\varphi\|_\infty^2 \cdot \int |f|^2 d\mu = \|\varphi\|_\infty^2 \cdot \|f\|^2,$$

it follows that $\|A\| \leq \|\varphi\|_\infty$. In the proof of the reverse inequality a pathological snag is possible.

A sensible way to begin the proof is to note that if $\epsilon > 0$, then $|\varphi(x)| > \|\varphi\|_\infty - \epsilon$ on a set, say M , of positive measure. If f is the characteristic function of M , then

$$\|f\|^2 = \int_M 1 \cdot d\mu = \mu(M),$$

and

$$\|Af\|^2 = \int_M |\varphi|^2 d\mu \geq (\|\varphi\|_\infty - \epsilon)^2 \mu(M).$$

It follows that $\|Af\| \geq (\|\varphi\|_\infty - \epsilon)\|f\|$, and hence that $\|A\| \geq \|\varphi\|_\infty - \epsilon$; since this is true for all ϵ , it follows that $\|A\| \geq \|\varphi\|_\infty$. The proof is over, but it is wrong.

What is wrong is that M may have infinite measure. The objection may not seem very strong. After all, even if the measurable set $\{x: |\varphi(x)| \geq \|\varphi\|_\infty - \epsilon\}$ has infinite measure, the reasoning above works perfectly well if M is taken to be a measurable subset of finite positive measure. This is true. The difficulty, however, is that the measure space may be pathological enough to admit measurable sets of positive measure (in fact infinite measure) with the property that the measure of each of their measurable subsets is either 0 or ∞ . There is no way out of this difficulty. If, for instance, X consists of two points x_1 and x_2 , and if $\mu(\{x_1\}) = 1$ and $\mu(\{x_2\}) = \infty$, then $L^2(\mu)$ is the one-dimensional space consisting of all those functions on X that vanish at x_2 . If φ is the characteristic function of the singleton $\{x_2\}$, then $\|\varphi\|_\infty = 1$, but the norm of the induced multiplication operator is 0.

Conclusion: if the measure is locally finite (meaning that every measurable set of positive measure has a measurable subset of finite positive measure), then the norm of each multiplication is the essential

supremum of the multiplier; otherwise the best that can be asserted is the inequality $\|A\| \leq \|\varphi\|_\infty$. Every finite or σ -finite measure is locally finite. The most practical way to avoid excessive pathology with (usually) hardly any loss in generality is to assume σ -finiteness. If that is done, the solution (as stated above) is complete.

Solution 50. Measurability is easy. Since the measure is σ -finite, there exists an element f of L^2 that does not vanish anywhere; since $\varphi \cdot f$ is in L^2 , it is measurable, and, consequently, so is its quotient by f .

To prove boundedness, observe that

$$\|\varphi^n \cdot f\| = \|A^n f\| \leq \|A\|^n \|f\|$$

for every positive integer n . If $A = 0$, then $\varphi = 0$, and there is nothing to prove; otherwise write $\psi = \varphi/\|A\|$, and rewrite the preceding inequality in the form

$$\int |\psi|^{2n} |f|^2 d\mu \leq \int |f|^2 d\mu.$$

(Here μ is, of course, the given σ -finite measure.) From this it follows that if $f \neq 0$ on some set of positive measure, then $|\psi| \leq 1$ (i.e., $|\varphi| \leq \|A\|$) almost everywhere on that set. If f is chosen (as above) so that $f \neq 0$ almost everywhere, then the conclusion is that $|\varphi| \leq \|A\|$ almost everywhere.

This proof is quick, but a little too slick; it is not the one that would suggest itself immediately. A more natural (and equally quick) approach is this: to prove that $|\varphi| \leq \|A\|$ almost everywhere, let M be a measurable set of finite measure on which $|\varphi| > \|A\|$, and prove that M must have measure 0. Indeed, if f is the characteristic function of M , then either $f = 0$ almost everywhere, or

$$\|Af\|^2 = \int |\varphi \cdot f|^2 d\mu = \int_M |\varphi|^2 d\mu > \|A\|^2 \mu(M) = \|A\|^2 \|f\|^2;$$

the latter possibility is contradictory. The proof in the preceding paragraph has, however, an advantage other than artificial polish: unlike the more natural proof, it works in a certain curious but useful situation.

The situation is this: suppose that \mathbf{H} is a subspace of \mathbf{L}^2 , suppose that an operator A on \mathbf{H} is such that $Af = \varphi \cdot f$ for all f in \mathbf{H} , and suppose that \mathbf{H} contains a nowhere vanishing function. Conclusion, as before: φ is measurable and bounded (by $\|A\|$). Proof: as above.

Solution 51. *If φ is a complex-valued function such that $\varphi \cdot f \in \mathbf{L}^2$ (for a σ -finite measure) whenever $f \in \mathbf{L}^2$, then φ is essentially bounded.*

Proof. One way to proceed is to generalize the discrete (diagonal) construction (Solution 47). If φ is not bounded, then there exists a disjoint sequence $\{M_n\}$ of measurable sets of positive finite measure such that $\varphi(x) \geq n$ whenever $x \in M_n$. (There is no trouble in proving that φ is measurable; cf. Solution 50.) Define a function f as follows: if $x \in M_n$ for some n , then

$$f(x) = \frac{1}{\sqrt{\mu(M_n) \cdot \varphi(x)}};$$

otherwise $f(x) = 0$. Since

$$\begin{aligned} \int |f|^2 d\mu &= \sum_n \int_{M_n} |f|^2 d\mu \\ &= \sum_n \int_{M_n} \frac{d\mu}{\mu(M_n) |\varphi|^2} \\ &\leq \sum_n \int_{M_n} \frac{d\mu}{\mu(M_n) \cdot n^2} = \sum_n \frac{1}{n^2} < \infty, \end{aligned}$$

the function f is in \mathbf{L}^2 ; since

$$\int |\varphi \cdot f|^2 d\mu = \sum_n \int_{M_n} \frac{d\mu}{\mu(M_n)},$$

the function $\varphi \cdot f$ is not.

For another proof, let A be the linear transformation that multiplies each element of \mathbf{L}^2 by φ , and prove that A is closed, as follows. Suppose that $\langle f_n, g_n \rangle$ belongs to the graph of A (i.e., $g_n = \varphi \cdot f_n$), and suppose

that $\langle f_n, g_n \rangle \rightarrow \langle f, g \rangle$ (i.e., $f_n \rightarrow f$ and $g_n \rightarrow g$). There is no loss of generality in assuming that $f_n \rightarrow f$ almost everywhere and $g_n \rightarrow g$ almost everywhere; if this is not true for the sequence $\{f_n\}$, it is true for a suitable subsequence. Since $f_n \rightarrow f$ almost everywhere, it follows that $\varphi \cdot f_n \rightarrow \varphi \cdot f$ almost everywhere; since, at the same time, $\varphi \cdot f_n \rightarrow g$ almost everywhere, it follows that $g = \varphi \cdot f$ almost everywhere, i.e., that $\langle f, g \rangle$ is in the graph of A . Conclusion (via the closed graph theorem): A is bounded, and therefore (by Problem 50) φ is bounded.

The second proof is worth a second glance. The concept of multiplication operator can be profitably generalized to unbounded multipliers. If φ is an arbitrary (not necessarily bounded) measurable function, let \mathbf{M} be the set (linear manifold) of all those f in L^2 for which $\varphi \cdot f \in L^2$. The second proof above proves that the linear transformation (from \mathbf{M} into L^2) that maps each f in \mathbf{M} onto $\varphi \cdot f$ is a closed transformation. (This sort of thing is the operator analogue of a vague, but well-known and correct, measure-theoretic principle. In measure theory, every function that can be written down is measurable; in operator theory, every transformation that can be written down is closed.) Briefly: multiplications (bounded or not) are closed. The closed graph theorem can then be invoked to prove that if, in addition, a multiplication has all L^2 for its domain, then it must be bounded.

Solution 52. For invertibility: if $\varphi \cdot \psi = 1$, then the multiplication operator induced by ψ acts as the inverse of A . Suppose, conversely, that A is invertible. This implies that φ can vanish on a set of measure 0 at most. (Otherwise take for f the characteristic function of a set of positive finite measure on which φ vanishes.) Since $\varphi \cdot A^{-1}f = f$, it follows that $A^{-1}f = (1/\varphi) \cdot f$ whenever $f \in L^2$. Conclusion (from Solution 50): $|1/\varphi| \leq \|A^{-1}\|$, and therefore $|\varphi| \geq 1/\|A^{-1}\|$ almost everywhere.

The assertion about the spectrum reduces to the one about invertibility. The beginner is advised to examine the reduction in complete detail. The concept of essential range is no more slippery than other measure-theoretic concepts in which alterations on null sets are gratis, but on first acquaintance it frequently appears to be.

Solution 53. (a) *A multiplication transformation on a functional Hilbert space is necessarily bounded.*

Proof. A proof can be based on the closed graph theorem. Suppose, indeed, that $\langle f_n, g_n \rangle$ is in the graph of A , $n = 1, 2, 3, \dots$, and suppose that $\langle f_n, g_n \rangle \rightarrow \langle f, g \rangle$ (i.e., $f_n \rightarrow f$ and $g_n \rightarrow g$). Since convergence in \mathbf{H} implies pointwise convergence (if $f_n \rightarrow f$ strongly, then $f_n \rightarrow f$ weakly) it follows that $f_n(x) \rightarrow f(x)$ and $g_n(x) \rightarrow g(x)$ for all x . Since $g_n = Af_n = \varphi \cdot f_n$, and since $\varphi(x)f_n(x) \rightarrow \varphi(x)f(x)$ for all x , it follows that $g = Af$. Conclusion: A is closed and therefore bounded.

The answer to (b) is not quite yes. The trouble is that there is nothing in the definition of a functional Hilbert space to prevent the existence of points x in X such that $f(x) = 0$ for all f in \mathbf{H} . The situation can be produced at will; given \mathbf{H} and X , enlarge X arbitrarily, and extend each function in \mathbf{H} so as to be 0 at the new points. At the same time, “null-points” are as easy to eliminate as they are to produce; omit them all from X and restrict each function in \mathbf{H} to the remaining set. As long as infinitely many null-points are present, however, the answer to (b) must be no. Reason: any function on X can be redefined at the null-points so as to become unbounded, without changing the effect that multiplication by it has on the elements of \mathbf{H} . Null-points play the same role for functional Hilbert spaces as atoms of infinite measure play for L^2 spaces (cf. Solution 49).

(b) *If \mathbf{H} is a functional Hilbert space with no null-points, then every (necessarily bounded) multiplication on \mathbf{H} is induced by a bounded multiplier.*

Proof. Note that

$$\|\varphi^n \cdot f\| = \|A^n f\| \leq \|A\|^n \|f\|$$

whenever n is a positive integer and f is in \mathbf{H} (cf. Solution 50). If $A = 0$, then $\varphi = 0$, and there is nothing to prove; otherwise write $\psi = \varphi/\|A\|$, and rewrite the preceding inequality in the form

$$\|\psi^n \cdot f\| \leq \|f\|.$$

From this it follows that if $f(x) \neq 0$, then $|\psi(x)| \leq 1$ (i.e., $|\varphi(x)| \leq \|A\|$). Reason: $(\psi^n \cdot f)(x)$ is bounded by some multiple of $\|\psi^n \cdot f\|$.

Since for each x there is an f such that $f(x) \neq 0$, it follows that $|\varphi| \leq \|A\|$ everywhere.

Here is an alternative proof that is more in the usual spirit of functional Hilbert spaces; it is due to A. L. Shields. Let K be the kernel function of the space (cf. Problem 30). Since $AK_x = \varphi \cdot K_x$ for each x , and since, at the same time, $(AK_x)(y) = (AK_x, K_y)$, it follows that

$$|\varphi(x)K(x,x)| = |(AK_x, K_x)| \leq \|A\| \cdot \|K_x\|^2.$$

Since $\|K_x\|^2 = (K_x, K_x)$, and since always $(K_x, K_y) = K_x(y)$, so that $\|K_x\|^2 = K(x,x)$, it follows that

$$|\varphi(x)K(x,x)| \leq \|A\| \cdot |K(x,x)|.$$

The relation $K(x,y) = K_y(x) = (K_y, K_x)$ implies that the “matrix” K is positive definite and hence, in particular, that $|K(x,y)| \leq \sqrt{K(x,x)} \sqrt{K(y,y)}$. It follows that if $K(x,x) = 0$ for some x , then $K(x,y) = 0$ for all y , i.e., $K_x = 0$, and hence $f(x) = (f, K_x) = 0$ for all f . The assumption that there are no null-points guarantees that this does not happen. Conclusion: $|\varphi(x)| \leq \|A\|$.

The proof used the Schwarz inequality for sesquilinear forms that are not known to be strictly positive. Some of the standard proofs of the Schwarz inequality work in this case; the one in Halmos [1951, p. 15] does not. The problem is to prove that if φ is a positive, symmetric, sesquilinear form, then

$$|\varphi(f,g)|^2 \leq \varphi(f,f) \cdot \varphi(g,g).$$

(Two alphabetic customs are in temporary collision here: the letter φ in this paragraph is not, as it was above, a multiplier on X , but a sesquilinear form on \mathbf{H} .) For the proof, let φ_+ be an arbitrary strictly positive, symmetric, sesquilinear form on the same space, and write, for each positive number ϵ ,

$$\varphi_\epsilon = \varphi + \epsilon\varphi_+.$$

The form φ_ϵ is strictly positive; apply the Schwarz inequality to it and let ϵ tend to 0. As for finding a strictly positive form φ_+ (on every real or complex vector space): just use Hamel bases. If $\{e_j\}$ is one, write

$\varphi_+(\sum_j \alpha_j e_j, \sum_j \beta_j e_j) = \sum_j \alpha_j \beta_j^*$. The sums are formally infinite but only finitely non-zero.

All this about the Schwarz inequality is a digression, but it is amusing; here is one more addition to it. The proof of the Schwarz inequality for inner products (strictly positive forms) consists of the verification of one line, namely:

$$\|f\|^2 \cdot \|g\|^2 - |(f, g)|^2 = \frac{1}{\|g\|^2} \left\| \|g\|^2 f - (f, g)g \right\|^2.$$

It might perhaps be more elegant to multiply through by $\|g\|^2$, so that the result should hold for $g = 0$ also, but the identity seems to be more perspicuous in the form given. This one line proves also that if the inequality degenerates to an equality, then f and g are linearly dependent. The converse is trivial: if f and g are linearly dependent, then one of them is a scalar multiple of the other, say $g = \alpha f$, and then both $|(f, g)|^2$ and $(f, f) \cdot (g, g)$ are equal to $|\alpha|^2 (f, f)^2$.

Solution 54. Let \mathbf{H} be the set of all those absolutely continuous (complex-valued) functions on $[0, 1]$ whose derivatives belong to \mathbf{L}^2 ; define inner product in \mathbf{H} by $(f, g) = f(0)g(0)^* + \int_0^1 f'(x)g'(x)^* dx$. If $\|f\| = 0$, then $\int_0^1 |f'(x)|^2 dx = 0$, so that $f'(x) = 0$ almost everywhere, and therefore f is a constant; since, however, $f(0) = 0$, it follows that $f = 0$. This proves that the inner product is strictly positive. If $\{f_n\}$ is a Cauchy sequence in \mathbf{H} , then $\{f_n(0)\}$ is a numerical Cauchy sequence and $\{f_n'\}$ is Cauchy in \mathbf{L}^2 . It follows that $f_n(0) \rightarrow \alpha$ and $f_n' \rightarrow g$, for some complex number α and for some g in \mathbf{L}^2 ; put $f(x) = \alpha + \int_0^x g(t) dt$, and thus obtain an f such that $f_n \rightarrow f$ in \mathbf{H} . This proves that \mathbf{H} is complete. If $0 \leq x \leq 1$, then

$$|f(x)|^2 = \left| f(0) + \int_0^x f'(t) dt \right|^2 \leq 2 \left(|f(0)|^2 + \int_0^1 |f'(t)|^2 dt \right) = 2 \|f\|^2;$$

this proves that evaluations are bounded and hence that \mathbf{H} is a functional Hilbert space.

If f and g are in \mathbf{H} , then f and g are bounded; it follows that $(fg)' (=fg' + f'g)$ belongs to \mathbf{L}^2 and hence that $fg \in \mathbf{H}$. Since 1 obviously belongs to \mathbf{H} , all the requirements are satisfied.

This example is due to A. L. Shields.

Chapter 7. Operator matrices

Solution 55. Write $\Delta = AD - BC$. If Δ is invertible, then the product of

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} D\Delta^{-1} & -B\Delta^{-1} \\ -C\Delta^{-1} & A\Delta^{-1} \end{pmatrix},$$

in either order, is

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix};$$

this proves the sufficiency of the condition. (Note that in this context 0 and 1 are not numbers but operators on the appropriate Hilbert space.)

Suppose now that

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

is invertible, and hence, in particular, bounded from below. This means that

$$\|Af + Bg\|^2 + \|Cf + Dg\|^2 \geq \delta(\|f\|^2 + \|g\|^2)$$

for some positive number δ . Two special cases of this relation can be usefully combined. First, set one of the two coordinates of $\langle f, g \rangle$ equal to zero (and, in both cases, call the other one f) to get

$$\|Af\|^2 + \|Cf\|^2 \geq \delta\|f\|^2,$$

$$\|Bf\|^2 + \|Df\|^2 \geq \delta\|f\|^2.$$

Second, take $\langle Df, -Cf \rangle$ and $\langle -Bf, Af \rangle$ instead of $\langle f, g \rangle$ (compare this with the two columns of the candidate for the inverse) to get

$$\|(AD - BC)f\|^2 \geq \delta(\|Cf\|^2 + \|Df\|^2),$$

$$\|(AD - BC)f\|^2 \geq \delta(\|Af\|^2 + \|Bf\|^2).$$

Add the latter pair of inequalities, divide by 2, and combine with the former pair to get

$$\|(AD - BC)f\|^2 \geq \delta \|f\|^2.$$

Conclusion: $AD - BC$ is bounded from below.

Since A, B, C , and D are pairwise commutative, the same is true of A^*, B^*, C^* , and D^* ; since

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

is invertible, the same is true of the adjoint

$$\begin{pmatrix} A^* & C^* \\ B^* & D^* \end{pmatrix}.$$

The result of the preceding paragraph implies that $A^*D^* - B^*C^*$ is bounded from below, and hence that its kernel is trivial, and this, in turn, implies that the range of $AD - BC$ is dense.

Since $AD - BC$ is bounded from below and has a dense range, it follows that $AD - BC$ is invertible, and the proof is complete.

Solution 56. Since

$$\begin{pmatrix} 1 & 0 \\ T & 1 \end{pmatrix}$$

is always invertible, with inverse

$$\begin{pmatrix} 1 & 0 \\ -T & 1 \end{pmatrix},$$

it follows that

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} 1 & 0 \\ T & 1 \end{pmatrix}$$

are invertible together. The product works out to

$$\begin{pmatrix} A + BT & B \\ C + DT & D \end{pmatrix};$$

set $T = -D^{-1}C$ and conclude that

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

is invertible if and only if

$$\begin{pmatrix} A - BD^{-1}C & B \\ 0 & D \end{pmatrix}$$

is invertible.

Introduce the temporary abbreviation $E = A - BD^{-1}C$ and proceed to consider the invertibility of

$$\begin{pmatrix} E & B \\ 0 & D \end{pmatrix}.$$

The assumption that D is invertible is still in force. If E also is invertible, then so is

$$\begin{pmatrix} E & B \\ 0 & D \end{pmatrix},$$

with inverse

$$\begin{pmatrix} E^{-1} & -E^{-1}BD^{-1} \\ 0 & D^{-1} \end{pmatrix}.$$

The converse is also true: if

$$\begin{pmatrix} E & B \\ 0 & D \end{pmatrix}$$

is invertible, then so is E . The proof is an easy computation. Suppose that

$$\begin{pmatrix} P & Q \\ R & S \end{pmatrix}$$

is the inverse of

$$\begin{pmatrix} E & B \\ 0 & D \end{pmatrix};$$

then

$$\begin{pmatrix} PE & PB + QD \\ RE & RB + SD \end{pmatrix} = \begin{pmatrix} EP + BR & EQ + BS \\ DR & DS \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Since $DR = 0$ and D is invertible, it follows that $R = 0$; since $PE = 1$ and $EP + BR = 1$, it follows that E is invertible (and, in fact, $E^{-1} = P$).

Now unabbreviate and conclude that

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

is invertible if and only if $A - BD^{-1}C$ is invertible. Since D is invertible, multiplication by D does not affect any statement of invertibility; it follows that

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

is invertible if and only if $AD - BD^{-1}CD$ is invertible. Up to this point the assumed commutativity of C and D was not needed; it comes in now and serves to make the statement more palatable. Since C and D commute, it follows that C and D^{-1} commute, and hence that $BD^{-1}CD = BC$. Conclusion:

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

is invertible if and only if $AD - BC$ is invertible.

The unsymmetry of the hypothesis (why C and D ? and why D^{-1} ?) is not so ugly as first it may seem. The point is that the conclusion is equally unsymmetric. What rights does (1) $AD - BC$ have that (2) $DA - BC$, or (3) $DA - CB$, or (4) $AD - CB$ do not have? Symmetry is restored not by changing the statement but by enlarging the context. The theorem is one of four. To get a conclusion about all possible versions of the formal determinant, assume that D is invertible and make the commutativity hypothesis about (1) C and D , or (2) B and D , or, alternatively, assume that A is invertible, and make the commutativity hypothesis about (3) A and B , or (4) A and C .

It is well known and obvious that if the underlying Hilbert space is finite-dimensional, then the invertible operators are dense in the metric space of all operators. This remark (together with the result proved above) implies that in the finite-dimensional case the invertibility assumption is superfluous: if C and D commute, then a necessary and sufficient condition that

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

be invertible is that $AD - BC$ be invertible. Actually the proof proves more than this: since multiplication by

$$\begin{pmatrix} 1 & 0 \\ T & 1 \end{pmatrix}$$

leaves unchanged not only the property of invertibility, but even the numerical value of the determinant, what the proof proves is that

$$\det \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \det(AD - BC).$$

As for the counterexamples, an efficient place to find them is ℓ^2 . Define A and D by

$$A \langle \xi_0, \xi_1, \xi_2, \dots \rangle = \langle \xi_1, \xi_2, \xi_3, \dots \rangle$$

and

$$D \langle \xi_0, \xi_1, \xi_2, \dots \rangle = \langle 0, \xi_0, \xi_1, \xi_2, \dots \rangle,$$

and put $B = C = 0$. It follows that $AD - BC = 1$, but

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

has a non-trivial kernel. (Look at $\langle f, g \rangle$, where $f = \langle 1, 0, 0, \dots \rangle$, and $g = 0$.) If, on the other hand, B is defined by

$$B \langle \xi_0, \xi_1, \xi_2, \dots \rangle = \langle \xi_0, 0, 0, 0, \dots \rangle,$$

then

$$\begin{pmatrix} D & B \\ 0 & A \end{pmatrix}$$

is invertible, with inverse

$$\begin{pmatrix} A & 0 \\ B & D \end{pmatrix},$$

but the formal determinant DA has a non-trivial kernel.

Solution 57. It is convenient to begin with a lemma of some independent interest: if a finite-dimensional subspace is invariant under an invertible operator, then it is invariant under the inverse too. (Easy examples show that the assumption of finite-dimensionality is indispensable here.) To avoid the introduction of extra notation, let $\mathbf{H} \oplus \mathbf{K}$ be the space, \mathbf{H} the subspace, and M the operator. (To be sure, \mathbf{H} is not really a subspace of $\mathbf{H} \oplus \mathbf{K}$, but it becomes one by an obvious identification.) Since $M\mathbf{H} \subset \mathbf{H}$, and since (by invertibility) M preserves linear independence, it follows that $\dim M\mathbf{H} = \dim \mathbf{H}$, and hence (by finite-dimensionality) that $M\mathbf{H} = \mathbf{H}$. This implies that $M^{-1}\mathbf{H} = \mathbf{H}$, and the proof of the lemma is complete.

The lemma applies to the case at hand. If

$$M = \begin{pmatrix} A & B \\ 0 & D \end{pmatrix},$$

then \mathbf{H} is invariant under M ; it follows from the lemma that if M is invertible, then M^{-1} has the form

$$\begin{pmatrix} A' & B' \\ 0 & D' \end{pmatrix}.$$

Finite-dimensionality has served its purpose by now; the rest of the argument is universally valid. Once it is known that a triangular matrix has a triangular inverse, then, regardless of the sizes of the entries, each diagonal entry in the matrix is invertible, and its inverse is the corresponding entry of the inverse matrix. Proof: multiply the two matrices in both possible orders and look.

Chapter 8. Properties of spectra

Solution 58. *If A is an operator, then $\Pi_0(A^*) = \Gamma(A)^*$ and $\Pi(A^*) \cup \Pi(A)^* = \Lambda(A^*)$.*

Proof. If $\lambda \in \Pi_0(A^*)$, then $A^* - \lambda$ has a non-zero kernel, and therefore the range of $A - \lambda^*$ has a non-zero orthogonal complement; both these implications are reversible.

The second equation is the best that can be said about the relation between Π and conjugation. The assertion is that if $A^* - \lambda$ is not invertible, then one of $A^* - \lambda$ and $A - \lambda^*$ is not bounded from below. Equivalently (with an obvious change of notation) it is to be proved that if both A^* and A are bounded from below, then A^* is invertible. The proof is trivial: if A is bounded from below, then its kernel is trivial, and therefore the range of A^* is dense; this, together with the assumption that A^* is bounded from below, implies that A^* is invertible. (This argument has been used already, to prove a special case of the present assertion; cf. Solution 55.)

Corollary. $\Pi_0(A) = \Gamma(A^*)^*$ and $\Pi(A) \cup \Pi(A^*)^* = \Lambda(A)$.

Proof. Replace A by A^* .

Solution 59. *If A is an operator and p is a polynomial, then $\Pi_0(p(A)) = p(\Pi_0(A))$, $\Pi(p(A)) = p(\Pi(A))$, and $\Gamma(p(A)) = p(\Gamma(A))$; the same equations are true if A is an invertible operator and $p(z) = 1/z$ for $z \neq 0$.*

Proof. It is convenient to make three elementary observations before the proof really begins. If the product of a finite number of operators (1) has a non-zero kernel, or (2) is not bounded from below, or (3) has a range that is not dense, then at least one factor must have the same property; if the factors commute, then the converse of each of these statements is true. The idea of the proofs is perhaps best suggested by the following sentences. If AB sends (1) a non-zero vector onto 0, or

(2) a sequence of unit vectors onto a null sequence, then argue from the right: if B does not already do so, then A must. (3) If the range of AB is not dense, argue from the left: if the range of A is dense, then the range of B cannot be.

Now for the proofs of the spectral mapping theorems. Assume, with no loss of generality, that the polynomial p has positive degree and leading coefficient 1. Since $p(\lambda) - p(\lambda_0)$ is divisible by $\lambda - \lambda_0$, it follows, by (1), that if $\lambda_0 \in \Pi_0(A)$, then $p(\lambda_0) \in \Pi_0(p(A))$, and hence that $p(\Pi_0(A)) \subset \Pi_0(p(A))$. (This part of the statement can be proved much more simply: if $Af = \lambda_0 f$, then $p(A)f = p(\lambda_0)f$. The longer sentence is adaptable to the other cases, and therefore saves time later.) If, on the other hand, $\alpha \in \Pi_0(p(A))$, then express $p(\lambda) - \alpha$ as a product of factors such as $\lambda - \lambda_0$, and apply (1); the conclusion is that $\alpha = p(\lambda_0)$ for some number λ_0 in $\Pi_0(A)$. This means that $\alpha \in p(\Pi_0(A))$, and hence that $\Pi_0(p(A)) \subset p(\Pi_0(A))$. The arguments for Π , or Γ , are exactly the same, except that (2), or (3), are used instead of (1). An alternative method is available for Γ : apply the result for Π_0 to A^* , conjugate, and apply Solution 58.

Turn now to inversion. If A is invertible and $Af = \lambda f$ with $f \neq 0$, then $\lambda \neq 0$. Apply A^{-1} to both sides of the equation, divide by λ , and obtain

$$A^{-1}f = \frac{1}{\lambda}f.$$

Conclusion:

$$\frac{1}{\Pi_0(A)} \subset \Pi_0(A^{-1}).$$

Replace A by A^{-1} and form reciprocals to get the reverse inclusion. Use the same method, but starting with $Af_n - \lambda f_n \rightarrow 0$, $\|f_n\| = 1$, to get the inversion spectral mapping theorem for Π . Derive the result for Γ by applying the result for Π_0 to the adjoint.

Solution 60. (1) If $A - \lambda$ is invertible, then so is $P^{-1}(A - \lambda)P = P^{-1}AP - \lambda$.

(2) If $Af = \lambda f$, then $P^{-1}AP(P^{-1}f) = \lambda(P^{-1}f)$.

(3) If $Af_n - \lambda f_n \rightarrow 0$, where $\|f_n\| = 1$, then $P^{-1}AP(P^{-1}f_n) -$

$\lambda(P^{-1}f_n) = P^{-1}(Af_n - \lambda f_n) \rightarrow 0$. The norms $\|P^{-1}f_n\|$ are bounded from below by $1/\|P\|$, and, consequently, division by $\|P^{-1}f_n\|$ does not affect convergence to 0. This implies that

$$P^{-1}AP\left(\frac{P^{-1}f_n}{\|P^{-1}f_n\|}\right) - \lambda\left(\frac{P^{-1}f_n}{\|P^{-1}f_n\|}\right) \rightarrow 0.$$

(4) If g belongs to the range of $P^{-1}AP - \lambda$ ($= P^{-1}(A - \lambda)P$), then g belongs to the image under P^{-1} of the range of $A - \lambda$; this implies that if the closure of the range of $A - \lambda$ is not \mathbf{H} , then the closure of the range of $P^{-1}(A - \lambda)P$ is not \mathbf{H} either.

The four proofs just given show that each named part of the spectrum of A is included in the corresponding part for $P^{-1}AP$. This assertion applied to $P^{-1}AP$ and P^{-1} (in place of A and P) implies its own converse.

Solution 61. It is to be proved that if $\lambda \neq 0$, then $AB - \lambda$ and $BA - \lambda$ are invertible or non-invertible together. Division by $-\lambda$ reduces the theorem to the general ring-theoretic assertion: if $1 - AB$ is invertible, then so is $1 - BA$. The motivation for the proof of this assertion (but not the proof itself) comes from pretending that the inverse, say C , of $1 - AB$ can be written in the form $1 + AB + ABAB + \cdots$, and that, similarly, the inverse of $1 - BA$ is $1 + BA + BABA + \cdots = 1 + B(1 + AB + ABAB + \cdots)A = 1 + BCA$. The proof itself consists of verifying that if

$$C(1 - AB) = (1 - AB)C = 1,$$

then

$$(1 + BCA)(1 - BA) = (1 - BA)(1 + BCA) = 1.$$

The verification is straightforward. It is a little easier to see if the assumption on C is rewritten in the form

$$CAB = ABC = C - 1.$$

Solution 62. For each operator A , the approximate point spectrum $\Pi(A)$ is closed.

Proof. A convenient attack is to prove that the complement of $\Pi(A)$ is open. If λ_0 is not in $\Pi(A)$, then $A - \lambda_0$ is bounded from below; say

$$\|Af - \lambda_0 f\| \geq \delta \|f\|$$

for all f . Since

$$\|Af - \lambda_0 f\| \leq \|Af - \lambda f\| + \|\lambda f - \lambda_0 f\|$$

for all λ , it follows that

$$(\delta - |\lambda - \lambda_0|)\|f\| \leq \|Af - \lambda f\|.$$

This implies that if $|\lambda - \lambda_0|$ is sufficiently small, then $A - \lambda$ is bounded from below.

Solution 63. It is convenient (but not compulsory) to prove the following slightly more general assertion: if $\{A_n\}$ is a sequence of invertible operators and if A is a non-invertible operator such that $\|A_n - A\| \rightarrow 0$ as $n \rightarrow \infty$, then $0 \in \Pi(A)$. Since A is not invertible, either $0 \in \Pi(A)$ or $0 \in \Gamma(A)$. If $0 \in \Pi(A)$, there is nothing to prove. It is therefore sufficient to prove that A is not bounded from below (i.e., that $0 \in \Pi(A)$) under the assumption that $\text{ran } A$ is not dense. Suppose then that f is a non-zero vector orthogonal to $\text{ran } A$, and write

$$f_n = \frac{A_n^{-1}f}{\|A_n^{-1}f\|}.$$

Since $\|f_n\| = 1$, it follows that $\|(A_n - A)f_n\| \leq \|A_n - A\| \rightarrow 0$. Since, however, $Af_n \in \text{ran } A$ and $A_n f_n \perp \text{ran } A$, it follows that

$$\|A_n f_n - A f_n\|^2 = \|A_n f_n\|^2 + \|A f_n\|^2 \geq \|A f_n\|^2,$$

and hence that $\|A f_n\| \rightarrow 0$.

To derive the original spectral assertion, suppose that λ is on the boundary of $\Lambda(A)$. It follows that there exist numbers λ_n not in $\Lambda(A)$ such that $\lambda_n \rightarrow \lambda$. The operators $A - \lambda_n$ are invertible and $A - \lambda$ is not; since

$$\|(A - \lambda_n) - (A - \lambda)\| = |\lambda_n - \lambda| \rightarrow 0,$$

it follows from the preceding paragraph that $\lambda \in \Pi(A)$.

Chapter 9. Examples of spectra

Solution 64. Normality says that $\|Af\| = \|A^*f\|$ for every vector f . It follows that $\|(A - \lambda)f\| = \|(A^* - \lambda^*)f\|$ for every λ , and hence that $\Pi_0(A) = (\Pi_0(A^*))^*$. The conclusion follows from Solution 58.

Solution 65. *If A is a diagonal operator, then both $\Pi_0(A)$ and $\Gamma(A)$ are equal to the diagonal, and $\Pi(A) (= \Lambda(A))$ is the closure of the diagonal.*

Proof. Suppose that $\{e_j\}$ is an orthonormal basis such that $Ae_j = \alpha_j e_j$. The first assertion is that a number is an eigenvalue of A if and only if it is equal to one of the α_j 's. "If" is trivial: each α_j is an eigenvalue of A . By an obvious subtraction, the "only if" is equivalent to this: if A has a non-zero kernel, then at least one of the α_j 's vanishes. Contrapositively: if $\alpha_j \neq 0$ for all j , then $Af = 0$ implies $f = 0$. Indeed: if $f = \sum_j \xi_j e_j$, then $Af = \sum_j \alpha_j \xi_j e_j$, so that $Af = 0$ is equivalent to $\alpha_j \xi_j = 0$ for all j ; since no α_j vanishes, every ξ_j must.

Now that $\Pi_0(A)$ is known, the result of Problem 64 applies. Since a diagonal operator is normal, it follows that $\Gamma(A)$ also is equal to the diagonal, and that the approximate point spectrum is the same as the entire spectrum.

Solution 66. *If A is the multiplication induced by a multiplier φ (over a σ -finite measure space), then both $\Pi_0(A)$ and $\Gamma(A)$ are equal to the set of those complex numbers λ for which $\varphi^{-1}(\{\lambda\})$ has positive measure, and $\Pi(A) (= \Lambda(A))$ is the essential range of φ .*

Proof. If $f \in L^2$ and $\varphi(x)f(x) = \lambda f(x)$ almost everywhere, then $\varphi(x) = \lambda$ whenever $f(x) \neq 0$. This implies that in order for λ to be an eigenvalue of A , the function φ must take the value λ on a set of positive measure. If, conversely, $\varphi(x) = \lambda$ on a set M of positive measure, and if f is the characteristic function of a measurable subset of M of positive finite measure, then $f \in L^2$, $f \neq 0$, and $Af = \lambda f$, so that λ is an eigenvalue of A .

The remaining assertions are proved just as in Solution 65.

Solution 67. If U is the unilateral shift, then $\Lambda(U) = D$ (= the closed unit disc), $\Pi_0(U) = \emptyset$, $\Pi(U) = C$ (= the unit circle), and $\Gamma(U) = D - C$ (= the interior of the unit disc). For the adjoint: $\Lambda(U^*) = D$, $\Pi_0(U^*) = D - C$, $\Pi(U^*) = D$, and $\Gamma(U^*) = \emptyset$.

Proof. It is wise to treat U and U^* together; each gives information about the other. To treat U^* , whether together with U or separately, it is advisable to know what it is. Since (for $i, j = 0, 1, 2, \dots$)

$$(U^*e_i, e_j) = (e_i, Ue_j) = (e_i, e_{j+1}) = \delta_{i, j+1},$$

it follows that

$$U^*e_0 = 0;$$

if $i > 0$, then

$$\delta_{i, j+1} = \delta_{i-1, j} = (e_{i-1}, e_j),$$

and therefore

$$U^*e_i = e_{i-1} \quad (i = 1, 2, 3, \dots).$$

In terms of coordinates the result is that

$$U^* \langle \xi_0, \xi_1, \xi_2, \dots \rangle = \langle \xi_1, \xi_2, \xi_3, \dots \rangle.$$

The functional representation of U (i.e., its representation as a multiplication on \mathbf{H}^2) is deceptive; since the adjoint of a multiplication operator is multiplication by the complex conjugate function, it is tempting to think that if $f \in \mathbf{H}^2$, then $(U^*f)(z) = z^*f(z)$. This is not only false, it is nonsense; \mathbf{H}^2 is not invariant under multiplication by e_{-1} . The correspondence between adjunction and conjugation works for \mathbf{L}^2 , but there is no reason to assume that it will work for a subspace of \mathbf{L}^2 . The correct expression for U^* on \mathbf{H}^2 is given by

$$(U^*f)(z) = z^*(f(z) - (f, e_0)).$$

Now for the spectrum and its parts. Since U is an isometry, so that $\|U\| = 1$, it follows that the spectrum of U is included in the closed unit disc, and hence the same is true of U^* .

If $Uf = \lambda f$, where $f = \langle \xi_0, \xi_1, \xi_2, \dots \rangle$, then

$$\langle 0, \xi_0, \xi_1, \xi_2, \dots \rangle = \langle \lambda \xi_0, \lambda \xi_1, \lambda \xi_2, \dots \rangle,$$

so that $0 = \lambda \xi_0$, and $\xi_n = \lambda \xi_{n+1}$ for all n . This implies that $\xi_n = 0$ for all n (look separately at the cases $\lambda = 0$ and $\lambda \neq 0$), and hence that $\Pi_0(U) = \emptyset$. Consequence: $\Gamma(U^*) = \emptyset$.

Here is an alternative proof that U has no eigenvalues, which has some geometric merit. It is a trivial fact, true for every operator A , that if f is an eigenvector belonging to a non-zero eigenvalue, then f belongs to $\text{ran } A^n$ for every positive integer n . (Proof by induction. Trivial for $n = 0$; if $f = A^n g$, then $f = (1/\lambda) A f = (1/\lambda) A^{n+1} g$.) The range of U^n consists of all vectors orthogonal to all the e_j 's with $0 \leq j < n$, and, consequently, $\bigcap_n \text{ran } U^n$ consists of 0 alone. This proves that U has no eigenvalues different from 0. The eigenvalue 0 is excluded by the isometric property of U : if $Uf = 0$, then $0 = \|Uf\| = \|f\|$.

If $U^*f = \lambda f$, then

$$\langle \xi_1, \xi_2, \xi_3, \dots \rangle = \langle \lambda \xi_0, \lambda \xi_1, \lambda \xi_2, \dots \rangle,$$

so that $\xi_{n+1} = \lambda \xi_n$, or $\xi_{n+1} = \lambda^n \xi_0$, for all n . If $\xi_0 = 0$, then $f = 0$; otherwise a necessary and sufficient condition that the resulting ξ 's be the coordinates of a vector (i.e., that they be square-summable) is that $|\lambda| < 1$. Conclusion: $\Pi_0(U^*)$ is the open unit disc (and consequently $\Gamma(U)$ is the open unit disc). Each λ in that disc is a simple eigenvalue of U^* (i.e., it has multiplicity 1); the corresponding eigenvector f_λ (normalized so that $(f_\lambda, e_0) = 1$) is given by

$$f_\lambda = \langle 1, \lambda, \lambda^2, \lambda^3, \dots \rangle.$$

Since spectra are closed, it follows that both $\Lambda(U)$ and $\Lambda(U^*)$ include the closed unit disc, and hence that they are equal to it. All that remains is to find $\Pi(U)$ and $\Pi(U^*)$. Since the boundary of the spectrum of every operator is included in the approximate point spectrum, it follows that both $\Pi(U)$ and $\Pi(U^*)$ include the unit circle. If $|\lambda| < 1$, then

$$\|Uf - \lambda f\| \geq \left| \|Uf\| - \|\lambda f\| \right| = |1 - |\lambda|| \cdot \|f\|$$

for all f , so that $U - \lambda$ is bounded from below; this proves that $\Pi(U)$ is exactly the unit circle. For U^* the situation is different: since Π_0 is always included in Π , and since $\Pi_0(U^*)$ is the open unit disc, it follows that $\Pi(U^*)$ is the closed unit disc.

