

W. I. SMIRNOW

LEHRGANG
DER
HÖHEREN
MATHEMATIK

TEIL V



VEB DEUTSCHER VERLAG DER WISSENSCHAFTEN

W. I. Smirnow
Lehrgang der höheren Mathematik · Teil V

Hochschulbücher für Mathematik

Herausgegeben von H. Grell, K. Maruhn und W. Rinow

Band 6

Lehrgang der höheren Mathematik

Teil V

von W. I. Smirnow

Mitglied der Akademie der Wissenschaften der UdSSR

Mit 3 Abbildungen

Elfte Auflage



VEB Deutscher Verlag der Wissenschaften
Berlin 1991

Titel der Originalausgabe

В. И. Смирнов
Курс высшей математики. Том 5
«Физматгиз», Москва 1959

Die Ausgabe in deutscher Sprache besorgten:
Renate Helle und Brigitte Mai

ISBN 3-326-00667-5

ISSN 0073-2842

Verlagslektoren: Erika Arndt, Brigitte Mai
Verlagshersteller: Norma Braun
Umschlaggestaltung: Hartwig Hoefmann
© der deutschsprachigen Ausgabe:
1962 Deutscher Verlag der Wissenschaften, O - 1080 Berlin, Postfach 1216
Lizenz-Nr. 206
Printed in Germany
Gesamtherstellung: Druckerei »G. W. Leibniz« GmbH, O - 4450 Gräfenhainichen
LSV 1034
Bestellnummer: 569 424 5

Aus dem Vorwort zur russischen Ausgabe

In der modernen mathematischen Physik sind die Funktionen einer reellen Veränderlichen, verschiedene Funktionenräume und die allgemeine Theorie der Operatoren von großer Bedeutung. Diesen Fragen ist im wesentlichen das vorliegende Buch gewidmet, das aus der im Jahre 1947 erschienenen ersten Auflage entstanden ist.

Die Theorie der Funktionen einer reellen Veränderlichen in diesem Buch umfaßt die Theorie des klassischen Stieltjesschen Integrals, des Lebesgue-Stieltjesschen Integrals und die Theorie der volladditiven Mengenfunktionen.

Im ersten Kapitel ist die Theorie des klassischen Stieltjesschen Integrals dargestellt. Ferner wird eine allgemeinere Definition des Stieltjesschen Integrals über ein Intervall beliebiger Art angegeben; diese Definition basiert auf der Übereinstimmung des entsprechenden oberen und unteren Darbouxschen Integrals bei der Zerlegung des Grundintervalls in beliebige Teilintervalle. Als Beispiel für das klassische Stieltjessche Integral werden die Fourier-Stieltjesschen und die Cauchy-Stieltjesschen Integrale betrachtet und für diese Integrale Umkehrformeln aufgestellt. Auch wird das Stieltjessche Integral für den Fall einer Ebene definiert. Wir studieren ferner den Raum C der stetigen Funktionen und geben die allgemeine Form der linearen Funktionale auf diesem Raum an.

Das zweite Kapitel enthält Grundzüge der metrischen Theorie der Funktionen einer reellen Veränderlichen und die Grundlagen des Lebesgue-Stieltjesschen Integrals. Die ganze Theorie wird am Fall der Ebene erklärt, und es wird gezeigt, wie diese Theorie auf den n -dimensionalen euklidischen Raum übertragen werden kann. Die Maßtheorie wird für eine beliebige, nichtnegative, additive und normale Funktion aufgebaut, die auf halboffenen zweidimensionalen Intervallen gegeben ist. Das Lebesgue-Stieltjessche Integral einer beschränkten Funktion wird definiert mit Hilfe der Übereinstimmung des oberen und des unteren Darbouxschen Integrals bei Zerlegung der zugrunde gelegten meßbaren Menge in meßbare Teilmengen. Am Schluß des zweiten Kapitels wird ausführlich auf den Prozeß der Mittelung von Funktionen und auf die Eigenschaften der Mittelfunktionen unter gewissen Forderungen für den Mittelungskern eingegangen. Dieser Prozeß wird im weiteren Verlauf der Ausführungen oft benutzt.