Solution 68. *If W is the bilateral shift, then $\Lambda(W) = C$ (= the unit circle), $\Pi_0(W) = \emptyset$, $\Pi(W) = C$, and $\Gamma(W) = \emptyset$. The same equations are true for the adjoint W^* .*

Proof. The determination of the spectrum of W , and of the fine structure of that spectrum, can follow the pattern indicated in the study of the unilateral shift U (Solution 67), but there is also another way to do it, a better way. Corresponding to the functional representation of U on \mathbf{H}^2 , the bilateral shift W has a natural functional representation on $\mathbf{L}^2(\mu)$ (where μ is normalized Lebesgue measure on the unit circle; see Problem 26). Since the functions e_n defined by $e_n(z) = z^n$ ($n = 0, \pm 1, \pm 2, \dots$) form an orthonormal basis for \mathbf{L}^2 , and since the effect on them of shifting forward by one index is the same as the effect of multiplying by e_1 , it follows that the bilateral shift is the same as the multiplication operator on \mathbf{L}^2 defined by

$$(Wf)(z) = zf(z).$$

This settles everything for W ; everything follows from Solution 66.

As for W^* , its study can be reduced to that of W . Indeed, since W is unitary, its adjoint is the same as its inverse. The calculation of W^{-1} takes no effort at all; clearly W^{-1} shifts backward the same way as W shifts forward. There is a thoroughgoing symmetry between W and W^* ; to obtain one from the other, just replace n by $-n$. In more pedantic language: W and W^* are unitarily equivalent, and, in particular, the unitary operator R determined by the conditions $Re_n = e_{-n}$ ($n = 0, \pm 1, \pm 2, \dots$) transforms W onto W^* (i.e., $R^{-1}WR = W^*$). Conclusion: the spectrum of W^* is equal to the spectrum of W , and the same is true, part for part, for each of the usual parts of the spectrum.

Solution 69. Suppose first that the eigenvectors of A^* span \mathbf{H} . Let X be an index set such that corresponding to each x in X there is an

eigenvector K_x of A^* , and such that the K_x 's span \mathbf{H} ; denote the eigenvalue corresponding to K_x by $\varphi(x)^*$. (The conjugation has no profound significance here; it is just a notational convenience.) It follows that $A^*K_x = \varphi(x)^*K_x$. For each f in \mathbf{H} , let \tilde{f} be the function on X defined by $\tilde{f}(x) = (f, K_x)$. The correspondence $f \rightarrow \tilde{f}$ is linear. If $\tilde{f} = 0$, i.e., if $(f, K_x) = 0$ for all x , then $f = 0$ (since the K_x 's span \mathbf{H}). This justifies the definition $(\tilde{f}, \tilde{g}) = (f, g)$. With this definition of inner product, the set $\tilde{\mathbf{H}}$ of all functions of the form \tilde{f} (with f in \mathbf{H}) becomes a functional Hilbert space. [Note: $|\tilde{f}(x)| = |(f, K_x)| \leq \|f\| \cdot \|K_x\| = \|\tilde{f}\| \cdot \|K_x\|$.] Let \tilde{A} be the image of A under the isomorphism $f \rightarrow \tilde{f}$ (i.e., $\tilde{A}\tilde{f} = (Af)^{\sim}$); then

$$\begin{aligned} (\tilde{A}\tilde{f})(x) &= (Af)^{\sim}(x) = (Af, K_x) = (f, A^*K_x) \\ &= (f, \varphi(x)^*K_x) = \varphi(x)(f, K_x) \\ &= \varphi(x)\tilde{f}(x), \end{aligned}$$

so that \tilde{A} is a multiplication.

The converse is proved by retracing the steps of the last computation. In detail, if A is a multiplication (with multiplier φ , say) on a functional Hilbert space \mathbf{H} with domain X and kernel function K , so that $(Af)(x) = \varphi(x)f(x)$, then $(Af, K_x) = \varphi(x)(f, K_x)$ (where $K_x(y) = K(y, x)$), and therefore $(f, A^*K_x - \varphi(x)^*K_x) = 0$ for all f . It follows that $A^*K_x = \varphi(x)^*K_x$; since in a functional Hilbert space the set of K_x 's always spans the space, the proof is complete.

Compare the construction with what is known about the unilateral shift (Solution 67).

Solution 70. *The relative spectrum of the unilateral shift is the unit circle.*

Proof. The proof can be made to depend on two simple lemmas. (1) For an operator with a trivial kernel, relative invertibility is the same as left invertibility. (2) For all operators, left invertibility is the same as boundedness from below.

The proof of (1) in one direction is trivial; left invertibility always implies relative invertibility. To prove the converse, suppose that

$ABA = A$, so that $A(1 - BA) = 0$. If the kernel of A is trivial, then it follows that $1 - BA = 0$, and hence that A is left invertible.

To prove (2), suppose that A is left invertible, say $BA = 1$; it follows that $\|f\| = \|BAf\| \leq \|B\| \cdot \|Af\|$ for every f , and hence that A is bounded from below. If, conversely, A is bounded from below, then the mapping A has a uniquely determined inverse mapping B that sends the (closed) range of A onto the whole space. The mapping B is a bounded linear transformation; extend it to an operator by, for instance, defining it to be 0 on the orthogonal complement of the range of A . The extended B is a left inverse of A .

The lemmas (1) and (2) imply that if the point spectrum of an operator is empty, then for that operator the relative spectrum is equal to the approximate point spectrum. The assertion for the unilateral shift is now immediate (see Solution 67).

Solution 71. *If A is an operator on a Hilbert space \mathbf{H} , if α is a non-zero scalar, and if M is the operator matrix*

$$\begin{pmatrix} \alpha & 0 \\ 1 & A \end{pmatrix},$$

then a necessary and sufficient condition that M be relatively invertible is that A be such.

The scalars α and 1 appearing in M are to be interpreted as operators on \mathbf{H} .

Proof. Suppose first that M is relatively invertible. If

$$N = \begin{pmatrix} P & Q \\ R & S \end{pmatrix}$$

is a relative inverse of M , i.e., if $MNM = M$, then $\alpha QA = 0$ and $QA + ASA = A$. (Carry out the indicated matrix multiplication; all that is needed is the second column of the product.) Since $\alpha \neq 0$, it follows that $QA = 0$ and hence that $ASA = A$; this proves that A is

relatively invertible. The converse is another easy computation: if A is relatively invertible, say $ABA = A$, then put

$$N = \begin{pmatrix} (1/\alpha)AB & 1 - AB \\ -(1/\alpha)B & B \end{pmatrix},$$

and verify that $MNM = M$.

The result just proved implies the existence of operators whose relative spectrum is not closed. To prove this, a glance at the case $\alpha = 0$ is in order. The fact is that if

$$M = \begin{pmatrix} 0 & 0 \\ 1 & A \end{pmatrix},$$

then M is relatively invertible no matter what A is. Reason: Write

$$N = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},$$

and verify that $MNM = M$.

The preceding two paragraphs together imply that the relative spectrum of the operator matrix

$$M = \begin{pmatrix} 1 & 0 \\ 1 & A \end{pmatrix}$$

is almost the same as the relative spectrum of the operator A ; the only possible difference between them is the single number 1. More precisely: if the relative spectra of M and of A are Φ and Ψ , respectively, then $\Phi = \Psi - \{1\}$. This in turn implies that if A is chosen so that its relative spectrum contains 1 as a cluster point, then the relative spectrum of M is not closed. For a concrete example, let A be the unilateral shift; see Solution 70.

Chapter 10. Spectral radius

Solution 72. A standard trick for proving operator functions analytic is the identity

$$(1 - A)^{-1} = 1 + A + A^2 + \cdots.$$

If $\|A\| < 1$, then the series converges (with respect to the operator norm), and obvious algebraic manipulations prove that its sum indeed acts as the inverse of $1 - A$. (Replace A by $1 - A$ and recapture the assertion that if $\|1 - A\| < 1$, then A is invertible. Cf. Halmos [1951, p. 52]; see also Problem 83.)

Suppose now that λ_0 is not in the spectrum of A , so that $A - \lambda_0$ is invertible. To prove that $(A - \lambda)^{-1}$ is analytic in λ , for λ near λ_0 , express $A - \lambda$ in terms of $A - \lambda_0$:

$$\begin{aligned} A - \lambda &= (A - \lambda_0) - (\lambda - \lambda_0) \\ &= (A - \lambda_0)(1 - (A - \lambda_0)^{-1}(\lambda - \lambda_0)). \end{aligned}$$

If $|\lambda - \lambda_0|$ is sufficiently small, then $\|(A - \lambda_0)^{-1}(\lambda - \lambda_0)\| < 1$, and the series trick can be applied. The conclusion is that if $|\lambda - \lambda_0|$ is sufficiently small, then $A - \lambda$ is invertible, and

$$(A - \lambda)^{-1} = (A - \lambda_0)^{-1} \sum_{n=0}^{\infty} ((A - \lambda_0)^{-1}(\lambda - \lambda_0))^n.$$

It follows that if f and g are in \mathbf{H} , then

$$(\rho(\lambda)f, g) = \sum_{n=0}^{\infty} ((A - \lambda_0)^{-n-1}f, g)(\lambda - \lambda_0)^n$$

in a neighborhood of λ_0 , and hence that ρ is analytic at λ_0 .

As for $\lambda = \infty$, note that

$$A - \frac{1}{\lambda} = -\frac{1}{\lambda}(1 - \lambda A)$$

whenever $\lambda \neq 0$, and hence that $A - 1/\lambda$ is invertible whenever $|\lambda|$ is sufficiently small (but different from 0). Since

$$\rho\left(\frac{1}{\lambda}\right) = \left(A - \frac{1}{\lambda}\right)^{-1} = -\lambda(1 - \lambda A)^{-1},$$

the series trick applies again:

$$\tau(\lambda) = \rho\left(\frac{1}{\lambda}\right) = -\lambda(1 + \lambda A + \lambda^2 A^2 + \cdots).$$

The parenthetical series converges for small λ , and the factor $-\lambda$ in front guarantees that $\tau(0) = 0$.

Solution 73. Proceed by contradiction. If the spectrum of A is empty, then $(\rho_A f, g)$ (i.e., the function $\lambda \rightarrow ((A - \lambda)^{-1} f, g)$) is an entire function for each f and g ; since $\rho_A(\infty) = 0$, the function $(\rho_A f, g)$ is bounded in a neighborhood of ∞ and therefore in the whole plane. Liouville's theorem implies that $(\rho_A f, g)$ is a constant; since $\rho_A(\infty) = 0$, it follows that

$$((A - \lambda)^{-1} f, g) = 0$$

identically in f, g , and λ . Since this is absurd (replace f and g by $(A - \lambda)f$ and f), the assumption of empty spectrum is untenable.

Solution 74. Since $(r(A))^n = r(A^n) \leq \|A^n\|$, so that $r(A) \leq \|A^n\|^{1/n}$ for all n , it follows that

$$r(A) \leq \liminf_n \|A^n\|^{1/n}.$$

The reverse inequality leans on the analytic character of the resolvent

(Problem 72). If

$$\tau(\lambda) = \rho\left(\frac{1}{\lambda}\right) = \left(A - \frac{1}{\lambda}\right)^{-1},$$

then $\tau(\lambda) = -\lambda(1 - \lambda A)^{-1}$ whenever $\lambda \neq 0$ and $1/\lambda$ is not in the spectrum of A . Since, for each f and g , the numerical function $(\tau f, g)$ is analytic as long as $|1/\lambda| > r(A)$ (i.e., $|\lambda| < 1/r(A)$), it follows that its Taylor series

$$-\lambda \sum_{n=0}^{\infty} \lambda^n (A^n f, g)$$

converges whenever $|\lambda| < 1/r(A)$. This implies that the sequence $\{((\lambda A)^n f, g)\}$ is bounded for each such λ . The principle of uniform boundedness yields the conclusion that the sequence $\{|\lambda|^n \|A^n\|\}$ is bounded. If $|\lambda|^n \|A^n\| \leq \alpha$ for all n , then

$$|\lambda| \cdot \|A^n\|^{1/n} \leq \alpha^{1/n},$$

and therefore

$$|\lambda| \cdot \limsup_n \|A^n\|^{1/n} \leq 1.$$

Since this is true whenever $|\lambda| < 1/r(A)$, it follows that

$$\limsup_n \|A^n\|^{1/n} \leq r(A).$$

The proof is complete.

Solution 75. The asserted unitary equivalence can be implemented by a diagonal operator. To see which diagonal operator to use, work backwards. Assume that D is a diagonal operator with diagonal $\{\delta_n\}$, and assume that $AD = DB$. It follows (apply both sides to e_n) that

$$\alpha_n \delta_n = \beta_n \delta_{n+1}$$

for each n . Put $\delta_0 = 1$, and determine the other δ 's by recursion. Consider first the positive n 's. If $\beta_n \neq 0$, put $\delta_{n+1} = (\alpha_n/\beta_n)\delta_n$. If $\beta_n = 0$, then $\alpha_n = 0$ (since, by assumption, $|\alpha_n| = |\beta_n|$); in that case put $\delta_{n+1} = 1$. For negative n 's (if there are any) apply the same process in the other

direction. That is, if $\alpha_n \neq 0$, put $\delta_n = (\beta_n/\alpha_n)\delta_{n+1}$; if $\alpha_n = 0$, then put $\delta_n = 1$. The result is a sequence $\{\delta_n\}$ of complex numbers of modulus 1. The steps leading to this sequence are reversible. Given the sequence, let it induce a diagonal operator D ; note that since $|\delta_n| = 1$ for all n , the operator D is unitary; and, finally, note that since $ADe_n = DBe_n$ for all n , the operator D transforms A onto B .

Solution 76. Suppose first that S is an invertible operator such that $A = S^{-1}BS$. It follows that $A^* = S^*B^*S^{*-1}$, and hence that $A^{*n} = S^*B^{*n}S^{*-1}$. Use the argument that worked for unitary equivalence to infer that S^* sends $\ker B^{*n}$ onto $\ker A^{*n}$. This implies that the matrix $\{\sigma_{ij}\}$ of S is lower triangular. Consider the equation $SA = BS$, and evaluate the matrix entries in row $n+1$, column n ($n = 0, 1, 2, \dots$) for both sides. The result is that $\sigma_{n+1,n+1}\alpha_n = \beta_n\sigma_{n,n}$, and hence that

$$\left| \frac{\beta_0 \cdots \beta_n}{\alpha_0 \cdots \alpha_n} \right| = \left| \frac{\sigma_{n+1,n+1}}{\sigma_{0,0}} \right| = \left| \frac{(Se_{n+1}, e_{n+1})}{\sigma_{0,0}} \right| \leq \frac{\|S\|}{|\sigma_{0,0}|}.$$

Consequence: $\{|\alpha_0 \cdots \alpha_n / \beta_0 \cdots \beta_n|\}$ is bounded away from 0. To get boundedness (away from ∞), work with S^{-1} (instead of S) and with the equation $AS^{-1} = S^{-1}B$ (instead of $SA = BS$).

If, conversely, the boundedness conditions are satisfied, then write $\sigma_0 = 1$, $\sigma_{n+1} = \beta_0 \cdots \beta_n / \alpha_0 \cdots \alpha_n$, let S be the (invertible) diagonal operator with diagonal sequence $\{\sigma_0, \sigma_1, \sigma_2, \dots\}$, and verify that $SA = BS$.

Solution 77. If A is a weighted shift with weights α_n , then $\|A\| = \sup_n |\alpha_n|$, and $r(A) = \lim_k \sup_n \left| \prod_{i=0}^{k-1} \alpha_{n+i} \right|^{1/k}$.

The expression for r looks mildly complicated, but there are cases when it can be used to compute something.

Proof. Since A is the product of a shift and the diagonal operator with diagonal $\{\alpha_n\}$, and since a shift is an isometry, it follows that the norm of A is equal to the norm of the associated diagonal operator.

To prove the assertion about the spectral radius, evaluate the powers of A . If $Ae_n = \alpha_n e_{n+1}$, then $A^2e_n = \alpha_n \alpha_{n+1} e_{n+2}$, $A^3e_n = \alpha_n \alpha_{n+1} \alpha_{n+2} e_{n+3}$,

etc. What this shows is that A^k is the product of an isometry (namely the k -th power of the associated shift) and a diagonal operator (namely the one whose n -th diagonal term is the product of k consecutive α 's starting with α_n). Conclusion: the norm of A^k is the supremum of the moduli of the "sliding products" of length k , or, explicitly,

$$\|A^k\| = \sup_n \left| \prod_{i=0}^{k-1} \alpha_{n+i} \right| \quad (k = 1, 2, 3, \dots).$$

The expression for the spectral radius follows immediately.

Solution 78. If A is the unilateral weighted shift with weights $\{\alpha_0, \alpha_1, \alpha_2, \dots\}$, and if $\alpha_n \neq 0$ for $n = 0, 1, 2, \dots$, then $\Pi_0(A) = \emptyset$, and $\Pi_0(A^*)$ is a disc with center 0 and radius $\liminf_n \left| \prod_{i=0}^{n-1} \alpha_i \right|^{1/n}$. The disc may be open or closed, and it may degenerate to the origin only.

Proof. The proof for A is the same as for the unweighted unilateral shift (Solution 67). In sequential (coordinate) notation, if $Af = \lambda f$, where $f = \langle \xi_0, \xi_1, \xi_2, \dots \rangle$, then $Af = \langle 0, \alpha_0 \xi_0, \alpha_1 \xi_1, \alpha_2 \xi_2, \dots \rangle$, so that $0 = \lambda \xi_0$ and $\alpha_n \xi_n = \lambda \xi_{n+1}$ for all n . This implies that $\xi_n = 0$ for all n ; look separately at the cases $\lambda = 0$ and $\lambda \neq 0$.

To treat A^* , it is advisable to know what it is. That can be learned by looking at matrices (the diagonal just below the main one flips over to the one just above), by imitating the procedure used to find U^* (Solution 67), or by writing A as the product of U and a diagonal operator and applying the known result for U^* . The answer is that $A^*e_n = 0$ if $n = 0$ and $A^*e_n = \alpha_{n-1}^* e_{n-1}$ if $n > 0$. Sequentially: if $f = \langle \xi_0, \xi_1, \xi_2, \dots \rangle$, then $A^*f = \langle \alpha_0^* \xi_1, \alpha_1^* \xi_2, \alpha_2^* \xi_3, \dots \rangle$. It follows that $A^*f = \lambda f$ if and only if

$$\alpha_n^* \xi_{n+1} = \lambda \xi_n$$

for all n . This implies that if $n > 1$, then ξ_n is the product of ξ_0 by

$$\frac{\lambda^n}{\prod_{i=0}^{n-1} \alpha_i^*}.$$

Since a sequence of numbers defines a vector if and only if it is square-summable, it follows that $\lambda \in \Pi_0(A^*)$ if and only if

$$\sum_{n=1}^{\infty} \left| \frac{\lambda^n}{\prod_{i=0}^{n-1} \alpha_i} \right|^2 < \infty.$$

The condition is that a certain power series in λ^2 be convergent; that proves that the λ 's satisfying it form a disc. The radius of the disc can be obtained from the formula for the radius of convergence of a power series.

If $\alpha_n = 1$ ($n = 0, 1, 2, \dots$), then $\prod_{i=0}^{n-1} \alpha_i = 1$ ($n = 1, 2, 3, \dots$), and therefore the power series converges in the open unit disc; cf. Solution 67. If

$$\alpha_n = \left(1 + \frac{1}{n+1}\right)^2 \quad \left(= \left(\frac{n+2}{n+1}\right)^2\right),$$

then $\prod_{i=0}^{n-1} \alpha_i = (n+1)^2$, and therefore the power series converges in the closed unit disc, which in this case happens to be the same as the spectrum; cf. Solution 77. If $\alpha_n = 1/(n+1)$, then $\prod_{i=0}^{n-1} \alpha_i = 1/n!$, and therefore the power series converges at the origin only.

Solution 79. If $p = \{p_n\}$ is a sequence of positive numbers such that $\{p_{n+1}/p_n\}$ is bounded, then the shift S on $l^2(p)$ is unitarily equivalent to the weighted shift A , with weights $\{\sqrt{p_{n+1}/p_n}\}$, on l^2 .

Proof. If $f = \langle \xi_0, \xi_1, \xi_2, \dots \rangle \in l^2(p)$ write

$$Uf = \langle \sqrt{p_0}\xi_0, \sqrt{p_1}\xi_1, \sqrt{p_2}\xi_2, \dots \rangle.$$

The transformation U maps $l^2(p)$ into l^2 ; it is clearly linear and isometric. If $\langle \eta_0, \eta_1, \eta_2, \dots \rangle \in l^2$, and if $\xi_n = \eta_n/\sqrt{p_n}$, then $\sum_{n=0}^{\infty} p_n |\xi_n|^2 = \sum_{n=0}^{\infty} |\eta_n|^2$; this proves that U maps $l^2(p)$ onto l^2 .

Assertion: U transforms S onto A . Computation:

$$\begin{aligned}
 USU^{-1} \langle \eta_0, \eta_1, \eta_2, \dots \rangle &= US \langle \eta_0 / \sqrt{p_0}, \eta_1 / \sqrt{p_1}, \eta_2 / \sqrt{p_2}, \dots \rangle \\
 &= U \langle 0, \eta_0 / \sqrt{p_0}, \eta_1 / \sqrt{p_1}, \eta_2 / \sqrt{p_2}, \dots \rangle \\
 &= \langle 0, \sqrt{p_1/p_0} \eta_0, \sqrt{p_2/p_1} \eta_1, \sqrt{p_3/p_2} \eta_2, \dots \rangle \\
 &= A \langle \eta_0, \eta_1, \eta_2, \dots \rangle.
 \end{aligned}$$

Conclusion: the transform of the ordinary shift on a weighted sequence space is a weighted shift on the ordinary sequence space.

In view of this result, all questions about weighted sequence spaces can be answered in terms of weighted shifts. The spectral radius of S , for instance, is $\lim_k \sup_n (\prod_{i=0}^{k-1} \sqrt{p_{n+i+1}/p_{n+i}})^{1/k}$ (see Solution 77).

Solution 80. If A is a unilateral weighted shift with positive weights α_n such that $\alpha_n \rightarrow 0$, then $\Lambda(A) = \{0\}$ and $\Pi_0(A) = \emptyset$.

Proof. Use Solution 77 to evaluate $r(A)$. In many special cases that is quite easy to do. If, for instance, $\alpha_n = 1/2^n$, then the supremum (over all n) of $(\prod_{i=0}^{k-1} 1/2^{n+i})^{1/k}$ is attained when $n = 0$. It follows that that supremum is

$$\left(\prod_{i=0}^{k-1} \frac{1}{2^i} \right)^{1/k} = \frac{1}{2^m},$$

where

$$m = \frac{1}{k} \sum_{i=0}^{k-1} i = \frac{1}{2}(k-1).$$

This implies that the supremum tends to 0 as k tends to ∞ .

In the general case, observe first that $(\prod_{i=0}^{k-1} \alpha_i)^{1/k} \rightarrow 0$ as $k \rightarrow \infty$. (This assertion is the multiplicative version of the one according to which convergence implies Cesaro convergence. The additive version is that if $\alpha_n \rightarrow 0$, then $(1/k) \sum_{i=0}^{k-1} \alpha_i \rightarrow 0$. The proofs are easy and similar. It is also easy to derive the multiplicative one from the additive one.)

Since $\alpha_{n+1} \rightarrow 0$, it follows equally that $(\prod_{i=0}^{k-1} \alpha_{1+i})^{1/k} \rightarrow 0$; more generally, $(\prod_{i=0}^{k-1} \alpha_{n+i})^{1/k} \rightarrow 0$ as $k \rightarrow \infty$ for each n .

The problem is to prove that $\sup_n (\prod_{i=0}^{k-1} \alpha_{n+i})^{1/k}$ is small when k is large. Given $\varepsilon (>0)$ and given $n (= 0, 1, 2, \dots)$, find $k_0(n, \varepsilon)$ so that $(\prod_{i=0}^{k-1} \alpha_{n+i})^{1/k} < \varepsilon$ whenever $k \geq k_0(n, \varepsilon)$. If n_0 is such that $\alpha_n < \varepsilon$ for $n \geq n_0$, then $\max(k_0(0, \varepsilon), k_0(1, \varepsilon), \dots, k_0(n_0-1, \varepsilon))$ is "large" enough; if k is greater than or equal to this maximum, then $\sup_n (\prod_{i=0}^{k-1} \alpha_{n+i})^{1/k} < \varepsilon$. Indeed, if $n < n_0$, then $(\prod_{i=0}^{k-1} \alpha_{n+i})^{1/k} < \varepsilon$ just because $k \geq k_0(n, \varepsilon)$; if $n \geq n_0$, then $(\prod_{i=0}^{k-1} \alpha_{n+i})^{1/k} < \varepsilon$ because each factor in the product is less than ε .

To see that $\Pi_0(A)$ is empty, apply Solution 78.

Solution 81. *There exists a countable set of operators, each with spectrum $\{0\}$, whose direct sum has spectral radius 1.*

Proof. Here is an example described in terms of weighted shifts. Consider the (unilateral) sequence

$$\{1, 0, 1, 1, 0, 1, 1, 1, 0, 1, 1, 1, 1, 0, \dots\},$$

and let A be the unilateral weighted shift with these weights. The 0's in the sequence guarantee that A is the direct sum of the operators given by

$$\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \quad \dots,$$

and hence it is the direct sum of operators each of which has spectrum $\{0\}$. Since, however, the sequence of weights has arbitrarily long blocks of 1's, the formula for the spectral radius of a weighted shift (Solution 77) implies that $r(A) = 1$.

What makes such examples possible is the misbehavior of the approximate point spectrum. For the point spectrum the best possible assertion is true (and easy to prove): the point spectrum of a direct sum is the

union of the point spectra of the summands. Passage to adjoints implies that the same best possible assertion is true for the compression spectrum.

Solution 82. The main step in the proof is the inequality

$$|(ABf, f)|^{2^n} \leq (AB^{2^n}f, f) \cdot (Af, f)^{2^n-1}.$$

To prove this, proceed by induction. For $n = 0$, the assertion is trivial. For the induction step, apply the Schwarz inequality to the inner product defined by the operator A , as follows:

$$\begin{aligned} |(ABf, f)|^{2^{n+1}} &= (|(ABf, f)|^{2^n})^2 \\ &\leq ((AB^{2^n}f, f)(Af, f)^{2^n-1})^2 \\ &\leq (AB^{2^n}f, B^{2^n}f) \cdot (Af, f) \cdot (Af, f)^{2^{n+1}-2} \\ &\leq (B^{*2^n}AB^{2^n}f, f) \cdot (Af, f)^{2^{n+1}-1}. \end{aligned}$$

The desired inequality would now follow if it were known that powers of B^* can be shifted from the left of A to become powers of B on the right. Precisely: $B^{*k}A = AB^k$ for all k . The induction argument that proves this is easy; only the case $k = 1$ is worth a second glance. That is where the assumption that AB is Hermitian comes in; indeed, $AB = (AB)^* = B^*A$.

The proof of the inequality involving the spectral radius of B is now immediate. The inequality just established implies that

$$|(ABf, f)|^{2^n} \leq \|A\| \cdot \|B^{2^n}\| \cdot \|f\|^2 \cdot (Af, f)^{2^n-1}.$$

In this latter inequality take the 2^n -th root of both sides and pass to the limit as $n \rightarrow \infty$.

Chapter 11. Norm topology

Solution 83. *The metric space of operators on an infinite-dimensional Hilbert space is not separable.*

Proof. Since every infinite-dimensional Hilbert space has a separable infinite-dimensional subspace, and since every separable infinite-dimensional space is isomorphic to $L^2(0,1)$, there is no loss of generality in assuming that the underlying Hilbert space is $L^2(0,1)$ to begin with. That granted, let φ_t be the characteristic function of $[0,t]$, and let P_t be the multiplication operator induced by φ_t , $0 \leq t \leq 1$. If $s < t$, then $P_t - P_s$ is the multiplication operator induced by the characteristic function of $(s,t]$, and therefore $\|P_t - P_s\| = 1$. Conclusion: there exists an uncountable set of operators such that the distance between any two of them is 1; the existence of such a set is incompatible with separability. For an alternative example of the same thing, consider diagonal operators whose diagonals consist of 0's and 1's only.

Solution 84. *The set of invertible operators is open and inversion is continuous.*

Proof. Recall first that if $\|1 - A\| < 1$, then A is invertible and $A^{-1} = \sum_{n=0}^{\infty} (1 - A)^n$ (cf. Solution 72); it follows that

$$\|A^{-1}\| \leq \sum_{n=0}^{\infty} \|1 - A\|^n = \frac{1}{1 - \|1 - A\|}.$$

Suppose now that A_0 is an invertible operator. Since

$$1 - AA_0^{-1} = (A_0 - A)A_0^{-1}$$

for each A , it follows that if $\|A_0 - A\| < 1/\|A_0^{-1}\|$, then $\|1 - AA_0^{-1}\| < 1$. This implies that if $\|A_0 - A\| < 1/\|A_0^{-1}\|$,

then A is invertible (because AA_0^{-1} is) and

$$\begin{aligned} \|A^{-1}\| &= \|((AA_0^{-1})A_0)^{-1}\| \leq \|A_0^{-1}\| \cdot \|A_0A^{-1}\| \\ &\leq \frac{\|A_0^{-1}\|}{1 - \|A_0 - A\| \cdot \|A_0^{-1}\|}. \end{aligned}$$

Conclusion: not only is the set of invertible operators open, but so long as an operator stays in a sufficiently small neighborhood of one of them, it is not only invertible, but its inverse remains bounded.

The result of the preceding paragraph makes the continuity proof accessible. Observe that

$$\|A_0^{-1} - A^{-1}\| = \|A_0^{-1}(A - A_0)A^{-1}\| \leq \|A_0^{-1}\| \cdot \|A - A_0\| \cdot \|A^{-1}\|.$$

If A_0 is fixed and if A is sufficiently near to A_0 , then the middle factor on the right makes the outcome small, and the other two factors remain bounded.

Solution 85. The sequence of weights for A_k is

$$\left\{ \dots, 1, 1, 1, \left(\frac{1}{k}\right), 1, 1, 1, \dots \right\}.$$

Since $1/k \leq 1$, it follows that the supremum of the sliding products that enter the formula for the spectral radius of a weighted shift (see Solution 77) is equal to 1, and hence that $r(A_k) = 1$. Conclusion: the spectrum of A_k is included in the closed unit disc, and this is true for $k = 1, 2, 3, \dots, \infty$.

If $k < \infty$, then A_k is invertible, and, in fact, A_k^{-1} itself is a weighted shift. Since $A_k^{-1}e_n$ is e_{n-1} or ke_{n-1} according as $n \neq 1$ or $n = 1$, it follows that A_k^{-1} shifts the e_n 's backwards (and weights them as just indicated). Backwards and forwards are indistinguishable to within unitary equivalence (cf. Solution 68), and, consequently, the theory of weighted shifts is applicable to A_k^{-1} . The sequence of weights for A_k^{-1} is

$$\{\dots, 1, 1, 1, (1), k, 1, 1, 1, \dots\}.$$

The supremum of the sliding products of length m is now equal to k ; it follows that $r(A_k^{-1}) = \lim_m k^{1/m} = 1$. Conclusion: the spectrum of A_k^{-1} is included in the closed unit disc, and this is true for $k = 1, 2, 3, \dots$ (but not for ∞).

The conclusions of the preceding two paragraphs, together with the spectral mapping theorem for inverses, imply that the spectrum of A_k (and also the spectrum of A_k^{-1}) is included in the unit circle (perimeter). This, together with the circular symmetry of the spectra of weighted shifts (see Problem 75), implies that the spectrum of A_k is equal to the unit circle ($k = 1, 2, 3, \dots$).

The spectrum of A_∞ is clearly not the unit circle; since $A_\infty e_0 = 0$, the spectrum of A_∞ contains the origin. This shows that the spectrum of A_∞ is discontinuously different from the spectra of the other A_k 's. (Note that $A_k \rightarrow A_\infty$, i.e., $\|A_k - A_\infty\| \rightarrow 0$, as $k \rightarrow \infty$.) The spectrum of A_∞ is, in fact, equal to the unit disc. The quickest way to prove this is to note that the span of the e_n 's with $n > 0$ reduces A_∞ (both it and its orthogonal complement are invariant under A_∞), and that the restriction of A_∞ to that span is the unilateral shift. Since the spectrum of every operator includes the spectrum of each direct summand, the proof is complete.

This example is due to G. Lumer.

Solution 86. Let \mathbf{T} be the set of all singular operators (on a fixed Hilbert space), and, given an operator A , fixed from now on, let $\varphi(\lambda)$ be the distance (in the metric space of operators) from $A - \lambda$ to \mathbf{T} . The function φ is continuous. (This is an elementary fact about metric spaces; it does not even depend on \mathbf{T} being closed.) If Λ_0 is an open set that includes $\Lambda(A)$, if Δ is the closed disc with center 0 and radius $1 + \|A\|$, and if $\lambda \in \Delta - \Lambda_0$, then $\varphi(\lambda) > 0$. (This does depend on \mathbf{T} being closed; if $\varphi(\lambda) = 0$, i.e., $d(A - \lambda, \mathbf{T}) = 0$, then $A - \lambda \in \mathbf{T}$, i.e., $\lambda \in \Lambda(A)$.) Since $\Delta - \Lambda_0$ is compact, there exists a positive number ε such that $\varphi(\lambda) \geq \varepsilon$ for all λ in $\Delta - \Lambda_0$; there is clearly no loss of generality in assuming that $\varepsilon < 1$. Suppose now that $\|A - B\| < \varepsilon$. It follows that if $\lambda \in \Delta - \Lambda_0$, then

$$\|(A - \lambda) - (B - \lambda)\| < \varepsilon \leq d(A - \lambda, \mathbf{T}).$$

This implies that $B - \lambda$ is not in \mathbf{T} , and hence that λ is not in $\Lambda(B)$. Conclusion: $\Lambda(B)$ is disjoint from $\Delta - \Lambda_0$. At the same time, if $\lambda \in \Lambda(B)$, then

$$|\lambda| \leq \|B\| \leq \|A\| + \|A - B\| < 1 + \|A\|,$$

so that $\Lambda(B) \subset \Delta$. These two properties of $\Lambda(B)$ say exactly that $\Lambda(B) \subset \Lambda_0$; the proof is complete.

A different proof can be based on the known properties of resolvents. If $\varphi(\lambda) = \|(A - \lambda)^{-1}\|$, then φ is defined and continuous outside Λ_0 ; since it vanishes at ∞ , it is bounded (cf. Problem 72). If, say, $\varphi(\lambda) < \alpha$ whenever $\lambda \notin \Lambda_0$, put $\varepsilon = 1/\alpha$. If $\|A - B\| < \varepsilon$ and $\lambda \notin \Lambda_0$, then

$$\|(A - \lambda) - (B - \lambda)\| = \|A - B\| < \varepsilon < \frac{1}{\|(A - \lambda)^{-1}\|};$$

it follows as in Solution 84 that $B - \lambda$ is invertible.

The metric space proof is due to C. Wasiutynski; the resolvent proof is due to E. A. Nordgren.

Solution 87. *There exists a convergent sequence of nilpotent operators such that the spectral radius of the limit is positive.*

Proof. The construction is based on a sequence $\{\varepsilon_n\}$ of positive numbers converging to 0. The question of what the ε 's can be will be answered after the question of what they are expected to do. Begin by defining a sequence $\{\alpha_n\}$ as follows: every second α is equal to ε_0 (i.e., $\alpha_0 = \varepsilon_0, \alpha_2 = \varepsilon_0, \alpha_4 = \varepsilon_0, \dots$); every second one of the remaining α 's is equal to ε_1 (i.e., $\alpha_1 = \varepsilon_1, \alpha_5 = \varepsilon_1, \alpha_9 = \varepsilon_1, \dots$); every second one of the still remaining α 's is equal to ε_2 ; and so on ad infinitum. The sequence of α 's looks like this:

$$\varepsilon_0, \varepsilon_1, \varepsilon_0, \varepsilon_2, \varepsilon_0, \varepsilon_1, \varepsilon_0, \varepsilon_3, \varepsilon_0, \varepsilon_1, \varepsilon_0, \varepsilon_2, \varepsilon_0, \varepsilon_1, \varepsilon_0, \dots$$

Let A be the weighted unilateral shift whose weights are the α 's, and, for each non-negative integer k , let A_k be the weighted unilateral shift whose weights are what the α 's become when each ε_k is replaced

by 0. Thus, for instance, the sequence of weights for A_2 is

$$\varepsilon_0, \varepsilon_1, \varepsilon_0, 0, \varepsilon_0, \varepsilon_1, \varepsilon_0, \varepsilon_3, \varepsilon_0, \varepsilon_1, \varepsilon_0, 0, \varepsilon_0, \varepsilon_1, \varepsilon_0, \dots$$

Two things are obvious from this construction: A_k is nilpotent of index 2^{k+1} , and the norm of $A_k - A$ (which is a weighted shift) is ε_k .

All that remains is to prove that the ε 's can be chosen so as to make $r(A) > 0$. For this purpose note that

$$\alpha_0 = \varepsilon_0,$$

$$\alpha_0 \alpha_1 \alpha_2 = \varepsilon_0^2 \varepsilon_1,$$

$$\alpha_0 \alpha_1 \alpha_2 \alpha_3 \alpha_4 \alpha_5 \alpha_6 = \varepsilon_0^4 \varepsilon_1^2 \varepsilon_2,$$

and, in general, if $n = 2^p - 2$ ($p = 1, 2, 3, \dots$), then

$$\alpha_0 \cdots \alpha_n = \varepsilon_0^{2^{p-1}} \cdots \varepsilon_{p-1}.$$

Hence

$$\log(\alpha_0 \cdots \alpha_n) = \sum_{k=0}^{p-1} 2^{p-1-k} \log \varepsilon_k = 2^p \sum_{k=0}^{p-1} \frac{\log \varepsilon_k}{2^{k+1}},$$

or

$$\log(\alpha_0 \cdots \alpha_n)^{1/n+1} = \frac{2^p}{2^p - 1} \sum_{k=0}^{p-1} \frac{\log \varepsilon_k}{2^{k+1}}.$$

This implies that if the series

$$\sum_{k=0}^{\infty} \frac{\log \varepsilon_k}{2^{k+1}}$$

is convergent (which happens if, for instance, $\varepsilon_k = 1/2^k$), then

$$\liminf_n \log(\alpha_0 \cdots \alpha_n)^{1/n+1} > -\infty,$$

and therefore

$$\liminf_n (\alpha_0 \cdots \alpha_n)^{1/n+1} > 0.$$

The desired conclusion follows from Solution 78.

This example is due to S. Kakutani; see Rickart [1960, p. 282].

Chapter 12.

Strong and weak topologies

Solution 88. The first assertion involving uniformity has nothing to do with operators; it is just a special case of Problem 16. To prove the second assertion, assume $A = 0$; this loses no generality. The assumption in this case is that, for each positive number ε , if n is sufficiently large, then

$$\|A_n f\| < \varepsilon \quad \text{whenever } \|f\| = 1;$$

the uniformity manifests itself in that the size of n does not depend on f . It follows that if n is sufficiently large, then

$$\left\| A_n \frac{f}{\|f\|} \right\| < \varepsilon \quad \text{whenever } f \neq 0,$$

and hence that

$$\|A_n f\| \leq \varepsilon \|f\| \quad \text{for all } f.$$

This implies that if n is sufficiently large, then

$$\|A_n\| \leq \varepsilon.$$

The argument is general; it applies to all nets, not only to sequences.

Solution 89. *The norm is continuous with respect to the uniform topology and discontinuous with respect to the strong and weak topologies.*

Proof. The proof for the uniform topology is contained in the inequality

$$|\|A\| - \|B\|| \leq \|A - B\|.$$

This is just a version of the subadditivity of the norm, and it implies that the norm is a uniformly continuous function in the norm topology. The proof says nothing about the continuity of the norm in any other topology. A norm is always continuous with respect to the topology it defines; other topologies take their chances.

To show that the norm is not continuous with respect to the strong topology (not even sequentially), and, a fortiori, it is not continuous with respect to the weak topology, consider the following example. Let $\{\mathbf{M}_n\}$ be a decreasing sequence of non-zero subspaces with intersection $\{0\}$, and let $\{P_n\}$ be the corresponding sequence of projections. The sequence $\{P_n\}$ converges to 0 strongly. (To see this, form an orthonormal basis for \mathbf{M}_1^\perp , one for $\mathbf{M}_1 \cap \mathbf{M}_2^\perp$, another for $\mathbf{M}_2 \cap \mathbf{M}_3^\perp$, etc., and string them together to make a basis for the whole space. Cf. also Solution 94.) The sequence $\{\|P_n\|\}$ of norms does not converge to the number 0; indeed $\|P_n\| = 1$ for all n .

Solution 90. *The adjoint is continuous with respect to the uniform and the weak topologies and discontinuous with respect to the strong topology.*

Proof. The proof for the uniform topology is contained in the identity

$$\|A^* - B^*\| = \|A - B\|.$$

If a function from one space to another is continuous, then it remains so if the topology of the domain is made larger, and it remains so if the topology of the range is made smaller. (This is the reason why the strong discontinuity of the norm implies its weak discontinuity.) If, however, a mapping from a space to itself is continuous, then there is no telling how it will behave when the topology is changed; every change works both ways. In fact, everything can happen, and the adjoint proves it. As the topologies march down (from norm to strong to weak), the adjoint changes from being continuous to being discontinuous, and back again.

To prove the strong discontinuity of the adjoint, let U be the unilateral shift, and write $A_k = U^{*k}$, $k = 1, 2, 3, \dots$. Assertion: $A_k \rightarrow 0$

strongly, but the sequence $\{A_k^*\}$ is not strongly convergent to anything. Indeed:

$$\begin{aligned} \|A_k \langle \xi_0, \xi_1, \xi_2, \dots \rangle\|^2 &= \|\langle \xi_k, \xi_{k+1}, \xi_{k+2}, \dots \rangle\|^2 \\ &= \sum_{n=k}^{\infty} \|\xi_n\|^2, \end{aligned}$$

so that $\|A_k f\|^2$ is, for each f , the tail of a convergent series, and therefore $A_k f \rightarrow 0$. The negative assertion about $\{A_k^*\}$ can be established by proving that if $f \neq 0$, then $\{A_k^* f\}$ is not a Cauchy sequence. Indeed:

$$\begin{aligned} \|A_{m+n}^* f - A_n^* f\|^2 &= \|U^{m+n} f - U^n f\|^2 = \|U^m f - f\|^2 \\ &= \|U^m f\|^2 - 2 \operatorname{Re}(U^m f, f) + \|f\|^2 \\ &= 2(\|f\|^2 - \operatorname{Re}(f, U^{*m} f)). \end{aligned}$$

Since $\|U^{*m} f\| \rightarrow 0$ as $m \rightarrow \infty$, it follows that $\|A_{m+n}^* f - A_n^* f\|$ refuses to become small as m and n become large; in fact if m is large, then $\|A_{m+n}^* f - A_n^* f\|$ is nearly equal to $\sqrt{2} \|f\|$.

As for the weak continuity of the adjoint, that is implied by the identity

$$|(A^* f, g) - (B^* f, g)| = |(f, Ag) - (f, Bg)| = |(Ag, f) - (Bg, f)|.$$

Solution 91. The proof for the uniform topology is contained in the inequalities

$$\begin{aligned} \|AB - A_0 B_0\| &\leq \|AB - AB_0\| + \|AB_0 - A_0 B_0\| \\ &\leq \|A\| \|B - B_0\| + \|A - A_0\| \|B_0\| \\ &\leq (\|A - A_0\| + \|A_0\|) \|B - B_0\| + \|A - A_0\| \|B_0\|. \end{aligned}$$

An elegant counterexample for the strong topology depends on the assertion that the set of all nilpotent operators of index 2 (i.e., the set of all operators A such that $A^2 = 0$) is strongly dense. (The idea is due

to Arnold Lebow.) To prove this, suppose that

$$\{A: \|A_0 f_i - A f_i\| < \epsilon, i = 1, \dots, k\}$$

is an arbitrary basic strong neighborhood. There is no loss of generality in assuming that the f 's are linearly independent (or even orthonormal); otherwise replace them by a linearly independent (or even orthonormal) set with the same span, and, at the same time, make ϵ as much smaller as is necessary. For each i ($= 1, \dots, k$) find a vector g_i such that $\|A_0 f_i - g_i\| < \epsilon$ and such that the span of the g 's has only 0 in common with the span of the f 's; so long as the underlying Hilbert space is infinite-dimensional, this is possible. Let A be the operator such that

$$A f_i = g_i \text{ and } A g_i = 0 \quad (i = 1, \dots, k)$$

and

$$A h = 0 \quad \text{whenever } h \perp f_i \text{ and } h \perp g_i \quad (i = 1, \dots, k).$$

Clearly A is nilpotent of index 2, and, just as clearly, A belongs to the prescribed neighborhood.

If squaring were strongly continuous, then the set of nilpotent operators of index 2 would be strongly closed, and therefore every operator would be nilpotent of index 2, which is absurd.

This result implies, of course, that multiplication is not jointly strongly continuous. Since the strong topology is larger than the weak, so that a strongly dense set is necessarily weakly dense, the auxiliary assertion about nilpotent operators holds for the weak topology as well as for the strong. Conclusion: squaring is not weakly continuous, and, consequently, multiplication is not jointly weakly continuous.

Solution 92. The easiest proof uses convergence. The convergence of sequences is sometimes misleading in general topology, but the convergence of nets (generalized sequences) is good enough. Suppose therefore that $A_j \rightarrow A$ strongly, i.e., that $A_j f \rightarrow A f$ for each f . It follows, in particular, that $A_j B f \rightarrow A B f$ for each f , and this settles strong continuity in A . If, on the other hand, $B_j \rightarrow B$ strongly, i.e., if $B_j f \rightarrow B f$ for each f , then apply A to conclude that $A B_j f \rightarrow A B f$ for each f ; this settles strong continuity in B . Weak continuity can be treated the same way. If

$(A_{ij}f, g) \rightarrow (Af, g)$ for each f and g , then, in particular, $(A_{ij}Bf, g) \rightarrow (ABf, g)$ for each f and g ; if $(B_{ij}f, g) \rightarrow (Bf, g)$ for each f and g , then, in particular, $(AB_{ij}f, g) = (B_{ij}A^*g) \rightarrow (Bf, A^*g) = (ABf, g)$ for each f and g .

Solution 93. (a) The crux of the matter is boundedness. Assume first that the sequence $\{\|A_n\|\}$ of norms is bounded. (The boundedness of $\{\|B_n\|\}$ would do just as well.) Since, for each f ,

$$\begin{aligned} \|A_n B_n f - ABf\| &\leq \|A_n B_n f - A_n Bf\| + \|A_n Bf - ABf\| \\ &\leq \|A_n\| \cdot \|(B_n - B)f\| + \|(A_n - A)Bf\|, \end{aligned}$$

the assumed boundedness implies, as desired, that $A_n B_n f \rightarrow ABf$.

Now what about the boundedness assumption? The answer is that it need not be assumed at all; it can be proved. It is, in fact, an immediate consequence of the principle of uniform boundedness for operators: if a sequence of operators is weakly convergent (and all the more if it is strongly convergent), then it is weakly bounded, and therefore bounded.

(b) *Multiplication is not weakly sequentially continuous.*

Proof. Let U be the unilateral shift, and write $A_n = U^{*n}$, $B_n = U^n$, $n = 1, 2, 3, \dots$. Since $A_n \rightarrow 0$ strongly, it follows that $A_n \rightarrow 0$ weakly, and hence that $B_n \rightarrow 0$ weakly (cf. Solution 90). Since, however, $A_n B_n = 1$ for all n , it is not true that $A_n B_n \rightarrow 0$ weakly.

Solution 94. *A bounded increasing sequence of Hermitian operators is always convergent with respect to the strong topology, but not necessarily with respect to the uniform topology.*

Proof. One way to prove the assertion about the strong topology is to make use of the weak version. Let $\{A_n\}$ be a bounded increasing sequence of Hermitian operators, and let A be its weak limit. Since $A_n \leq A$, the operator $A - A_n$ is positive, and therefore it has a positive square root, say B_n (see Problem 95). Since

$$\|B_n f\|^2 = (B_n f, B_n f) = (B_n^2 f, f) = ((A - A_n)f, f) \rightarrow 0,$$

the sequence $\{B_n\}$ tends strongly to 0. Since $\{\|A - A_n\|\}$ is bounded, so is $\{\|B_n\|\}$; say $\|B_n\| \leq \beta$ for all n . The asserted strong convergence now follows from the relation

$$\|(A - A_n)f\| = \|B_n^2 f\| \leq \beta \|B_n f\|.$$

As once before (cf. Solution 1) sequences play no essential role here; nets would do just as well.

There is sometimes a technical advantage in not using the theorem about the existence of positive square roots. The result just obtained can be proved without that theorem, if it must be, but the proof with square roots shows better what really goes on. Here is how a proof without square roots goes. Assume, with no loss of generality, that $A \leq 1$. If $m < n$, then

$$\begin{aligned} \|(A_n - A_m)f\|^4 &= ((A_n - A_m)f, (A_n - A_m)f)^2 \\ &\leq ((A_n - A_m)f, f)((A_n - A_m)^2 f, (A_n - A_m)f), \end{aligned}$$

by the Schwarz inequality for the inner product determined by the positive operator $A_n - A_m$. Since $A_n - A_m \leq 1$, so that $\|A_n - A_m\| \leq 1$, it follows that

$$\|A_n f - A_m f\|^4 \leq ((A_n f, f) - (A_m f, f))\|f\|^2.$$

A frequently used consequence of the strong convergence theorem is about projections. If $\{\mathbf{M}_n\}$ is an increasing sequence of subspaces, then the corresponding sequence $\{P_n\}$ of projections is an increasing (and obviously bounded) sequence of Hermitian operators. It follows that there exists a Hermitian operator P such that $P_n \rightarrow P$ strongly. Assertion: P is the projection onto the span, say \mathbf{M} , of all the \mathbf{M}_n 's (cf. Solution 89). Reason: if f belongs to some \mathbf{M}_n , then $Pf = f$, and if f is orthogonal to all \mathbf{M}_n 's, then $Pf = 0$; these two comments together imply that there is a dense set on which P agrees with the projection onto \mathbf{M} .

Increasing sequences of projections serve also to show that the monotone convergence assertion is false for the uniform topology. Indeed, if the sequence $\{\mathbf{M}_n\}$ is strictly increasing, then the sequence $\{P_n\}$ cannot converge to P (or, for that matter, to anything at all) in the norm,

because it is not even a Cauchy sequence. In fact, a monotone sequence of projections can be a Cauchy sequence in trivial cases only; $\|P_n - P_m\| = 1$ unless $P_n = P_m$.

Solution 95. It is convenient (for purposes of reference) to break up the proof into small steps, as follows.

(1) All the positive integral powers of a positive operator are positive. Indeed $(A^{2n}f, f) = \|A^n f\|^2$ and $(A^{2n+1}f, f) = (A \cdot A^n f, A^n f)$; the former is positive because norms are, and the latter is positive because A is. In the sequel the result is needed not for A but for $1 - A$. (Note: the assertion is a trivial consequence of the spectral theorem.)

(2) Each B_n is a polynomial in $1 - A$ with positive coefficients (by induction), and hence (by (1)) each B_n is a positive operator.

(3) By (2), all the B_n 's commute with one another, and it follows that

$$B_{n+2} - B_{n+1} = \frac{1}{2}(B_{n+1}^2 - B_n^2) = \frac{1}{2}(B_{n+1} - B_n)(B_{n+1} + B_n).$$

This implies (by (2) and induction) that $B_{n+1} - B_n$ is a polynomial in $1 - A$ with positive coefficients, and hence positive; it follows that the sequence $\{B_n\}$ is increasing.

(4) The definition of B_{n+1} in terms of B_n implies (induction) that $\|B_n\| \leq 1$ for all n ; the sequence $\{B_n\}$ is bounded.

(5) By (3) and (4), $\{B_n\}$ is a bounded increasing sequence of positive operators, and therefore it is strongly convergent to some (necessarily positive) operator B . Note that the argument needs Solution 94. Since the point of what is now going on is to avoid square roots, it is necessary to use the version of Solution 94 that does not use square roots.

Convergence is proved; it remains only to evaluate the limit. This is easy from Problem 93; since $B_n \rightarrow B$ (strongly), it follows that $B_n^2 \rightarrow B^2$ (strongly), and hence that

$$B = \frac{1}{2}((1 - A) + B^2).$$

This says that

$$A = 1 - 2B + B^2 = (1 - B)^2,$$

and the proof is complete.

The proof is standard; cf. Riesz-Nagy [1952, §104].

Solution 96. Even a small amount of experience with non-commutative projections shows that the familiar algebraic operations are not likely to suffice to express $E \wedge F$ in terms of E and F . The following quite pretty and geometrical consideration shows how topology comes in, and motivates the actual proof. Suppose that the underlying Hilbert space \mathbf{H} is two-dimensional real Euclidean space, and suppose that \mathbf{M} and \mathbf{N} are two distinct but not orthogonal lines through the origin. Take an arbitrary point f in \mathbf{H} , project it on \mathbf{M} (i.e., form Ef), project the result on \mathbf{N} (FEf), then project on \mathbf{M} ($EFEf$), and continue so on ad infinitum; it looks plausible that the sequence so obtained converges to 0, which, in this case, is $(E \wedge F)f$. This suggests the formation of the sequence

$$E, FE, EFE, FEFE, EFEFE, \dots$$

The proof itself works with the subsequence

$$EFE, EF EFE, EF EF EFE, \dots;$$

this is a matter of merely technical convenience.

Since $\|EFE\| \leq 1$, the powers of EFE form a decreasing (and even commutative) sequence of positive operators. It follows that $(EFE)^n$ is weakly convergent to, say, G ; since in this case weak and strong convergence are equivalent, G belongs to the given von Neumann algebra. Assertion: $G = E \wedge F$. Clearly G is Hermitian. Since $(EFE)^m G = G$ for all m , therefore $G^2 = G$, so that G is a projection. Since $(EFE)^m FG = G$ for all m , therefore $GFG = G$; this implies that $G \leq F$. (Proof: $0 = G - GFG = G(1 - F)G = G(1 - F)(1 - F)G$, and $(1 - F)G = (G(1 - F))^*$.) Since $E(EFE)^n = (EFE)^n$ for all n , therefore $EG = G$ or $G \leq E$. If, finally, G_0 is a projection such that $G_0 \leq E$ and $G_0 \leq F$, then $G_0(EFE)^n = G_0$, whence $G_0G = G_0$, so that $G_0 \leq G$. The proof is complete.

The theorem has its own dual for an easy corollary. The assertion is that the projection $E \vee F$ on the subspace $\mathbf{M} \vee \mathbf{N}$ belongs to any von Neumann algebra containing E and F . Since

$$E \vee F = 1 - ((1 - E) \wedge (1 - F)),$$

the proof is immediate.

An examination of the proof shows that not all the defining properties of von Neumann algebras were used; all that was needed was a sequentially strongly closed set of operators such that if A and B are in the set, then so is ABA (for the theorem about $E \wedge F$) or

$$1 - (1 - A)(1 - B)(1 - A)$$

(for the theorem about $E \vee F$). Observe that even in the latter case it is not required that 1 belong to the set; an expression such as

$$1 - (1 - A)(1 - B)(1 - A)$$

is a convenient way of writing something that can obviously (though clumsily) be written without 1 if so desired.

Chapter 13. Partial isometries

Solution 97. Use the spectral theorem to represent A as a multiplication by, say, φ . If $\lambda \in \Lambda(A)$ and if N is an arbitrary neighborhood of $F(\lambda)$, then $F^{-1}(N)$ is a neighborhood of λ , and therefore $\varphi^{-1}(F^{-1}(N))$ has positive measure. Since $\varphi^{-1}(F^{-1}(N)) = (F \circ \varphi)^{-1}(N)$, it follows that λ is in the essential range of $F \circ \varphi$, so that $\lambda \in \Lambda(F(A))$. This proves that $F(\Lambda(A)) \subset \Lambda(F(A))$.

To prove the reverse inclusion is the same as to prove that if $\lambda \notin F(\Lambda(A))$, then $\lambda \notin \Lambda(F(A))$. The set $F(\Lambda(A))$ is compact. (It is the image under the continuous function F of the compact set $\Lambda(A)$.) Since λ is not in it, λ has a neighborhood disjoint from it. If N is such a neighborhood, $N \cap F(\Lambda(A)) = \emptyset$, then $F^{-1}(N) \cap \Lambda(A) = \emptyset$, and therefore $\varphi^{-1}(F^{-1}(N)) \cap \varphi^{-1}(\Lambda(A)) = \emptyset$. Since $\varphi^{-1}(\Lambda(A))$ can differ from the whole underlying measure space by a set of measure zero at most, it follows that $(F \circ \varphi)^{-1}(N)$ has measure zero, and hence that λ does not belong to the spectrum of $F(A)$. This completes the proof.

Solution 98. Suppose that \mathbf{H} and \mathbf{K} are Hilbert spaces and suppose that U is a partial isometry from \mathbf{H} into \mathbf{K} with initial space \mathbf{M} . (For a discussion of such transformations and their adjoints, see Problem 40.) If E is the projection from \mathbf{H} onto \mathbf{M} , and if $f \in \mathbf{M}$, then

$$(U^*Uf, f) = \|Uf\|^2 = \|f\|^2 = (Ef, f);$$

if $f \perp \mathbf{M}$, then

$$(U^*Uf, f) = 0 = (Ef, f).$$

It follows that $(U^*Uf, f) = (Ef, f)$ for all f in \mathbf{H} , and this implies that $U^*U = E$.

Suppose, conversely, that U is a bounded linear transformation from \mathbf{H} into \mathbf{K} such that U^*U is a projection with domain \mathbf{H} and range \mathbf{M} , say. It follows that

$$\|Uf\|^2 = (U^*Uf, f) = (Ef, f) = \|Ef\|^2$$

for all f , and hence that $\|Uf\| = \|f\|$ or $Uf = 0$ according as $f \in \mathbf{M}$ or $f \perp \mathbf{M}$.

To prove Corollary 1, observe that $\ker U^*U = \ker U$ (this is true for every bounded linear transformation U). The proof of Corollary 2 is a trick. If U^*U is idempotent, then $(UU^*)^3 = U(U^*UU^*U)U^* = (UU^*)^2$; the spectral theorem implies that a Hermitian operator A with $A^3 = A^2$ is idempotent. The assertion about initial and final spaces follows from the observation that $(\ker UU^*)^\perp = (\ker U^*)^\perp = \text{ran } U$ (since $\text{ran } U$ is closed). As for Corollary 3: if U is a partial isometry, then the product of U and the projection U^*U agrees with U on both $\ker U$ and its orthogonal complement; if, conversely, $U = UU^*U$, then premultiply by U^* and conclude that U^*U is idempotent.

Solution 99. If U is an isometry and if $U\mathbf{M} = \mathbf{M}$, then \mathbf{M} reduces U ; if U is a co-isometry and if \mathbf{M} reduces U , then $U\mathbf{M} = \mathbf{M}$. The first implication is false for co-isometries; the second implication is false for isometries.

Proof. If $U^*U = 1$ and $U\mathbf{M} = \mathbf{M}$, then $U^*\mathbf{M} = U^*U\mathbf{M} = \mathbf{M}$. If $UU^* = 1$ and both $U\mathbf{M} \subset \mathbf{M}$ and $U^*\mathbf{M} \subset \mathbf{M}$, then apply U to the second inclusion to obtain the reverse of the first.

The first implication is false if U is the adjoint of the unilateral shift and \mathbf{M} is the (one-dimensional) subspace of eigenvectors belonging to a non-zero eigenvalue (see Solution 67). In that case $U\mathbf{M} = \mathbf{M}$, but \mathbf{M} does not reduce U . The second implication is false if U is the unilateral shift and \mathbf{M} is the whole space. In that case \mathbf{M} reduces U , but $U\mathbf{M} \neq \mathbf{M}$.

Solution 100. The assertion about closure is obvious; the reason is that (1) the mapping $A \rightarrow AA^*A$ is continuous, and (2) the equation $A = AA^*A$ characterizes partial isometries.

An even easier version of the same proof shows that the set of all isometries is closed; consider the mapping (1) $A \rightarrow A^*A$ and the equation (2) $A^*A = 1$. This comment is pertinent to the question concerning the connectedness of the set of all non-zero partial isometries. One way to prove that the answer to that question is no is to prove that the set of all isometries is not only closed but also open in the set of all partial isometries (in the relative topology of the latter). The fact is that if a partial isometry is sufficiently near to an isometry, then it is an isometry.

More precisely, if U is a partial isometry, if V is an isometry, and if $\|U - V\| < 1$, then U is an isometry. It is sufficient to prove that if $Uf = 0$, then $f = 0$. Indeed, since $\|f\| = \|Vf\| = \|Uf - Vf\| \leq \|U - V\| \cdot \|f\|$, it follows that if $f \neq 0$, then $\|U - V\| \geq 1$, which contradicts the assumption that $\|U - V\| < 1$.

The same argument shows that if the underlying Hilbert space is infinite-dimensional, then the set of all isometries is not connected. Reason: the set of all unitary operators is a non-empty proper (!) subset that is simultaneously open and closed.

Solution 101. The kernel of U and the initial space of V can have only 0 in common. Indeed, if f is a non-zero vector such that $Uf = 0$ and $\|Vf\| = \|f\|$, then $\|Uf - Vf\| = \|f\|$, and this contradicts the hypothesis $\|U - V\| < 1$. It follows that the restriction of U to the initial space of V is one-to-one, and hence (Problem 42) the dimension of the initial space of V is less than or equal to the dimension of the entire range of U . In other words, the result is that $\rho(V) \leq \rho(U)$; the assertion about ranks follows by symmetry.

The assertion about nullities can be phrased this way: if $\nu(U) \neq \nu(V)$, then $\|U - V\| \geq 1$. Indeed, if $\nu(U) \neq \nu(V)$, say, for definiteness, $\nu(U) < \nu(V)$, then there exists at least one unit vector f in the kernel of V that is orthogonal to the kernel of U . To say that f is orthogonal to the kernel of U is the same as to say that f belongs to the initial space of U . It follows that $1 = \|f\| = \|Uf\| = \|Uf - Vf\| \leq \|U - V\|$, and the proof of the assertion about nullities is complete.

The assertion about co-ranks is an easy corollary: if $\|U - V\| < 1$, then $\|U^* - V^*\| < 1$, and therefore $\rho'(U) = \nu(U^*) = \nu(V^*) = \rho'(V)$.

The result appears in Riesz-Nagy [1952, §105] for the special case of projections (which is, in fact, Problem 43). The present statement is a generalization, and, at the same time, the proof is a considerable simplification. The proof in Riesz-Nagy is, however, more constructive; it not only proves that two subspaces have the same dimension, but it exhibits a partial isometry that has the first for initial space and the second for final space. The generalization appears in Halmos-McLaughlin [1962].

Solution 102. Suppose that V_1 and V_2 are partial isometries with the same rank, co-rank, and nullity; let \mathbf{N}_1 and \mathbf{N}_2 be their kernels, \mathbf{M}_1 and \mathbf{M}_2 their initial spaces, and \mathbf{R}_1 and \mathbf{R}_2 their ranges. Let U be an arbitrary

unitary operator that maps \mathbf{N}_1 onto \mathbf{N}_2 and \mathbf{M}_1 onto \mathbf{M}_2 . Let W be a linear transformation that maps \mathbf{R}_1^\perp isometrically onto \mathbf{R}_2^\perp ; for f in \mathbf{R}_1 , define $Wf = V_2UV_1^*f$. Since it is easy to verify that this definition yields a linear transformation W that maps \mathbf{R}_1 isometrically onto \mathbf{R}_2 , it follows that there exists a unitary operator W that maps \mathbf{R}_1 onto \mathbf{R}_2 and \mathbf{R}_1^\perp onto \mathbf{R}_2^\perp as indicated. If $g \in \mathbf{N}_1$, then

$$WV_1g = 0 = V_2Ug;$$

if $g \in \mathbf{M}_1$, then

$$WV_1g = V_2UV_1^*V_1g = V_2Ug.$$

It follows that $WV_1 = V_2U$, or $WV_1U^* = V_2$. If $t \rightarrow W_t$ and $t \rightarrow U_t$ are continuous curves of unitary operators that join 1 to W and to U , then $t \rightarrow W_tV_1U_t^*$ is a continuous curve of partial isometries all with the same rank, co-rank, and nullity, that joins V_1 to V_2 .

This proof is a simplification of the one in Halmos-McLaughlin [1962]; it is due to R. G. Douglas.

Solution 103. Suppose that A and B are unitarily equivalent. If U is a unitary operator that transforms A onto B , then U transforms A^* onto B^* , and therefore U transforms $A' = \sqrt{1 - AA^*}$ onto $B' = \sqrt{1 - BB^*}$; it follows that

$$\begin{pmatrix} U & 0 \\ 0 & U \end{pmatrix}$$

transforms $M(A)$ onto $M(B)$.

Suppose next that A and B are invertible and that $M(A)$ and $M(B)$ are unitarily equivalent. The range of $M(A)$ consists of all ordered pairs of the form $\langle Af + A'g, 0 \rangle$. This set is included in the set of all ordered pairs with vanishing second coordinate; the invertibility of A implies that the range of $M(A)$ consists exactly of all ordered pairs with vanishing second coordinate. Since the same is true for $M(B)$, it follows that every unitary operator matrix that transforms $M(A)$ onto $M(B)$ maps the subspace of all vectors of the form $\langle f, 0 \rangle$ onto itself. This implies that that subspace reduces every such unitary operator matrix (cf. Solution 99), and hence that every such unitary operator matrix is

diagonal. Since the diagonal entries of a diagonal unitary matrix are unitary operators, it follows that if $M(A)$ and $M(B)$ are unitarily equivalent, then so also are A and B .

Solution 104. *If a compact subset Λ of the closed unit disc contains the origin, then there exists a partial isometry with spectrum Λ .*

Proof. Let A be a contraction whose spectrum is Λ (see Problem 48). If, as in Problem 103,

$$M = \begin{pmatrix} A & A' \\ 0 & 0 \end{pmatrix},$$

where $A' = \sqrt{1 - AA^*}$, then M is a partial isometry; what is its spectrum? The question reduces to this: for which values of λ is the operator matrix

$$M - \lambda = \begin{pmatrix} A - \lambda & A' \\ 0 & -\lambda \end{pmatrix}$$

not invertible? Since M^* annihilates every ordered pair whose first coordinate is 0, it follows that 0 is in the spectrum of M^* and hence of M . If $\lambda \neq 0$, then Problem 56 applies. The conclusion is that λ is in the spectrum of M if and only if λ is in the spectrum of A . Summary: $\Lambda(M) = \Lambda \cup \{0\} = \Lambda$.

In the finite-dimensional case more can be said. If Λ is a finite subset of the closed unit disc, with 0 in Λ , and if each element of Λ is assigned a positive integral multiplicity, then there exists a partial isometry with spectrum Λ whose eigenvalues have exactly the prescribed algebraic multiplicities; see Halmos-McLaughlin [1962].

Solution 105. Begin with the construction of P . Since A^*A is a positive operator on \mathbf{H} , it has a (unique) positive square root; call it P . Since

$$\|Pf\|^2 = (Pf, Pf) = (P^2f, f) = (A^*Af, f) = \|Af\|^2$$

for all f in \mathbf{H} , it follows that the equation

$$UPf = Af$$

unambiguously defines a linear transformation U from the range \mathbf{R} of P into the space \mathbf{K} , and that U is isometric on \mathbf{R} . Since U is bounded on \mathbf{R} , it has a unique bounded extension to the closure $\bar{\mathbf{R}}$, and, from there, a unique extension to a partial isometry from \mathbf{H} to \mathbf{K} with initial space $\bar{\mathbf{R}}$. The equation $A = UP$ holds by construction. The kernel of a partial isometry is the orthogonal complement of its initial space, and the orthogonal complement of the range of a Hermitian operator is its kernel. This implies that $\ker U = \ker P$ and completes the existence proof.

To prove uniqueness, suppose that $A = UP$, where U is a partial isometry, P is positive, and $\ker U = \ker P$. It follows that $A^* = PU^*$ and hence that

$$A^*A = PU^*UP = PEP,$$

where E is the projection from \mathbf{H} onto the initial space of U . Since that initial space is equal to $(\ker U)^\perp$, and hence to $\overline{\text{ran } P}$, it follows that $EP = P$, and hence that $A^*A = P^2$. Since the equation $UPf = Af$ uniquely determines U for f in $\text{ran } P$, and since $Uf = 0$ when f is in $\ker P$, it follows that U also is uniquely determined by the stated conditions.

To deduce Corollary 1, multiply $A = UP$ on the left by U^* , and use the equation $U^*U = E$; cf. Solution 98. For Corollary 2, observe that $\ker U = \ker P = \ker A^*A = \ker A$, and $\ker U^* = (\text{ran } U)^\perp = (\text{ran } A)^\perp$.

Solution 106. Suppose that A is a bounded linear transformation from a Hilbert space \mathbf{H} to a Hilbert space \mathbf{K} , let $A = UP$ be the polar decomposition of A , let $\mathbf{M}(\subset \mathbf{H})$ be the initial space of the partial isometry U , and let $\mathbf{R}(\subset \mathbf{K})$ be the range of U (or, equivalently, the closure of the range of A). If $\dim \mathbf{M}^\perp \leq \dim \mathbf{R}^\perp$, then there exist isometries from \mathbf{H} into \mathbf{K} that agree with U on \mathbf{M} (many of them); all that is needed is to map \mathbf{M}^\perp isometrically into \mathbf{R}^\perp and to combine such a mapping with what U does on \mathbf{M} . If, on the other hand, $\dim \mathbf{R}^\perp \leq \dim \mathbf{M}^\perp$, then there exist isometries from \mathbf{K} into \mathbf{H} that agree with U^* on \mathbf{R} ; the adjoint of each such isometry is a co-isometry from \mathbf{H} into \mathbf{K} that agrees with U on \mathbf{M} . In either case there exists a linear transformation V from \mathbf{H} into \mathbf{K} such that either V or V^* is an isometry and such

that V agrees with U on \mathbf{M} . Since the range of P is included in \mathbf{M} , it follows that $VP = UP = A$.

Solution 107. *The extreme points of the unit ball in the space of operators on a Hilbert space are the maximal partial isometries.*

Proof. Suppose first that U is an isometry and that $U = \alpha A + \beta B$, with $\alpha > 0, \beta > 0, \alpha + \beta = 1, \|A\| \leq 1$, and $\|B\| \leq 1$. If f is a unit vector, then so is Uf , and $Uf = \alpha Af + \beta Bf$, where $\|Af\| \leq 1$ and $\|Bf\| \leq 1$. Since the closed unit ball of a Hilbert space is strictly convex (Problem 3), it follows that $Af = Bf = Uf$, and hence that $A = B = U$. Conclusion: isometries are extreme points. The result for co-isometries is an immediate consequence.

The converse can be proved by showing that every operator A , with $\|A\| \leq 1$, is equal to a convex combination (in fact, to the average) of two extreme points of the kind already found. Here the theory of polar decompositions (or, rather, a consequence of it) is useful. By Problem 106, it is possible to write $A = VP$, where V is a maximal partial isometry and $0 \leq P \leq 1$. (The justification for the upper bound on P is that $\|A\| \leq 1$.) Assertion: there exists a unitary operator W such that $P = \frac{1}{2}(W + W^*)$. (The assertion is true and the proof below is valid whenever $-1 \leq P \leq 1$; in case the underlying Hilbert space is one-dimensional, then both the assertion and its proof make simple geometric sense.) To prove the assertion, just write

$$W = P + i\sqrt{1 - P^2},$$

and verify that everything works. Now, since $A = VP$ and $P = \frac{1}{2}(W + W^*)$, it follows that $A = \frac{1}{2}(VW + VW^*)$. Since the product of a maximal partial isometry and a unitary operator is a maximal partial isometry, the proof is complete.

Kadison [1951] has proved that, for certain operator algebras, the extreme points in the unit ball of the algebra are those partial isometries U that satisfy the identity

$$(1 - U^*U)A(1 - UU^*) = 0$$

for all A in the algebra. For the algebra of all operators on a Hilbert

space this is consistent with what was just proved. It is, indeed, clear that if either U or U^* is an isometry, then the Kadison condition is satisfied. Suppose, conversely, that the condition is satisfied, and assume that $1 - UU^* \neq 0$; it is to be proved that $1 - U^*U = 0$. In other words, it is to be proved that if $(1 - UU^*)f \neq 0$ for some f , then $(1 - U^*U)g = 0$ for each g . That is easy: given g , find an operator A such that $A(1 - UU^*)f = g$.

Solution 108. Write $UP = A$. If U commutes with P , then U commutes with P^2 ; since P also commutes with P^2 , it follows that $A (= UP)$ commutes with $A^*A (= P^2)$.

The converse is harder. If A is quasinormal, then A commutes with $P^2 (= A^*A)$. It follows from the most elementary aspects of the functional calculus that A commutes with P . (Compare Problem 95, which shows that the positive square root of a positive operator is the weak limit of a sequence of polynomials in that operator. Alternatively, apply the Weierstrass theorem on the approximation of continuous functions by polynomials to prove that “weak” can be replaced by “uniform”.) This says that $(UP - PU)P = 0$, so that $UP - PU$ annihilates $\text{ran } P$. Since $\ker P = \ker U$, it is trivial that $UP - PU$ annihilates $\ker P$ also, and it follows that $UP - PU = 0$.

Solution 109. By Problem 106, every operator has the form VP , where V is a maximal partial isometry and P is positive. Given a positive number ϵ , find an invertible operator Q (which can be made positive, if so desired) such that $\|P - Q\| < \epsilon$. It follows that $\|VP - VQ\| < \epsilon$. The proof of the density theorem for unilaterally invertible operators is completed by observing that if V is a maximal partial isometry, then V is unilaterally invertible (left invertible if V is an isometry and right invertible if V^* is one), and that the product of a unilaterally invertible operator and an invertible operator is unilaterally invertible.

To obtain the negative conclusion for invertible operators, consider an operator A that is unilaterally invertible but not invertible. (Example: the unilateral shift.) Assertion: there is a neighborhood of A that contains no invertible operators. Assume (with no loss of generality) that A has a left inverse B , with $\|B\| \leq 1$. In the presence of this normalization, the assertion can be made more precise: the open ball

with center A and radius 1 contains no invertible operators. Now, for the proof, observe first that B cannot be invertible, for if it were, then it would follow that $A = B^{-1}BA = B^{-1}$, and hence that A is invertible. It is to be proved that if $\|A - T\| < 1$, then T is not invertible. Indeed:

$$\|1 - BT\| = \|B(A - T)\| \leq \|A - T\| < 1,$$

and hence BT is invertible; this implies that if T were invertible, then B would be, but it is not.

Solution 110. One way to approach the proof is to show that for each invertible operator A there is a continuous curve that connects it to the identity. For this purpose, consider the polar decomposition UP of A . Since A is invertible, so also are U and P . It follows that U is unitary and P is strictly positive. Join U to 1 by a continuous curve $t \rightarrow U_t$ of unitary operators (cf. Problem 102), and, similarly, join P to 1 by a continuous curve $t \rightarrow P_t$ of strictly positive operators. (The latter does not even need the spectral theorem; consider $tP + (1 - t)$, $0 \leq t \leq 1$.) If $A_t = U_t P_t$, then $t \rightarrow A_t$ is a continuous curve of invertible operators that joins A ($= A_1$) to 1 ($= A_0$).

Chapter 14. Unilateral shift

Solution 111. If \mathbf{H} is not separable, then it is the direct sum of separable infinite-dimensional subspaces that reduce A , and, consequently, there is no loss of generality in assuming that \mathbf{H} is separable in the first place. In a separable Hilbert space all infinite-dimensional subspaces have the same dimension; the assertion, therefore, is just that \mathbf{H} is the direct sum of \aleph_0 infinite-dimensional subspaces that reduce A . It is sufficient to prove the assertion for 2 in place of \aleph_0 . In other words, it is sufficient to prove that *for each normal operator on a separable infinite-dimensional Hilbert space there exists a reducing subspace such that both it and its orthogonal complement are infinite-dimensional*. Indeed, if this is true, then there exists a reducing subspace \mathbf{H}_1 of \mathbf{H} such that both \mathbf{H}_1 and \mathbf{H}_1^\perp are infinite-dimensional. Another application of the same result (consider the restriction of A to \mathbf{H}_1^\perp) implies that there exists a reducing subspace \mathbf{H}_2 of \mathbf{H}_1^\perp such that both \mathbf{H}_2 and $\mathbf{H}_1^\perp \cap \mathbf{H}_2^\perp$ are infinite-dimensional. Proceed inductively to obtain an infinite sequence $\{\mathbf{H}_n\}$ of pairwise orthogonal infinite-dimensional reducing subspaces. If the intersection $\bigcap_{n=1}^{\infty} \mathbf{H}_n^\perp$ is not zero, adjoin it to, say, \mathbf{H}_1 .

It remains to prove the assertion italicized above. The spectral theorem shows that there is no loss of generality in restricting attention to a multiplication operator A induced by a bounded measurable function φ on some measure space. For each Borel subset M of the complex plane, let $E(M)$ be the multiplication operator induced by the characteristic function of $\varphi^{-1}(M)$; the operator $E(M)$ is the projection onto the subspace of functions that vanish outside $\varphi^{-1}(M)$. Clearly each $E(M)$ commutes with A , i.e., the range of each $E(M)$ reduces A . If, for some M , both $E(M)$ and $1 - E(M)$ have infinite-dimensional ranges, the desired assertion is true.

In the contrary case what must happen is that for each M either $E(M)$ or $1 - E(M)$ has finite rank. Draw a sequence of finer and finer square grids on the plane, and let each square in each grid play the role of M ; it follows that if $E(M)$ has positive rank, then M contains at least one point λ such that $E(\{\lambda\})$ has positive rank. There cannot be more than finitely many λ 's like that, for then they could be separated

into two infinite subsets, and that would contradict the main assumption of this paragraph. Conclusion: there exists at least one point λ such that the dimension of the range of $E(\{\lambda\})$ is infinite; let \mathbf{M} be that range. The restriction of A to \mathbf{M} is a scalar and is therefore reduced by every subspace of \mathbf{M} . Split \mathbf{M} into two infinite-dimensional subspaces \mathbf{M}_0 and \mathbf{M}_1 ; if $\mathbf{H}_0 = \mathbf{M}_0$ and $\mathbf{H}_1 = \mathbf{M}_1 \vee \mathbf{M}^\perp$, then \mathbf{H}_0 and \mathbf{H}_1 do everything that is required.

Solution 112. *Every unitary operator on an infinite-dimensional Hilbert space is the product of four symmetries; three is not always enough.*

If the underlying Hilbert space \mathbf{H} is finite-dimensional, then the concept of determinant makes sense. Since the determinant of a symmetry is ± 1 , it follows that no unitary operator with a non-real determinant can be the product of symmetries.

Proof. Suppose that \mathbf{H}_1 is an infinite-dimensional Hilbert space, and begin by representing \mathbf{H} as the direct sum of a sequence $\{\mathbf{H}_n\}$ of equi-dimensional subspaces each of which reduces the given unitary operator U (Problem 111). It is convenient to let the index n run through all integers, positive, negative, and zero.

Relative to the fixed direct sum decomposition $\mathbf{H} = \sum_n \mathbf{H}_n$, define a *right shift* as a unitary operator S such that $S\mathbf{H}_n = \mathbf{H}_{n+1}$, and define a *left shift* as a unitary operator T such that $T\mathbf{H}_n = \mathbf{H}_{n-1}$, $n = 0, \pm 1, \pm 2, \dots$. The equi-dimensionality of all the \mathbf{H}_n 's guarantees the existence of shifts. If S is an arbitrary right shift, write $T = S^*U$. Since $T\mathbf{H}_n = S^*U\mathbf{H}_n = S^*\mathbf{H}_n = \mathbf{H}_{n-1}$ for all n , it follows that T is a left shift. Since $U = ST$, it follows that every unitary operator is the product of two shifts; the proof will be completed by showing that every shift is the product of two symmetries.

Since the inverse (equivalently, the adjoint) of a left shift is a right shift, it is sufficient to consider right shifts. Suppose then that S is a right shift; let P be the operator that is equal to S^{1-2n} on \mathbf{H}_n and let Q be the operator that is equal to S^{-2n} on \mathbf{H}_n ($n = 0, \pm 1, \pm 2, \dots$). If $f \in \mathbf{H}_n$, then $Qf = S^{-2nf} \in S^{-2n}\mathbf{H}_n = \mathbf{H}_{-n}$, so that $PQf = PS^{-2nf} = S^{1-2(-n)}S^{-2nf} = Sf$. The existence proof is complete.

To prove that on every Hilbert space there exists a unitary operator that is not the product of three symmetries, let ω be a non-real cube root of unity, and let U be $\omega 1$. The operator U belongs to the center of the group of all unitary operators; the order of U in that group is exactly three. The remainder of the proof has nothing to do with operator theory; the point is that in no group can a central element of order 3 be the product of three elements of order 2. Suppose indeed that u is central and that $u = xyz$, where $x^2 = y^2 = z^2 = 1$; it follows that

$$\begin{aligned} u^4 &= uxuyuz = u(xu)y(uz) = u(yz)y(xy) \\ &= y(uz)y(xy) = yxy \cdot yxy = 1. \end{aligned}$$

Reference: Halmos-Kakutani [1958].

Solution 113. (a) *The unilateral shift is not the product of a finite number of normal operators.* (b) *The norm of both the real and the imaginary part of the unilateral shift is 1.* (c) *The distance from the unilateral shift to the set of normal operators is 1.*

Proof. (a) The principal tool is the observation that if a normal operator has a one-sided inverse, then it has an inverse. (Proof: for every operator, left invertibility is the same as boundedness from below, cf. (2), Solution 70; boundedness from below for a normal operator is the same as boundedness from below for its adjoint.) Suppose, indeed, that $U = A_1 \cdots A_n$, where U is the unilateral shift and A_1, \dots, A_n are normal. Since $U^* = A_n^* \cdots A_1^*$, it follows that $A_n^* \cdots A_1^* A_1 \cdots A_n = 1$, and hence that A_n is left invertible. In view of the preceding comments, this implies that A_n is invertible, and therefore so is A_n^* . Invertible operators can be peeled off either end of a product without altering its invertibility character. It follows by an obvious inductive repetition of the argument that each of the A 's is invertible, and so therefore is U . This is a contradiction, and the proof is complete.