Im dritten Kapitel werden die volladditiven Mengenfunktionen behandelt. Nach dem Beweis der ersten Sätze dieser Theorie wird (zunächst ohne Beweis) ein Satz angegeben, der die Zerlegung einer volladditiven Mengenfunktion in einen singulären und in einen absolut stetigen Summanden betrifft, und die mit dieser Zerlegung verknüpften grundlegenden Tatsachen werden erläutert. Der Fall einer unabhängigen Veränderlichen wird eingehend untersucht. Im allgemeinen Fall betrachten wir eine absolut stetige Mengenfunktion und geben eine Formel für die Variablensubstitution im mehrdimensionalen Lebesgue-Stieltjesschen Integral an. Am Schluß des dritten Kapitels wird dann der Satz über die Zerlegung einer volladdi-

tiven Mengenfunktion in zwei Summanden bewiesen. Für den mehrdimensionalen Fall führen wir dann den Begriff des Hellingerschen Integrals ein und untersuchen seine Eigenschaften. Insbesondere wird der Zusammenhang zwischen dem Hellingerschen und dem Lebesgue-Stieltjesschen Integral erwähnt und auf den Fall des eindimensionalen Hellingerschen Integrals ausführlich eingegangen. Sämtliche am Schluß dieses Kapitels geführten Beweise stützen sich auf die vorhergehenden Untersuchungen der Eigenschaften volladditiver Mengenfunktionen [78, 79].

Das vierte Kapitel enthält die Grundzüge der Theorie der metrischen und der normierten Räume und außerdem eine ausführliche Darlegung der verallgemeinerten Ableitungen, der Einbettungssätze für verschiedene Funktionenräume und der Theorie der Funktionale im Raum der stetig differenzierbaren Funktionen. Diese Fragen sind mit den bekannten Sobolewschen Untersuchungen verknüpft.¹⁾ Die verallgemeinerten Ableitungen werden zweimal definiert, und zwar mit Hilfe der Formel für die partielle Integration und durch das Abschließen der Funktionen mit stetigen Ableitungen, wobei dann die Äquivalenz dieser beiden Definitionen nachgewiesen wird. Der Fall sternförmiger Gebiete wird gesondert betrachtet. Ferner werden die vollständigen normierten Funktionenräume $\tilde{W}_p^{(l)}(D)$ und $W_p^{(l)}(D)$ eingeführt; der Raum $\tilde{W}_p^{(l)}(D)$ besteht aus denjenigen Funktionen $\varphi(x)$, die in einem Gebiet D definiert sind und verallgemeinerte l -te Ableitungen haben, wobei $\varphi(x)$ nebst allen Ableitungen zu $L_p(D)$ gehört; der Raum $W_p^{(l)}(D)$ enthält diejenigen Funktionen $\varphi(x)$, die alle verallgemeinerten Ableitungen bis einschließlich l -ter Ordnung besitzen. Dann wird gezeigt, daß die Räume $\tilde{W}_p^{(l)}(D)$ und $W_p^{(l)}(D)$ für eine umfangreiche Klasse von Gebieten D ein und dieselben Funktionen enthalten und daß die in ihnen eingeführten Normen äquivalent sind. Ferner lassen sich für den Raum $W_p(D)$ verhältnismäßig einfache Sätze beweisen, die Spezialfälle der Einbettungssätze für $W_p^{(l)}(D)$ sind. Nach der Formulierung der Einbettungssätze wird (in Kleindruck) ihr Beweis mit Hilfe der Sobolewschen Integraldarstellung angegeben. Der hier behandelte Stoff findet sich auch in der in der Fußnote zitierten Monographie.

Das fünfte und letzte Kapitel bringt die allgemeine Theorie des Hilbertschen Raumes (zunächst nur für beschränkte Operatoren). Hier werden die Fredholmischen Sätze für lineare Gleichungen mit vollstetigen Operatoren bewiesen; für normierte Räume sind diese Sätze ohne Beweis notiert. Ferner werden für selbstadjungierte Operatoren auf einem kontinuierlichen Spektrum die entsprechenden Integraldarstellungen durch Differentiallösungen mit Hilfe von Hellingerschen Integralen angegeben und Beispiele für die Anwendung der allgemeinen Theorie der beschränkten Operatoren aus l_2 und L_2 betrachtet. Der letzte Paragraph dieses Kapitels ist der Theorie der nicht beschränkten Operatoren im Hilbertschen Raum gewidmet. Nachdem die allgemeinen Sätze bewiesen sind, werden zahlreiche Beispiele für Differentialoperatoren mit einer und mit mehreren unabhängigen Veränderlichen angegeben. Nach der allgemeinen Theorie der Erweiterungen abgeschlossener symmetrischer Operatoren werden speziell halbbeschränkte Operatoren und insbesondere ihre Friedrichssche Erweiterung untersucht.

Bei der Zusammenstellung dieses Buches habe ich außer speziellen Arbeiten mehrere Bücher benutzt, in der Hauptsache: W. I. GLIWENKO, Das Stieltjessche Inte-

¹⁾ Vgl. S. L. SOBOLEW. Einige Anwendungen der Funktionalanalysis in der mathematischen Physik. In dieser Monographie werden die hier erwähnten Fragen behandelt. (Bezüglich des Erscheinungsortes und -jahres dieses Buches und weiterer im folgenden zitierter Werke vgl. das Literaturverzeichnis am Schluß des Buches.)

gral; I. P. NATANSON, Theorie der Funktionen einer reellen Veränderlichen; S. SAKS, Théorie de l'intégrale; CH. DE LA VALLÉE-POUSSIN, Intégrales de Lebesgue, fonctions d'ensembles, classes de Baire; M. H. STONE, Linear transformation in Hilbert space and their applications to analysis; N. I. ACHIESER und I. M. GLAS-MANN, Theorie der linearen Operatoren im Hilbert-Raum; A. I. PLESSNER, Spektraltheorie der linearen Operatoren I; N. I. ACHIESER, Unendliche Jacobische Matrizen und das Momentenproblem; S. L. SOBOLEW, Einige Anwendungen der Funktionalanalysis in der mathematischen Physik.

Abschließend möchte ich Herrn S. M. LOSINSKI danken, der das ursprüngliche Manuskript gelesen und mir eine Reihe wertvoller Hinweise gegeben hat. Mehrere Darlegungen im zweiten Teil dieses Buches stammen von Frau O. A. LADYSHENSKAJA, die bei diesem Teil mitgearbeitet hat und mit der ich den Aufbau des Buches eingehend erörtert habe. Eine große Hilfe bei der Zusammenstellung des zweiten Teils erwies mir auch Herr M. S. BIRMAN. Von ihm stammen die Ausführungen über die Einbettungssätze [114–118] und die Theorie der kleinen Störungen des Spektrums [198]. Wertvolle Ratschläge gab er mir für die Spektren symmetrischer Operatoren und deren Erweiterungen sowie für die Darlegung des Stoffes im vierten Kapitel. Meinen tiefen Dank möchte ich hiermit Frau LADYSHENSKAJA und Herrn BIRMAN aussprechen; ohne ihre Hilfe hätte ich die Arbeit nicht durchführen können. Die ersten drei Kapitel sind auch von Herrn G. P. AKILOW gelesen worden, von dem ich viele Hinweise erhielt und dem ich hiermit sehr dafür danke.

Leningrad, 20. Juli 1959

W. SMIRNOW

Hinweis

Die Rückverweise auf die Teile III/2, IV/1 und IV/2 beziehen sich auf neue überarbeitete russische Ausgaben in den siebziger und achtziger Jahren. Deutschsprachige Ausgaben der Teile IV/1 und IV/2 erschienen 1988 bzw. 1989, die des Teiles III/2 ist in Vorbereitung.

Inhalt

I. Das Stieltjessche Integral

1. Mengen und ihre Mächtigkeit	15
2. Das Stieltjessche Integral und seine Haupteigenschaften	17
3. Darboux'sche Summen	21
4. Das Stieltjessche Integral einer stetigen Funktion	24
5. Das uneigentliche Stieltjessche Integral	27
6. Die Sprungfunktion	29
7. Physikalische Interpretation	32
8. Funktionen von beschränkter Variation	33
9. Integrierbare Funktionen von beschränkter Variation	38
10. Existenz des Stieltjesschen Integrals	39
11. Grenzübergang unter dem Stieltjesschen Integral	40
12. Satz von HELLY	42
13. Das Auswahlprinzip	45
14. Der Raum der stetigen Funktionen	46
15. Die allgemeine Form der Funktionale in C	48
16. Lineare Operatoren in C	52
17. Intervallfunktionen	53
18. Das allgemeine Stieltjessche Integral	55
19. Eigenschaften des (allgemeinen) Stieltjesschen Integrals	57
20. Existenz des allgemeinen Stieltjesschen Integrals	60
21. Intervallfunktionen in der Ebene	62
22. Übergang zur Punktfunktion	65
23. Das Stieltjessche Integral in der Ebene	67
24. Funktionen von beschränkter Variation in der Ebene	68
25. Der Raum der stetigen Funktionen mehrerer Veränderlicher	71
26. Das Fourier-Stieltjessche Integral	71
27. Die Umkehrformel	74
28. Der Faltungssatz	76
29. Das Cauchy-Stieltjessche Integral	78