(b) If U is the unilateral shift, and if $A = \frac{1}{2}(U + U^*)$, then it is clear that $\|A\| \leq 1$. Since 1 is an approximate eigenvalue of U , there exists a sequence $\{f_n\}$ of unit vectors such that $Uf_n - f_n \rightarrow 0$. Apply U^* and change sign to get $U^*f_n - f_n \rightarrow 0$. Add and divide by 2 to get

$Af_n - f_n \rightarrow 0$. Conclusion: 1 is an approximate eigenvalue of A , and therefore $\|A\| \geq 1$. To get the result for the imaginary part, note that if $U = A + iB$, then $-iU = B - iA$, and $-iU$ is unitarily equivalent to U (cf. Problem 75).

(c) It is trivial that there is a normal operator (namely 0) within 1 of U ; the less trivial part of the assertion is that if A is normal, then $\|U - A\| \geq 1$. If A is invertible, this follows from Solution 109; the assertion there implies that the open ball with center U and radius 1 contains no invertible operators. The general case is now immediate: the set of invertible normal operators is dense in the set of all normal operators.

Solution 114. *The unilateral shift has no square root.*

Proof. It turns out that U^* is easier to treat than U , and, of course, it comes to the same thing. Suppose therefore that $V^2 = U^*$, and let \mathbf{N}_0 be the (one-dimensional) kernel of U^* . Since $\ker V \subset \ker V^2 = \mathbf{N}_0$, it follows that $\dim \ker V \leq 1$. If the kernel of V were trivial (zero-dimensional), then the same would be true of U^* ; it follows that $\dim \ker V = 1$, and hence that $\ker V = \mathbf{N}_0$. Since U^* maps the underlying Hilbert space onto itself, the same must be true of V . It follows in particular that \mathbf{N}_0 is included in the range of V , and hence that there exists a vector f such that Vf is a non-zero element of \mathbf{N}_0 . Since \mathbf{N}_0 is the kernel of V , this implies that $V^2f = 0$, i.e., that $U^*f = 0$, and hence that $f \in \mathbf{N}_0$. Do it again: since \mathbf{N}_0 is the kernel of V , this implies that $Vf = 0$, in contradiction to the way f was chosen in the first place. Conclusion: there is no such V .

Similar negative results were first obtained by Halmos-Lumer-Schäffer [1953]; the techniques used there would serve here too. The very much simpler proof given above is due to J. G. Thompson. Further interesting contributions to the square root problem were made by Deckard-Pearcy [1963] and Schäffer [1965].

Solution 115. It is obvious that every multiplication operator on L^2 commutes with W . If A is the multiplication operator induced by a bounded measurable function φ , then

$$Ae_0 = \varphi \cdot e_0 = \varphi.$$

This shows that in any attempt to prove that some operator A is a multiplication on L^2 there is no choice in the determination of the multiplier; if there is one, it must be Ae_0 .

Suppose now that $AW = WA$, and put $\varphi = Ae_0$. The first (and in fact the major) difficulty is to prove that φ is bounded; one way to do it is this. If ψ is an arbitrary bounded measurable function, and if B is the multiplication operator it induces, then, in the usual sense of the functional calculus for normal operators, $B = \psi(W)$. Since W commutes with A , every function of W commutes with A , and hence, in particular, B commutes with A ; it follows that

$$\varphi \cdot \psi = \psi \cdot \varphi = B\varphi = BAe_0 = AB e_0 = A\psi.$$

The statement that every function of W commutes with A is not trivial; it is the Fuglede commutativity theorem for normal operators. (See Halmos [1951, p. 68] and Problem 152.) It is not necessary in this argument to use all bounded measurable functions; it would be sufficient to use trigonometric polynomials (i.e., finite linear combinations of the e_n 's). The Fuglede theorem would still come in; it is needed to show that if W commutes with A , then W^* commutes with A .

At this point Problem 50 is almost applicable. The hypothesis there was that A is an operator on L^2 such that $Af = \varphi \cdot f$ for all f in L^2 ; the situation here is that A is an operator on L^2 such that $A\psi = \varphi \cdot \psi$ for all bounded measurable ψ . The difference is large enough to invalidate one of the proofs that worked there, but not large enough to invalidate the second, more "natural" proof. Conclusion: φ is bounded.

The rest of the proof is trivial. Since φ is bounded, it induces a multiplication operator; since that multiplication operator agrees with A on the dense set of all bounded functions, it agrees with A everywhere.

To prove the corollary, note that if a multiplication is a projection, then the multiplier is a characteristic function.

Solution 116. As in Solution 115, it is inevitable to put $\varphi = Ae_0$ and to try to prove that φ is the desired multiplier. Since, for each n , multiplication by e_n leaves H^2 invariant ($n = 0, 1, 2, \dots$), it follows that $\varphi \cdot e_n \in H^2$. Since, moreover,

$$\varphi \cdot e_n = e_n \cdot \varphi = U^n \varphi = U^n Ae_0 = AU^n e_0 = Ae_n,$$

it follows that, for each polynomial p , the product $\varphi \cdot p$ belongs to \mathbf{H}^2 and $\varphi \cdot p = Ap$. If φ were known to be bounded, the proof would be over (the multiplication operator induced by φ agrees with A on a dense set), and, if it were known that $\varphi \cdot f = Af$ for all f in \mathbf{H}^2 , then φ would be bounded (cf. the last comment in Solution 50). Since at the moment neither of these ifs is known, there is nothing for it but to prove something. The least troublesome way seems to be to adapt (or, to put it bluntly, to repeat) the second proof used in Solution 51.

If $f \in \mathbf{H}^2$, then there exist polynomials p_n such that $p_n \rightarrow f$ in \mathbf{H}^2 ; it follows, of course, that $Ap_n \rightarrow Af$ in \mathbf{H}^2 . There is no loss of generality in assuming that $p_n \rightarrow f$ almost everywhere and $Ap_n \rightarrow Af$ almost everywhere; if this is not true for the sequence $\{p_n\}$, it is true for a suitable subsequence. Since $p_n \rightarrow f$ almost everywhere, it follows that $\varphi \cdot p_n \rightarrow \varphi \cdot f$ almost everywhere; since, at the same time, $\varphi \cdot p_n \rightarrow Af$ almost everywhere, it follows that $\varphi \cdot f = Af$ almost everywhere.

There are two ideas in this twice used proof: (1) if a closed transformation agrees with a bounded one on a dense set, then it is bounded, and (2) multiplications are always closed.

The corollary is equivalent to this: if E is a projection that commutes with U , then $E = 0$ or $E = 1$. The result proved above implies that E is the restriction to \mathbf{H}^2 of a multiplication, where the multiplier itself is in \mathbf{H}^∞ . Since an idempotent multiplication on \mathbf{H}^2 must be induced by an idempotent multiplier (apply to e_0), the multiplier must be the characteristic function of a set, and hence, in particular, real; the desired conclusion follows from Problem 26.

The corollary, incidentally, does not have to be deduced from the main assertion; for an easy direct proof see Halmos [1951, p. 41].

Solution 117. Let U be the unilateral shift, represented as the restriction to \mathbf{H}^2 of the multiplication induced by e_1 ; see, for instance, Problem 116. If A commutes with U , then (by Problem 116) there exists a function φ in \mathbf{H}^∞ such that $Af = \varphi \cdot f$ for all f in \mathbf{H}^2 . The crucial tool is that φ is the limit almost everywhere of a sequence $\{p_n\}$ of polynomials such that $\|p_n\|_\infty \leq \|\varphi\|_\infty$ for every n ; cf. Solution 33. It follows that if $f \in \mathbf{H}^2$, then

$$|p_n(z)f(z)| \leq \|\varphi\|_\infty \cdot |f(z)|$$

almost everywhere. Since $p_n \cdot f \rightarrow \varphi \cdot f$ almost everywhere, the Lebesgue dominated convergence theorem applies; the conclusion is that $p_n \cdot f \rightarrow \varphi \cdot f$ in \mathbf{H}^2 . Since multiplication by p_n is a polynomial in U (namely $p_n(U)$), the proof is complete.

Solution 118. The only way an isometry V on a Hilbert space \mathbf{H} can fail to be unitary is to map \mathbf{H} onto a proper subspace of \mathbf{H} . This suggests that the extent to which $V\mathbf{H}$ differs from \mathbf{H} is a useful measure of the non-unitariness of V . One application of V compresses \mathbf{H} into $V\mathbf{H}$, another application of V compresses $V\mathbf{H}$ into $V^2\mathbf{H}$, and so on. The incompressible core of \mathbf{H} seems to be what is common to all the $V^n\mathbf{H}$'s. This is true, and it is the crux of the matter: the main thing to prove is that that incompressible core reduces the operator V . A slightly sharper result is sometimes useful; it is good to know exactly what the orthogonal complement of that core is. Write $\mathbf{N} = (V\mathbf{H})^\perp$; in terms of \mathbf{N} the main result is that

$$\bigcap_{n=0}^{\infty} V^n\mathbf{H} = \bigcap_{n=0}^{\infty} (V^n\mathbf{N})^\perp.$$

Both the statement (and the proof below) become intuitively obvious if orthogonal complements are replaced by ordinary set-theoretic complements. (A picture helps.)

Begin with the observation that $V\mathbf{M}^\perp \subset (V\mathbf{M})^\perp$ for all subspaces \mathbf{M} . (Indeed, if $f \in \mathbf{M}^\perp$, so that Vf is a typical element of $V\mathbf{M}^\perp$, and if $g \in \mathbf{M}$, so that Vg is a typical element of $V\mathbf{M}$, then $Vf \perp Vg$ follows, since V is an isometry, from $f \perp g$.) This implies that

$$V^{n+1}\mathbf{H} = V^n(V\mathbf{H}) = V^n\mathbf{N}^\perp \subset (V^n\mathbf{N})^\perp,$$

and that settles half the proof. For the reverse inclusion, assume that $f \in \bigcap_{n=0}^{\infty} (V^n\mathbf{N})^\perp$ and prove by induction that $f \in V^n\mathbf{H}$ for all n . If $n = 0$, this is trivial. If $f \in V^n\mathbf{H}$, so that $f = V^n g$ for some g , then $V^n g \perp V^n\mathbf{N}$ (since $f \in (V^n\mathbf{N})^\perp$), and therefore $g \perp \mathbf{N}$. This implies that $g \in V\mathbf{H}$, and

hence that $f \in V^{n+1}\mathbf{H}$, as desired. The proof of the asserted equation is complete.

The rest is easy. Obviously $\bigcap_{n=0}^{\infty} V^n\mathbf{H}$ is invariant under V ; since, by the result just proved, its orthogonal complement is equal to $\bigvee_{n=0}^{\infty} V^n\mathbf{N}$, which is also invariant under V , it follows that $\bigcap_{n=0}^{\infty} V^n\mathbf{H}$ reduces V . The restriction of V to this reducing subspace is unitary (because it is an isometry whose range is equal to its domain). The restriction of V to the orthogonal complement $\bigvee_{n=0}^{\infty} V^n\mathbf{N}$ is a direct sum of copies of the unilateral shift; the number of copies is $\dim \mathbf{N}$.

Solution 119. *If U is the unilateral shift, then $\|U - V\| = 2$ for each unitary operator V .*

Proof. The proof begins with the observation that if -1 belongs to the spectrum of an operator A , then -2 belongs to the spectrum of $A - 1$. It follows that if A is a non-normal (i.e., non-unitary) isometry, then $r(A - 1) \geq 2$, and hence $\|A - 1\| \geq 2$. (Use Problem 118, and recall that the spectrum of the unilateral shift is the closed unit disc.) If V is unitary, then $\|U - V\| = \|V^*U - 1\|$. Since V^*U is a non-normal isometry, it follows that $\|U - V\| \geq 2$; the reverse inequality is trivial.

This is a geometrically very peculiar result. The unilateral shift is on the unit sphere of the space of operators, and so also is each unitary operator. What was just proved can be expressed in geometric language by saying that if V is unitary, then U and V are diametrically opposite; they are as far from each other as if they were at the opposite ends of a diameter. What is peculiar is that this is true for every V .

Solution 120. *There exist commutative isometries U_0 and V_0 such that the domain of the unitary component of U_0 does not reduce V_0 .*

Proof. Let U_0 be the direct sum of the unilateral shift and infinitely many copies of the bilateral shift; let V_0 be an isometry that transplants the unilateral component into the first bilateral one and shifts the bilateral components forward. To make this description computationally explicit, let U be the unilateral shift, write $E = 1 - UU^*$, and define

U_0 and V_0 as the infinite operator matrices

$$U_0 = \begin{pmatrix} U & 0 & 0 & 0 & 0 \\ 0 & U^* & 0 & 0 & 0 \\ 0 & E & U & 0 & 0 \\ 0 & 0 & 0 & U^* & 0 \\ 0 & 0 & 0 & E & U \\ & & & & \ddots \\ & & & & & \ddots \\ & & & & & & \ddots \end{pmatrix}, \quad V_0 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ & & & & \ddots \\ & & & & & \ddots \\ & & & & & & \ddots \end{pmatrix}.$$

(Note: the 0's and 1's in V_0 are operators, the same size as the entries in U_0 .) Computation proves that $U_0 V_0 = V_0 U_0$. The domain of the unitary part of U_0 is the tail (all vectors with vanishing first coordinate); clearly the orthogonal complement of the tail (all vectors that vanish after the first coordinate) is not invariant under V_0 .

The 3×3 northwest corners of U_0 and V_0 serve the same purpose, but in that case V_0 is not an isometry, but only a partial isometry.

Solution 121. Given: a Hilbert space \mathbf{H} and on it a contraction A such that $A^n \rightarrow 0$ strongly. To construct: a Hilbert space $\tilde{\mathbf{H}}$ and on it a shift U with the stated unitary equivalence property. The construction is partially motivated by the following observation: if a vector f in \mathbf{H} is replaced by Af , then the sequence

$$\langle f, Af, A^2f, \dots \rangle$$

is shifted back by one step, i.e., it is replaced by

$$\langle Af, A^2f, A^3f, \dots \rangle.$$

What this suggests is that $\tilde{\mathbf{H}}$ be something like the direct sum

$$\mathbf{H} \oplus \mathbf{H} \oplus \mathbf{H} \oplus \dots$$

That does not work. There is no reason why the sequence $\langle f, Af, A^2f, \dots \rangle$ should belong to the direct sum (the series $\sum_{n=0}^{\infty} \|A^n f\|^2$ need not converge), and, even if it does, the correspondence between f and $\langle f, Af, A^2f, \dots \rangle$ may fail to be norm-preserving (even if $\sum_{n=0}^{\infty} \|A^n f\|^2$ converges, its sum will be equal to $\|f\|^2$ only in case $Af = 0$).

The inspiration that removes these difficulties is to transform each term of the sequence $\langle f, Af, A^2f, \dots \rangle$ by an operator T so that the resulting series of square norms converges to $\|f\|^2$ the easy way, by telescoping. That is: replace $\langle f, Af, A^2f, \dots \rangle$ by $\langle Tf, TAf, TA^2f, \dots \rangle$, so that

$$\|Tf\|^2 = \|f\|^2 - \|Af\|^2,$$

$$\|TAf\|^2 = \|Af\|^2 - \|A^2f\|^2,$$

$$\|TA^2f\|^2 = \|A^2f\|^2 - \|A^3f\|^2,$$

etc.

The first of these equations alone, if required to hold for all f , implies that $T^*T = 1 - A^*A$, and, conversely, if $T^*T = 1 - A^*A$, then all the equations hold.

The preceding paragraphs were intended as motivation. For the proof itself, proceed as follows. Since A is a contraction, $1 - A^*A$ is positive; write $T = \sqrt{1 - A^*A}$, and let \mathbf{R} be the closure of the range of T . Let $\tilde{\mathbf{H}}$ be the direct sum $\mathbf{R} \oplus \mathbf{R} \oplus \mathbf{R} \oplus \dots$. If $f \in \mathbf{H}$, then $TA^n f \in \mathbf{R}$ for all n , and

$$\begin{aligned} \sum_{n=0}^k \|TA^n f\|^2 &= \sum_{n=0}^k ((1 - A^*A)A^n f, A^n f) \\ &= \sum_{n=0}^k (\|A^n f\|^2 - \|A^{n+1} f\|^2) \\ &= \|f\|^2 - \|A^{k+1} f\|^2. \end{aligned}$$

Since $\|A^{k+1} f\| \rightarrow 0$ by assumption, it follows that if $f \in \mathbf{H}$, and if the mapping V is defined by

$$Vf = \langle Tf, TAf, TA^2f, \dots \rangle,$$

then V is an isometric embedding of \mathbf{H} into $\tilde{\mathbf{H}}$. If U is the obvious shift on $\tilde{\mathbf{H}}$ ($U\langle f_0, f_1, f_2, \dots \rangle = \langle 0, f_0, f_1, \dots \rangle$), then, clearly, $VAf = U^*Vf$ for all f . Since the V image of \mathbf{H} in $\tilde{\mathbf{H}}$ is invariant under U^* , the proof is complete.

Note that the multiplicity of the shift that the proof gives is equal to the rank of $1 - A^*A$, where "rank" is interpreted to mean the dimension of the closure of the range.

Solution 122. Suppose that A is an operator on a Hilbert space \mathbf{H} such that $r (= r(A)) < 1$. Since $r = \lim_n \|A^n\|^{1/n}$, it follows that the power series $\sum_{n=0}^{\infty} \|A^n\| z^n$ converges in a disc with center 0 and radius ($= 1/r$) greater than 1. This implies that $\sum_{n=0}^{\infty} \|A^n\| < \infty$, and hence, all the more, that $\sum_{n=0}^{\infty} \|A^n\|^2 < \infty$. Let \mathbf{H}_0 be the Hilbert space obtained from \mathbf{H} by redefining the inner product; the new inner product is given by

$$(f, g)_0 = \sum_{n=0}^{\infty} (A^n f, A^n g).$$

Since $|(A^n f, A^n g)| \leq \|A^n f\| \cdot \|A^n g\| \leq \|A^n\|^2 \cdot \|f\| \cdot \|g\|$, there is no difficulty about convergence. If $\|f\|_0^2 = (f, f)_0$, then

$$\|f\|^2 \leq \|f\|_0^2 \leq \left(\sum_{n=0}^{\infty} \|A^n\|^2 \right) \cdot \|f\|^2,$$

and that implies that the identity mapping I from \mathbf{H} to \mathbf{H}_0 is an invertible bounded linear transformation. (This, incidentally, is what guarantees that \mathbf{H}_0 is complete.) If $A_0 = IAI^{-1}$, then A_0 is an operator on \mathbf{H}_0 , similar to A . If $f \neq 0$, then

$$\begin{aligned} \frac{\|A_0 f\|_0^2}{\|f\|_0^2} &= \frac{\sum_{n=1}^{\infty} \|A^n f\|^2}{\|f\|^2 + \sum_{n=1}^{\infty} \|A^n f\|^2} = \frac{\sum_{n=1}^{\infty} (\|A^n f\| / \|f\|)^2}{1 + \sum_{n=1}^{\infty} (\|A^n f\| / \|f\|)^2} \\ &\leq \frac{\sum_{n=1}^{\infty} \|A^n\|^2}{1 + \sum_{n=1}^{\infty} \|A^n\|^2} < 1, \end{aligned}$$

so that A_0 is a proper contraction. This implies that the powers of A_0 tend to zero not only strongly, but in the norm.

Corollary 1 is immediate from Problem 121, and Corollary 2 is implied by the proof (given above) that $\|A_0\| < 1$. For Corollary 3: if A is quasinilpotent, then

$$r\left(\frac{1}{\varepsilon}A\right) < 1$$

for every positive number ε ; if $(1/\varepsilon)A$ is similar to the contraction C , then A is similar to εC .

Corollary 4 requires a little more argument. Clearly $r(A) = r(S^{-1}AS) \leq \|S^{-1}AS\|$ and therefore $r(A) \leq \inf_S \|S^{-1}AS\|$. To prove the reverse inequality, let t be a number in the open unit interval and write

$$B = \frac{t}{r(A)}A.$$

(If $r(A) = 0$, apply Corollary 3 instead.) Corollary 2 implies that $\|S^{-1}BS\| < 1$ for some S , so that $t \cdot \|S^{-1}AS\| < r(A)$. Infer that $t \cdot \inf_S \|S^{-1}AS\| \leq r(A)$, and then let t tend to 1.

Solution 123. The restriction of U to \mathbf{M} is an isometry. If $\mathbf{N} = \mathbf{M} \cap (U\mathbf{M})^\perp$, then \mathbf{N} is the orthogonal complement of the range of that restriction. Apply the result obtained in Solution 118 to that restriction to obtain

$$\bigcap_{n=0}^{\infty} U^n \mathbf{M} = \mathbf{M} \cap \bigcap_{n=0}^{\infty} (U^n \mathbf{N})^\perp.$$

Since $\bigcap_{n=0}^{\infty} U^n \mathbf{H}^2 = \{0\}$, it follows that

$$\mathbf{M}^\perp \vee \bigvee_{n=0}^{\infty} U^n \mathbf{N} = \mathbf{H}^2.$$

Since, on the other hand, $U^n \mathbf{N} \subset U^n \mathbf{M} \subset \mathbf{M}$, it follows that

$$\bigvee_{n=0}^{\infty} U^n \mathbf{N} \subset \mathbf{M}.$$

The span of \mathbf{M}^1 and a proper subspace of \mathbf{M} can never be the whole space. Conclusion:

$$\bigvee_{n=0}^{\infty} U^n \mathbf{N} = \mathbf{M}.$$

It remains to prove that $\dim \mathbf{N} = 1$. For this purpose it is convenient to regard the unilateral shift U as the restriction to \mathbf{H}^2 of the bilateral shift W on the larger space \mathbf{L}^2 . If f and g are orthogonal unit vectors in \mathbf{N} , then the set of all vectors of either of the forms $W^n f$ or $W^m g$ ($n, m = 0, \pm 1, \pm 2, \dots$) is an orthonormal set in \mathbf{L}^2 . (This assertion leans on the good behavior of wandering subspaces for unitary operators.) It follows that

$$\begin{aligned} 2 &= \|f\|^2 + \|g\|^2 = \sum_n |(f, e_n)|^2 + \sum_m |(g, e_m)|^2 \\ &= \sum_n |(f, W^n e_0)|^2 + \sum_m |(g, W^m e_0)|^2 \\ &= \sum_n |(W^{*n} f, e_0)|^2 + \sum_m |(W^{*m} g, e_0)|^2 \\ &\leq \|e_0\|^2 = 1. \end{aligned}$$

(The inequality is Bessel's.) This absurdity shows that f and g cannot co-exist. The dimension of \mathbf{N} cannot be as great as 2; since it cannot be 0 either, the proof is complete.

The last part of the proof is due to I. Halperin; see Nagy-Foiaş [1962, p. 108]. It is geometric; the original proof in Halmos [1961] was analytic. See also Robertson [1965].

Solution 124. Since $\mathbf{M}_k^1(\lambda)$ is spanned by $f_\lambda, \dots, U^{k-1}f_\lambda$, it is clear that $\dim \mathbf{M}_k^1(\lambda) \leq k$. To prove equality, note first that $\mathbf{M}_k^1(\lambda)$ is invariant under U^* . (Indeed $U^*f_\lambda = \lambda f_\lambda$ and, if $j \geq 1$, $U^*U^j f_\lambda = U^*UU^{j-1}f_\lambda = U^{j-1}f_\lambda$. Note that this proves the invariance of $\mathbf{M}_k(\lambda)$ under U .) If $\dim \mathbf{M}_k^1(\lambda) < k$, then $\sum_{i=0}^{k-1} \alpha_i U^i f_\lambda = 0$ for suitable scalars α_i , or, in other words, there exists a polynomial p of degree less than k

such that $p(U)f_\lambda = 0$. This implies that $U^n f_\lambda$ is a linear combination of $f_\lambda, \dots, U^{k-1}f_\lambda$ for all n , and hence that $\mathbf{M}_k^\perp(\lambda)$ is invariant under U also. This is impossible, and therefore $\dim \mathbf{M}_k^\perp(\lambda) = k$.

Since $f_\lambda - \lambda U f_\lambda = \sum_{n=0}^{\infty} \lambda^n e_n - \lambda \sum_{n=1}^{\infty} \lambda^{n-1} e_n = e_0$, it follows that $e_0 \in \mathbf{M}_k^\perp(\lambda)$ as soon as $k > 1$. This implies that $U^j f_\lambda - \lambda U^{j+1} f_\lambda = U^j e_0 = e_j \in \mathbf{M}_k^\perp(\lambda)$ as soon as $k > j + 1$, and, consequently, $\bigvee_{k=1}^{\infty} \mathbf{M}_k^\perp(\lambda)$ contains all e_j 's.

Solution 125. If $\mathbf{M} = \varphi \cdot \mathbf{H}^2$, then $U\mathbf{M} = e_1 \cdot \mathbf{M} = e_1 \cdot \varphi \cdot \mathbf{H}^2 = \varphi \cdot e_1 \cdot \mathbf{H}^2 \subset \varphi \cdot \mathbf{H}^2 = \mathbf{M}$; this proves the "if". For another proof of the same implication, use the theory of wandering subspaces. If \mathbf{N} is the (one-dimensional) subspace spanned by φ , then \mathbf{N} is wandering; the reason is that $(U^n \varphi, U^m \varphi) = \int e_n e_m^* d\mu = \delta_{nm}$. To prove "only if", suppose that \mathbf{M} is invariant under U and use Problem 123 to represent \mathbf{M} in the form $\bigvee_{n=0}^{\infty} U^n \mathbf{N}$, where \mathbf{N} is a wandering subspace for U . Take a unit vector φ in \mathbf{N} . Since, by assumption, $(U^n \varphi, \varphi) = 0$ when $n > 0$, or $\int e_n |\varphi|^2 d\mu = 0$ when $n > 0$, it follows (by the formation of complex conjugates) that $\int e_n |\varphi|^2 d\mu = 0$ when $n < 0$, and hence that $|\varphi|^2$ is a function in \mathbf{L}^1 such that all its Fourier coefficients with non-zero index vanish. Conclusion: $|\varphi|$ is constant almost everywhere, and, since $\int |\varphi|^2 d\mu = 1$, the constant modulus of φ must be 1. (Note that the preceding argument contains a proof, different from the one used in Solution 123, that every non-zero wandering subspace of U is one-dimensional.) Since φ , by itself, spans \mathbf{N} , the functions $\varphi \cdot e_n$ ($n = 0, 1, 2, \dots$) span \mathbf{M} . Equivalently, the set of all functions of the form $\varphi \cdot p$, where p is a polynomial, spans \mathbf{M} . Since multiplication by φ (restricted to \mathbf{H}^2) is an isometry, its range is closed; since \mathbf{M} is the span of the image under that isometry of a dense set, it follows that \mathbf{M} is in fact equal to the range of that isometry, and hence that $\mathbf{M} = \varphi \cdot \mathbf{H}^2$.

To prove the first statement of Corollary 1, observe that if $\varphi \cdot \mathbf{H}^2 \subset \psi \cdot \mathbf{H}^2$, then $\varphi = \varphi \cdot e_0 = \psi \cdot f$ for some f in \mathbf{H}^2 ; since $f = \varphi \cdot \psi^*$, it follows that $|f| = 1$, so that f is an inner function. To prove the second statement, it is sufficient to prove that if both θ and θ^* are inner functions, then θ is a constant. To prove that, observe that both $\operatorname{Re} \theta$ and $\operatorname{Im} \theta$ are real functions in \mathbf{H}^2 , and therefore (Problem 26) both $\operatorname{Re} \theta$ and $\operatorname{Im} \theta$ are constants. As for Corollary 2: if $\mathbf{M} = \varphi \cdot \mathbf{H}^2$ and $\mathbf{N} = \psi \cdot \mathbf{H}^2$, then $\varphi \cdot \psi \in \mathbf{M} \cap \mathbf{N}$.

Solution 126. Consider first the simple unilateral shift U . Let $\langle \xi_0, \xi_1, \xi_2, \dots \rangle$ be a sequence of complex numbers such that

$$\lim_k \frac{1}{|\xi_k|^2} \sum_{n=1}^{\infty} |\xi_{n+k}|^2 = 0.$$

(Concrete example: $\xi_n = 1/n!$.) Assertion: $f = \langle \xi_0, \xi_1, \xi_2, \dots \rangle$ is a cyclic vector for U^* . For the proof, observe first that $U^{*k}f = \langle \xi_k, \xi_{k+1}, \xi_{k+2}, \dots \rangle$, and hence that

$$\begin{aligned} \left\| \frac{1}{\xi_k} U^{*k}f - e_0 \right\|^2 &= \left\| \left\langle 1, \frac{\xi_{k+1}}{\xi_k}, \frac{\xi_{k+2}}{\xi_k}, \dots \right\rangle - \langle 1, 0, 0, \dots \rangle \right\|^2 \\ &= \left\| \left\langle 0, \frac{\xi_{k+1}}{\xi_k}, \frac{\xi_{k+2}}{\xi_k}, \dots \right\rangle \right\|^2 \\ &= \frac{1}{|\xi_k|^2} \sum_{n=1}^{\infty} |\xi_{n+k}|^2 \rightarrow 0. \end{aligned}$$

Consequence: e_0 belongs to the span of $f, U^*f, U^{*2}f, \dots$. This implies that

$$U^{*k-1}f - \xi_{k-1}e_0 = \langle 0, \xi_k, \xi_{k+1}, \xi_{k+2}, \dots \rangle$$

belongs to that span ($k = 1, 2, 3, \dots$). Since

$$\left\| \frac{1}{\xi_k} (U^{*k-1}f - \xi_{k-1}e_0) - e_1 \right\|^2 = \frac{1}{|\xi_k|^2} \sum_{n=1}^{\infty} |\xi_{n+k}|^2 \rightarrow 0,$$

it follows that e_1 belongs to the span of $f, U^*f, U^{*2}f, \dots$. An obvious inductive repetition of this twice-used argument proves that e_n belongs to the span of $f, U^*f, U^{*2}f, \dots$ for all n ($= 0, 1, 2, \dots$), and hence that f is cyclic.

Once this is settled, the cases of higher multiplicity turn out to be trivial. For multiplicity 2 consider the same sequence $\{\xi_n\}$ and form the vector

$$\langle \langle \xi_0, \xi_2, \xi_4, \dots \rangle, \langle \xi_1, \xi_3, \xi_5, \dots \rangle \rangle.$$

Proof. Let g be any non-zero element of L^2 that vanishes on a set of positive measure, and write $f(z) = zg(z^3)$. Clearly f is a non-zero element of L^2 that vanishes on a set of positive measure. If $g = \sum_n \beta_n e_n$, then $f = \sum_n \beta_n e_{3n+1}$ (this needs justification), so that $(f, e_n) = 0$ unless $n \equiv 1 \pmod{3}$. If $n \equiv 1 \pmod{3}$, then $-n \equiv 2 \pmod{3}$, and therefore $(f, e_n)(f, e_{-n}) = 0$ for all n .

Solution 129. Suppose first that $\{\alpha_n\}$ is periodic of period p ($= 1, 2, 3, \dots$), and let \mathbf{M}_j ($j = 0, \dots, p-1$) be the span of all those basis vectors e_n for which $n \equiv j \pmod{p}$. Each vector f has a unique representation in the form $f_0 + \dots + f_{p-1}$ with f_j in \mathbf{M}_j . Consider the functional representation of the two-sided shift, and, using it, make the following definition. For each measurable subset E of the circle, let \mathbf{M} ($= \mathbf{M}_E$) be the set of all those f 's for which $f_j(z) = 0$ whenever $j = 0, \dots, p-1$ and $z \notin E$. If $f = \sum_{j=0}^{p-1} f_j$ (with f_j in \mathbf{M}_j), then

$$Af = \sum_{j=0}^{p-1} \alpha_j W f_j$$

and

$$A^*f = \sum_{j=0}^{p-1} \alpha_{j-1} W^* f_j;$$

this proves that \mathbf{M} reduces A . (Note that $W\mathbf{M}_j = \mathbf{M}_{j+1}$ and $W^*\mathbf{M}_j = \mathbf{M}_{j-1}$, where addition and subtraction are interpreted modulo p .)

To show that this construction does not always yield a trivial reducing subspace, let E_0 be a measurable set, with measure strictly between 0 and $1/p$, and let E be its inverse image under the mapping $z \rightarrow z^p$. The set E is a measurable set, with measure strictly between 0 and 1. If g is a function that vanishes on the complement of E_0 , and if $f_0(z) = g(z^p)$, then f_0 vanishes on the complement of E . If, moreover, $f_j(z) = z^j f_0(z)$, $j = 0, \dots, p-1$, then the same is true of each f_j . Clearly $f_j \in \mathbf{M}_j$, and $f_0 + \dots + f_{p-1}$ is a typical non-trivial example of a vector in \mathbf{M} . This completes the proof of the sufficiency of the condition.

Necessity is the surprising part. To prove it, suppose first that B is

an operator, with matrix $\{\beta_{ij}\}$, that commutes with A . Observe that

$$\begin{aligned}\beta_{i+1,j+1} &= (Be_{j+1}, e_{i+1}) = \left(B \frac{1}{\alpha_j} A e_j, e_{i+1} \right) \\ &= \frac{1}{\alpha_j} (Be_j, A^* e_{i+1}) = \frac{\alpha_i}{\alpha_j} \beta_{ij}.\end{aligned}$$

Consequence 1: the main diagonal of $\{\beta_{ij}\}$ is constant (put $i = j$).

Consequence 2: if $\beta_{ij} = 0$ for some i and j , then $\beta_{i+k,j+k} = 0$ for all k .

If B happens to be Hermitian, then it commutes with A^* also, and hence with A^*A . Since $A^*Ae_n = \alpha_n^2 e_n$, it follows that

$$\begin{aligned}\beta_{ij} &= (Be_j, e_i) = \left(B \frac{1}{\alpha_j^2} A^* A e_j, e_i \right) \\ &= \frac{1}{\alpha_j^2} (Be_j, A^* A e_i) = \frac{\alpha_i^2}{\alpha_j^2} \beta_{ij}.\end{aligned}$$

Consequence 3: if $\alpha_i \neq \alpha_j$, then $\beta_{ij} = 0$.

Assume now that the sequence $\{\alpha_n\}$ is not periodic; it is sufficient to prove that every Hermitian B that commutes with A is a scalar. The assumption implies that if m and n are distinct positive integers, then there exist integers i and j such that $\alpha_i \neq \alpha_j$ and $i - j = m - n$. It follows that

$$\begin{aligned}0 &= \beta_{ij} && \text{(by Consequence 3)} \\ &= \beta_{i-j+n, j-j+n} && \text{(by Consequence 2),}\end{aligned}$$

i.e., that $\beta_{mn} = 0$ whenever $m \neq n$. This says that the matrix of B is diagonal; by Consequence 1 it follows that B is a scalar.

Chapter 15. Compact operators

Solution 130. If A is $(s \rightarrow s)$ continuous, and if $\{f_j\}$ is a net w -convergent to f , then $(Af_j, g) = (f_j, A^*g) \rightarrow (f, A^*g) = (Af, g)$ for all g , so that $Af_j \rightarrow Af$ (w). This proves that A is $(w \rightarrow w)$ continuous. Note that the assumption of $(s \rightarrow s)$ continuity was tacitly, but heavily, used via the existence of the adjoint A^* .

If A is $(w \rightarrow w)$ continuous, and if $\{f_j\}$ is a net s -convergent to f , then, a fortiori, $f_j \rightarrow f$ (w), and the assumption implies that $Af_j \rightarrow Af$ (w). This proves that A is $(s \rightarrow w)$ continuous.

To prove that if A is $(s \rightarrow w)$ continuous, then A is bounded, assume the opposite. That implies the existence of a sequence $\{f_n\}$ of unit vectors such that $\|Af_n\| \geq n$. Since

$$\frac{1}{n}f_n \rightarrow 0 \quad (s),$$

the assumption implies that

$$\frac{1}{n}Af_n \rightarrow 0 \quad (w),$$

and hence that

$$\left\{ \frac{1}{n}Af_n \right\}$$

is a bounded sequence; this is contradicted by

$$\left\| \frac{1}{n}Af_n \right\| \geq n.$$

Suppose, finally, that A is $(w \rightarrow s)$ continuous. It follows that the inverse image under A of the open unit ball is a weak open set, and hence that it includes a basic weak neighborhood of 0. In other words, there exist vectors f_1, \dots, f_k and there exists a positive number ε such

that if $|(f, f_i)| < \epsilon, i = 1, \dots, k$, then $\|Af\| < 1$. If f is in the orthogonal complement of the span of $\{f_1, \dots, f_k\}$, then certainly $|(f, f_i)| < \epsilon, i = 1, \dots, k$, and therefore $\|Af\| < 1$. Since this conclusion applies to all scalar multiples of f too, it follows that Af must be 0. This proves that A annihilates a subspace of finite co-dimension, and this is equivalent to the statement that A has finite rank. (If there is an infinite-dimensional subspace on which A is one-to-one, then the range of A is infinite-dimensional; to prove the converse, note that the range of A is equal to the image under A of the orthogonal complement of the kernel.)

To prove the corollary, use the result that an operator (i.e., a linear transformation that is continuous ($s \rightarrow s$)) is continuous ($w \rightarrow w$). Since the closed unit ball is weakly compact, it follows that its image is weakly compact, therefore weakly closed, and therefore strongly closed.

Solution 131. The proof that \mathbf{C} is an ideal is elementary. The proof that \mathbf{C} is self-adjoint is easy via the polar decomposition. Indeed, if $A \in \mathbf{C}$ and $A = UP$, then $P = U^*A$ (see Corollary 1, Problem 105), so that $P \in \mathbf{C}$; since $A^* = PU^*$, it follows that $A^* \in \mathbf{C}$.

Suppose now that $A_n \in \mathbf{C}$ and $\|A_n - A\| \rightarrow 0$; it is to be proved that $Af_j \rightarrow Af$ whenever $\{f_j\}$ is a bounded net converging weakly to f . Note that

$$\|Af_j - Af\| \leq \|Af_j - A_nf_j\| + \|A_nf_j - A_nf\| + \|A_nf - Af\|.$$

The first term on the right is dominated by $\|A - A_n\| \cdot \|f_j\|$; since $\{\|f_j\|\}$ is bounded, it follows that the first term is small for all large n , uniformly in j . The last term is dominated by $\|A_n - A\| \cdot \|f\|$, and, consequently, it too is small for large n . Fix some large n ; the compactness of A_n implies that the middle term is small for "large" j . This completes the proof that \mathbf{C} is closed.

Solution 132. Let A be an operator with diagonal $\{\alpha_n\}$, and, for each positive integer n , consider the diagonal operator A_n with diagonal $\{\alpha_0, \dots, \alpha_{n-1}, 0, 0, 0, \dots\}$. Since $A - A_n$ is a diagonal operator with diagonal $\{0, \dots, 0, \alpha_n, \alpha_{n+1}, \dots\}$, so that $\|A - A_n\| = \sup_k |\alpha_{n+k}|$, it is clear that the assumption $\alpha_n \rightarrow 0$ implies the conclusion

$\|A - A_n\| \rightarrow 0$. Since the limit (in the norm) of compact operators is compact, it follows that if $\alpha_n \rightarrow 0$, then A is compact.

To prove the converse, consider the orthonormal basis $\{e_n\}$ that makes A diagonal. If A is compact then $Ae_n \rightarrow 0$ strongly (because $e_n \rightarrow 0$ weakly; cf. Solution 13). In other words, if A is compact, then $\|\alpha_n e_n\| \rightarrow 0$, and this says exactly that $\alpha_n \rightarrow 0$.

If $Se_n = e_{n+1}$, then each of A and SA is a multiple of the other (recall that $S^*S = 1$), which implies that A and SA are simultaneously compact or not compact. This remark proves the corollary

Solution 133. With a sufficiently powerful tool (the spectral theorem) the proof becomes easy. Begin with the observation that a compact operator on an infinite-dimensional Hilbert space cannot be invertible (Proof: the image of the unit ball under an invertible operator is strongly compact if and only if the unit ball itself is strongly compact). Since the restriction of a compact operator to an invariant subspace is compact, it follows that if the restriction of a compact operator to an invariant subspace is invertible, then the subspace is finite-dimensional.

Suppose now that A is a compact normal operator; by the spectral theorem there is no loss of generality in assuming that A is a multiplication operator induced by a bounded measurable function φ on some measure space. For each positive number ϵ , let M_ϵ be the set $\{x: |\varphi(x)| > \epsilon\}$, and let \mathbf{M}_ϵ be the subspace of L^2 consisting of the functions that vanish outside M_ϵ . Clearly each \mathbf{M}_ϵ reduces A , and the restriction of A to \mathbf{M}_ϵ is bounded from below; it follows that \mathbf{M}_ϵ is finite-dimensional.