II. Mengenfunktionen und das Lebesguesche Integral

§ 1. Mengenfunktionen und Maßtheorie	82
30. Mengenoperationen	82
31. Punktmengen	85
32. Eigenschaften abgeschlossener und offener Mengen	86
33. Elementarbereiche	89
34. Das äußere Maß und seine Eigenschaften	92
35. Meßbare Mengen	94
36. Meßbare Mengen (Fortsetzung)	101

37. Meßbarkeitskriterien	102
38. Mengenkörper	104
39. Unabhängigkeit von der Wahl des Koordinatensystems	106
40. Der Körper B	107
41. Der Fall einer Veränderlichen	108
 § 2. Meßbare Funktionen	 109
42. Die Definition der meßbaren Funktionen	109
43. Eigenschaften der meßbaren Funktionen	112
44. Der Grenzwert meßbarer Funktionen	113
45. Die Eigenschaft C	117
46. Stückweise konstante Funktionen	117
47. Die Baireschen Klassen	119
 § 3. Das Lebesguesche Integral	 120
48. Das Integral beschränkter Funktionen	120
49. Eigenschaften des Integrals	123
50. Das Integral von unbeschränkten nichtnegativen Funktionen	126
51. Eigenschaften des Integrals	129
52. Funktionen beliebigen Vorzeichens	131
53. Komplexe summierbare Funktionen	135
54. Der Grenzübergang unter dem Integralzeichen	136
55. Die Klasse L_2	139
56. Konvergenz im Mittel	141
57. Der Hilbertsche Funktionenraum	144
58. Orthogonale Funktionensysteme	146
59. Der Raum l_2	150
60. Lineare Mannigfaltigkeiten in L_2	152
61. Beispiele für vollständige Systeme	155
62. Die Höldersche und die Minkowskische Ungleichung	156
63. Das Integral über eine Menge unendlichen Maßes	160
64. Die Klasse L_2 auf einer Menge unendlichen Maßes	164
65. Integrierbare Funktionen von beschränkter Variation	166
66. Die Reduktion mehrfacher Integrale	168
67. Der Fall einer charakteristischen Funktion	170
68. Der Satz von FUBINI	173
69. Das Vertauschen der Reihenfolge der Integration	176
70. Die Stetigkeit im Mittel	177
71. Mittelfunktionen	179

III. Mengenfunktionen. Absolute Stetigkeit. Verallgemeinerung des Integralbegriffs

72. Additive Mengenfunktionen	185
73. Die singuläre Funktion	188
74. Der Fall einer Veränderlichen	190
75. Absolut stetige Mengenfunktionen	194
76. Beispiel	199
77. Absolut stetige Funktionen mehrerer Veränderlicher	201
78. Hilfssätze	203
79. Hilfssätze (Fortsetzung)	207
80. Hauptsatz	211

81. Das Hellingersche Integral	214
82. Der Fall einer Veränderlichen	217
83. Eigenschaften des Hellingerschen Integrals	220

IV. Metrische und normierte Räume

84. Der metrische Raum	224
85. Vervollständigung eines metrischen Raumes	226
86. Operatoren und Funktionale. Das Prinzip der kontrahierenden Abbildungen	230
87. Beispiele	231
88. Beispiele für die Anwendung des Prinzips der kontrahierenden Abbildungen	234
89. Kompaktheit	236
90. Kompaktheit im Raum C	238
91. Kompaktheit im Raum L_p	238
92. Kompaktheit im Raum l_p	241
93. Funktionale auf in sich kompakten Mengen	242
94. Separabilität	243
95. Lineare normierte Räume	245
96. Beispiele für normierte Räume	248
97. Operatoren in normierten Räumen	248
98. Lineare Funktionale	251
99. Konjugierte Räume	254
100. Schwache Konvergenz von Funktionalen	257
101. Schwache Konvergenz von Elementen	259
102. Lineare Funktionale in den Räumen C , L_p und l_p	262
103. Schwache Konvergenz in den Räumen C , L_p und l_p	268
104. Der Raum der linearen Operatoren und die Konvergenz von Operatorfolgen	269
105. Adjungierte Operatoren	271
106. Vollstetige Operatoren	271
107. Operatorengleichungen	272
108. Vollstetige Operatoren in den Räumen C , L_p und l_p	274
109. Verallgemeinerte Ableitungen	277
110. Verallgemeinerte Ableitungen (Fortsetzung)	281
111. Sternförmige Gebiete	283
112. Die Räume $\dot{W}_p^{(l)}$ und $W_p^{(l)}$	284
113. Eigenschaften der Funktionen aus der Klasse $W_p^{(l)}(D)$	286
114. Einbettungssätze	292
115. Integraloperatoren mit polaren Kernen	295
116. Sobolewsche Integraldarstellungen	300
117. Einbettungssätze	304
118. Gebiete allgemeinerer Art	306
119. Der Raum $\dot{C}^{(l)}(D)$	307

V. Der Hilbertsche Raum

§ 1. Die Theorie der beschränkten Operatoren	317
120. Die Axiome des Raumes	317
121. Orthogonalität und Orthogonalsysteme von Elementen	319
122. Die Projektion	322

123. Lineare Funktionale	324
124. Lineare Operatoren	326
125. Bilineare und quadratische Funktionale	329
126. Die Grenzen eines selbstadjungierten Operators	330
127. Der inverse Operator	332
128. Das Spektrum eines Operators	335
129. Das Spektrum eines selbstadjungierten Operators	337
130. Die Resolvente	341
131. Folgen von Operatoren	341
132. Schwache Konvergenz	342
133. Vollstetige Operatoren	344
134. Die Räume H und l_2	346
135. Lineare Gleichungen mit vollstetigen Operatoren	348
136. Vollstetige selbstadjungierte Operatoren	352
137. Unitäre Operatoren	356
138. Die absolute Norm eines Operators	359
139. Operatoren über Unterräumen	361
140. Projektionsoperatoren	362
141. Zerlegung der Einheit. Das Stieltjessche Integral	366
142. Die Spektralschar eines selbstadjungierten Operators	371
143. Stetige Funktionen eines selbstadjungierten Operators	372
144. Eine Formel für die Resolvente. Die Charakterisierung der regulären Werte von λ	375
145. Eigenwerte und Eigenelemente	377
146. Das reine Punktspektrum	378
147. Das einfache kontinuierliche Spektrum	379
148. Invariante Unterräume	384
149. Der allgemeine Fall eines kontinuierlichen Spektrums	386
150. Das gemischte Spektrum	388
151. Differentiallösungen	389
152. Die Multiplikation mit der unabhängigen Veränderlichen	392
153. Unitär äquivalente selbstadjungierte Operatoren	394
154. Die Spektralzerlegung unitärer Operatoren	395
155. Funktionen eines selbstadjungierten Operators	395
156. Vertauschbare Operatoren	399
157. Störung des Spektrums eines selbstadjungierten Operators	400
158. Normale Operatoren	402
159. Hilfssätze	404
160. Die Potenzreihe eines Operators	407
161. Die Spektralschar	408
§ 2. Die Räume l_2 und L_2	411
162. Lineare Operatoren im Raum l_2	411
163. Beschränkte Operatoren	412
164. Unitäre Matrizen und Projektionsmatrizen	416
165. Selbstadjungierte Matrizen	417
166. Das kontinuierliche Spektrum	419
167. Jacobische Matrizen	423
168. Differentiallösungen	425
169. Beispiele	427
170. Schwache Konvergenz im Raum l_2	430
171. Vollstetige Operatoren im Raum l_2	430
172. Integraloperatoren im Raum L_2	431
173. Der adjungierte Operator	432
174. Vollstetige Operatoren	434

175. Die Spektralschar	435
176. Die Spektralschar (Fortsetzung)	436
177. Unitäre Transformationen im Raum L_2	438
178. Fouriertransformationen	440
179. Fouriertransformation und Hermitesche Funktionen	443
180. Die Operation der Multiplikation	444
181. Differenzkerne	446
182. Schwache Konvergenz	450
183. Andere Realisierungen des Raumes H	450