The spectrum of A is the essential range of φ . The preceding paragraph implies that, for each positive integer n , the part of the spectrum that lies outside the disc $\{\lambda: |\lambda| \leq 1/n\}$ can contain nothing but a finite number of eigenvalues each of finite multiplicity; from this everything follows.

Solution 134. Suppose that the identity operator is an integral operator, with kernel K say. This means that if $f \in L^2$ (over a measure space X , with σ -finite measure μ), then

$$\int K(x,y)f(y)d\mu(y) = f(x)$$

for almost every x , and it follows that if f and g are in L^2 , then

$$(f, g) = \iint K(x, y) f(y) g(x) d\mu(x) d\mu(y).$$

If, in particular, f and g are the characteristic functions of measurable sets, F and G say, then the equation becomes

$$\mu(F \cap G) = \iint_{F \times G} K(x, y) d\mu(x) d\mu(y).$$

In the latter form the equation is one between two set functions. The right side is the indefinite integral of K evaluated at $F \times G$. For a useful description of the left side, let T be the diagonal mapping from X to $X \times X$ ($Tx = \langle x, x \rangle$), and let ν be the measure defined for measurable subsets E of $X \times X$ by

$$\nu(E) = \mu(T^{-1}E).$$

If F and G are measurable subsets of X , then $T^{-1}(F \times G) = F \cap G$, and therefore

$$\nu(F \times G) = \mu(F \cap G).$$

Conclusion: the indefinite integral of K agrees with ν on all “rectangles” and hence on all measurable sets, and therefore, in particular, the indefinite integral of K is concentrated on the diagonal. The last assertion means that if D is the diagonal of $X \times X$, $D = \{ \langle x, y \rangle : x = y \}$, then the indefinite integral of K vanishes on every measurable subset of the complement of D (because ν does). It follows that $K(x, y) = 0$ for almost every point in the complement of D .

The preceding reasoning is valid for general (σ -finite) measures; it used no special property of Lebesgue measure. It applies, for instance, to the counting measure on a countable set, and it implies, in that case, that the matrix of the identity operator is a diagonal matrix—no surprise. Since, however, the reasoning applies to Lebesgue measure too, and since the Lebesgue measure of the diagonal in the plane is 0, it follows that if μ is Lebesgue measure, then $K = 0$ almost everywhere. In view of the expression for (f, g) in terms of K , this is absurd, and the proof is complete.

Solution 135. Recall that a simple function is a measurable function with a finite range; equivalently, a simple function is a finite linear combination of characteristic functions of measurable sets. A simple function belongs to L^2 if and only if the inverse image of the complement of the origin has finite measure; an equivalent condition is that it is a finite linear combination of characteristic functions of measurable sets of finite measure. The simple functions in $L^2(\mu)$ are dense in $L^2(\mu)$. It follows that the finite linear combinations of characteristic functions of measurable rectangles of finite measure are dense in $L^2(\mu \times \mu)$. In view of these remarks it is sufficient to prove that if A is an integral operator with kernel K , where

$$K(x, y) = \sum_{i=1}^n g_i(x) h_i(y),$$

and where each g_i and each h_i is a scalar multiple of a characteristic function of a measurable set of finite measure, then A is compact. It is just as easy to prove something much stronger: as long as each g_i and each h_i belongs to $L^2(\mu)$, the operator A has finite rank. In fact the range of A is included in the span of the g 's. The proof is immediate: if $f \in L^2(\mu)$, then

$$(Af)(x) = \sum_{i=1}^n g_i(x) \int h_i(y) f(y) d\mu(y).$$

Solution 136. *If A is a Hilbert-Schmidt operator, then the sum of the eigenvalues of A^*A is finite.*

Proof. To say that A is a Hilbert-Schmidt operator means, of course, that A is an integral operator on, say, $L^2(\mu)$, induced by a kernel K in $L^2(\mu \times \mu)$. Since A^*A is a compact normal operator, there exists an orthonormal basis $\{f_j\}$ consisting of eigenvectors of A^*A (Problem 133); write $A^*A f_j = \lambda_j f_j$. The useful way to put the preceding two statements together is to introduce a suitable basis for $L^2(\mu \times \mu)$ and, by Parseval's equality, express the $L^2(\mu \times \mu)$ norm of K (which is finite, of course) in terms of that basis. There is only one sensible looking basis in sight, the one consisting of the functions g_{ij} , where $g_{ij}(x, y) = f_i(x) f_j(y)$.

It turns out, however, that a slightly less sensible looking basis is algebraically slightly more convenient; it consists of the functions g_{ij} defined by $g_{ij}(x, y) = f_i(x)f_j(y)^*$.

The rest is simple computation:

$$\begin{aligned}
 \|K\|^2 &= \sum_i \sum_j |(K, g_{ij})|^2 \text{ (by Parseval)} \\
 &= \sum_i \sum_j \left| \iint K(x, y) f_i(x)^* f_j(y) d\mu(x) d\mu(y) \right|^2 \\
 &= \sum_j \sum_i \left| \int \left(\int K(x, y) f_j(y) d\mu(y) \right) f_i(x)^* d\mu(x) \right|^2 \\
 &= \sum_j \sum_i \left| \int (Af_j)(x) f_i(x)^* d\mu(x) \right|^2 \\
 &= \sum_j \sum_i |(Af_j, f_i)|^2 = \sum_j \|Af_j\|^2 \text{ (by Parseval)} \\
 &= \sum_j (Af_j, Af_j) = \sum_j (A^* Af_j, f_j) = \sum_j \lambda_j.
 \end{aligned}$$

The proof is over. The construction of a concrete compact operator that does not satisfy the Hilbert-Schmidt condition is now easy. Consider an infinite matrix (i.e., a “kernel” on ℓ^2). By definition, if the sum of the squares of the moduli of the entries is finite, the matrix defines a Hilbert-Schmidt operator. This is true, in particular, if the matrix is diagonal. The theorem just proved implies that in that case the finiteness condition is not only sufficient but also necessary for the result to be a Hilbert-Schmidt operator. Thus, in the diagonal case, the difference between compact and Hilbert-Schmidt is the difference between a sequence that tends to 0 and a sequence that is square-summable.

Solution 137. If A is compact and UP is its polar decomposition, then $P (= U^*A)$ is compact. By Problem 133, P is the direct sum of 0 and a diagonal operator on a separable space, and the sequence of diagonal terms of the diagonal operator tends to 0. This implies that P is the limit (in the norm) of a sequence $\{P_n\}$ of operators of finite rank,

and hence that $A = U \cdot \lim_n P_n = \lim_n UP_n$. Since UP_n has finite rank for each n , the proof is complete.

Solution 138. Suppose that \mathbf{I} is a non-zero closed ideal of operators. The first step is to show that \mathbf{I} contains every operator of rank 1. To prove this, observe that if u and v are non-zero vectors, then the operator A defined by $Af = (f, u)v$ has rank 1, and every operator of rank 1 has this form. To show that each such operator belongs to \mathbf{I} , take a non-zero operator A_0 in \mathbf{I} , and let u_0 and v_0 be non-zero vectors such that $A_0 u_0 = v_0$. Let B be the operator defined by $Bf = (f, u)u_0$, and let C be an arbitrary operator such that $Cv_0 = v$. It follows that

$$CA_0 Bf = CA_0 (f, u)u_0 = (f, u)Cv_0 = (f, u)v = Af,$$

i.e., that $CA_0 B = A$. Since \mathbf{I} is an ideal, it follows that $A \in \mathbf{I}$, as promised.

Since \mathbf{I} contains all operators of rank 1, it contains also all operators of finite rank, and, since \mathbf{I} is closed, it follows that \mathbf{I} contains every compact operator. (Note that separability was not needed yet.)

The final step is to show that if \mathbf{I} contains an operator A that is not compact, then \mathbf{I} contains every operator. If UP is the polar decomposition of A , then $P \in \mathbf{I}$ (because $P = U^*A$), and P is not compact (because $A = UP$). Since P is Hermitian, there exists an infinite-dimensional subspace \mathbf{M} , invariant under P , on which P is bounded from below, by ϵ say. (If not, P would be compact.) Let V be an isometry from \mathbf{H} onto \mathbf{M} . (Here is where the separability of \mathbf{H} comes in.) Since $P\mathbf{M} = \mathbf{M}$, it follows that $V^*PV\mathbf{H} = V^*P\mathbf{M} = V^*\mathbf{M} = \mathbf{H}$. Since, moreover, $Vf \in \mathbf{M}$ for all f , it follows that

$$\|V^*PVf\| = \|PVf\| \geq \epsilon \|Vf\| = \epsilon \|f\|.$$

These two assertions imply that V^*PV is invertible. Since $V^*PV \in \mathbf{I}$, the proof is complete; an ideal that contains an invertible element contains everything.

Solution 139. If A is normal and if A^n is compact for some positive integer n , then A is compact.

Proof. Represent A as a multiplication operator, induced by a bounded measurable function φ on a suitable measure space, and note that this automatically represents A^n as the multiplication operator induced by φ^n . Since, by Problem 133, A^n is the direct sum of 0 and a diagonal operator with diagonal terms converging to 0, it follows that the essential range of φ^n is a countable set that can cluster at 0 alone. This implies that A is the direct sum of 0 and a diagonal operator with diagonal terms converging to 0, and hence that A is compact.

Solution 140. Given the compact operator C on a Hilbert space \mathbf{H} , write $A = 1 - C$. It is to be proved that if $\ker A = \{0\}$, then A is invertible. It is convenient to approach the proof via the following two lemmas. (1) If $\text{ran } A = \mathbf{H}$, then $\ker A = \{0\}$. (2) A is bounded from below on $(\ker A)^\perp$.

(1) Put $\mathbf{K}_n = \ker A^n$, $n = 1, 2, 3, \dots$. If $\mathbf{K}_1 \neq 0$, let f_1 be a non-zero vector in \mathbf{K}_1 , and then, inductively, find f_{n+1} so that $Af_{n+1} = f_n$. It follows that $f_n \in \mathbf{K}_n$ for all n , and that, in fact, the smallest power of A that annihilates f_n is the n -th. This implies that the sequence $\{\mathbf{K}_1, \mathbf{K}_2, \mathbf{K}_3, \dots\}$ is strictly increasing, and hence that there exists an orthonormal sequence $\{e_1, e_2, e_3, \dots\}$ such that $e_n \in \mathbf{K}_n$ for all n . Since $Ae_{n+1} \in \mathbf{K}_n$, so that $Ae_{n+1} \perp e_{n+1}$, it follows that

$$\|Ce_{n+1}\|^2 = \|e_{n+1} - Ae_{n+1}\|^2 = \|e_{n+1}\|^2 + \|Ae_{n+1}\|^2 \geq 1.$$

Since $e_n \rightarrow 0$ weakly, this contradicts the assumed compactness of C . Conclusion: $\mathbf{K}_1 = \{0\}$.

(2) If A is not bounded from below on $(\ker A)^\perp$, then there exist unit vectors f_n in $(\ker A)^\perp$ such that $Af_n \rightarrow 0$. In view of the compactness of C , there is no loss of generality in assuming that the sequence $\{Cf_n\}$ is (strongly) convergent, to f say. Since $f_n = Af_n + Cf_n \rightarrow f$, it follows that $f \in (\ker A)^\perp$ and that $\|f\| = 1$; since, however, $Af_n \rightarrow 0$, so that $Af = 0$, it follows that $f \in \ker A$, and hence that $f = 0$. This contradiction completes the proof that A is bounded from below on $(\ker A)^\perp$.

The results of (1) and (2) apply to C^* and A^* just as well as to C and A . Since $\text{ran } A = A\mathbf{H} = A((\ker A)^\perp)$, it follows from (2) that $\text{ran } A$ is always closed, and hence, by the comment just made, that $\text{ran } A^*$ is always closed. Suppose now that $\ker A = \{0\}$; it follows, of

course, that $\text{ran } A^*$ is dense in \mathbf{H} . Since $\text{ran } A^*$ is closed, this implies that $\text{ran } A^* = \mathbf{H}$, and hence that (1) is applicable (to A^*). Conclusion: $\ker A^* = \{0\}$, and therefore $(\ker A^*)^\perp = \mathbf{H}$. Apply (2) (to A^*) and infer that A^* is bounded from below. Together with the density of $\text{ran } A^*$, already established, this implies that A^* is invertible, and hence that A is invertible; the proof is complete.

Solution 141. Suppose that A is compact and suppose that \mathbf{M} is a subspace such that $\mathbf{M} \subset \text{ran } A$. The inverse image of \mathbf{M} under A is a subspace, say \mathbf{N} , and so is the intersection $\mathbf{N} \cap (\ker A)^\perp$. The restriction of A to that intersection is a one-to-one bounded linear transformation from that intersection onto \mathbf{M} , and therefore (Problem 41) it is invertible. The image of the closed unit ball of \mathbf{N} (i.e., of the intersection of the closed unit ball with \mathbf{N}) is a strongly compact subset of \mathbf{M} , and, by invertibility, it includes a closed ball in \mathbf{M} (i.e., it includes the intersection of a closed ball with \mathbf{M}). This implies that \mathbf{M} is finite-dimensional; the proof is complete.

Solution 142. (1) implies (2). It is always true that A maps $(\ker A)^\perp$ one-to-one onto $\text{ran } A$, and hence that the inverse mapping maps $\text{ran } A$ one-to-one onto $(\ker A)^\perp$. In the present case $\text{ran } A$ is closed, and therefore, by the closed graph theorem, the inverse mapping is bounded. Let B be the operator that is equal to that inverse on $\text{ran } A$ and equal to 0 on $(\text{ran } A)^\perp$. Let P be the projection on $\ker A$, and let Q be the projection on $(\text{ran } A)^\perp$. Note that both P and Q have finite rank. Since $BA = 1 - P$ on both $(\ker A)^\perp$ and $\ker A$, and since $AB = 1 - Q$ on both $\text{ran } A$ and $(\text{ran } A)^\perp$, it follows that both $1 - BA$ and $1 - AB$ have finite rank.

(2) implies (3). Trivial: an operator of finite rank is compact.

(3) implies (1). If $C = 1 - AB$ and $D = 1 - BA$, with C and D compact, then both $\ker B^*A^*$ and $\ker BA$ are finite-dimensional. It follows that both $\ker A^*$ and $\ker A$ are finite-dimensional, and hence that both $(\text{ran } A)^\perp$ and $\ker A$ are finite-dimensional. To prove that $\text{ran } A$ is closed, note first that BA is bounded from below on $(\ker BA)^\perp$. (See Solution 140.) Since $\|BAf\| \leq \|B\| \cdot \|Af\|$ for all f , it follows that A is bounded from below on $(\ker BA)^\perp$, and hence that the image under A of $(\ker BA)^\perp$ is closed. Since $\ker BA$ is finite-dimensional, the

image under A of $\ker BA$ is finite-dimensional and hence closed, and $\text{ran } A$ is the sum $A(\ker BA) + A(\ker BA)^\perp$. (Recall that the sum of two subspaces, of which one is finite-dimensional, is always a subspace; see Problem 8.)

Solution 143. Translate by λ and reduce the assertion to this: if A is not invertible but $\ker A = \{0\}$, then B is not invertible. Contrapositively: if B is invertible, then either $\ker A \neq \{0\}$ or A is invertible. For the proof, assume that B is invertible and write

$$A = B + (A - B) = B(1 + B^{-1}(A - B)).$$

The operator $B^{-1}(A - B)$ is compact along with $A - B$. It follows that either -1 is an eigenvalue of $B^{-1}(A - B)$ (in which case $\ker A \neq \{0\}$), or $1 + B^{-1}(A - B)$ is invertible (in which case A is invertible).

Solution 144. The bilateral shift is an example. Suppose that $\{e_n: n = 0, \pm 1, \pm 2, \dots\}$ is the basis that is being shifted ($We_n = e_{n+1}$), and let C be the operator defined by $Cf = (f, e_{-1})e_0$. The operator C has rank 1 (its range is the span of e_0), and it is therefore compact. What is the operator $W - C$? Since \mathbf{H}^2 (the span of the e_n 's with $n \geq 0$) is invariant under both W and C , it is invariant under $W - C$ also. The orthogonal complement of \mathbf{H}^2 (the span of the e_n 's with $n < 0$) is invariant under neither W nor C (since $We_{-1} = Ce_{-1} = e_0$), but it is invariant under $W - C$. (Reason: if $n < 0$, then $W - C$ maps e_n onto e_{n+1} or 0, according as $n < -1$ or $n = -1$.) Conclusion: \mathbf{H}^2 reduces $W - C$. This conclusion makes it easy to describe $W - C$; it agrees with the unilateral shift on \mathbf{H}^2 and it agrees with the adjoint of the unilateral shift on the orthogonal complement of \mathbf{H}^2 . In other words, $W - C$ is the direct sum $U^* \oplus U$, and, consequently, its spectrum is the union of the spectra of U^* and U .

It helps to look at all this via matrices. The matrix of W (with respect to the shifted basis) has 1's on the diagonal just below the main one and 0's elsewhere; the effect of subtracting C is to replace one of the 1's, the one in row 0 and column -1 , by 0.

Solution 145. *No perturbation makes the unilateral shift normal.*

Proof. The technique is to examine the spectrum and to use the relative stability of the spectrum under perturbation.

If $U = B - C$, with B normal and C compact, then

$$U^*U = B^*B - D,$$

where

$$D = C^*B + B^*C - C^*C,$$

so that D is compact. Since $U^*U = 1$, and since $\Lambda(1) = \{1\}$, it follows (Problem 143) that every number in $\Lambda(B^*B)$, except possibly 1, must in fact be an eigenvalue of B^*B . (Alternatively, use the Fredholm alternative.) Since a Hermitian operator on a separable Hilbert space can have only countably many eigenvalues, it follows that the spectrum of B^*B must be countable. Since $\Lambda(U)$ is the closed unit disc, and since U has no eigenvalues, another consequence of Problem 143 is that the spectrum of B can differ from that disc by the set of eigenvalues of B only. A normal operator on a separable Hilbert space can have only countably many eigenvalues. Conclusion: modulo countable sets, $\Lambda(B)$ is the unit disc, and therefore (Problem 97), modulo countable sets, $\Lambda(B^*B)$ is the interval $[0,1]$. This contradicts the countability of $\Lambda(B^*B)$.

Solution 146. If H and K are Volterra kernels, then so is their “product” (matrix composition). Reason:

$$(HK)(x,y) = \int_0^1 H(x,z)K(z,y)dz,$$

and if $x < y$, then, for all z , either $x < z$ (in which case $H(x,z) = 0$), or $z < y$ (in which case $K(z,y) = 0$). In other words $(HK)(x,y) = 0$ if $x < y$; if $x \geq y$, then

$$(HK)(x,y) = \int_y^x H(x,z)K(z,y)dz,$$

because unless z is between y and x one of $H(x,z)$ and $K(z,y)$ must vanish. It follows that if K is a bounded Volterra kernel, with, say,

$|K(x, y)| \leq c$, and if $x \geq y$, then

$$|K^2(x, y)| = \left| \int_y^x K(x, z) K(z, y) dz \right| \leq c^2 \cdot (x - y).$$

(In this context symbols such as K^2 , K^3 , etc., refer to the “matrix products” KK , KKK , etc.) From this in turn it follows that if $x \geq y$, then

$$|K^3(x, y)| = \left| \int_y^x K(x, z) K^2(z, y) dz \right| \leq c^3 \int_y^x (z - y) dz = \frac{c^3}{2} (x - y)^2.$$

These are the first two steps of an obvious inductive procedure; the general result is that if $n \geq 1$ and $x \geq y$, then

$$|K^n(x, y)| \leq \frac{c^n}{(n-1)!} (x - y)^{n-1}.$$

This implies, a fortiori, that

$$|K^n(x, y)| \leq \frac{c^n}{(n-1)!},$$

and hence that if A is the induced integral operator, then

$$\|A^n\| \leq \|K^n\| \leq \frac{c^n}{(n-1)!}.$$

Since

$$\left(\frac{1}{(n-1)!} \right)^{1/n} \rightarrow 0 \quad \text{as } n \rightarrow \infty$$

(recall that the radius of convergence of the exponential series is ∞), the proof that A is quasinilpotent is complete.

Solution 147. *Every Volterra operator is quasinilpotent.*

Proof. It is natural to try to prove the theorem by approximation. Given a kernel that vanishes above the diagonal, redefine it to be 0 on a

thin strip parallel to the diagonal, and prove that the approximating kernel so obtained induces a nilpotent operator. It turns out that this works, but only if the approximation is handled with care. Recall (Solution 87) that a limit of quasinilpotent operators may fail to be quasinilpotent. The necessary care can be formulated as follows.

Lemma. *If A is a Volterra operator, and if ϵ is a positive number, then there exist Volterra operators B and C and there exists a positive integer k such that (1) $A = B + C$, (2) $\|B\| < \epsilon$, and (3) every product of B 's and C 's in which k or more factors are equal to C is equal to 0.*

Proof of lemma. (1) Let K be the Volterra kernel that induces A . The natural way to break K into a small part M and a nilpotent part N is to break the lower right triangle $E (= \{\langle x, y \rangle : x > y\})$ into a diagonal strip $D(\delta) (= \{\langle x, y \rangle : 0 \leq x - y \leq \delta\})$ and a similar parallel triangle $E(\delta) (= \{\langle x, y \rangle : x > y + \delta\})$. The natural way works. Using the absolute continuity of indefinite integrals, choose δ so that $\iint_{D(\delta)} |K(x, y)|^2 dx dy < \epsilon^2$. Let M be K in $D(\delta)$ and 0 elsewhere, and let N be K in $E(\delta)$ and 0 elsewhere. If B and C are the integral operators induced by M and N , respectively, then it is clear that $A = B + C$.

(2) The proof that $\|B\| < \epsilon$ is immediate; since the L^2 norm of the kernel of B is less than ϵ , it follows that the operator norm of B is less than ϵ .

(3) The proof that C has the unusually strong nilpotence property depends on some simple calculations with integrals. The kernel of BC is given by $\int_y^x M(x, z)N(z, y)dz$ when $x \geq y$. (When $x < y$ it vanishes, of course.) Assertion: if $\langle x, y \rangle \in D(\delta)$, this integral vanishes. Indeed: if $y \leq z \leq x$, then $0 \leq z - y \leq x - y \leq \delta$, and therefore $N(z, y) = 0$. More generally (same proof): if B and C are Volterra operators (not necessarily the ones constructed above), and if the kernel of C vanishes on $D(\delta)$, then the kernel of BC vanishes on $D(\delta)$. Next: the kernel of CB is given by $\int_y^x N(x, z)M(z, y)dz$ when $x \geq y$. Assertion: if $\langle x, y \rangle \in D(\delta)$, this integral vanishes. Indeed: if $y \leq z \leq x$, then $0 \leq x - z \leq x - y \leq \delta$, and therefore $N(x, z) = 0$. More generally (same proof): if B and C are Volterra operators, and if the kernel of C vanishes on $D(\delta)$, then the kernel of CB vanishes on $D(\delta)$. Summary: for each positive

number δ , the Volterra operators whose kernel vanishes on $D(\delta)$ form a two-sided ideal in the algebra of all Volterra operators. Finally, the kernel of C^2 is given by $\int_y^x N(x,z)N(z,y)dz$ when $x \geq y$. Assertion: if $\langle x,y \rangle \in D(2\delta)$, this integral vanishes. Indeed, if $0 \leq x - y \leq 2\delta$, then, for all z , either $x - z \leq \delta$ or $z - y \leq \delta$, and therefore either $N(x,z) = 0$ or $N(z,y) = 0$. More generally (same proof): if C_1 and C_2 are Volterra operators whose kernels vanish on $D(\delta_1)$ and $D(\delta_2)$, respectively, then the kernel of C_1C_2 vanishes on $D(\delta_1 + \delta_2)$. These three algebraic relations (BC , CB , and C_1C_2) imply what is wanted. Consider a collection of B 's and C 's and start multiplying them. Each time that a factor C is used, the strip on which the kernel vanishes grows by δ ; when a B is used, the strip at least does not shrink. Conclusion: if k is the smallest integer such that $k\delta > 1$, then the product vanishes as soon as k of the factors are equal to C . The proof of the lemma is complete.

What was proved so far was an approximation lemma; it says that every Volterra operator can be approximated by "highly nilpotent" Volterra operators. The alleged consequence of this approximation lemma says that, for each positive number ε , the inequality $\|A^n\|^{1/n} < \varepsilon$ holds when n is sufficiently large. For the proof, apply the approximation lemma, but, for convenience, with $\varepsilon/2$ in place of ε . Since $A^n = (B + C)^n$, it follows that if $n > k$ (this is the k mentioned in the approximation lemma), then

$$\|A^n\| \leq \sum_{i=0}^{k-1} \binom{n}{i} \left(\frac{\varepsilon}{2}\right)^{n-i} \|C\|^i.$$

Reason: the other terms that the binomial theorem contributes have k or more factors equal to C , and therefore (by (3)) they vanish. Now if $0 \leq i \leq k - 1$, then (by a wastefully generous estimate)

$$\binom{n}{i} \leq n^k,$$

and therefore

$$\|A^n\|^{1/n} \leq \frac{\varepsilon}{2} \cdot n^{k/n} \cdot \left(\sum_{i=0}^{k-1} \left(\frac{\varepsilon}{2}\right)^{-i} \|C\|^i \right)^{1/n}.$$

The second and third factors on the right side of this inequality tend to

1 as n becomes large; the proof is completed by choosing n large enough to make their contribution less than 2.

Solution 148. $\|V\| = 2/\pi$.

Proof. A direct attack on the problem seems not to lead anywhere. Here is a not unnatural indirect attack: evaluate $\|V^*V\|$ and take its square root. The reason this is promising is that V^*V is not only compact (as is V), but it is also Hermitian. It follows that V^*V is diagonal; the obvious way to find its norm is to find its largest eigenvalue. (Note that V^*V is positive, so that its eigenvalues are positive.)

Since V^* is given by

$$(V^*f)(x) = \int_x^1 f(y) dy,$$

it is easy to find the integral kernel that induces V^*V . A simple computation shows that that kernel, K say, is given by

$$K(x,y) = 1 - \max(x,y) = \begin{cases} 1 - x & \text{if } 0 \leq y \leq x \leq 1, \\ 1 - y & \text{if } 0 \leq x < y \leq 1. \end{cases}$$

It follows that

$$(V^*Vf)(x) = \int_0^1 f(y) dy - x \int_0^x f(y) dy - \int_x^1 y f(y) dy$$

for almost every x , whenever $f \in L^2(0,1)$. This suggests that the eigenvalues of V^*V can be explicitly determined by setting $V^*Vf = \lambda f$, differentiating (twice, to get rid of all integrals), and solving the resulting differential equation. There is no conceptual difficulty in filling in the steps. The outcome is that if

$$c_k(x) = \frac{1}{\sqrt{2}}(e^{i\pi(k+\frac{1}{2})x} + e^{-i\pi(k+\frac{1}{2})x}),$$

for $k = 0, 1, 2, \dots$, then the c_k 's form an orthonormal basis for L^2 , and each c_k is an eigenvector of V^*V , with corresponding eigenvalue $1/(k + \frac{1}{2})^2\pi^2$. The largest of these eigenvalues is the one with $k = 0$.

The outline above shows how the eigenvalues and eigenvectors can be *discovered*. If all that is wanted is an answer to the question (how much is $\|V\|$?), it is enough to *verify* that the c_k 's are eigenvectors of V^*V , with the eigenvalues as described above, and that the c_k 's form an orthonormal basis for L^2 . The first step is routine computation. The second step is necessary in order to guarantee that V^*V has no other eigenvalues, possibly larger than any of the ones that go with the c_k 's.

Here is a way to prove that the c_k 's form a basis. For each f in L^2 write

$$(Uf)(x) = \frac{1}{\sqrt{2}}(f(x)e^{i\pi x/2} + f(1-x)e^{-i\pi x/2}).$$

It is easy to verify that U is a unitary operator. If $e_n(x) = e^{2\pi i n x}$, $n = 0, \pm 1, \pm 2, \dots$, then $Ue_n = c_{2n}$ for $n = 0, 1, 2, \dots$, and $Ue_n = c_{-(2n+1)}$ for $n = -1, -2, -3, \dots$.

Solution 149. $\Lambda(V_0) = \{0\}$, $\|V_0\| = 4/\pi$.

Proof. The most illuminating remark about V_0 is that its range is included in the set of all odd functions in $L^2(-1, +1)$. (Recall that f is even if $f(x) = f(-x)$, and f is odd if $f(x) = -f(-x)$.) The second most illuminating remark (suggested by the first) is that if f is odd, then $V_0 f = 0$. These two remarks imply that V_0 is nilpotent of index 2, and hence that the spectrum of V_0 consists of 0 only.

One way to try to find the norm of V_0 is to identify $L^2(-1, +1)$ with $L^2(0,1) \oplus L^2(0,1)$, determine the two-by-two operator matrix of V_0 corresponding to such an identification, and hope that the entries in the matrix are simple and familiar enough to make the evaluation of the norm feasible. One natural way to identify $L^2(-1, +1)$ with $L^2(0,1) \oplus L^2(0,1)$ is to map f onto $\langle g, h \rangle$, where $g(x) = f(x)$ and $h(x) = f(-x)$ whenever $x \in (0,1)$. This gives something, but it is not the best thing to do. For present purposes another identification of $L^2(-1, +1)$ with $L^2(0,1) \oplus L^2(0,1)$ is more pertinent; it is the one that maps f onto $\langle g, h \rangle$, where

$$g(x) = \frac{1}{2}(f(x) - f(-x)) \quad \text{and} \quad h(x) = \frac{1}{2}(f(x) + f(-x))$$

whenever $x \in (0,1)$. The inverse map sends $\langle g, h \rangle$ onto f , where

$$f(x) = h(x) + g(x) \quad \text{and} \quad f(-x) = h(x) - g(x)$$

whenever $x \in (0,1)$. Since

$$(V_0 f)(x) = 2 \int_0^x \frac{1}{2} (f(y) + f(-y)) dy,$$

it follows that if $x \in (0,1)$, then

$$(V_0 f)(x) = 2(Vh)(x) \quad \text{and} \quad (V_0 f)(-x) = -2(Vh)(x).$$

The conclusion can be expressed in the form

$$V_0 \langle g, h \rangle = \langle 2Vh, 0 \rangle.$$

From this form the matrix of V_0 can be read off; it is

$$\begin{pmatrix} 0 & 0 \\ 2V & 0 \end{pmatrix}.$$

This proves, again, that $V_0^2 = 0$, and it shows, moreover, that $\|V_0\| = 2\|V\|$.

Solution 150. If V is the Volterra integration operator, and if $A = (1 + V)^{-1}$, then $\Lambda(A) = \{1\}$ and $\|A\| = 1$.

Proof. The example is simple, but it is the sort that takes either inspiration or experience to produce; reason alone does not seem to be enough. To prove that the example works, begin by recalling that $\Lambda(V) = \{0\}$ (cf. Problems 146 and 74); it follows that $\Lambda(1 + V) = \{1\}$, so that $1 + V$ is invertible, and hence that the definition of A makes sense. Since $\Lambda(1 + V) = \{1\}$, it follows that $\Lambda(A) = \{1\}$. Since $r(A) = 1$, it follows that $\|A\| \geq 1$. Clearly $A \neq 1$. This settles all properties except one; everything except the inequality $\|A\| \leq 1$ is obvious.

One way to prove that $\|Af\| \leq \|f\|$ for all f , i.e., that A is bounded from above by 1, is to prove that A^{-1} is bounded from below by 1. Since

$$\begin{aligned}\|A^{-1}f\|^2 &= \|(1+V)f\|^2 = (f+Vf, f+Vf) \\ &= \|f\|^2 + (Vf, f) + (V^*f, f) + \|Vf\|^2,\end{aligned}$$

it is sufficient to prove that $((V+V^*)f, f) \geq 0$ (i.e., that the real part of V is positive). This is true and already known; the operator $V+V^*$ is, in fact, the projection onto the (one-dimensional) space of constant functions (see Problem 148).

Solution 151. Let $\{\alpha_n\}$ be the weight sequence, so that $\alpha_0 \geq \alpha_1 \geq \alpha_2 \geq \dots$, $\alpha_n \neq 0$, and $\sum_{n=0}^{\infty} \alpha_n^2 < \infty$; the operator A is given by $Ae_n = \alpha_{n-1}e_{n-1}$ when $n > 0$ and $Ae_0 = 0$.

Each non-zero vector f in the given Hilbert space \mathbf{H} has a “degree”, namely the largest index n (or ∞ if there is no largest) such that the Fourier coefficient (f, e_n) is not zero. Suppose that $f \in \mathbf{M}$ and that $\deg f = n < \infty$. It is easy to see that the vectors $f, \dots, A^n f$ are linearly independent; the point is that the non-vanishing of the α ’s implies that $\deg A^i f = n - i$, $i = 0, \dots, n$. Since $A^i f \in \mathbf{M}_n$, $i = 0, \dots, n$, it follows that the span of $\{f, \dots, A^n f\}$ is \mathbf{M}_n , and hence that $\mathbf{M}_n \subset \mathbf{M}$.

The degrees of the non-zero vectors in \mathbf{M} are either bounded or not. If they are, and if their maximum is n , then $\mathbf{M} \subset \mathbf{M}_n$, and the preceding paragraph implies that $\mathbf{M} = \mathbf{M}_n$. It remains to show that if \mathbf{M} is a subspace invariant under A and if the degrees of the non-zero vectors in \mathbf{M} are not bounded, then $\mathbf{M} = \mathbf{H}$. If \mathbf{M} contains vectors of arbitrarily large finite degree, then, by the preceding paragraph, $\mathbf{M}_n \subset \mathbf{M}$ for infinitely many n , and hence $\mathbf{M} = \mathbf{H}$. The only remaining case is the one in which \mathbf{M} contains a vector of infinite degree.

Consider the following lemma: if \mathbf{M} is a subspace invariant under A , and if \mathbf{M} contains a vector of infinite degree, then \mathbf{M} contains e_0 . Assertion: the lemma implies the theorem. To prove this, it is sufficient to prove that $\mathbf{M}_k \subset \mathbf{M}$ for all k . The idea of the proof is that nothing changes if the first few terms of the basis are omitted. In precise language, the proof is induction on k . The initial step is the lemma itself. Suppose now that $\mathbf{M}_k \subset \mathbf{M}$, let P_k be the projection onto \mathbf{M}_k^\perp , and let A_k be the

operator on \mathbf{M}_k^\perp defined by $A_k f = P_k A f$ for each f in \mathbf{M}_k^\perp . The induction hypothesis implies that $P_k f \in \mathbf{M}$ for all f , and hence that $\mathbf{M} \cap \mathbf{M}_k^\perp$ is invariant under A_k . Since A_k is a weighted shift on \mathbf{M}_k^\perp (with respect to the orthonormal basis $\{e_k, e_{k+1}, \dots\}$), satisfying exactly the same conditions as A on \mathbf{H} , and since the image under P_k of a vector of infinite degree in \mathbf{H} has infinite degree in \mathbf{M}_k^\perp (with respect to the basis $\{e_k, e_{k+1}, \dots\}$), the lemma is applicable. The conclusion is that $\mathbf{M} \cap \mathbf{M}_k^\perp$ contains e_k (so that, in particular, $e_k \in \mathbf{M}$, whence $\mathbf{M}_k \subset \mathbf{M}$), and the derivation of the theorem from the lemma is complete.

Turn now to the proof of the lemma. Suppose that $f \in \mathbf{M}$ and $\deg f = \infty$. If $f = \sum_{i=0}^{\infty} \xi_i e_i$, then

$$A^n f = \sum_{i=n}^{\infty} \xi_i \alpha_{i-1} \cdots \alpha_{i-n} e_{i-n}.$$

If n is such that $\xi_n \neq 0$, then

$$\frac{1}{\xi_n \alpha_{n-1} \cdots \alpha_0} A^n f = e_0 + f_n$$

where

$$f_n = \sum_{i=n+1}^{\infty} \frac{\xi_i}{\xi_n} \cdot \frac{\alpha_{i-1} \cdots \alpha_{i-n}}{\alpha_{n-1} \cdots \alpha_0} e_{i-n}.$$

It is sufficient to prove that for each positive number ε the integer n can be chosen so that $\|f_n\| < \varepsilon$. To do this, first choose k so that

$$\sum_{i=k}^{\infty} \alpha_i^2 < \varepsilon^2 \alpha_0^2,$$

and then choose n so that $n \geq k$ and so that

$$|\xi_n| \geq \max\{|\xi_i| : i \geq k\}.$$

With this choice, $\xi_n \neq 0$, and, if $i \geq n$, then $|\xi_i/\xi_n| \leq 1$. Note also that if $i \geq n+1$, then $\alpha_{i-2} \leq \alpha_{n-1}$, \dots , $\alpha_{i-n} \leq \alpha_1$ (here is where mo-

notoneness is used). Conclusion:

$$\begin{aligned}
 \|f_n\|^2 &= \sum_{i=n+1}^{\infty} \left| \frac{\xi_i}{\xi_n} \right|^2 \left(\frac{\alpha_{i-1} \cdots \alpha_{i-n}}{\alpha_{n-1} \cdots \alpha_0} \right)^2 \\
 &\leq \sum_{i=n+1}^{\infty} \left(\frac{\alpha_{i-1}}{\alpha_0} \right)^2 \\
 &= \sum_{i=n}^{\infty} \left(\frac{\alpha_i}{\alpha_0} \right)^2 < \varepsilon^2.
 \end{aligned}$$

Chapter 16. Subnormal operators

Solution 152. Every known proof of Fuglede's theorem can be modified so as to yield this generalized conclusion. Alternatively, there is a neat derivation, via operator matrices, of the statement for two normal operators from the statement for one. Write

$$\hat{A} = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix} \quad \text{and} \quad \hat{B} = \begin{pmatrix} 0 & B \\ 0 & 0 \end{pmatrix}.$$

The operator \hat{A} is normal, and a straightforward verification proves that \hat{B} commutes with it. The Fuglede theorem implies that \hat{B} commutes with \hat{A}^* also, and (multiply the matrices \hat{A}^* and \hat{B} in both orders and compare corresponding entries) this implies the desired conclusion.

The corollary takes a little more work. If B is invertible, and if UP is its polar decomposition, then U is unitary and P is, as always, the positive square root of B^*B . If A_1 and A_2 are normal and $A_1B = BA_2$, then

$$A_2(B^*B) = (A_2B^*)B = (B^*A_1)B = B^*(A_1B) = B^*(BA_2) = (B^*B)A_2,$$

so that

$$A_2P^2 = P^2A_2;$$

it follows that

$$A_2P = PA_2.$$

(Compare Solution 108.) Since $A_1UP = UPA_2$ (by assumption) = UA_2P (by what was just proved), it follows that $A_1U = UA_2$, and the proof of the corollary is complete.

There is a breathtakingly elegant and simple proof of the Putnam-Fuglede theorem in Rosenblum [1958]. The original proof is in Fuglede [1950]; a variant is in Halmos [1951, §41] or Halmos [1963 b]; the two-operator generalization first appeared in Putnam [1951 b]. The ingenious matrix derivation of Putnam from Fuglede is due to Berberian [1959].

Solution 153. If $\chi_{\varphi^{-1}(D)}f = f$, then $\{x: f(x) \neq 0\} \subset \varphi^{-1}(D)$, and therefore

$$\begin{aligned} \|A^n f\|^2 &= \int |\varphi^n f|^2 d\mu = \int_{\varphi^{-1}(D)} |\varphi^n f|^2 d\mu \\ &\leq \int_{\varphi^{-1}(D)} |f|^2 d\mu = \|f\|^2. \end{aligned}$$

If, conversely, $\|A^n f\| \leq \|f\|$ for all n , and if $M_r = \{x: |\varphi(x)| \geq r > 1\}$, then

$$\|f\|^2 \geq \int |\varphi^n f|^2 d\mu \geq \int_{M_r} r^{2n} |f|^2 d\mu.$$

Unless f vanishes on M_r , the last written integral becomes infinite with n . Conclusion: f vanishes on M_r , for every r , and therefore $\{x: f(x) \neq 0\} \subset \varphi^{-1}(D)$.

Solution 154. It is convenient to begin with the observation that if A is quasinormal, then $\ker A$ reduces A . Reason: $\ker A = \ker A^*A$ for every operator A ; since quasinormality implies that A^* commutes with A^*A , it follows that $\ker A^*A$ is invariant under A^* .

In view of the preceding paragraph every quasinormal operator is the direct sum of 0 and an operator with trivial kernel. Since the direct summands can be treated separately, there is no loss of generality in assuming that $\ker A = \{0\}$ in the first place. If, in that case, UP is the polar decomposition of A , then U is an isometry, and (by Problem 108) $UP = PU$ and $U^*P = PU^*$. The isometric character of U implies that if E is the projection UU^* , then $(1 - E)U = U^*(1 - E) = 0$. In view of these algebraic relations, A can be shown to be subnormal by explicitly constructing a normal extension for it. If A acts on \mathbf{H} , then a normal extension B can be constructed that acts on $\mathbf{H} \oplus \mathbf{H}$. (If \mathbf{H} is identified with $\mathbf{H} \oplus \{0\}$, then \mathbf{H} is a subspace of $\mathbf{H} \oplus \mathbf{H}$.) An operator on $\mathbf{H} \oplus \mathbf{H}$ is given by a two-by-two matrix whose entries are operators on \mathbf{H} . If, in particular,

$$V = \begin{pmatrix} U & 1 - E \\ 0 & U^* \end{pmatrix} \quad \text{and} \quad Q = \begin{pmatrix} P & 0 \\ 0 & P \end{pmatrix},$$

then V is unitary, Q is positive, V and Q commute, and therefore

$$B = \begin{pmatrix} UP & (1 - E)P \\ 0 & U^*P \end{pmatrix}$$

is a normal extension of A .