§ 3. Nichtbeschränkte Operatoren 451

184. Abgeschlossene Operatoren	451
185. Der adjungierte Operator	453
186. Der Graph eines Operators	455
187. Symmetrische und selbstadjungierte Operatoren	457
188. Beispiele für nichtbeschränkte Operatoren	459
189. Das Spektrum eines selbstadjungierten Operators	468
190. Das Punktspektrum	471
191. Invariante Unterräume und Reduzibilität eines Operators	472
192. Zerlegung der Einheit. Das Stieltjesses Integral	475
193. Stetige Funktionen eines selbstadjungierten Operators	480
194. Die Resolvente	481
195. Eigenwerte	483
196. Das gemischte Spektrum	483
197. Funktionen eines selbstadjungierten Operators	485
198. Kleine Störungen des Spektrums	487
199. Der Operator der Multiplikation	489
200. Integraloperatoren	492
201. Erweiterung eines abgeschlossenen symmetrischen Operators	495
202. Defektindizes	498
203. Der adjungierte Operator	501
204. Maximale Operatoren	503
205. Erweiterung symmetrischer halbbeschränkter Operatoren	504
206. Vergleich halbbeschränkter Operatoren	508
207. Beispiele aus der Theorie der Erweiterungen	510
208. Das Spektrum eines symmetrischen Operators	512
209. Einige Sätze über Erweiterungen und ihre Spektren	514
210. Die Unabhängigkeit der Defektindizes von λ	516
211. Über die Invarianz des kontinuierlichen Teils des Kerns bei symmetrischen Erweiterungen	518
212. Über die Spektren selbstadjungierter Erweiterungen	518
213. Beispiele	519
214. Unendliche Matrizen	520
215. Jacobische Matrizen	522
216. Matrizen und Operatoren	526
217. Unitäre Äquivalenz von C -Matrizen	528
218. Die Existenz einer Spektralschar	530

Literaturhinweise 534

Namen- und Sachverzeichnis 540

I. Das Stieltjessche Integral

1. Mengen und ihre Mächtigkeit. Bei der Anwendung der Analysis in der modernen Naturwissenschaft spielen mehrere Integralbegriffe eine große Rolle. In den ersten beiden Kapiteln dieses Bandes wird die Integrationstheorie unter einem allgemeineren Gesichtspunkt als früher behandelt. In diesem Abschnitt bringen wir zur Vorbereitung einige grundsätzliche Ausführungen aus der Mengenlehre. Sie bilden in gewisser Weise eine Ergänzung zu dem in [IV/1, 15] Gesagten.

Gegeben seien zwei aus irgendwelchen Objekten (Elementen) bestehende Mengen A_1 und A_2 . Die Mengen heißen von *gleicher Mächtigkeit*, wenn sich die Elemente von A_1 und die Elemente von A_2 eineindeutig einander zuordnen lassen, d. h., wenn jedem Element von A_1 ein einziges Element von A_2 zugeordnet werden kann und umgekehrt jedem Element aus A_2 ein und nur ein Element aus A_1 entspricht. Eine unendliche Menge (d. h. eine Menge, die unendlich viele Elemente enthält) heißt *abzählbar*, wenn sie die gleiche Mächtigkeit hat wie die Menge aller natürlichen Zahlen, d. h., wenn die Elemente dieser Menge mit Hilfe der natürlichen Zahlen numeriert werden können: $a_1, a_2, a_3, \dots$ Zwei abzählbare Mengen haben die gleiche Mächtigkeit.

Wir leiten nun einige Eigenschaften abzählbarer Mengen her. Es sei $p_1, p_2, p_3, \dots$ eine monoton wachsende Folge natürlicher Zahlen. Wir wählen aus einer abzählbaren Menge eine unendliche Teilmenge aus, die aus den Elementen $a_{p_1}, a_{p_2}, a_{p_3}, \dots$ besteht. Die Elemente dieser neuen Menge kann man ebenfalls numerieren. Nummer eines jeden Elementes ist die bei p stehende Indexzahl. Mit anderen Worten, die Elemente werden mit den wachsenden Indizes $p_1, p_2, \dots$ numeriert. *Jede unendliche Teilmenge einer abzählbaren Menge ist also wieder eine abzählbare Menge.* Wir betrachten jetzt zwei abzählbare Mengen: die aus den Elementen $a_1, a_2, a_3, \dots$ bestehende Menge $A(a_1, a_2, a_3, \dots)$ und die aus den Elementen $b_1, b_2, b_3, \dots$ bestehende Menge $B(b_1, b_2, b_3, \dots)$.

Wir bilden die Summe der beiden Mengen, d. h., wir vereinigen zu einer einzigen Menge C alle Elemente, die zu einer der beiden obigen Mengen gehören. Die so erhaltene neue Menge C nennt man gewöhnlich *Summe*¹⁾ der Mengen A und B ; sie ist ebenfalls abzählbar. Man braucht die Elemente von C beispielsweise nur in der Reihenfolge $a_1, b_1, a_2, b_2, \dots$ anzuordnen, um sich von der Abzählbarkeit zu überzeugen. Wenn gleiche Elemente a_k, b_l auftreten, kann man eines von ihnen festhalten und die übrigen streichen. Analoges gilt auch für die Summe endlich vieler abzählbarer Mengen, d. h., *die Summe endlich vieler abzählbarer Mengen ist wieder abzählbar.*

Wir nehmen nun an, es seien abzählbar viele abzählbare Mengen gegeben. Die Elemente aller dieser Mengen bezeichnet man mit einem Buchstaben mit zwei

¹⁾ In der deutschen Literatur spricht man auch häufig von der *Vereinigung* der Mengen A und B (Anm. d. Red.).

ganzzahligen Indizes: $a_p^{(q)}$. Der obere Index gibt die Nummer der Menge an, zu der das Element gehört, der untere Index die Nummer des Elementes in dieser (abzählbaren) Menge, zu der es gehört. Es ist nicht schwer, alle Elemente $a_p^{(q)}$ zu numerieren. Als erstes Element nehmen wir dasjenige, bei dem beide Indizes gleich 1 sind: $a_1^{(1)}$. Dann nehmen wir die Elemente, bei denen die Summe der Indizes gleich 3 ist, und ordnen sie nach wachsendem oberem Index. Wir erhalten so das zweite Element $a_2^{(1)}$ und das dritte Element $a_1^{(2)}$. Wir wählen nun die Elemente, bei denen die Summe der Indizes gleich 4 ist, und ordnen sie nach wachsendem oberem Index: $a_3^{(1)}$, $a_2^{(2)}$, $a_1^{(3)}$. Dadurch erhalten wir das vierte, fünfte und sechste Element. Dieses Verfahren lehrt, daß *die Summe von abzählbar vielen abzählbaren Mengen wieder eine abzählbare Menge ist*. Diese Behauptung gilt offenbar auch, wenn einige der zu summierenden Mengen nicht abzählbare, sondern endliche Mengen sind.

Es sei jetzt eine unendliche Menge A gegeben. Wir wählen aus A ein beliebiges Element und ordnen ihm die Nummer 1 zu. Die übrigbleibende Menge ist nach wie vor unendlich. Wir wählen aus ihr ein weiteres Element aus und ordnen ihm die Nummer 2 zu. Wenn wir dieses Verfahren fortsetzen, sehen wir, daß *auf diese Weise von jeder unendlichen Menge eine abzählbare Menge abgespalten werden kann*. Die nach einer solchen Abspaltung übrigbleibende Menge kann entweder leer sein (d. h., sie braucht kein einziges Element zu enthalten), oder sie kann endlich bzw. unendlich sein. Wir zeigen: Ist die übrigbleibende Menge unendlich, so hat sie die gleiche Mächtigkeit wie die ursprüngliche Menge, d. h., es gilt die folgende Behauptung: *Wenn beim Abspalten einer abzählbaren Menge P von der unendlichen Menge A eine unendliche Menge B übrigbleibt, dann haben die Mengen A und B die gleiche Mächtigkeit*. Wir spalten von der unendlichen Menge B nochmals eine abzählbare Menge Q ab, und es sei C die übrigbleibende Menge. Die ursprüngliche Menge A wird also in drei Mengen zerlegt, $A = P + Q + C$, von denen die Menge C auch leer oder unendlich sein kann und die Mengen P und Q abzählbar sind. Vor der zweiten Abspaltung war $A = P + B$. Es ist nicht schwierig, die Elemente von A und B eineindeutig einander zuzuordnen. In der Tat gilt $A = P + Q + C$ und $B = Q + C$. Die Menge $P + Q$ ist als Summe abzählbarer Mengen wieder abzählbar, daher kann man die Elemente von $P + Q$ und Q eineindeutig einander zuordnen. Jedes Element der Menge C wird sich selbst zugeordnet. Auf diese Weise entsteht eine eineindeutige Zuordnung der Elemente von A und B . Aus dem soeben Bewiesenen folgt unmittelbar: *Wenn man zu einer unendlichen Menge eine abzählbare Menge addiert, dann hat die so erhaltene Menge die gleiche Mächtigkeit wie die ursprüngliche Menge*. Diese beiden Behauptungen über die Subtraktion und Addition einer abzählbaren Menge gelten auch, wenn man die abzählbare Menge durch eine endliche Menge ersetzt. Das kann man genauso beweisen.