Solution 155. Let \mathbf{M}_1 be the set of all finite sums of the form $\sum_j B_1^* f_j$, where $f_j \in \mathbf{H}$ for all j ($= 0, 1, 2, \dots$). The set \mathbf{M}_1 is a linear manifold; since $B_1(\sum_j B_1^* f_j) = \sum_j B_1^* (B_1 f_j)$ and $B_1^*(\sum_j B_1^* f_j) = \sum_j B_1^{*j+1} f_j$, the closure of \mathbf{M}_1 reduces B_1 . Since \mathbf{H} itself is included in \mathbf{M}_1 , the minimality of B_1 implies that $\mathbf{K}_1 = \bar{\mathbf{M}}_1$. Similarly, of course, the set \mathbf{M}_2 of all finite sums of the form $\sum_j B_2^* f_j$, where each f_j is in \mathbf{H} , is dense in \mathbf{K}_2 .

It is tempting to try to complete the proof by setting $U(\sum_j B_1^* f_j) = \sum_j B_2^* f_j$. This works, but it takes a little care. First: does this equation really define anything? That is: if $\sum_j B_1^* f_j = \sum_j B_1^* g_j$ (with f_j and g_j in \mathbf{H}), does it follow that $\sum_j B_2^* f_j = \sum_j B_2^* g_j$? Equivalently (subtract): if $\sum_j B_1^* f_j = 0$, does it follow that $\sum_j B_2^* f_j = 0$? The answer is yes; the reason is contained in the following computation:

$$\begin{aligned} \|\sum_j B_1^* f_j\|^2 &= (\sum_j B_1^* f_j, \sum_k B_1^* f_k) \\ &= \sum_j \sum_k (B_1^* f_j, B_1^* f_k) = \sum_j \sum_k (A^* f_j, A^* f_k). \end{aligned}$$

This computation accomplishes much more than the proof that U is unambiguously defined; it implies that U is an isometry (from \mathbf{M}_1 onto \mathbf{M}_2), that therefore U has a unique isometric extension that maps \mathbf{K}_1 onto \mathbf{K}_2 , and that U is the identity on \mathbf{H} . The proof that $UB_1 = B_2U$ is another computation. It suffices to verify that UB_1 agrees with B_2U on \mathbf{M}_1 , and this is implied by

$$\begin{aligned} UB_1(\sum_j B_1^* f_j) &= U(\sum_j B_1^* B_1 f_j) = \sum_j B_2^* A f_j \\ &= \sum_j B_2^* B_2 f_j = B_2 \sum_j B_2^* f_j = B_2 U(\sum_j B_1^* f_j). \end{aligned}$$

Solution 156. *There exist two subnormal operators that are similar but not unitarily equivalent.*

Proof. Consider the measure space consisting of the unit circle together with its center, with measure ν defined so as to be normalized Lebesgue measure in the circle and a unit mass at the center. Let B be the position operator on $L^2(\nu)$ (that is, $(Bf)(z) = zf(z)$), and let A be its restriction to the closure $H^2(\nu)$ of all polynomials. Clearly B is normal and A is subnormal.

An orthonormal basis for $H^2(\nu)$ consists of the functions e_n ($n = 1, 2, 3, \dots$), defined by $e_n(z) = z^n$, together with the function e_0 , defined by $e_0(z) = 1/\sqrt{2}$. The action of A on this basis is easy to describe: $Ae_0 = (1/\sqrt{2})e_1$ and $Ae_n = e_{n+1}$ for $n = 1, 2, 3, \dots$. In other words, A is a unilateral weighted shift, with weight sequence $\{1/\sqrt{2}, 1, 1, 1, \dots\}$. It follows from Problem 76 (but it is just as easy to verify directly) that A is similar to the ordinary unweighted unilateral shift U . There are several ways of proving that U and A are not unitarily equivalent. One way is to recall that two unilateral weighted shifts are unitarily equivalent only if corresponding weights have equal moduli (cf. Problem 76); the simplest way, however, is to observe that U is an isometry and A is not.

It is worth noting that B is the minimal normal extension of A (see Problem 155). This is not obvious at a glance, but it is quite easy to prove. From this it follows again that U and A are not unitarily equivalent. Reason: their minimal normal extensions are not.

This example is due to D. E. Sarason.

Solution 157. It is to be proved that if λ is a complex number such that $B - \lambda$ is not invertible, then neither is $A - \lambda$. By simple geometry (translate) and equally simple logic (form the contrapositive), the assertion reduces to this: if A is invertible, then so is B . Suppose therefore that A is invertible; without loss of generality normalize so that $\|A^{-1}\| = 1$. Let ϵ be an arbitrary number in the open interval $(0, 1)$, fixed from now on, and write $E = \{f: \|B^n f\| \leq \epsilon^n \|f\|, n = 1, 2, 3, \dots\}$. If H and K are the domains of A and B , and if $f \in E$ and $g \in H$, then

$$\begin{aligned} |(f, g)| &= |(f, A^n A^{-n} g)| = |(f, B^n A^{-n} g)| \\ &= |(B^{*n} f, A^{-n} g)| \leq \|B^{*n} f\| \cdot \|A^{-n} g\| \\ &= \|B^n f\| \cdot \|A^{-n} g\| \leq \epsilon^n \cdot \|f\| \cdot \|g\| \end{aligned}$$

for all n , and, consequently, $(f, g) = 0$. In other words, $\mathbf{E} \perp \mathbf{H}$, and therefore $\mathbf{H} \subset \mathbf{E}^\perp$. Since (Problem 153) \mathbf{E} is a reducing subspace for B , it follows that $\mathbf{E}^\perp = \mathbf{K}$, so that $\mathbf{E} = \{0\}$; from this in turn (see Problem 153) it follows that B is invertible.

Solution 158. The proof depends on the one non-trivial relation between the spectra of A and B , namely the spectral inclusion theorem,

$$\Lambda(B) \subset \Lambda(A),$$

and on the trivial fact about spectral inclusion,

$$\Pi(A) \subset \Pi(B).$$

The conclusion holds for a pair of operators A and B whenever their spectra and approximate point spectra are so related; no deeper or more special properties of subnormal and normal operators are needed.

Consider the sets $\Delta^- = \Delta - \Lambda(A)$ and $\Delta^+ = \Delta \cap \Lambda(A)$. Since Δ is open and $\Lambda(A)$ is closed, the set Δ^- is open. Assertion: Δ^+ is also open. To prove this, consider an arbitrary point λ in Δ^+ . Since $\lambda \in \Delta$ and Δ is a hole of $\Lambda(B)$, the point λ cannot belong to $\Lambda(B)$. This implies, of course, that λ is not in $\Pi(B)$, hence that λ is not in $\Pi(A)$, and hence that λ is not on the boundary of $\Lambda(A)$ (see Problem 63). Since, however, $\lambda \in \Delta^+$ and $\Delta^+ \subset \Lambda(A)$, it follows that the only place λ can be is in the interior of $\Lambda(A)$. This argument proves that Δ^+ is, in fact, the intersection of Δ with the interior of $\Lambda(A)$, and it follows, as asserted, that Δ^+ is open.

Since Δ is the union of the disjoint open sets Δ^- and Δ^+ , the connectedness of Δ implies that one of them is empty.

The result is due to Bram [1955]; for a generalization see Itô [1958]. The simple proof above is due to S. K. Parrott.

Solution 159. *Every finite-dimensional subspace invariant under a normal operator B reduces B .*

Proof. Since on a finite-dimensional space every operator has an eigenvalue, it is sufficient to prove that each one-dimensional invariant subspace of B reduces B . This is easy: in fact each eigenvector of B is

an eigenvector of B^* too. (If $Bf = \lambda f$, then, by normality,

$$0 = \|(B - \lambda)f\| = \|(B^* - \lambda^*)f\|.)$$

Corollary. *On finite-dimensional spaces every subnormal operator is normal.*

Proof. The restriction of a normal operator to a reducing subspace is normal.

From the result thus proved it follows that the answer to the dimension question is no. Reason: if B is a normal extension of A to \mathbf{K} , then $\mathbf{K} \cap \mathbf{H}^\perp$ is invariant under the normal operator B^* , and, therefore, if $\dim(\mathbf{K} \cap \mathbf{H}^\perp)$ is finite, \mathbf{H} reduces B . Since A was assumed to be non-normal, this is impossible.

Solution 160. The difficulty is to prove that something is *not* subnormal. Since subnormality was defined by requiring the existence of something, what is wanted here is a non-existence theorem. The best way to prove such a theorem (the only way?) is to assume existence, derive a usable “constructive” necessary condition from it (with luck it will be sufficient as well), and then look for something that violates the condition.

If B (on \mathbf{K}) is a normal extension of A (on \mathbf{H}), and if f_0, \dots, f_n are vectors in \mathbf{H} , then $\|\sum_j B^{*j} f_j\| \geq 0$. This triviality can be rewritten in a non-trivial way, as follows:

$$\begin{aligned} 0 &\leq \left(\sum_j B^{*j} f_j, \sum_i B^{*i} f_i \right) = \sum_j \sum_i (B^{*j} f_j, B^{*i} f_i) \\ &= \sum_j \sum_i (B^i B^{*j} f_j, f_i) \\ &= \sum_j \sum_i (B^{*j} B^i f_j, f_i) \quad (\text{because } B \text{ is normal}) \\ &= \sum_j \sum_i (B^i f_j, B^j f_i) = \sum_j \sum_i (A^i f_j, A^j f_i). \end{aligned}$$

Replace each f_j by some scalar multiple $\xi_j f_j$, and conclude that $\sum_j \sum_i (A^i f_j, A^j f_i) \xi_j^* \xi_i \geq 0$, i.e., that the finite matrix $\langle (A^i f_j, A^j f_i) \rangle$ is

positive definite. This is a “constructive” intrinsic necessary condition that follows from subnormality; it will be used to exhibit a hyponormal operator that is not subnormal. First, however, it is pertinent to comment that the condition is sufficient, as well as necessary, for subnormality. To put it precisely: if, for each finite set of vectors f_0, \dots, f_n , the corresponding matrix $\langle (A^i f_j, A^i f_i) \rangle$ is positive definite, then the operator A is subnormal. The proof is somewhat involved; the fact will not be used in the sequel.

The desired counterexample can be found among weighted shifts. When is a weighted shift S , with weights $\{\alpha_0, \alpha_1, \alpha_2, \dots\}$ hyponormal? Since both S^*S and SS^* are diagonal, there is an easy answer in terms of the α 's. The diagonal of S^*S is $\{|\alpha_0|^2, |\alpha_1|^2, |\alpha_2|^2, \dots\}$, and the diagonal of SS^* is $\{0, |\alpha_0|^2, |\alpha_1|^2, |\alpha_2|^2, \dots\}$; it follows that S is hyponormal if and only if the sequence $\{|\alpha_n|\}$ is monotone increasing.

With this much information available, the construction of a counterexample along these lines (if it is possible at all) should be easy. A finite amount of experimentation might lead to the weighted shift S with weights $\{\alpha, \beta, 1, 1, 1, \dots\}$, where $0 < \alpha < \beta < 1$. The preceding paragraph implies that S is hyponormal. To prove that S is not subnormal, examine the matrix $\langle (S^i e_j, S^j e_i) \rangle$, where $\{e_0, e_1, e_2, \dots\}$ is the orthonormal basis that S shifts, and where i and j take the values 0, 1, 2. Written explicitly, the matrix is

$$\begin{pmatrix} 1 & \alpha & \alpha\beta \\ \alpha & \beta^2 & \beta \\ \alpha\beta & \beta & 1 \end{pmatrix}.$$

Its determinant is $-\alpha^2(1 - \beta^2)^2$, which is negative.

Examples of this type have been studied by J. G. Stampfli.

Solution 161. The “if” for normal partial isometries is trivial and for subnormal ones is a consequence of Problem 118. (The point is that the typical non-unitary isometry, the unilateral shift, is subnormal.) To prove “only if”, suppose that U is a partial isometry, so that U^*U is the projection on the initial space (the orthogonal complement of the kernel), and UU^* is the projection on the final space (the range). If U

is subnormal, then it is hyponormal, and consequently the initial space includes the range. This implies that the initial space is invariant under U , and hence that it reduces U ; clearly the restriction of U to the initial space is an isometry. If, moreover, U is normal, then the initial space is equal to the range, and therefore the restriction of U to the initial space is unitary.

It is interesting to note (as a consequence of the proof) that a partial isometry is subnormal if and only if it is hyponormal.

Solution 162. For $n = 1$, the equality is trivial; proceed by induction. Since

$$\begin{aligned} \|A^n f\|^2 &= (A^n f, A^n f) = (A^* A^n f, A^{n-1} f) \\ &\leq \|A^* A^n f\| \cdot \|A^{n-1} f\| \leq \|A^{n+1} f\| \cdot \|A^{n-1} f\| \\ &\leq \|A^{n+1}\| \cdot \|A^{n-1}\| \cdot \|f\|^2 \end{aligned}$$

for every vector f , it follows that

$$\|A^n\|^2 \leq \|A^{n+1}\| \cdot \|A^{n-1}\|.$$

In view of the induction hypothesis ($\|A^k\| = \|A\|^k$ whenever $1 \leq k \leq n$), this can be rewritten as

$$\|A\|^{2n} \leq \|A^{n+1}\| \cdot \|A\|^{n-1},$$

from which it follows that

$$\|A\|^{n+1} \leq \|A^{n+1}\|.$$

Since the reverse inequality is universal, the induction step is accomplished.

Reference: Andô [1963], Stampfli [1962]. The proof above is a slight simplification of Stampfli's simple proof.

Solution 163. Suppose that A is hyponormal. The program is to prove that the span of the eigenvectors of A reduces A ; compactness

does not enter here. In the presence of compactness the orthogonal complement of that span becomes amenable; an application of Problem 162 will yield the conclusion.

(1) For each complex number λ ,

$$\{f: Af = \lambda f\} \subset \{f: A^*f = \lambda^*f\}.$$

The reason is that $A - \lambda$ is just as hyponormal as A , and that, on general grounds that have nothing to do with hyponormality, a necessary and sufficient condition that $(A^* - \lambda^*)f = 0$ is that

$$(A - \lambda)(A^* - \lambda^*)f = 0.$$

(2) For each complex number λ , the subspace $\{f: Af = \lambda f\}$ reduces A . Indeed: invariance under A is trivial, and invariance under A^* follows from (1).

(3) If $\lambda_1 \neq \lambda_2$, then

$$\{f: Af = \lambda_1 f\} \perp \{f: Af = \lambda_2 f\}.$$

A straightforward and often-used argument: if $Af_1 = \lambda_1 f_1$ and $Af_2 = \lambda_2 f_2$, then

$$\lambda_1(f_1, f_2) = (Af_1, f_2) = (f_1, A^*f_2) = \lambda_2(f_1, f_2).$$

(4) The span of all the eigenvectors of A reduces A and the restriction of A to that span is normal. Proof: use (2) and observe that, by (3), the restriction of A to each eigenspace is normal (in fact equal to a scalar).

(5) Now assume that A is compact, and consider the restriction of A to the orthogonal complement of the span of all the eigenvectors. The resulting operator is still hyponormal (by the reduction assertion of (4)), and still compact. Since the point spectrum of this compact operator is empty, it is quasinilpotent (Problem 140); an application of Problem 162 implies that it must be 0. If the orthogonal complement on which all this action is taking place is not $\{0\}$, then there is a contradiction: the non-zero vectors in it both must be and cannot be eigenvectors of eigenvalue 0.

Solution 164. Let V be a Hilbert space and let H be the direct sum of countably many copies of V indexed by the set of all integers (positive,

negative, or zero). Explicitly, \mathbf{H} is the set of all sequences

$$f = \langle \cdots, f_{-1}, (f_0), f_1, \cdots \rangle$$

of vectors in \mathbf{V} such that $\sum_n \|f_n\|^2 < \infty$; the inner product of f and g is defined by $(f, g) = \sum_n (f_n, g_n)$. If $\{S_n: n = 0, \pm 1, \pm 2, \cdots\}$ is a sequence of positive operators on \mathbf{V} such that the sequence $\{\|S_n\|\}$ of norms is bounded, then the equations $(Sf)_n = S_n f_n$ define an operator S on \mathbf{H} . If W is the shift defined by $(Wf)_n = f_{n-1}$, then W is an operator on \mathbf{H} . The adjoints are easy to compute: $(S^*f)_n = S_n^* f_n = S_n f_n$ (so that S is Hermitian, and, in fact, positive), and $(W^*f)_n = f_{n+1}$ (so that W is invertible, and, in fact, unitary).

If $A = WS$, then $(Af)_n = S_{n-1} f_{n-1}$; since $A^* = SW^*$, it follows that $(A^*f)_n = S_n f_{n+1}$. These relations imply that $(A^*Af)_n = S_n^2 f_n$ and $(AA^*f)_n = S_{n-1}^2 f_n$, and consequently that A is hyponormal if and only if the sequence $\{S_n^2\}$ is increasing. On the other hand, $(A^2f)_n = S_{n-1}S_{n-2}f_{n-2}$, and $(A^{*2}f)_n = S_n S_{n+1} f_{n+2}$, and therefore $(A^{*2}A^2f)_n = S_n S_{n+1} S_n f_n$ and $(A^2A^{*2}f)_n = S_{n-1} S_{n-2}^2 S_{n-1} f_n$, so that A^2 is hyponormal if and only if $S_{n-1} S_{n-2}^2 S_{n-1} \leq S_n S_{n+1}^2 S_n$ for all n .

It remains to choose \mathbf{V} and the S_n 's so that A is hyponormal but A^2 is not. The construction is based on the existence of positive operators C and D such that $C \leq D$ is true but $C^2 \leq D^2$ is false. If, for instance, \mathbf{V} is two-dimensional, so that operators on \mathbf{V} may be identified with two-by-two matrices, and if

$$C = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad D = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix},$$

then

$$D - C = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \geq 0,$$

but

$$D^2 - C^2 = \begin{pmatrix} 5 & 3 \\ 3 & 2 \end{pmatrix} - \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix},$$

which has a negative determinant. If S_n is defined to be the positive square root of C (which is equal to C) whenever $n \leq 0$ and the positive

square root of D whenever $n > 0$, then $S_n^2 \leq S_{n+1}^2$ for all n , so that A is hyponormal, but, if $n = 1$, then $S_{n-1}S_{n-2}^2S_{n-1} = C^2$ and $S_nS_{n+1}^2S_n = D^2$, so that A^2 is not hyponormal.

Solution 165. *If U is unitary, P is positive and invertible, $C = P^{-1}UP$, and $A = UP$, then a necessary and sufficient condition that C be a contraction is that A be hyponormal.*

Proof. Under the stated assumptions, the following assertions are mutually equivalent:

$$CC^* \leq 1,$$

$$P^{-1}UP^2U^*P^{-1} \leq 1,$$

$$UP^2U^* \leq P^2,$$

$$(UP)(UP)^* \leq (UP)^*(UP),$$

$$AA^* \leq A^*A.$$

The answer to the question formulated in Problem 165 is now relatively easy. If W is unitary, S is invertible, and $C = S^{-1}WS$, then consider the polar decomposition VP of S , and observe that $C = P^{-1}UP$, where $U (= V^{-1}WV)$ is unitary and P is positive. Consequence: it is sufficient to consider the transforms of unitary operators by positive operators, and to them the statement just proved is applicable.

Observe next that if " \leq " is replaced by " $=$ " in each of the five displayed relations above, the resulting relations are still mutually equivalent. Consequence 1: on a finite-dimensional space a contraction that is similar to a unitary operator is unitary. Reason: on a finite-dimensional space every hyponormal operator is normal. Consequence 2: on an infinite-dimensional space a contraction that is similar to a unitary operator need not be unitary. Reason: there exist invertible hyponormal operators that are not normal. (Note that if A is hyponormal, then so is $A + \lambda$, for every scalar λ .)

The argument is due to R. G. Douglas.

Chapter 17. Numerical Range

Solution 166. Suppose that A is an operator on a Hilbert space, and that $\xi = (Af, f)$, $\eta = (Ag, g)$, where f and g are unit vectors. The problem is to prove that every point of the segment joining ξ and η is in $W(A)$. Two preliminary reductions will simplify the proof.

If $\xi = \eta$, the problem is trivial. If $\xi \neq \eta$, then there exist complex numbers α and β such that $\alpha\xi + \beta = 1$ and $\alpha\eta + \beta = 0$. It is sufficient to prove that the unit interval $[0, 1]$ is included in $W(\alpha A + \beta)$ ($= \alpha W(A) + \beta$). Reason: if $\alpha(Ah, h) + \beta = t$, then

$$\begin{aligned}\alpha(Ah, h) + \beta &= t(\alpha\xi + \beta) + (1 - t)(\alpha\eta + \beta) \\ &= \alpha(t\xi + (1 - t)\eta) + \beta.\end{aligned}$$

Consequence: there is no loss of generality in assuming that $\xi = 1$ and $\eta = 0$ in the first place.

Write $A = B + iC$ with B and C Hermitian. Since (Af, f) ($= 1$) and (Ag, g) ($= 0$) are real, it follows that (Cf, f) and (Cg, g) vanish. If f is replaced by λf , where $|\lambda| = 1$, then (Af, f) remains the same and (Cf, g) becomes $\lambda(Cf, g)$. Consequence: there is no loss of generality in assuming that (Cf, g) is purely imaginary.

With these reductions agreed on, put $h(t) = tf + (1 - t)g$, $0 \leq t \leq 1$. Assertion: $h(t)$ is never 0; in fact, the vectors f and g are linearly independent. This is a consequence of $(Af, f) \neq (Ag, g)$. If, indeed, f and g were linearly dependent, then, since they are unit vectors, either one could be written as a multiple of the other. Since, moreover, the factor would have to have absolute value 1, it would then follow that $(Af, f) = (Ag, g)$.

Since

$$(Ch(t), h(t)) = t^2(Cf, f) + t(1 - t)((Cf, g) + (Cf, g)^*) + (1 - t)^2(Cg, g),$$

the relations $(Cf, f) = (Cg, g) = 0$ and $\operatorname{Re}(Cf, g) = 0$ imply that $(Ch(t), h(t)) = 0$ for all t , and hence that $(Ah(t), h(t))$ is real for all t .

That is all that is needed. The function

$$t \rightarrow (Ah(t), h(t)) / \|h(t)\|^2$$

is real-valued and continuous on the closed unit interval; its values at 0 and 1, respectively, are 0 and 1. Conclusion: the range of the function contains every number in the unit interval.

This arrangement of the proof is due to C. W. R. de Boor.

Solution 167. For every operator A and for every positive integer k , the k -numerical range $W_k(A)$ is convex.

Proof. Suppose to begin with that \mathbf{M} and \mathbf{N} are k -dimensional Hilbert spaces and that T is a linear transformation from \mathbf{M} into \mathbf{N} . There is a useful sense in which T and T^* (from \mathbf{N} into \mathbf{M}) can be simultaneously diagonalized. The assertion is that there exist orthonormal bases $\{f_1, \dots, f_k\}$ for \mathbf{M} and $\{g_1, \dots, g_k\}$ for \mathbf{N} , and there exist positive (≥ 0) scalars $\alpha_1, \dots, \alpha_k$ such that $Tf_i = \alpha_i g_i$ and $T^*g_i = \alpha_i f_i$, $i = 1, \dots, k$. To prove this, let UP be the polar decomposition of T , and diagonalize P . That is: find an orthonormal basis $\{f_1, \dots, f_k\}$ for \mathbf{M} and find positive scalars $\alpha_1, \dots, \alpha_k$ such that $Pf_i = \alpha_i f_i$. If the partial isometry U is not an isometry from \mathbf{M} onto \mathbf{N} , it can be replaced by one (since $\dim \mathbf{M} = \dim \mathbf{N} = k$); assume that that has been done. Then put $g_i = Uf_i$, $i = 1, \dots, k$, and reap the consequences: $Tf_i = UPf_i = U(\alpha_i f_i) = \alpha_i g_i$, and $T^*g_i = PU^*g_i = Pf_i = \alpha_i f_i$, $i = 1, \dots, k$.

That is a lemma; now for the theorem. Suppose that P and Q are projections of rank k , with respective ranges \mathbf{M} and \mathbf{N} . If T is the restriction of QP to \mathbf{M} , then the preceding lemma is applicable. For each i ($= 1, \dots, k$), let \mathbf{L}_i be the span of f_i and g_i . Assertion: the subspaces \mathbf{L}_i are pairwise orthogonal. Suppose, indeed, that $i \neq j$; since $f_i \perp f_j$ and $g_i \perp g_j$, it is sufficient to prove that $f_i \perp g_j$ (for then $f_j \perp g_i$ follows by symmetry). The proof is easy:

$$(f_i, g_j) = (Pf_i, Qg_j) = (QPf_i, g_j) = (\alpha_i g_i, g_j).$$

The desired convexity proof is now near at hand. If $0 \leq t \leq 1$, use the classical Toeplitz-Hausdorff theorem k times to obtain a unit vector

h_i in L_i so that

$$(Ah_i, h_i) = t(Af_i, f_i) + (1 - t)(Ag_i, g_i).$$

Since $\{h_1, \dots, h_k\}$ is an orthonormal set, the projection R onto its span has rank k , and

$$\begin{aligned} t \cdot \operatorname{tr} PAP + (1 - t) \cdot \operatorname{tr} QAQ &= t \cdot \sum_i (Af_i, f_i) + (1 - t) \cdot \sum_i (Ag_i, g_i) \\ &= \sum_i (Ah_i, h_i) = \operatorname{tr} RAR. \end{aligned}$$

The proof of the theorem is complete.

The problem was raised by Halmos [1964]. The first solution, somewhat more complicated than the one above, is due to C. A. Berger.

Solution 168. *If A is an operator and λ is a complex number such that $|\lambda| = \|A\|$ and $\lambda \in W(A)$, then λ is an eigenvalue of A .*

Proof. If $\lambda = (Af, f)$ with $\|f\| = 1$, then

$$\|A\| = |\lambda| = |(Af, f)| \leq \|Af\| \cdot \|f\| \leq \|A\|,$$

so that equality holds everywhere. The known facts about when the Schwarz inequality becomes an equation imply that $Af = \lambda_0 f$ for some λ_0 , and this in turn implies that

$$\lambda_0 = \lambda_0(f, f) = (\lambda_0 f, f) = (Af, f) = \lambda,$$

so that λ is an eigenvalue of A .

It follows from this theorem that if λ is a number in $\overline{W(A)}$ such that $|\lambda| = \|A\|$ and λ is *not* an eigenvalue of A (and, in particular, if A has no eigenvalues), then λ does not belong to $W(A)$. In view of this comment it is easy to construct examples of operators whose numerical range is not closed.

(1) Observe that the eigenvalues of every operator A belong to $W(A)$. (Proof: if $Af = \lambda f$ with $\|f\| = 1$, then $(Af, f) = \lambda$.) If A is normal, then $\|A\| = \sup\{|\lambda| : \lambda \in W(A)\}$, so that there always exists

a λ in $\overline{W(A)}$ such that $|\lambda| = \|A\|$. It follows that if a normal operator has sufficiently many eigenvalues to approximate its norm, but does not have one whose modulus is as large as the norm, then its numerical range will not be closed. Concrete example: a diagonal operator such that the modulus of the diagonal terms does not attain its supremum. Another example, along slightly different lines: take A to be the diagonal operator with diagonal $\{1, \frac{1}{2}, \frac{1}{3}, \dots\}$. Since $A \geq 0$ and $\ker A = \{0\}$, it follows that $0 \notin W(A)$; in fact $W(A) = (0, 1]$. This shows, by the way, that the numerical range may fail to be closed even for compact operators.

(2) Take A to be the unilateral shift. Since every number in the open unit disc is an eigenvalue of A^* , it follows that the open unit disc is included in $W(A^*)$. Since $W(A^*)$ is always $(W(A))^*$ (proof: $(A^*f, f) = (Af, f)^*$), it follows that the open unit disc is included in $W(A)$. Since, finally, A has no eigenvalues, the theorem proved above implies that $W(A)$ cannot contain any number of modulus 1, so that $W(A)$ is equal to the open unit disc.

Solution 169. If λ is the compression spectrum of A , then λ^* is an eigenvalue of A^* , so that $\lambda^* \in W(A^*)$, and therefore $\lambda \in W(A)$. Conclusion: the numerical range includes the compression spectrum.

If λ is in the approximate point spectrum of A , then there exist unit vectors f_n such that $(A - \lambda)f_n \rightarrow 0$. Since

$$\begin{aligned} |(Af_n, f_n) - \lambda| &= |((A - \lambda)f_n, f_n)| \\ &\leq \|(A - \lambda)f_n\|, \end{aligned}$$

it follows that $(Af_n, f_n) \rightarrow \lambda$. Conclusion: the closure of the numerical range includes the approximate point spectrum.

These two paragraphs complete the proof. A slightly different proof can be obtained by combining the fact just proved for the approximate point spectrum with two other facts: the boundary of the spectrum is included in the approximate point spectrum, and the numerical range is convex.

Solution 170. If V is the Volterra integration operator and if $A = 1 - (1 + V)^{-1}$ ($= V(1 + V)^{-1}$), then A is quasinilpotent but $W(A)$ does not contain 0.

Proof. Since the quasinilpotence of A is obvious (Problem 146), it is sufficient to prove that if f is a vector such that $(Af, f) = 0$, then $f = 0$. If, indeed, $(Af, f) = 0$, then

$$\|f\|^2 = ((1 + V)^{-1}f, f) \leq \|(1 + V)^{-1}\| \cdot \|f\|^2 = \|f\|^2.$$

(See Solution 150. The trick of considering $(1 + V)^{-1}$ has more than one application.) It follows (by what is known about when the Schwarz inequality degenerates) that f must be an eigenvector of $(1 + V)^{-1}$. Since $\Lambda((1 + V)^{-1}) = \{1\}$, it follows that $(1 + V)^{-1}f = f$, or $f = (1 + V)f$, or $Vf = 0$. This implies that $f = 0$ (see Problem 148); the proof is complete.

Observe that the operator A is compact.

Solution 171. Suppose that A is a normal operator. Since $\overline{W(A)}$ is convex and $\Lambda(A) \subset \overline{W(A)}$ (Problems 166 and 169), it follows that $\text{conv } \Lambda(A) \subset \overline{W(A)}$. It remains to prove the reverse inclusion. In view of the characterization of convex hulls in terms of half planes, the desired result can be formulated this way: if a closed half plane includes $\Lambda(A)$, then it includes $W(A)$. If A is replaced by $\alpha A + \beta$ (where α and β are complex numbers), then Λ and W are replaced by $\alpha\Lambda + \beta$ and $\alpha W + \beta$. This remark makes it possible to “normalize” the problem. Its effect is to reduce the problem to the study of any one particular half plane, for instance the right half plane. The desired result now is this: if every number in the spectrum of A has a positive (≥ 0) real part, then the same is true of the numerical range of A . (Observe that the reduction to this point did not use normality; that assumption enters in the proof of the reduced statement.)

Use the spectral theorem to justify the assumption that A is a multiplication, induced by a bounded measurable function φ on a measure space with measure μ . If $f \in L^2(\mu)$, then $(Af, f) = \int \varphi |f|^2 d\mu$. In these terms, the reduced statement says that if $0 \leq \text{Re } \varphi$ almost everywhere (this says that the essential range of φ is included in the right half plane), then $0 \leq \text{Re} \int \varphi |f|^2 d\mu = \int (\text{Re } \varphi) |f|^2 d\mu$. This, finally, is obvious; if $d\nu = |f|^2 d\mu$, then ν is a positive measure, and the assertion is just that the integral of a positive function with respect to a positive measure is positive.

Solution 172. *The closure of the numerical range of a subnormal operator is the convex hull of its spectrum.*

Proof. If A is subnormal and B is its minimal normal extension (see Problem 155), then $\Lambda(B) \subset \Lambda(A)$ (Problem 157), and, trivially, $W(A) \subset W(B)$. It follows that

$$\begin{aligned}\overline{W(B)} &= \text{conv } \Lambda(B) \text{ (Problem 171)} \\ &\subset \text{conv } \Lambda(A) \\ &\subset \overline{W(A)} \text{ (Problems 166 and 169)} \\ &\subset \overline{W(B)},\end{aligned}$$

and hence that all the sets that enter are the same.

Note, as a corollary of the proof, that the closure of the numerical range of a subnormal operator is the same as the closure of the numerical range of its minimal normal extension.

Solution 173. (a) If A is not invertible, then $0 \in \Lambda(A)$, so that $1 \in \Lambda(1 - A)$; it follows that $1 \leq r(1 - A) \leq w(1 - A)$. (b) Assume, with no loss of generality, that $\|A\| = 1$. (Multiply by a suitable positive constant.) The hypothesis $w(A) = \|A\|$ then guarantees the existence of a sequence $\{f_n\}$ of unit vectors such that $|(Af_n, f_n)| \rightarrow 1$; assume with no loss of generality that $(Af_n, f_n) \rightarrow 1$. (Multiply by a suitable constant of modulus 1.) Since $|(Af_n, f_n)| \leq \|Af_n\| \leq 1$ and $(Af_n, f_n) \rightarrow 1$, it follows that $\|Af_n\| \rightarrow 1$. This implies that

$$\|Af_n - f_n\|^2 = \|Af_n\|^2 - 2\operatorname{Re}(Af_n, f_n) + 1 \rightarrow 0,$$

so that 1 is an approximate eigenvalue of A , and therefore $r(A)$ must be equal to 1.

Solution 174. *There exist convexoid operators that are not normaloid and vice versa.*

Proof. Write

$$M = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix},$$

and let N be a normal operator whose spectrum is the closed disc D with center 0 and radius $\frac{1}{2}$. If

$$A = \begin{pmatrix} M & 0 \\ 0 & N \end{pmatrix},$$

then $\Lambda(A) = \{0\} \cup D = D$, and $W(A) = \text{conv}(W(M) \cup W(N)) = D$. This shows that A is convexoid. Since $\|A\| = 1$ (in fact $\|M\| = 1$), A is not normaloid.

Next write

$$A = \begin{pmatrix} M & 0 \\ 0 & 1 \end{pmatrix}.$$

Since $\|A\| = 1$, and $W(A) = \text{conv}(D \cup \{1\})$, it follows that $w(A) = 1$ and hence that A is normaloid. Since, however, $\Lambda(A) = \{0\} \cup \{1\}$, so that $\text{conv } \Lambda(A)$ is the closed unit interval, A is not convexoid.

Many of these concepts were first studied by Wintner [1929]. The paper contains a small error; it asserts that every normaloid operator is convexoid.

Solution 175. *The function \bar{W} is continuous with respect to the uniform (norm) topology; if the underlying Hilbert space is infinite-dimensional, then the function w is discontinuous with respect to the strong topology (and hence with respect to the weak).*

Proof. If $\|A - B\| < \epsilon$, and if f is a unit vector, then

$$|((A - B)f, f)| < \epsilon,$$

and therefore

$$(Af, f) = (Bf, f) + ((A - B)f, f) \in W(B) + (\epsilon).$$

It follows that $W(A) \subset W(B) + (\epsilon)$; symmetrically, $W(B) \subset W(A) + (\epsilon)$. This proves the first assertion. (The proof is due to A. Brown.)

As for the second assertion, consider the unilateral shift U . The sequence $\{U^{*n}\}$ tends to 0 in the strong topology (more and more Fourier coefficients get lost as n increases), but $w(U^{*n}) = 1$ for all n .

Solution 176. If a is a complex number with $|a| \leq 1$ and if $|z| < 1$, then $\operatorname{Re}(1 - za) = 1 - \operatorname{Re}(za) \geq 1 - |z| > 0$. If conversely the complex number a is such that $\operatorname{Re}(1 - za) \geq 0$ for each z with $|z| < 1$, then this is true, in particular, if $za = t|a|$, $0 < t < 1$; since, therefore, $1 - t|a| = \operatorname{Re}(1 - t|a|) \geq 0$, it follows (let t tend to 1) that $|a| \leq 1$.

The operator fact corresponding to (and implied by) this numerical fact is that $w(A) \leq 1$ if and only if $\operatorname{Re}(1 - zA) \geq 0$. Indeed, the following assertions about A are pairwise equivalent:

$$w(A) \leq 1,$$

$$|(Af, f)| \leq 1 \quad \text{whenever} \quad \|f\| = 1,$$

$$(\operatorname{Re}(1 - zA)f, f) \geq 0 \quad \text{whenever} \quad \|f\| = 1 \text{ and } |z| < 1.$$

If $w(A) \leq 1$, then $r(A) \leq 1$, and therefore $1 - zA$ is invertible whenever $|z| < 1$. Since an invertible operator has positive real part if and only if its inverse has positive real part (if B is invertible, then $(B^{-1}f, f) = (B^{-1}f, BB^{-1}f) = (B(B^{-1}f), (B^{-1}f))^*$), it follows that $w(A) \leq 1$ if and only if $\operatorname{Re}(1 - zA)^{-1} \geq 0$ in the unit disc.

Observe next that if n is a positive integer and if ω is a primitive n -th root of unity (i.e., n is the smallest positive integer such that $\omega^n = 1$), then

$$\frac{1}{1 - z^n} = \frac{1}{n} \sum_{k=0}^{n-1} \frac{1}{1 - \omega^k z}$$

for all z other than the powers of ω . This identity is, in fact, the partial fraction expansion of the left side. For a direct verification, multiply through by $1 - z^n$, observe that the right side becomes a polynomial of degree $n - 1$ at most that is invariant under each of the n substitutions $z \rightarrow \omega^k z$ ($k = 0, \dots, n - 1$) and is therefore constant, and then evaluate the constant by setting z equal to 0.

The identity of the preceding paragraph implies that if $w(A) \leq 1$, then

$$(1 - z^n A^n)^{-1} = \frac{1}{n} \sum_{k=0}^{n-1} (1 - \omega^k z A)^{-1}$$

whenever $|z| < 1$. Since each summand on the right side has positive real part (because $w(\omega^k A) \leq 1$), it follows that the left side has positive real part, and that implies that $w(A^n) \leq 1$.

One step in the proof might be unfamiliar enough to deserve a second look. To prove an identity between operators by substitution into an identity between rational functions is to make use of the functional calculus for rational functions (cf. Problem 97). Explicitly: if φ_1 and φ_2 are rational functions whose poles are not in the spectrum of A , so that $\varphi_1(A)$ and $\varphi_2(A)$ make sense, then the same is true of each polynomial p in φ_1 and φ_2 ; if $\varphi(\lambda) = p(\varphi_1(\lambda), \varphi_2(\lambda))$, then $\varphi(A) = p(\varphi_1(A), \varphi_2(A))$. The proof is obvious.

The equivalence of $w(A) \leq 1$ and $\operatorname{Re}(1 - zA)^{-1} \geq 0$ for $|z| < 1$ is elementary, but basic for the argument; it was Berger's main new idea. That idea is visible in some form in all subsequent proofs. The proof given above is a simplification of a simplification discovered by Percy [1966].

Chapter 18. Unitary dilations

Solution 177. (a) As a heuristic guide to the proof, consider the very special case in which the given Hilbert space \mathbf{H} is one-dimensional real Euclidean space and the dilation space \mathbf{K} is the plane. In that case the given contraction A is a scalar α (with $|\alpha| \leq 1$), and, in geometric terms, the assertion is that multiplication (on the line) by α can be achieved by a suitable rotation (in the plane), followed by projection (back to the line). A picture makes all this crystal clear; simple analytic geometry shows that the matrix of the rotation is

$$\begin{pmatrix} \alpha & \sqrt{1 - \alpha^2} \\ \sqrt{1 - \alpha^2} & -\alpha \end{pmatrix}.$$

The proof itself is the most direct possible imitation of the technique that worked for the plane. A few experiments are needed, to see whether the role of α^2 should be played by A^2 , or AA^* , or A^*A , or sometimes one and sometimes another. The result can be described as follows. Given \mathbf{H} , write $\mathbf{K} = \mathbf{H} \oplus \mathbf{H}$ and identify \mathbf{H} with the first summand; then each operator on \mathbf{K} is a two-rowed matrix of operators on \mathbf{H} , and, in particular,

$$P = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

Given A , write

$$S = \sqrt{1 - AA^*} \quad \text{and} \quad T = \sqrt{1 - A^*A},$$

where the positive square roots are meant, of course; note that since $\|A\| \leq 1$, it follows that $1 - AA^*$ and $1 - A^*A$ are positive. The

desired dilation B can be defined by

$$B = \begin{pmatrix} A & S \\ T & -A^* \end{pmatrix}.$$

That B is a dilation of A is clear. Since

$$B^* = \begin{pmatrix} A^* & T \\ S & -A \end{pmatrix},$$

it follows by direct computation that

$$\begin{aligned} B^*B &= \begin{pmatrix} A^*A + T^2 & A^*S - TA^* \\ SA - AT & S^2 + AA^* \end{pmatrix}, \\ BB^* &= \begin{pmatrix} AA^* + S^2 & AT - SA \\ TA^* - A^*S & T^2 + A^*A \end{pmatrix}. \end{aligned}$$

It remains only to prove that $AT = SA$. Trivially $AT^2 = S^2A$, and it follows, by induction, that $AT^{2n} = S^{2n}A$ for $n = 0, 1, 2, \dots$. This implies that $Ap(T^2) = p(S^2)A$ for every polynomial p (cf. Solution 108), and hence that $AT = SA$, as desired.

(b) The proof is similar to that of (a), and simpler. Given A , with $0 \leq A \leq 1$, let R be the positive square root of $A(1 - A)$, and write

$$B = \begin{pmatrix} A & R \\ R & 1 - A \end{pmatrix}.$$

The verification that B is a projection is painless. (The result (b) is due to E. A. Michael; see Halmos [1950 a].)

Solution 178. The proof is constructive. Given \mathbf{H} , let \mathbf{K} be the direct sum of countably infinitely many copies of \mathbf{H} , indexed by all integers

(positive, negative, zero); then each operator on \mathbf{K} is an infinite operator matrix, and, in particular, the projection P from \mathbf{K} to \mathbf{H} is given by

$$P = \begin{pmatrix} \cdot & & & \cdot \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ & 0 & 0 & 0 \\ & 0 & (1) & 0 \\ & 0 & 0 & 0 \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ \cdot & & & \cdot \end{pmatrix}.$$

(The parentheses indicate the entry in position $\langle 0,0 \rangle$.) Given A , put

$$B = \begin{pmatrix} \cdot & & & & & & \cdot \\ \cdot & & & & & & \cdot \\ & 0 & 0 & 0 & 0 & 0 & 0 \\ & 1 & 0 & 0 & 0 & 0 & 0 \\ & 0 & 1 & 0 & 0 & 0 & 0 \\ & 0 & 0 & S & (A) & 0 & 0 & 0 \\ & 0 & 0 & -A^* & T & 0 & 0 & 0 \\ & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ \cdot & & & & & & \cdot \\ \cdot & & & & & & \cdot \\ \cdot & & & & & & \cdot \end{pmatrix},$$

where S and T are as in Solution 177. Since B is triangular, its powers are triangular, and the diagonal entries of the powers are the corresponding powers of the diagonal entries of B . This makes it obvious that

B is a power dilation of A . The proof that B is unitary is an obvious computation (which uses the results of Solution 177).

Although this may not be the most revealing proof of the theorem, it is certainly the shortest; it is due to Schäffer [1955].

Solution 179. The theorem can be proved directly, but the proof via unitary operators and dilation theory has an elegance that is hard to surpass. As for the theorem for unitary operators, it can be proved by relatively elementary and widely generalizable geometric methods (cf. Halmos [1958, p. 185]), but the parochial Hilbert space proof via the spectral theorem is more transparent.

If U is a unitary operator on \mathbf{H} , then the spectral theorem justifies the assumption that $\mathbf{H} = \mathbf{L}^2(\mu)$ for some measure μ , on some suitable measure space, in such a way that U is the multiplication induced by a measurable function φ of constant modulus 1 almost everywhere. If $f \in \mathbf{H} (= \mathbf{L}^2(\mu))$, then

$$\frac{1}{n} \sum_{j=0}^{n-1} U^j f = \left(\frac{1}{n} \sum_{j=0}^{n-1} \varphi^j \right) f.$$

Since $|\varphi| = 1$ almost everywhere, it follows that

$$\left| \frac{1}{n} \sum_{j=0}^{n-1} \varphi^j \right| \leq 1$$

almost everywhere. Since, moreover, the assumption that $|\varphi| = 1$ almost everywhere implies that the averages

$$\frac{1}{n} \sum_{j=0}^{n-1} \varphi^j$$

form a convergent sequence almost everywhere (whose limit is the characteristic function of the set where $\varphi = 1$), it follows that the Lebesgue dominated convergence theorem (not necessarily the bounded convergence theorem) is applicable to the sequence

$$\left\{ \left(\frac{1}{n} \sum_{j=0}^{n-1} \varphi^j \right) f \right\}.$$

This completes the proof of convergence; a quick second glance will even reveal what the limit is.

If A is an arbitrary contraction on \mathbf{H} , then let U be a unitary power dilation of it on a Hilbert space \mathbf{K} , say, and let P be the projection from \mathbf{K} to \mathbf{H} . This means that if $f \in \mathbf{H}$, then $A^n f = P U^n f$, $n = 0, 1, 2, \dots$, and it follows that

$$\frac{1}{n} \sum_{j=0}^{n-1} A^j f = P \left(\frac{1}{n} \sum_{j=0}^{n-1} U^j f \right).$$

Since

$$\frac{1}{n} \sum_{j=0}^{n-1} U^j f$$

has a limit as $n \rightarrow \infty$ for each f , and since P is continuous (i.e., bounded), it follows that

$$\frac{1}{n} \sum_{j=0}^{n-1} A^j f$$

has a limit as $n \rightarrow \infty$ for each f .

The mean ergodic theorem for unitary operators was first proved by von Neumann [1932]. The extension to contractions is due to Riesz-Nagy [1943]; the proof via dilation theory is due to Nagy [1955]. A good recent reference to ergodic theory in general is Jacobs [1960].

Solution 180. Relatively hard analytic proofs can be given; with dilation theory all becomes simple (Nagy [1955]). Given A on \mathbf{H} , let U on \mathbf{K} be a unitary power dilation of it, and let P be the projection from \mathbf{K} to \mathbf{H} . If p is a polynomial and if f is in \mathbf{H} , then

$$\begin{aligned} \|p(A)f\| &= \|Pp(U)f\| && \text{(by the definition of power dilation)} \\ &\leq \|p(U)\| \cdot \|f\| && \text{(because } \|P\| = 1) \\ &\leq \|p\|_{\mathcal{D}} \cdot \|f\| && \text{(because } U \text{ is normal),} \end{aligned}$$

and it follows, as stated, that $\|p(A)\| \leq \|p\|_{\mathcal{D}}$.

Solution 181. Let \mathbf{H}_0 be the vector space of all finitely non-zero sequences of vectors in \mathbf{H} (with coordinatewise vector operations), and let \mathbf{K}_0 be the vector space of all sequences of the form $\{u_n: n = 0, \pm 1, \pm 2, \dots\}$, where

$$u_j = \sum_i A_{i-j} f_i$$

for some sequence $\{f_n: n = 0, \pm 1, \pm 2, \dots\}$ in \mathbf{H}_0 (with, again, coordinatewise vector operations). If $u = \{u_n\}$ and $v = \{v_n\}$ are in \mathbf{K}_0 , with $u_j = \sum_i A_{i-j} f_i$ and $v_j = \sum_i A_{i-j} g_i$, write

$$[u, v] = \sum_j (u_j, g_j).$$

This definition appears to be ambiguous, because it appears to depend on the representation of v in terms of \mathbf{H}_0 . Since, however,

$$\begin{aligned} \sum_j (u_j, g_j) &= \sum_i \sum_j (A_{i-j} f_i, g_j) = \sum_i \sum_j (f_i, A_{j-i} g_j) \\ &= \sum_i (f_i, v_i), \end{aligned}$$

that dependence is illusory. The definition yields an inner product for \mathbf{K}_0 . The verifications of sesquilinearity and Hermitian symmetry are obvious, and positiveness follows from the assumed positive definiteness of $\{A_n\}$. Only strict positiveness needs a moment's pause, but that is easy too. By the Schwarz inequality for (not necessarily strictly positive) inner products,

$$|[u, v]|^2 \leq [u, v] \cdot [v, v].$$

It follows that if $[u, u] = 0$, then $[u, v] = 0$ for all v in \mathbf{K}_0 . Fix an index i , choose $\{g_n\}$ so that $g_i = u_i$ and $g_n = 0$ for $n \neq i$, infer that

$$\|u_i\|^2 = (u_i, g_i) = \sum_j (u_j, g_j) = [u, v] = 0,$$

and conclude that $u = 0$.

The vector space \mathbf{K}_0 with the inner product so obtained may fail to be complete; let \mathbf{K} be its completion. To each element f of \mathbf{H} there

corresponds a sequence $\{f_n\}$ in \mathbf{H}_0 defined by $f_0 = f$ and $f_n = 0$ for $n \neq 0$. To this sequence $\{f_n\}$, in turn, there corresponds a sequence $\{u_n\}$ in \mathbf{K}_0 ,

$$u_j = \sum_i A_{i-j} f_i = A_{-j} f.$$

If f and g are in \mathbf{H} , with corresponding sequences $u = \{u_n\}$ and $v = \{v_n\}$ in \mathbf{K}_0 , then the definitions imply that

$$[u, v] = (f, g).$$

It follows that the mapping $f \rightarrow u$ is an isometric embedding of \mathbf{H} into \mathbf{K} ; in the rest of the proof \mathbf{H} will be identified with its image in \mathbf{K} .

If P is the projection from \mathbf{K} to \mathbf{H} , if $u = \{u_n\}$ is an element of \mathbf{K}_0 , and if $g \in \mathbf{H}$, then

$$[Pu, g] = [u, g] = (u, g),$$

and therefore

$$Pu = u_0$$

for all such u 's.

If $u = \{u_n\}$ is in \mathbf{K}_0 , write $Uu = v = \{v_n\}$, where $v_n = u_{n-1}$. If $u_j = \sum_i A_{i-j} f_i$, with $\{f_n\}$ in \mathbf{H}_0 , then

$$v_j = u_{j-1} = \sum_i A_{i-(j-1)} f_i = \sum_i A_{i-j} f_{i-1}.$$

This implies that U is not only a one-to-one linear transformation, but one that maps \mathbf{K}_0 onto itself. If $\{f_n\}$ and $\{g_n\}$ are arbitrary sequences in \mathbf{H}_0 and if $u = \{u_n\}$ and $v = \{v_n\}$ are their correspondents in \mathbf{K}_0 , then

$$(Uu, Uv) = \sum_j (u_{j-1}, g_{j-1}) = \sum_j (u_j, g_j) = (u, v),$$

so that U is an isometry. It follows that U has a unique extension to a unitary operator on \mathbf{K} (which might as well be denoted by the same symbol as U). Since $PUnu = (U^nu)_0 = u_{-n} = \sum_i A_{i+n} f_i$, it follows that if $f \in \mathbf{H}$, then

$$PU^n f = A_n f,$$

and the proof is complete.

Chapter 19.

Commutators of operators

Solution 182. Wintners's proof. If $PQ - QP = \alpha$, replace P by $P + \lambda$, where λ is an arbitrary scalar, and observe that the new P satisfies the same commutation relation. There is, consequently, no loss of generality in assuming that P is invertible. Since, in that case, $QP = P^{-1}(PQ)P$, and therefore $\Lambda(QP) = \Lambda(PQ)$, the relation $PQ = QP + \alpha$ implies that

$$\Lambda(PQ) = \Lambda(QP + \alpha) = \Lambda(QP) + \alpha = \Lambda(PQ) + \alpha.$$

The only translation that can leave a non-empty compact subset (such as $\Lambda(PQ)$) of the complex plane invariant is the trivial translation (i.e., no translation at all); in other words, α must be 0.

Wielandt's proof. If $PQ - QP = \alpha$, then

$$\begin{aligned} P^2Q - QP^2 &= P^2Q - PQP + PQP - QP^2 \\ &= P(PQ - QP) + (PQ - QP)P = 2P\alpha, \end{aligned}$$

and more generally (induction)

$$P^nQ - QP^n = nP^{n-1}\alpha, \quad n = 1, 2, 3, \dots$$

If P is nilpotent, of index n , say, then $nP^{n-1}\alpha = 0$, and therefore $\alpha = 0$. If P is not nilpotent, then the inequality

$$n \|P^{n-1}\| \|\alpha\| \leq 2 \|P\| \|Q\| \|P^{n-1}\|,$$

true for $n = 1, 2, 3, \dots$, implies that

$$n \|\alpha\| \leq 2 \|P\| \|Q\|,$$

and hence that, again, $\alpha = 0$.

Solution 183. Given the Hilbert space \mathbf{H} , let \mathbf{B} be the normed vector space of all bounded sequences $f = \langle f_1, f_2, f_3, \dots \rangle$ of vectors in \mathbf{H} (coordinatewise vector operations, supremum norm), and let \mathbf{N} be the subspace of all null sequences in \mathbf{B} (i.e., sequences f with $\lim_n \|f_n\| = 0$). The quotient space $\hat{\mathbf{B}} = \mathbf{B}/\mathbf{N}$ is a normed vector space. Each bounded sequence $A = \langle A_1, A_2, A_3, \dots \rangle$ of operators on \mathbf{H} induces an operator on \mathbf{B} ; the image of $\langle f_1, f_2, f_3, \dots \rangle$ under $\langle A_1, A_2, A_3, \dots \rangle$ is $\langle A_1 f_1, A_2 f_2, A_3 f_3, \dots \rangle$. Since the subspace \mathbf{N} is invariant under each such induced operator, the sequence A also induces, in a natural manner, an operator on $\hat{\mathbf{B}}$; call it \hat{A} . Bounded sequences of operators on \mathbf{H} form a normed algebra (coordinatewise operations, supremum norm). The correspondence $A \rightarrow \hat{A}$ from such bounded sequences to operators on $\hat{\mathbf{B}}$ is a norm-decreasing homomorphism. If $P = \langle P_1, P_2, P_3, \dots \rangle$ and $Q = \langle Q_1, Q_2, Q_3, \dots \rangle$ are such that $P_n Q_n - Q_n P_n \rightarrow C$, then $\hat{P}\hat{Q} - \hat{Q}\hat{P}$ is a commutator on $\hat{\mathbf{B}}$; since that commutator cannot be equal to $\hat{1}$ (= the identity operator on $\hat{\mathbf{B}}$), the proof is complete.

Solution 184. Fix P and consider $C = \Delta Q = PQ - QP$ as a function of Q . The operation Δ is obviously a linear transformation on the vector space of operators; since

$$\|\Delta Q\| = \|PQ - QP\| \leq 2\|P\|\|Q\|,$$

that linear transformation is bounded (on the Banach space of operators), and

$$\|\Delta\| \leq 2\|P\|.$$

Mappings such as Δ often play an important algebraic role. The most important property of Δ is that it is a *derivation* in the sense that

$$\Delta(QR) = \Delta Q \cdot R + Q \cdot \Delta R.$$

Proof: $PQR - QRP = (PQR - QPR) + (QPR - QRP)$.

Derivations have many of the algebraic properties of differentiation, but, as is visible in the definition itself, they have them in a non-commutative way. First among those properties is the validity of the Leibniz formula for "differentiating" products with several factors. The assertion is that $\Delta(Q_1 \cdots Q_n)$ is the sum of n terms; to obtain the j -th term, replace Q_j by ΔQ_j in the product $Q_1 \cdots Q_n$. The proof is an obvious induction. For $n = 1$, there is nothing to do; for the step from n to

$n + 1$, write $Q_1 \cdots Q_{n+1}$ as $(Q_1 \cdots Q_n)Q_{n+1}$ and use the given (two-factor) product formula. The result is, of course, applicable to the special case in which all the Q_i 's coincide, but it does not become much more pleasant to contemplate.

A special property of the derivation Δ is that $\Delta^2 Q = 0$. Here Δ^2 is, of course, the composition of Δ with itself, so that

$$\Delta^2 Q = \Delta(\Delta Q) = P \cdot \Delta Q - \Delta Q \cdot P;$$

the vanishing of $\Delta^2 Q$ expresses exactly that ΔQ commutes with P .

The Leibniz formula and the vanishing of $\Delta^2 Q$ make it easy to evaluate higher order derivatives of higher powers of Q . The process begins with ΔQ^n : it is equal to the sum of the n possible products each of which has one factor equal to ΔQ and $n - 1$ factors equal to Q . When Δ is applied to one of these summands, the result is the sum of only $n - 1$ products. (Reason: when Δ is applied to ΔQ , the result is 0.) Each of the $n - 1$ products so obtained has two factors equal to ΔQ and $n - 2$ factors equal to Q . Consequence: $\Delta^2 Q^n$ is equal to the sum of the $n(n - 1)$ possible products of that kind. The argument continues from here on with no surprises and yields a description of $\Delta^k Q^n$. With $k = n$, the result is that $\Delta^n Q^n$ is the sum of $n!$ terms, each of which is $(\Delta Q)^n$; in other words

$$\Delta^n Q^n = n!(\Delta Q)^n.$$

The last equation is the crucial point of the proof; the desired result is a trivial consequence of it. Indeed, since

$$\|(\Delta Q)^n\| = \frac{1}{n!} \|\Delta^n Q^n\| \leq \frac{1}{n!} \|\Delta^n\| \cdot \|Q^n\| \leq \frac{1}{n!} \|\Delta\|^n \cdot \|Q\|^n,$$

it follows that

$$\|(\Delta Q)^n\|^{1/n} \leq \left(\frac{1}{n!}\right)^{1/n} \cdot \|\Delta\| \cdot \|Q\|,$$

and hence that ΔQ is quasinilpotent.

As a dividend, the equation for $\Delta^n Q^n$ yields a proof of Jacobson's original algebraic result. Statement: if an element Q of an algebra over a field of characteristic greater than $n!$ satisfies a polynomial equation of degree n , and if Δ is a derivation of that algebra such that $\Delta^2 Q = 0$,

then ΔQ is nilpotent of index n . Proof: from $\Delta(\Delta Q) = 0$ infer that $\Delta(\Delta Q)^k = 0$ for every positive integer k , and hence, from the equation for $\Delta^n Q^n$ infer that $\Delta^{n+1} Q^n = 0$. Consequence: $\Delta^n Q^i = 0$ whenever $n > i$. If $Q^n = \sum_{i=0}^{n-1} \alpha_i Q^i$ is the polynomial equation satisfied by Q , then it follows that $\Delta^n Q^n = 0$, and hence it follows, again from the equation for $\Delta^n Q^n$, that $n!(\Delta Q)^n = 0$. The conclusion follows from the assumption about the characteristic.

Solution 185. (a) The trick is to generalize the formula for the “derivative” of a power to the non-commutative case; cf. Solutions 182 and 184. The generalization that is notationally most convenient here says that

$$P^n Q - Q P^n = \sum_{i=0}^{n-1} P^{n-i-1} C P^i;$$

the proof is a straightforward induction. It follows that

$$P^n Q - Q P^n = n P^{n-1} - \sum_{i=0}^{n-1} P^{n-i-1} (1 - C) P^i,$$

and hence that

$$\begin{aligned} n \| P^{n-1} \| &\leq 2 \| P \| \cdot \| Q \| \cdot \| P^{n-1} \| \\ &\quad + \| 1 - C \| \cdot \sum_{i=0}^{n-1} \| P^{n-i-1} \| \cdot \| P^i \|. \end{aligned}$$

Up to now P could have been arbitrary. Since P was assumed hyponormal, the last written sum is equal to $n \| P^{n-1} \|$ (see Problem 162). Divide through by $n \| P^{n-1} \|$. (If $P = 0$, everything is trivial, and if $P \neq 0$, then $P^{n-1} \neq 0$.) The result is that

$$1 \leq \frac{2}{n} \| P \| \cdot \| Q \| + \| 1 - C \|,$$

and the conclusion follows.

(b) If $\| 1 - C \| < 1$, then C is invertible; since, by the Kleinecke-Shirokov theorem, C is quasinilpotent, that is impossible.

Solution 186. If A has a large kernel, then that kernel is the direct sum of \aleph_0 subspaces all of the same dimension. The orthogonal comple-

ment of the kernel may or may not be large. If, however, one of the direct summands of the kernel is adjoined to that orthogonal complement, the result is a representation of the underlying Hilbert space in the form of an infinite direct sum $\mathbf{H} \oplus \mathbf{H} \oplus \mathbf{H} \oplus \cdots$ in such a way that the direct sum of all the summands beginning with the second one is annihilated by A . If corresponding to this representation of the space the operator A is represented as a matrix, it will have the form

$$A = \begin{pmatrix} A_0 & 0 & 0 & 0 \\ A_1 & 0 & 0 & 0 \\ A_2 & 0 & 0 & 0 \\ A_3 & 0 & 0 & 0 \\ & & & \cdot \\ & & & \cdot \\ & & & \cdot \end{pmatrix}$$

where each A_n (and each 0) is an operator on \mathbf{H} . Write

$$P = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ & & & \cdot \\ & & & \cdot \\ & & & \cdot \end{pmatrix}$$

and

$$Q = \begin{pmatrix} A_1 & -A_0 & 0 & 0 \\ A_2 & 0 & -A_0 & 0 \\ A_3 & 0 & 0 & -A_0 \\ A_4 & 0 & 0 & 0 \\ & & & \cdot \\ & & & \cdot \\ & & & \cdot \end{pmatrix};$$

then P and Q are operators and (straightforward computation) $PQ - QP = A$. The proof of the main assertion is complete.

To prove Corollary 1, suppose that $\{f_1, \dots, f_n\}$ is a finite set of vectors in an infinite-dimensional Hilbert space \mathbf{H} , and let \mathbf{M} be their span. If A is an operator on \mathbf{H} , let C be the operator that is A on \mathbf{M} and 0 on \mathbf{M}^\perp ; by what was just proved C is a commutator, and C agrees with A on each f_i , $i = 1, \dots, n$. This implies that every basic strong neighborhood of A contains commutators.

The proof of Corollary 2 is similar. Given \mathbf{H} , let \mathbf{M} be a large subspace with a large orthogonal complement; given A , define C as in the preceding paragraph, and write $B = A - C$. Since $A = B + C$ and both B and C are commutators, the proof is complete.

Solution 187. Since A is not a scalar, there exists a vector f such that f and Af are linearly independent. Let T be an invertible operator such that $Tf = f$ and $TAf = -Af$. Since this implies that $(A + T^{-1}AT)f = Af - Af = 0$, it follows that the direct sum

$$S = (A + T^{-1}AT) \oplus (A + T^{-1}AT) \oplus (A + T^{-1}AT) \oplus \dots$$

has a large kernel. (What really follows is that the kernel is infinite-dimensional; since the whole space is separable, this implies that the kernel is large.) By Problem 186, the direct sum S is a commutator. If

$$B = A \oplus A \oplus A \oplus \dots$$

and

$$C = T^{-1}AT \oplus T^{-1}AT \oplus T^{-1}AT \oplus \dots,$$

then

$$S = B + C.$$

The next step is the following somewhat surprising lemma: whenever B and C are operators such that $B + C$ is a commutator, then $B \oplus C$ is a commutator. The proof is an inspired bit of elementary algebra. If $B + C = PQ - QP$, then write $R = C + QP = PQ - B$, and

compute the commutator of

$$\begin{pmatrix} 0 & P \\ 1 & 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 0 & R \\ Q & 0 \end{pmatrix};$$

$$\begin{pmatrix} 0 & P \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & R \\ Q & 0 \end{pmatrix} - \begin{pmatrix} 0 & R \\ Q & 0 \end{pmatrix} \begin{pmatrix} 0 & P \\ 1 & 0 \end{pmatrix}$$

$$= \begin{pmatrix} PQ & 0 \\ 0 & R \end{pmatrix} - \begin{pmatrix} R & 0 \\ 0 & QP \end{pmatrix} = \begin{pmatrix} B & 0 \\ 0 & C \end{pmatrix}.$$

Consequence: $B \oplus C$ is a commutator. Since, however, C is clearly similar to B , it follows that $B \oplus B$ is a commutator; since, finally, $B \oplus B$ is unitarily equivalent to B , the proof is complete.

Solution 188. Suppose that $C = A^*A - AA^* \geq 0$. The problem is to show that $0 \in \Lambda(C)$. It is sufficient to show that there exists a sequence $\{f_n\}$ of unit vectors such that $Cf_n \rightarrow 0$ (i.e., that $0 \in \Pi(C)$). For this purpose, take a complex number λ in the approximate point spectrum of A , and, corresponding to λ , find a sequence $\{f_n\}$ of unit vectors such that $(A - \lambda)f_n \rightarrow 0$. Since the self-commutator of $A - \lambda$ is equal to C , and since $C \geq 0$, it follows that

$$(A - \lambda)^*(A - \lambda) \geq (A - \lambda)(A - \lambda)^*.$$

Since $(A - \lambda)f_n \rightarrow 0$, it follows that $(A - \lambda)^*(A - \lambda)f_n \rightarrow 0$. The last two remarks imply that $(A - \lambda)(A - \lambda)^*f_n \rightarrow 0$. Since, therefore, C is the difference of two operators, each of which annihilates $\{f_n\}$, the operator C does so too.

Solution 189. (a) The program is to show that (1) A is quasinormal, (2) $\ker(1 - A^*A)$ reduces A , and (3) the orthogonal complement of $\ker(1 - A^*A)$ is included in $\ker(A^*A - AA^*)$.

(1) Write $P = A^*A - AA^*$. Since, for all f ,

$$\begin{aligned} \|f\|^2 &\geq \|Af\|^2 = (A^*Af, f) = (AA^*f, f) + (Pf, f) \\ &= \|A^*f\|^2 + \|Pf\|^2, \end{aligned}$$

it follows that if $Pf = f$, then $A^*P = 0$. (The norm condition was used at the first step.) This implies that $A^*P = 0$, and hence that $PA = 0$, or, equivalently, that $(A^*A)A = A(A^*A)$.

(2) Write $\mathbf{M} = \ker(1 - A^*A)$. If $f \in \mathbf{M}$, then $f - A^*Af = 0$. It follows that $Af - (A^*A)Af = Af - A(A^*A)f = A(f - A^*Af) = 0$, so that \mathbf{M} is invariant under A . Similarly (instead of replacing f by Af , replace f by A^*f) \mathbf{M} is invariant under A^* . (Cf. Solution 154.)

(3) Since P is idempotent, it follows that

$$A^*A - AA^* = A^*AA^*A - AA^*A^*A - A^*AAA^* + AA^*AA^*.$$

Since A^*A commutes with both A and A^* , this can be rewritten as

$$A^*A - AA^* = A^*A(A^*A - AA^*).$$

In other words,

$$P = A^*AP,$$

or

$$(1 - A^*A)P = 0.$$

It follows that

$$\text{ran } P \subset \mathbf{M},$$

or

$$\mathbf{M}^\perp \subset \ker P.$$

Now use the assumption that A is abnormal: it says exactly that $\ker P$ includes no non-zero subspace that reduces A . Conclusion: $\mathbf{M}^\perp = \{0\}$, and this means that A is an isometry.

(b) If A is the bilateral shift with weights $\{\alpha_n\}$ such that $\alpha_n = 1$ or $\sqrt{2}$ according as $n \leq 0$ or $n > 0$, then A is abnormal and the self-commutator of A is a projection.

Proof. The self-commutator of A is easy to compute; it turns out to be the projection of rank 1 whose range is spanned by e_1 . The abnormality of A follows from Problem 129: according to that result, A is irreducible, and hence as abnormal as can be.

Solution 190. An infinite-dimensional Hilbert space is the direct sum of Hilbert spaces of dimension \aleph_0 . To prove that every scalar of modulus 1 is a commutator, it is therefore sufficient to prove that on a Hilbert space of dimension \aleph_0 every scalar of modulus 1 is the commutator of

two *unitary* operators. (The unitary character of the factors guarantees that the possibly uncountable direct sum is bounded.) In a Hilbert space of dimension \aleph_0 there always exists an orthonormal basis $\{e_n: n = 0, \pm 1, \pm 2, \dots\}$. Given α with $|\alpha| = 1$, let P be the diagonal operator defined by $Pe_n = \alpha^n e_n$ and let Q be the bilateral shift, $Qe_n = e_{n+1}$. Both P and Q are unitary; a straightforward computation shows that $PQP^{-1}Q^{-1} = \alpha$.

The proof that if $\alpha = PQP^{-1}Q^{-1}$, then $|\alpha| = 1$ is an adaptation of the Wintner argument (Solution 182). Since $PQ = \alpha QP$, it follows that $\Lambda(PQ) = \alpha\Lambda(QP)$; since, however, PQ is similar to QP , it follows that $\Lambda(PQ) = \Lambda(QP)$, and hence that $\Lambda(QP) = \alpha\Lambda(QP)$. Since $\Lambda(QP)$ is a non-empty compact set different from $\{0\}$ (remember that QP is invertible), and since the only homothety that can leave such a set fixed is a rotation, the result follows.

Solution 191. The proof is an adaptation of Solution 190. The first step is to use Problem 111 to represent the given space as the direct sum of \aleph_0 subspaces, all of the same dimension, each of which reduces the given unitary operator U . The direct sum decomposition serves to represent U as a diagonal operator matrix whose n -th diagonal entry is U_n , say, for $n = 0, \pm 1, \pm 2, \dots$.

Solution 190 suggests that the multiplicative commutator of a diagonal operator and a bilateral shift may work here too. To avoid writing down large matrices, it is convenient to introduce some more notation. Think of the given Hilbert space as the set of all sequences $f = \{f_n: n = 0, \pm 1, \pm 2, \dots\}$ of vectors in some fixed Hilbert space (subject of course to the usual condition $\sum_n \|f_n\|^2 < \infty$). A typical diagonal operator matrix P is defined by

$$(Pf)_n = V_n f_n,$$

and the bilateral shift Q is defined by

$$(Qf)_n = f_{n-1}.$$

The commutator is easy to compute; the result is that

$$(PQP^{-1}Q^{-1}f)_n = V_n V_{n-1}^{-1} f_n.$$

The equations

$$U_n = V_n V_{n-1}^{-1}$$

can be solved for the V 's in terms of the U 's. If, for instance, V_0 is set equal to 1, then

$$V_n = U_n \cdots U_1 \quad \text{for } n \geq 1,$$

and

$$V_{-(n+1)} = U_{-n}^{-1} \cdots U_0^{-1} \quad \text{for } n \geq 0.$$

The unitary character of the U 's implies that the transformation P given by these V 's is a unitary operator, and all is well.

Solution 192. *On an infinite-dimensional Hilbert space, the commutator subgroup of the full linear group is the full linear group itself.*

Proof. The assertion is that every invertible operator is the product of a finite (but not necessarily bounded) number of multiplicative commutators. The fact is that every invertible operator is the product of two commutators (Brown-Pearcy [1966]). The proof of that fact takes more work than the present purpose is worth; it is sufficient to prove that every invertible operator is the product of three commutators, and that is much easier.

Given an arbitrary invertible operator A on an arbitrary infinite-dimensional Hilbert space, consider the infinite operator matrices

$$P = \begin{pmatrix} \cdot & & & & & & & \cdot \\ & \cdot & & & & & & \cdot \\ & & \cdot & & & & & \cdot \\ & & & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ & & & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ & & & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ & & & 0 & 0 & 0 & 0 & (0) & A & 0 & 0 \\ & & & 0 & 0 & 0 & 0 & 0 & 0 & A & 0 \\ & & & 0 & 0 & 0 & 0 & 0 & 0 & 0 & A \\ & & & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ & \cdot & & & & & & & & & \cdot \\ & \cdot & & & & & & & & & \cdot \\ & \cdot & & & & & & & & & \cdot \end{pmatrix}$$

and

$$Q = \begin{pmatrix} \cdot & & & & & & & \cdot \\ \cdot & & & & & & & \cdot \\ & \cdot & & & & & & \cdot \\ & & 0 & 0 & 0 & 0 & 0 & 0 \\ & & 1 & 0 & 0 & 0 & 0 & 0 \\ & & 0 & 1 & 0 & 0 & 0 & 0 \\ & & 0 & 0 & 1 & (0) & 0 & 0 \\ & & 0 & 0 & 0 & 1 & 0 & 0 \\ & & 0 & 0 & 0 & 0 & 1 & 0 \\ & & 0 & 0 & 0 & 0 & 0 & 1 \\ \cdot & & & & & & & \cdot \\ \cdot & & & & & & & \cdot \\ \cdot & & & & & & & \cdot \end{pmatrix}.$$

Routine computation proves that

$$PQP^{-1}Q^{-1} = \begin{pmatrix} \cdot & & & & & & & \cdot \\ \cdot & & & & & & & \cdot \\ & \cdot & & & & & & \cdot \\ & & 1 & 0 & 0 & 0 & 0 & 0 \\ & & 0 & 1 & 0 & 0 & 0 & 0 \\ & & 0 & 0 & 1 & 0 & 0 & 0 \\ & & 0 & 0 & 0 & (A) & 0 & 0 \\ & & 0 & 0 & 0 & 0 & 1 & 0 \\ & & 0 & 0 & 0 & 0 & 0 & 1 \\ & & 0 & 0 & 0 & 0 & 0 & 0 \\ \cdot & & & & & & & \cdot \\ \cdot & & & & & & & \cdot \\ \cdot & & & & & & & \cdot \end{pmatrix}.$$

Regard the direct sum on which these matrices act as the direct sum of the summand with index 0 and the others, and identify the direct sum of the others with one of them. With an obvious change of notation, the result of the above computations becomes this: every operator matrix of either of the forms

$$\begin{pmatrix} A & 0 \\ 0 & 1 \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} 1 & 0 \\ 0 & A \end{pmatrix}$$

is a multiplicative commutator (provided that the matrix entries operate on an infinite-dimensional space, and that A is invertible).

Every invertible normal operator on an infinite-dimensional Hilbert space has large reducing subspaces with large orthogonal complements (Problem 111), and is therefore representable as the product of two matrices of the indicated forms. Consequence: every invertible normal operator is the product of two commutators. Since every invertible operator is the product of a unitary operator and an invertible positive operator (polar decomposition), it follows, as stated, that every invertible operator is the product of three multiplicative commutators. (Apply Problem 191 to dispose of the unitary factor.)

Chapter 20. Toeplitz operators

Solution 193. If $\varphi = \sum_n \alpha_n e_n$, then the matrix entries of L_φ are given by

$$\begin{aligned}\lambda_{ij} &= (L_\varphi e_j, e_i) = \left(\sum_n \alpha_n e_{n+j}, e_i \right) \\ &= \sum_n \alpha_n \delta_{n+j, i} = \alpha_{i-j}.\end{aligned}$$

If, conversely, A is an operator on \mathbf{L}^2 such that

$$(Ae_{j+1}, e_{i+1}) = (Ae_j, e_i)$$

for all i and j , and if W is the bilateral shift (multiplication by e_1), then

$$\begin{aligned}(AWe_j, e_i) &= (Ae_{j+1}, e_i) = (Ae_j, e_{i-1}) \\ &= (Ae_j, W^*e_i) = (W Ae_j, e_i).\end{aligned}$$

This implies that A commutes with W , and hence (Problem 115) that A is a multiplication.

Solution 194. The proof of necessity is a simple computation: if $i, j = 0, 1, 2, \dots$, then

$$\begin{aligned}(T_\varphi e_j, e_i) &= (PL_\varphi e_j, e_i) = (L_\varphi e_j, e_i) = (L_\varphi e_{j+1}, e_{i+1}) \\ &= (PL_\varphi e_{j+1}, e_{i+1}) = (T_\varphi e_{j+1}, e_{i+1}).\end{aligned}$$

To prove sufficiency, assume that A is an operator on \mathbf{H}^2 such that

$$(Ae_{j+1}, e_{i+1}) = (Ae_j, e_i) \quad (i, j = 0, 1, 2, \dots);$$

it is to be proved that A is a Toeplitz operator. Consider for each non-negative integer n the operator on \mathbf{L}^2 given by

$$A_n = W^{*n} A P W^n$$

(where W is, as before, the bilateral shift). If $i, j \geq 0$, then

$$(A_n e_j, e_i) = (A_0 e_{j+n}, e_{i+n}) = (A e_j, e_i).$$

Something like this is true even for negative indices. Indeed, for n sufficiently large both $j + n$ and $i + n$ are positive, and from then on $(A_0 e_{j+n}, e_{i+n})$ is independent of n . Consequence: if p and q are trigonometric polynomials (finite linear combinations of the e_i 's, $i = 0, \pm 1, \pm 2, \dots$), then the sequence $\{(A_n p, q)\}$ is convergent. Since

$$\|A_n\| \leq \|A_0\| = \|A\|,$$

it follows on easy general grounds that the sequence $\{A_n\}$ of operators on L^2 is weakly convergent to an operator A_∞ on L^2 .

Since, for all i and j ,

$$\begin{aligned} (A_\infty e_j, e_i) &= \lim_n (W^{*n} A P W^n e_j, e_i) \\ &= \lim_n (W^{*n+1} A P W^{n+1} e_j, e_i) \\ &= \lim_n (W^{*n} A P W^n e_{j+1}, e_{i+1}) \\ &= (A_\infty e_{j+1}, e_{i+1}), \end{aligned}$$

it follows that the operator A_∞ has a Laurent matrix and hence that it is a Laurent operator (Problem 193). If f and g are in \mathbf{H}^2 , then

$$(P A_\infty f, g) = (A_\infty f, g) = \lim_n (W^{*n} A P W^n f, g) = (A f, g),$$

so that $P A_\infty f = A f$ for each f in \mathbf{H}^2 . Conclusion: A is the compression to \mathbf{H}^2 of a Laurent operator, and hence, by definition, A is a Toeplitz operator.

How can the function φ that induces A be recaptured from the matrix of A ? If $A = T_\varphi$, then $A_\infty = L_\varphi$, and therefore the Fourier coefficients of φ are the entries in the 0 column of the matrix of A_∞ . This is an answer, but not a satisfying one; it is natural to wish for an answer

expressed in terms of A instead of A_∞ . That turns out to be easy. If $i, j \geq 0$, then

$$(Ae_j, e_i) = (A_\infty e_j, e_i) = (\varphi, e_{i-j});$$

this implies that

$$(\varphi, e_i) = (Ae_0, e_i) \quad \text{for } i \geq 0,$$

and

$$(\varphi, e_{-j}) = (Ae_j, e_0) \quad \text{for } j \geq 0.$$

Conclusion: φ is the function whose forward Fourier coefficients (the ones with positive index) are the terms of the 0 column of the matrix of A and whose backward Fourier coefficients are the terms of the 0 row of that matrix.

To prove Corollary 1, observe that

$$(Ae_{j+1}, e_{i+1}) = (AUe_j, Ue_i) = (U^*AUe_j, e_i).$$

To prove Corollary 2, observe that if φ is a bounded measurable function, and if both n and $n + k$ are non-negative integers, then

$$(\varphi, e_k) = (T_\varphi e_n, e_{n+k}).$$

If T_φ is compact, then $\|T_\varphi e_n\| \rightarrow 0$ (since $e_n \rightarrow 0$ weakly); it follows that $(\varphi, e_k) = 0$ for all k (positive, negative, or zero), and hence that $\varphi = 0$.

Solution 195. Write $C = T_\varphi T_\psi$ and let $\langle \gamma_{ij} \rangle$ be the (not necessarily Toeplitz) matrix of C . If the Fourier expansions of φ and ψ are $\varphi = \sum_i \alpha_i e_i$ and $\psi = \sum_j \beta_j e_j$, so that the matrices of T_φ and T_ψ are $\langle \alpha_{i-j} \rangle$ and $\langle \beta_{i-j} \rangle$, respectively, then

$$\gamma_{i+1, j+1} = \gamma_{ij} + \alpha_{i+1} \beta_{-j-1}$$

whenever $i, j \geq 0$. The proof is straightforward. Since

$$\gamma_{ij} = \sum_{k=0}^{\infty} \alpha_{i-k} \beta_{k-j},$$

it follows that

$$\begin{aligned}
 \gamma_{i+1,j+1} &= \sum_{k=0}^{\infty} \alpha_{i+1-k} \beta_{k-j-1} \\
 &= \alpha_{i+1} \beta_{-j-1} + \sum_{k=1}^{\infty} \alpha_{i+1-k} \beta_{k-j-1} \\
 &= \alpha_{i+1} \beta_{-j-1} + \sum_{k=0}^{\infty} \alpha_{i-k} \beta_{k-j} \\
 &= \alpha_{i+1} \beta_{-j-1} + \gamma_{ij}.
 \end{aligned}$$

If now ψ is analytic, then

$$T_{\varphi} T_{\psi} f = T_{\varphi}(\psi \cdot f) = P(\varphi \cdot \psi \cdot f) = T_{\varphi \psi} f$$

for all f in \mathbf{H}^2 , so that $T_{\varphi} T_{\psi} = T_{\varphi \psi}$; if φ^* is analytic, then

$$T_{\varphi} T_{\psi} = (T_{\psi^*} T_{\varphi^*})^* = T_{\varphi \psi}.$$

This proves the sufficiency of the condition and the last assertion of the problem. If, conversely, the product $T_{\varphi} T_{\psi}$ is a Toeplitz operator, then its matrix is a Toeplitz matrix (Problem 194); the equation for $\gamma_{i+1,j+1}$ then implies that $\alpha_{i+1} \beta_{-j-1} = 0$ whenever $i, j \geq 0$. From this, in turn, it follows that either $\alpha_{i+1} = 0$ for all $i \geq 0$ or else $\beta_{-j-1} = 0$ for all $j \geq 0$, which is equivalent to the desired conclusion.

As for the corollary, sufficiency is trivial. If, conversely, $T_{\varphi} T_{\psi} = 0$, then, since 0 is a Toeplitz operator, it follows from Problem 195 that either φ^* or ψ is analytic and that $\varphi \psi = 0$. The F. and M. Riesz theorem applies (Problem 127) and proves that if φ^* is analytic, then $\psi = 0$, and if ψ is analytic, then $\varphi = 0$.

Solution 196. It is helpful to begin with some qualitative reflections. Consider a Laurent matrix written, as usual, so that the row index increases (from $-\infty$ to $+\infty$) as the rows go down, and the column index increases (from $-\infty$ to $+\infty$) as the columns go to the right.

Fix attention on any particular entry on the main diagonal, and look at the unilaterally infinite matrix that starts there and goes down and to the right. All the matrices obtained in this way from one fixed Laurent matrix look the same; they all look like the associated Toeplitz matrix. Intuition suggests that as the selected diagonal entry moves up and left, the resulting Toeplitz matrices swell and tend to the original Laurent matrix.

An efficient non-matrix way of describing the situation might go like this. If P_n is the projection onto the span of $\{e_{-n}, \dots, e_{-1}, e_0, e_1, e_2, \dots\}$, $n = 1, 2, 3, \dots$, then each Laurent operator L is the strong limit of the Toeplitz-like operators $P_n L P_n$. Since $P_n = W^{*n} P W^n$, and since W commutes with L (so that $W L W^* = L$), it follows that $W^{*n} P L P W^n \rightarrow L$ (strongly) as $n \rightarrow \infty$. This implies that if T is the Toeplitz operator corresponding to L , then $W^{*n} T P W^n \rightarrow L$ (strongly). It is instructive to compare this result with Solution 194 where weak convergence was enough.

The ground is now prepared for the proof of the spectral inclusion theorem for Toeplitz operators. It is sufficient to prove that if 0 is an approximate eigenvalue of L , then it is an approximate eigenvalue of T also; the non-zero values are recaptured by an obvious translation argument. Suppose therefore that to each positive number ε there corresponds a unit vector f_ε such that $\|L f_\varepsilon\| < \varepsilon$. The preceding paragraph implies that $W^{*n} P W^n f_\varepsilon \rightarrow f_\varepsilon$ and $W^{*n} T P W^n f_\varepsilon \rightarrow L f_\varepsilon$ (strongly). It follows that $\|P W^n f_\varepsilon\| \rightarrow 1$ and $\|T P W^n f_\varepsilon\| \rightarrow 0$. The first of these assertions says that $P W^n f_\varepsilon$ is, for large n , nearly a unit vector; the second one says that T nearly annihilates it. It follows, as promised, that 0 is an approximate eigenvalue of T .

Corollary 1 is now straightforward. Since L is normal, $\|L\| = r(L)$, and, by the result just proved, $r(L) \leq r(T)$. It follows that $\|L\| \leq \|T\|$. The reverse inequality was proved before, and the corollary follows from the known facts about the norm of a multiplication.

For Corollary 2: if the spectrum of T_φ consists of 0 alone, then the same is true of L_φ , and it follows that $\varphi = 0$.

The proof of Corollary 3 is similar to that of Corollary 2: if the spectrum of T_φ is real, then the same is true of L_φ , and it follows that φ is real.

The proof of Corollary 4 is the same as Solution 172: $\overline{W(L)} = \text{conv } \Lambda(L) \subset \text{conv } \Lambda(T) \subset \overline{W(T)} \subset \overline{W(L)}$.

Solution 197. It is useful to remember that $\tilde{\mathbf{H}}^2$ is a functional Hilbert space, and, as such, it has a kernel function (Problem 30); it is not, however, important to know what that kernel function is. Let \tilde{T}_φ be what T_φ becomes when it is transferred from \mathbf{H}^2 to $\tilde{\mathbf{H}}^2$; it follows from Solution 34 that $\tilde{T}_\varphi \tilde{f} = \tilde{\varphi} \cdot \tilde{f}$ for each \tilde{f} in $\tilde{\mathbf{H}}^2$. If y is a complex (!) number, with $|y| < 1$, then $\tilde{f}(y) = (\tilde{f}, K_y)$; this implies that $\tilde{f}(y) = 0$ if and only if $\tilde{f} \perp K_y$. Fix y , put $\lambda = \tilde{\varphi}(y)$, temporarily fix an element \tilde{f} in $\tilde{\mathbf{H}}^2$, and let \tilde{g} be the function defined by $\tilde{g}(z) = (\tilde{\varphi}(z) - \lambda)\tilde{f}(z)$. Since $\tilde{g}(y) = (\tilde{\varphi}(y) - \lambda)\tilde{f}(y) = 0$, it follows that $\tilde{g} \perp K_y$. This implies that $(\tilde{T}_\varphi - \lambda)\tilde{\mathbf{H}}^2$ is included in the orthogonal complement of K_y , so that it is a proper subspace of $\tilde{\mathbf{H}}^2$, and hence that λ belongs to the (compression) spectrum of T_φ . Conclusion: $\tilde{\varphi}(D) \subset \Lambda(T_\varphi)$, and therefore $\overline{\tilde{\varphi}(D)} \subset \Lambda(T_\varphi)$.

The converse is even easier. If $|\tilde{\varphi}(z) - \lambda| \geq \delta > 0$ whenever $|z| < 1$, then $1/(\tilde{\varphi} - \lambda)$ is a bounded analytic function in the open unit disc. It follows that its product with a function analytic in the disc and having a square-summable set of Taylor coefficients is another function with the same properties, i.e., that it induces a bounded multiplication operator on $\tilde{\mathbf{H}}^2$. Conclusion: $T_\varphi - \lambda$ is invertible, i.e., λ is not in $\Lambda(T_\varphi)$.

Solution 198. *A Hermitian Toeplitz operator that is not a scalar has no eigenvalues.*

Proof. It is sufficient to show that if φ is a real-valued bounded measurable function, and if $T_\varphi \cdot f = 0$ for some f in \mathbf{H}^2 , then either $f = 0$ or $\varphi = 0$. Since $\varphi \cdot f^* = \varphi^* \cdot f^* \in \mathbf{H}^2$ (because $P(\varphi \cdot f) = 0$), and since $f \in \mathbf{H}^2$, it follows that $\varphi \cdot f^* \cdot f \in \mathbf{H}^1$ (Problem 27). Since, however, $\varphi \cdot f^* \cdot f$ is real, it follows that $\varphi \cdot f^* \cdot f$ is a constant (Solution 26). Since $\int \varphi \cdot f^* \cdot f d\mu = (\varphi \cdot f, f) = (T_\varphi f, f) = 0$ (because $T_\varphi f = 0$), the constant must be 0. The F. and M. Riesz theorem (Problem 127) implies that either $f = 0$ or $\varphi \cdot f^* = 0$. If $f \neq 0$, then f^* can vanish on a set of measure 0 only, and therefore $\varphi = 0$.

Solution 199. *If φ is a real-valued bounded measurable function, and if its essential lower and upper bounds are α and β , then $\Lambda(T_\varphi)$ is the closed interval $[\alpha, \beta]$.*

Proof. If $\alpha = \beta$, then φ is constant, and everything is trivial. If $\alpha < \lambda < \beta$, it is to be proved that $T_\varphi - \lambda$ is not invertible. Assume the contrary, i.e., assume that $T_\varphi - \lambda$ is invertible, and, by an inessential change of notation, assume $\lambda = 0$. It follows, as an apparently very small consequence of invertibility, that e_0 belongs to the range of T_φ , and hence that there exists a (non-zero) function f in \mathbf{H}^2 such that $T_\varphi f = e_0$. This means that $\varphi \cdot f - e_0 \perp \mathbf{H}^2$. Equivalently (recall that $e_0(z) = 1$ for all z) the complex conjugate of $\varphi \cdot f$ is in \mathbf{H}^2 ; the next step is to deduce from this that $\text{sgn } \varphi$ is constant (so that either $\varphi > 0$ almost everywhere or $\varphi < 0$ almost everywhere).

Since φ is real, it follows that $(\varphi \cdot f)^* = \varphi \cdot f^*$. Since both $\varphi \cdot f^*$ and f are in \mathbf{H}^2 , Problem 27 implies that $\varphi \cdot f^* \cdot f \in \mathbf{H}^1$. Solution 26 becomes applicable and yields the information that $\varphi \cdot f^* \cdot f$ is a constant almost everywhere. Since $f \neq 0$, it follows that f is different from 0 almost everywhere (Problem 127), and consequently φ has almost everywhere the same sign as the constant value of $\varphi \cdot f \cdot f^*$.

In the original notation the result just obtained is that $\text{sgn}(\varphi - \lambda)$ is constant, and, since $\alpha < \lambda < \beta$, that is exactly what it is not. This contradiction proves that $[\alpha, \beta] \subset \Lambda(T_\varphi)$.

The reverse inclusion is easier. Since $\alpha \leq \varphi \leq \beta$, it follows that $\alpha \leq L_\varphi \leq \beta$; since $T_\varphi f = PL_\varphi f$ whenever $f \in \mathbf{H}^2$, it follows that $(T_\varphi f, f) = (PL_\varphi f, f) = (L_\varphi f, f)$, and hence that $\alpha \leq T_\varphi \leq \beta$. This of course implies that $\Lambda(T_\varphi) \subset [\alpha, \beta]$.

References

- L. V. Ahlfors [1953], *Complex analysis*, McGraw-Hill, New York.
- N. I. Akheizer, I. M. Glazman [1961], *Theory of linear operators in Hilbert space*, Ungar, New York.
- A. A. Albert, B. Muckenhoupt [1957], *On matrices of trace zero*, Michigan Math. J., vol. 4, pp. 1–3.
- T. Andô [1963], *On hyponormal operators*, Proc. A.M.S., vol. 14, pp. 290–291.
- N. Aronszajn [1950], *Theory of reproducing kernels*, Trans. A.M.S., vol. 68, pp. 337–404.
- N. Aronszajn, K. T. Smith [1954], *Invariant subspaces of completely continuous operators*, Ann. Math., vol. 60, pp. 345–350.
- E. Asplund [1958], *A non-closed relative spectrum*, Arkiv för Mat., vol. 3, pp. 425–427.
- F. V. Atkinson [1951], *The normal solubility of linear equations in normed spaces*, Mat. Sbornik, vol. 28, pp. 3–14.
- S. K. Berberian [1959], *Note on a theorem of Fuglede and Putnam*, Proc. A.M.S., vol. 10, pp. 175–182.
- S. K. Berberian [1962], *A note on hyponormal operators*, Pac. J. Math., vol. 12, pp. 1171–1175.
- S. Bergman [1947], *Sur les fonctions orthogonales de plusieurs variables complexes avec les applications à la théorie des fonctions analytiques*, Gauthier-Villars, Paris.
- A. R. Bernstein, A. Robinson [1966], *Solution of an invariant subspace problem of K. T. Smith and P. R. Halmos*, Pac. J. Math., vol. 16, pp. 421–431.
- A. Beurling [1949], *On two problems concerning linear transformations in Hilbert space*, Acta Math., vol. 81, pp. 239–255.
- G. Birkhoff, G. C. Rota [1960], *On the completeness of Sturm-Liouville expansions*, Amer. Math. Monthly, vol. 67, pp. 835–841.
- E. Bishop [1957], *Spectral theory for operators on a Banach space*, Trans. A.M.S., vol. 86, pp. 414–445.
- J. Bram [1955], *Subnormal operators*, Duke Math. J., vol. 22, pp. 75–94.
- M. S. Brodskii [1957], *On a problem of I. M. Gelfand*, Uspekhi Mat. Nauk, vol. 12, pp. 129–132.
- A. Brown [1953], *On a class of operators*, Proc. A.M.S., vol. 4, pp. 723–728.
- A. Brown, P. R. Halmos [1963], *Algebraic properties of Toeplitz operators*, J. reine angew. Math., vol. 231, pp. 89–102.
- A. Brown, P. R. Halmos, C. Pearcy [1965], *Commutators of operators on Hilbert space*, Can. J. Math., vol. 17, pp. 695–708.

- A. Brown, P. R. Halmos, A. L. Shields [1965], *Cesaro operators*, Acta Szeged, vol. 26, pp. 125–137.
- A. Brown, C. Pearcy [1965], *Structure of commutators of operators*, Ann. Math., vol. 82, pp. 112–127.
- A. Brown, C. Pearcy [1966], *Multiplicative commutators of operators*, Can. J. Math., vol. 18, pp. 737–749.
- L. de Branges, J. Rovnyak [1964], *The existence of invariant subspaces*, Bull. A.M.S., vol. 70, pp. 718–721.
- L. de Branges, J. Rovnyak [1965], *Correction to “The existence of invariant subspaces”*, Bull. A.M.S., vol. 71, p. 396.
- D. Deckard, C. Pearcy [1963], *Another class of invertible operators without square roots*, Proc. A.M.S., vol. 14, pp. 445–449.
- W. F. Donoghue [1957 a], *On the numerical range of a bounded operator*, Michigan Math. J., vol. 4, pp. 261–263.
- W. F. Donoghue [1957 b], *The lattice of invariant subspaces of a completely continuous quasi-nilpotent transformation*, Pac. J. Math., vol. 7, pp. 1031–1035.
- N. Dunford, J. T. Schwartz [1958], *Linear operators, Part I: General theory*, Interscience, New York.
- N. Dunford, J. T. Schwartz [1963], *Linear operators, Part II: Spectral theory*, Interscience, New York.
- C. Foiaş [1963], *A remark on the universal model for contractions of G. C. Rota*, Com. Acad. R. P. Romîne, vol. 13, pp. 349–352.
- J. M. Freeman [1965], *Perturbations of the shift operator*, Trans. A.M.S., vol. 114, pp. 251–260.
- B. Fuglede [1950], *A commutativity theorem for normal operators*, Proc. N.A.S., vol. 36, pp. 35–40.
- P. R. Halmos [1950 a], *Normal dilations and extensions of operators*, Summa Brasil., vol. 2, pp. 125–134.
- P. R. Halmos [1950 b], *Measure theory*, Van Nostrand, Princeton.
- P. R. Halmos [1951], *Introduction to Hilbert space and the theory of spectral multiplicity*, Chelsea, New York.
- P. R. Halmos [1952 a], *Spectra and spectral manifolds*, Ann. Soc. Pol. Math., vol. 25, pp. 43–49.
- P. R. Halmos [1952 b], *Commutators of operators*, Amer. J. Math., vol. 74, pp. 237–240.
- P. R. Halmos [1954], *Commutators of operators, II*, Amer. J. Math., vol. 76, pp. 191–198.
- P. R. Halmos [1958], *Finite-dimensional vector spaces*, Van Nostrand, Princeton.
- P. R. Halmos [1961], *Shifts on Hilbert spaces*, J. reine angew. Math., vol. 208, pp. 102–112.

- P. R. Halmos [1963 a], *A glimpse into Hilbert space*, Lectures on modern mathematics, vol. 1, Wiley, New York, pp. 1–22.
- P. R. Halmos [1963 b], *What does the spectral theorem say?*, Amer. Math. Monthly, vol. 70, pp. 241–247.
- P. R. Halmos [1964], *Numerical ranges and normal dilations*, Acta Szeged, vol. 25, pp. 1–5.
- P. R. Halmos [1966], *Invariant subspaces of polynomially compact operators*, Pac. J. Math., vol. 16, pp. 433–437.
- P. R. Halmos, S. Kakutani (1958), *Products of symmetries*, Bull. A.M.S., vol. 64, pp. 77–78.
- P. R. Halmos, G. Lumer [1954], *Square roots of operators, II*. Proc. A.M.S. vol. 5, pp. 589–595.
- P. R. Halmos, G. Lumer, J. J. Schäffer [1953], *Square roots of operators*, Proc. A.M.S., vol. 4, pp. 142–149.
- P. R. Halmos, J. E. McLaughlin [1962], *Partial isometries*, Pac. J. Math., vol. 13, pp. 585–596.
- G. H. Hardy, J. E. Littlewood, G. Pólya [1934], *Inequalities*, Cambridge University Press, Cambridge.
- P. Hartman, A. Wintner [1950], *On the spectra of Toeplitz's matrices*, Amer. J. Math., vol. 72, pp. 359–366.
- F. Hausdorff [1919], *Der Wertvorrat einer Bilinearform*, Math. Z., vol. 3, pp. 314–316.
- E. Heinz [1952], *Ein v. Neumannscher Satz über beschränkte Operatoren im Hilbertschen Raum*, Göttingen Nachr., pp. 5–6.
- H. Helson [1964], *Lectures on invariant subspaces*, Academic Press, New York.
- E. Hewitt, K. Stromberg [1965], *Real and abstract analysis*, Springer, New York.
- E. Hille, R. S. Phillips [1957], *Functional analysis and semi-groups*, A.M.S., Providence.
- K. Hoffman [1962], *Banach spaces of analytic functions*, Prentice-Hall, Englewood Cliffs.
- T. Itô [1958], *On the commutative family of subnormal operators*, J. Fac. Sci., Hokkaido University, vol. 14, pp. 1–15.
- K. Jacobs [1960], *Neuere Methoden und Ergebnisse der Ergodentheorie*, Springer, Berlin.
- N. Jacobson [1935], *Rational methods in the theory of Lie algebras*, Ann. Math. vol. 36, pp. 875–881.
- R. V. Kadison [1951], *Isometries of operator algebras*, Ann. Math., vol. 54, pp. 325–338.

- G. K. Kalisch [1957], *On similarity, reducing manifolds, and unitary equivalence of certain Volterra operators*, Ann. Math., vol. 66, pp. 481–494.
- I. Kaplansky [1953], *Products of normal operators*, Duke Math. J., vol. 20, pp. 257–260.
- T. Kato [1965], *Some mapping theorems for the numerical range*, Proc. Japan Acad., vol. 41, pp. 652–655.
- J. L. Kelley [1955], *General topology*, Van Nostrand, Princeton.
- D. C. Kleinecke [1957], *On operator commutators*, Proc. A.M.S., vol. 8, pp. 535–536.
- P. D. Lax [1959], *Translation invariant spaces*, Acta Math., vol. 101, pp. 163–178.
- A. Lebow [1963], *On von Neumann's theory of spectral sets*, J. Math. Anal. Appl., vol. 7, pp. 64–90.
- K. Löwner [1939], *Grundzüge einer Inhaltslehre im Hilbertschen Raume*, Ann. Math., vol. 40, pp. 816–833.
- W. Mlak [1965], *Unitary dilations of contraction operators*, Rozprawy Mat., vol. 46.
- B. Sz.-Nagy [1953], *Sur les contractions de l'espace de Hilbert*, Acta Szeged, vol. 15, pp. 87–92.
- B. Sz.-Nagy [1955], *Prolongements des transformations de l'espace de Hilbert qui sortent de cet espace*, Appendix to “Leçons d'analyse fonctionnelle” by F. Riesz, B. Sz.-Nagy, Akadémiai Kiadó, Budapest. ([1960], *Extensions of linear transformations in Hilbert space which extend beyond this space*, Appendix to “Functional analysis” by F. Riesz, B. Sz.-Nagy, Ungar, New York).
- B. Sz.-Nagy [1957], *Sur les contractions de l'espace de Hilbert. II.*, Acta Szeged, vol. 18, pp. 1–14.
- B. Sz.-Nagy, C. Foiaş [1962], *Sur les contractions de l'espace de Hilbert. V. Translations bilatérales*, Acta Szeged, vol. 23, pp. 106–129.
- B. Sz.-Nagy, C. Foiaş [1964], *Sur les contractions de l'espace de Hilbert. IX. Factorisations de la fonction caractéristique. Sousespaces invariants*, Acta Szeged, vol. 25, pp. 283–316.
- B. Sz.-Nagy, C. Foiaş [1966], *On certain classes of power-bounded operators in Hilbert space*, Acta Szeged, vol. 27, pp. 17–25.
- N. K. Nikolskii [1965], *Invariant subspaces of certain completely continuous operators*, Vestnik Leningrad Univ., pp. 68–77.
- C. Pearcy [1965], *On commutators of operators on Hilbert space*, Proc. A.M.S., vol. 16, pp. 53–59.
- C. Pearcy [1966], *An elementary proof of the power inequality for the numerical radius*, Mich. Math. J., vol. 13, pp. 289–291.

- C. R. Putnam [1951 a], *On commutators of bounded matrices*, Amer. J. Math., vol. 73, pp. 127–131.
- C. R. Putnam [1951 b], *On normal operators in Hilbert space*, Amer. J. Math., vol. 73, pp. 357–362.
- W. T. Reid [1951], *Symmetrizable completely continuous linear transformations in Hilbert space*, Duke Math. J., vol. 18, pp. 41–56.
- F. Riesz, B. Sz.-Nagy [1943], *Ueber Kontraktionen des Hilbertschen Raumes*, Acta Szeged, vol. 10, pp. 202–205.
- F. Riesz, B. Sz.-Nagy [1952], *Leçons d'analyse fonctionnelle*, Akadémiai Kiadó, Budapest. ([1955], *Functional analysis*, Ungar, New York.)
- C. E. Rickart [1960], *General theory of Banach algebras*, Van Nostrand, Princeton.
- J. R. Ringrose [1962], *On the triangular representation of integral operators*, Proc. London Math. Soc., vol. 12, pp. 385–399.
- J. B. Robertson [1965], *On wandering subspaces for unitary operators*, Proc. A.M.S., vol. 16, pp. 233–236.
- M. Rosenblum [1958], *On a theorem of Fuglede and Putnam*, J. London Math. Soc., vol. 33, pp. 376–377.
- G. C. Rota [1960], *On models for linear operators*, Comm. Pure Appl. Math., vol. 13, pp. 469–472.
- L. A. Sakhnovich [1957], *On the reduction of Volterra operators to the simplest form and on inverse problems*, Izv. Akad. Nauk SSSR, vol. 21, pp. 235–262.
- D. Sarason [1965], *A remark on the Volterra operator*, J. Math. Anal. Appl., vol. 12, pp. 244–246.
- H. H. Schaefer [1963], *Eine Bemerkung zur Existenz invarianter Teilräume linearer Abbildungen*, Math. Z., vol. 82, p. 90.
- J. J. Schäffer [1955], *On unitary dilations of contractions*, Proc. A.M.S., vol. 6, p. 322.
- J. J. Schäffer [1965], *More about invertible operators without roots*, Proc. A.M.S., vol. 16, pp. 213–219.
- R. Schatten [1960], *Norm ideals of completely continuous operators*, Springer, Berlin.
- M. Schreiber [1956], *Unitary dilations of operators*, Duke Math. J., vol. 24, pp. 579–594.
- J. Schur [1917], *Über Potenzreihen, die im Innern des Einheitskreises beschränkt sind*, J. reine angew. Math., vol. 147, pp. 205–232.
- I. E. Segal [1951], *Equivalences of measure spaces*, Amer. J. Math., vol. 73, pp. 275–313.
- I. E. Segal [1965], *Algebraic integration theory*, Bull. A.M.S., vol. 71, pp. 419–489.

- F. V. Shirokov [1956], *Proof of a conjecture of Kaplansky*, Uspekhi Mat. Nauk, vol. 11, pp. 167–168.
- K. Shoda [1936], *Einige Satze über Matrizen*, Japanese J. Math., vol. 13, pp. 361–365.
- J. G. Stampfli [1962], *Hyponormal operators*, Pac. J. Math., vol. 12, pp. 1453–1458.
- J. G. Stampfli [1965], *Perturbations of the shift*, J. London Math. Soc., vol. 40, pp. 345–347.
- M. H. Stone [1932], *Linear transformatins in Hilbert space and their applications to analysis*, A.M.S., New York.
- R. C. Thompson [1958], *On matrix commutators*, J. Wash. Acad. Sci., vol. 48, pp. 306–307.
- O. Toeplitz [1910], *Zur Theorie der quadratischen Formen von unendlichvielen Veränderlichen*, Göttingen Nachr., pp. 489–506.
- O. Toeplitz [1918], *Das algebraische Analogon zu einem Satze von Fejér*, Math. Z., vol. 2, pp. 187–197.
- F. A. Valentine [1964], *Convex sets*, McGraw-Hill, New York.
- J. von Neumann [1929], *Zur Algebra der Funktionaloperationen und Theorie der normalen Operatoren*, Math. Ann., vol. 102, pp. 370–427.
- J. von Neumann [1932], *Proof of the quasi-ergodic hypothesis*, Proc. N.A.S., vol. 18, pp. 70–82.
- J. von Neumann [1936], *On regular rings*, Proc. N.A.S., vol. 22, pp. 707–713.
- J. von Neumann [1950], *Functional operators*, Princeton University Press, Princeton.
- J. von Neumann [1951], *Eine Spektraltheorie für allgemeine Operatoren eines unitären Raumes*, Math. Nachr., vol. 4, pp. 258–281.
- H. Widom [1960], *Inversion of Toeplitz matrices, II*, Ill. J. Math., vol. 4, pp. 88–99.
- H. Widom [1964], *On the spectrum of a Toeplitz operator*, Pac. J. Math., vol. 14, pp. 365–375.
- N. A. Wiegmann [1948], *Normal products of matrices*, Duke Math. J., vol. 15, pp. 633–638.
- N. A. Wiegmann [1949], *A note on infinite normal matrix*, Duke Math. J., vol. 16, pp. 535–538.
- H. Wielandt [1949], *Ueber die Unbeschränktheit der Operatoren des Quantenmechanik*, Math. Ann., vol. 121, p. 21.
- A. Wintner [1929], *Zur Theorie der beschränkten Bilinearformen*, Math. Z., vol. 30, pp. 228–282.
- A. Wintner [1947], *The unboundedness of quantum-mechanical matrices*, Phys. Rev., vol. 71, pp. 738–739.

Index

- \mathbf{A}^2 , 15, 184
- Abnormal operator, 133
- Absolutely continuous function, 217
- Adjoint:
 - of a linear transformation, 24
 - of a weighted shift, 240
- Analytic:
 - characterization of \mathbf{H}^2 , 17, 192
 - element of \mathbf{L}^2 , 16
 - function, 44
 - Hilbert space, 15, 187
 - Toeplitz operator, 137, 140
- Analyticity of resolvent, 44, 236
- Approximate:
 - basis, 6, 172
 - point spectrum, 37
- Atkinson's theorem, 91, 294
- Atom of infinite measure, 211, 215

- Baire** category theorem, 13
- Ball, 4
- Basis, 6
 - for \mathbf{A}^2 , 15, 189
- Bendixson-Hirsch theorem, 111
- Bergman kernel, 19, 194
- Bilateral shift, 41, 232
- Bilinear functional, 3
- Borel set, 8
- Boundary of spectrum, 39, 228
- Bounded:
 - approximation, 20, 196
 - linear transformation, 24
 - Volterra kernel, 93, 296
- Boundedness:
 - from below, 26
 - of multiplications, 31, 213
 - of multipliers, 31, 212
 - on bases, 24, 204

- Cauchy** net and sequence, 14, 186
- Cesaro convergence, 196
- Characterization of isometries, 74, 274
- Chord, 5

- Closed:
 - graph, 28, 186, 206, 214, 215
 - ideal, 85, 287
 - transformation, 28
 - unit ball, 4
- Closure:
 - of approximate point spectrum, 39, 227
 - of numerical range, 111, 319
 - of partial isometries, 64, 260
 - of relative spectrum, 43, 234
 - of subnormal operators, 105
- Co-analytic:
 - element of \mathbf{L}^2 , 16
 - Toeplitz operator, 138
- Co-dimension, ix
- Co-hyponormal operator, 104
- Co-isometry, 62
- Column-finite matrix, 21, 201
- Commutant, 72
 - of the bilateral shift, 72, 271
 - of the unilateral shift, 73, 272
 - of the unilateral shift as limit, 74, 273
- Commutative operator determinants, 34, 218
- Commutator, 126, 333
 - subgroup, 134, 342
- Compact:
 - diagonal operator, 86, 287
 - hyponormal operator, 106, 313
 - normal operator, 86, 288
 - operator, 84, 287
 - versus Hilbert-Schmidt, 89, 290
- Complete continuity, 84
- Completeness of space of operators, 52
- Components of the space of partial isometries, 66, 261
- Compression, 118
 - spectrum, 37
- Conjugate:
 - function, 20, 200
 - linearity, 3

- Connectedness:**
 of invertible operators, 70, 267
 of partial isometries, 64, 260
Continuity:
 of adjoint, 56, 251
 of co-rank, 65, 261
 of extension, 19, 194
 of inversion, 52, 245
 of multiplication, 56, 252
 of norm, 56, 250
 of nullity, 65, 261
 of numerical range, 115, 323
 of rank, 65, 261
 of spectral radius, 54, 248
 of spectrum, 53, 248
Continuous:
 curve, 5, 169
 curve of unitary operators, 66, 267
 spectrum, 38
Contraction, 66
 similar to a unitary operator, 107, 316
Co-nullity, 65
Conv, 112
Convex hull, 112
Convexity, 4
Convexoid operator, 114, 322
Co-range, 75
Co-rank, 65
Co-subnormal operator, 104
Counting measure, 31
Cyclic vector, 79, 282

Degree of a vector, 303
Density of invertible operators, 70, 266
Derivation, 334
Diagonal, 29
 compact operator, 86, 287
 operator, 29, 209
Dilation, 118
 of a positive definite sequence, 124, 331
Dim, ix
Dimension, 5
 and local compactness, 8, 175
 and separability, 8, 176
Direct sum as commutator, 131, 338

Dirichlet:
 kernel, 197
 problem, 20, 198, 200
Distance:
 from a commutator to the identity, 129, 336
 from shift to normal operators, 270
 from shift to unitary operators, 74, 275
Distributive lattice, 7, 175
Donoghue lattice, 95, 303

Eigenvalue, x
 of Hermitian Toeplitz operator, 140, 350
 of weighted shift, 48, 240
Ergodic theorem, 121, 329
Essential range, 32
Evaluation functional, 18
Even function, 301
Extension:
 into the interior, 18
 of finite co-dimension, 103, 310
Extreme point, 4, 69, 265

Fejér kernel, 197
Fejér's theorem, 196
Field of values, 108
Filling in holes, 102, 310
Final space, 63
Form, 3
Fredholm:
 alternative, 90, 293
 operator, 91
Fuglede theorem, 98, 272
Full linear group, 134
Functional:
 calculus, 61
 Hilbert space, 18, 193

 Γ , 37
Gram-Schmidt process, 176
Graph, 28
Greatest lower bound, 7

 \mathbf{H}^1 , 16
 \mathbf{H}^2 , 16
 $\tilde{\mathbf{H}}^2$, 17

- \mathbf{H}^∞ , 16
- Hamel basis, 6
- Hausdorff metric, 115
- Heisenberg uncertainty principle, 126
- Herglotz's theorem, 125
- Hermitian Toeplitz operator, 140, 350
- Higher-dimensional numerical range, 110, 318
- Hilbert:
 - matrix, 23, 203
 - transform, 20, 200
- Hilbert-Schmidt operator, 87, 290
- Hole, 102
- Hyponormal:
 - compact operator, 106, 313
 - operator, 103, 104, 311
 - quasinilpotent operator, 106
- Ideal**, 85, 287
 - of operators, 90, 292
- Im, ix
- Increasing sequence of Hermitian operators, 57, 254
- Infimum, 7
 - of two projections, 58, 257
- Infinite Vandermonde, 6, 171
- Initial space, 63
- Inner function, 79
- Integral kernel and operator, 87
- Interior, 4
- Invariant subspace, 95
 - of the shift, 78, 281
- Invertible:
 - function, 32
 - modulo the ideal of compact operators, 91
 - sequence, 30
 - transformation, 25, 205
- Irreducibility of unilateral shift, 73
- Isometry, 62
- Jacobi matrix**, 201
- Joint continuity, 56
- Ker**, ix
- kernel:
 - function, 19, 193
 - of an integral operator, 87
 - of the identity, 87, 288
- Kleinecke-Shirokov theorem, 128, 334
- k -numerical range, 111, 318
- l^1 , 14
- l^2 , 6
- Λ , 37
- Large subspace, 71, 130, 268
- Lattice of subspaces, 7, 175
- Laurent operator and matrix, 135, 345
- Least upper bound, 7
- Left shift, 269
- Leibniz formula, 334
- Limit:
 - of commutators, 127, 132, 334
 - of operators of finite rank, 89, 291
 - of quadratic forms, 3, 167
- Linear:
 - basis, 6
 - dimension, 5, 170
 - functional, 3
 - functional on l^2 , 14, 185
- Liouville's theorem, 237
- Local compactness and dimension, 8, 175
- Locally finite measure, 211
- Matrix**, 21
- Maximal:
 - partial isometry, 64, 260
 - polar representation, 69, 264
- Maximum modulus principle, 196
- Mean ergodic theorem, 122, 330
- Measure in Hilbert space, 8, 176
- Metric:
 - space of operators, 52, 245
 - topology, 10
- Minimal normal extension, 101, 308
- Mixed continuity, 84, 286
- Modular lattice, 7, 175
- Monotone shift, 97
- Multiple, 79
- Multiplication:
 - operator, 31
 - on a functional Hilbert space, 32, 214
 - on l^2 , 29, 209

- Multiplicative commutator, 133, 340
- Multiplicativity of extension, 20, 198
- Multiplicity of shift, 75
- Multiplier of functional Hilbert space, 33, 217
- Non-closed:**
 - range, 26
 - vector sum, 26
- Non-emptiness of spectrum, 44, 237
- Non-overlapping chords, 5
- Norm:
 - of a linear transformation, 24
 - of a multiplication, 31, 210
 - of a weighted shift, 48, 239
 - of the Volterra integration operator, 95
 - power and power norm, 105, 313
 - topology, 10, 55
- Norm 1, spectrum $\{1\}$, 95, 302
- Normal:
 - and subnormal partial isometries, 105, 312
 - compact operator, 86, 288
 - dilation, 119
- Normality and numerical range, 112, 321
- Normaloid operator, 114, 322
- Normal partial isometry, 105, 312
- Normed algebra, 126
- ν , 65
- Nullity, 65
- Null-point, 215
- Null-space, ix
- Numerical radius, 114, 322
- Numerical range, 108
 - and normality, 112, 321
 - and quasinilpotence, 112, 320
 - and spectrum, 111, 320
 - and subnormality, 113, 321
- Odd function**, 301
- One-point spectrum, 49, 242
- Open unit ball, 4
- Openness of invertible operators, 53, 245
- Operator, ix
 - with large kernel, 129, 336
 - topologies, 55
- Operator determinant, 34, 35, 219
 - with finite entry, 36, 223
- Order for partial isometries, 63
- Orthogonal dimension, 5
- Orthonormal basis, 6
- Parseval's identity**, 191, 291
- Part of an operator, 75
- Partial isometry, 63, 259
- Perturbation, 91
- Perturbed spectrum, 92, 295
- Π , 37
- Π_0 , 37
- Point spectrum, 37
- Polar decomposition, 68, 69, 263
- Position operator, 309
- Positive definiteness, 124
- Positive self-commutator, 132, 339
- Power:
 - dilation, 121
 - inequality, 115, 117, 324
 - norm 105, 313
- Powers of a hyponormal operator, 106, 314
- Preservation of dimension, 27, 205
- Principle of uniform boundedness, 12, 13, 183, 186, 204, 254
- Product:
 - in \mathbf{H}^2 , 17, 191
 - of normal operators, 99
 - of symmetries, 71, 269
- Projection:
 - as self-commutator, 132, 339
 - dilation, 119, 327
- Projections of equal rank, 27, 206
- Proper:
 - contraction, 76
 - ideal, 90
 - value, x
- Pure isometry, 74
- Putnam-Fuglede theorem, 98, 306
- Quadratic form**, 3
- Quasinilpotence, 50
 - and numerical range, 112, 320

- Quasinilpotent:
 hyponormal operator, 106
 Toeplitz operator, 139
 Quasinormal operator, 69, 266
- r , 45
 Radial limit, 19, 195
 Ran, ix
 Range of a compact operator, 91, 294
 Rank, 65
 Re, ix
 Real function in \mathbf{H}^2 , 15, 190
 Reducible weighted shift, 82, 284
 Reducing subspace of a normal operator, 71, 269
 Reduction by the unitary part, 74, 275
 Region, 15
 Regular ring, 43
 Reid's inequality, 51, 244
 Relative spectrum, 43, 233
 Representation of linear functional, 3, 167
 Reproducing kernel, 19
 Residual spectrum, 38
 of a normal operator, 40, 229
 Resolvent, 44
 ρ , 165
 Riesz-Fischer theorem, 188
 Riesz representation theorem, 3, 167
 Riesz theorem, 82, 283
 Riesz theorem generalized, 82, 283
 Right shift, 269
 Ring of operators, 59
- Schur** test, 22, 202
 Schwarz inequality, 216, 331
 Segment, 4
 Self-adjoint ideal, 85, 287
 Self-commutator, 132
 Semicontinuity of spectrum, 53, 247
 Semilinear functional, 3
 Separability:
 and dimension, 8, 176
 and weak metrizability, 12, 183
 of space of operators, 52, 245
 Separate continuity of multiplication, 57, 253
- Sequential continuity of multiplication, 57, 254
 Sequential weak completeness, 14
 Sesquilinear form, 3
 Sgn, ix
 Shift:
 as universal operator, 75, 276
 modulo compact operators, 92, 295
 σ -field, 8
 Similar:
 normal operators, 99, 306
 operators, 38
 subnormal operators, 102, 308
 Similarity:
 and spectrum, 38, 226
 of weighted shifts, 47, 239
 to parts of shifts, 76, 278
 Simple:
 curve, 5
 unilateral shift, 75
 Singular operator, 53
 Skew-symmetric Volterra operator, 95, 301
 Sliding product, 240
 Special invariant subspaces of the shift, 78, 280
 Spectra and conjugation, 37, 225
 Spectral:
 set, 122, 330
 theorem, 60
 Spectral inclusion theorem, 102, 309
 for Toeplitz operators, 138, 348
 Spectral mapping theorem, 38, 225
 for normal operators, 60, 259
 Spectral measure, 169
 of the unit disc, 99, 307
 Spectral parts:
 of a diagonal operator, 40, 229
 of a multiplication, 40, 229
 Spectral radius, 45, 237
 of a weighted shift, 48, 239
 Spectraloid operator, 114
 Spectrum:
 and numerical range, 111, 320
 and similarity, 38, 226
 of a commutator, 132
 of a diagonal operator, 30, 210
 of a direct sum, 50, 243

- of a functional multiplication, 42, 232
- of a Hermitian Toeplitz operator, 140, 350
- of a multiplication, 32, 214
- of a partial isometry, 67, 263
- of a product, 39, 227
- Square root, 58, 256
 - of a compact operator, 90, 292
 - of shift, 72, 271
- Strict convexity, 4, 5, 168
- Strong:
 - boundedness, 25
 - convergence, 10
 - dilation, 121
 - operator topology, 55
 - topology, 10
- Subnormal:
 - operator, 100, 307
 - partial isometry, 105, 312
 - versus hyponormal, 105, 311
 - versus quasinormal, 101, 307
- Subnormality and numerical range, 113, 321
- Subspace, ix
- Support, 50
- Supremum, 7
- Symmetric transformation, 28
- Symmetry, 71, 269
- Szegő kernel, 19, 194

- Toeplitz:**
 - operator and matrix, 135, 345
 - product, 137, 347
- Toeplitz-Hausdorff theorem, 108, 317
- Topologies for operators, 55, 250
- Tr, xi, 111
- Trigonometric polynomial, 16
- Two-sided ideal, 85
- Tychonoff-Alaoglu theorem, 12, 180

- Unbounded:**
 - symmetric transformation, 28, 207
 - Volterra kernel, 94, 297
- Uniform:
 - strong convergence, 56, 250
 - topology, 55
 - weak convergence, 12, 56, 179, 250
- Uniform boundedness, 12, 56
 - of linear transformations, 24, 204
- Unilateral shift, 40, 230
 - of higher multiplicity, 75, 276
 - versus normal operators, 71, 270
- Unilaterally invertible operator, 70, 266
- Unit, 127
 - ball, 4
 - disc, 4
- Unitary:
 - dilation, 118, 326
 - multiplicative commutator, 134, 341
 - power dilation, 119, 327
- Unitary equivalence:
 - of partial isometries, 66, 262
 - of similar normal operators, 99, 306
 - of similar subnormal operators, 102, 308
- Upper semicontinuity, 53, 247

- Vandermonde**, 6, 171
- Vector sum, 7, 174
- Volterra:
 - integration operator, 94, 300
 - kernel, 93, 296, 297
 - operator, 93, 300
- Von Neumann algebra, 59, 258
- Von Neumann-Heinz theorem, 123
- W , 108
- w , 114
- Wandering subspace, 77, 279
- Weak:
 - analyticity, 44
 - boundedness, 13
 - Cauchy net and sequence, 14, 186
 - closure of a subspace, 10, 178
 - compactness of the unit ball, 12, 180
 - completeness, 14, 186
 - continuity, 10
 - continuity of norm and inner product, 11, 178
 - operator topology, 55
 - separability, 12, 179
 - topology, 10

Weak metrizability:
 and separability, 12, 183
 of Hilbert space, 13, 185
 of the unit ball, 12, 181

Weighted:
 sequence space, 48, 241
 shift 46, 238

Weyl's theorem, 91, 295
Wielandt's proof, 333
Wintner's proof, 333
 W_k , 111, 318

Zero-divisor, 82, 138, 283