Wir haben früher gezeigt [IV/1,15], daß die Menge der in einem Intervall $[a, b]$ liegenden rationalen Zahlen sowie die Menge aller rationalen Zahlen abzählbar ist. Das läßt sich auf dieselbe Art beweisen wie die Behauptung, daß die Summe von abzählbar vielen abzählbaren Mengen abzählbar ist. Den oberen Indizes entsprechen jetzt die Zähler, den unteren Indizes die Nenner der Brüche. Dabei betrachtet man zuerst nur die positiven Brüche. Wir bringen nun ein Beispiel für eine nicht abzählbare Menge. Wir betrachten alle reellen Zahlen im Intervall $[0, 1]$. Jede dieser Zahlen, mit Ausnahme der Null, läßt sich als unendlicher Dezimalbruch mit verschwindendem ganzem Anteil darstellen. Umgekehrt läßt sich jedem solchen Dezimalbruch eine gewisse reelle Zahl aus dem Intervall $[0, 1]$ zuordnen. Mit endlichen Dezimalbrüchen beschäftigen wir uns nicht, weil sie dieselben Zahlen dar-

stellen wie unendliche Dezimalbrüche mit der Periode 9, etwa $0,37 = 0,36999\dots$. Wir beweisen nun indirekt, daß die Menge dieser reellen Zahlen nicht abzählbar ist. Wir nehmen an, daß wir alle betrachteten Dezimalbrüche samt dem Dezimalbruch $0,00\dots$, der dem linken Randpunkt des Intervalls entspricht, numerieren können. Wir bilden nun folgendermaßen einen neuen Dezimalbruch mit verschwindendem ganzem Anteil. Als erste Ziffer nach dem Komma wählen wir irgendeine Ziffer, die von der ersten nach dem Komma stehenden Ziffer des ersten der numerierten Dezimalbrüche verschieden ist, als zweite Ziffer wählen wir irgendeine Ziffer, die von der zweiten Ziffer des zweiten der numerierten Dezimalbrüche verschieden ist, usw. Die Zahl 0 wird dabei nicht verwendet. Es ergibt sich ein unendlicher Dezimalbruch, der von allen numerierten Dezimalbrüchen verschieden ist. Also ist die ihm entsprechende reelle Zahl noch nicht numeriert worden, was im Widerspruch zu unserer Annahme steht, daß alle reellen Zahlen im Intervall $[0, 1]$ numeriert worden seien. Wir haben also bewiesen, daß *die Menge aller reellen Zahlen im Intervall $[0, 1]$ nicht abzählbar ist*. Man sagt, diese Menge habe die *Mächtigkeit des Kontinuums*. Man sieht leicht, daß die Menge der reellen Zahlen in einem beliebigen endlichen Intervall $[a, b]$ offenbar die gleiche Mächtigkeit wie die Menge der reellen Zahlen im Intervall $[0, 1]$ hat; denn die Elemente der beiden Mengen können durch die Relation $y = \frac{x-a}{b-a}$ einindeutig einander zugeordnet werden. Wenn die Veränderliche x das Intervall $[a, b]$ durchläuft, durchläuft die Veränderliche y das Intervall $[0, 1]$. Wenn wir die Relation $y = \tan\left(\pi x - \frac{\pi}{2}\right)$ benutzen, so durchläuft bei der Variation von x innerhalb von $[0, 1]$ die Veränderliche y die Menge aller reellen Zahlen, d. h., *die Menge aller reellen Zahlen hat ebenfalls die Mächtigkeit des Kontinuums*. Wenn die Randpunkte des Intervalls zur Menge gehören, ändert sich die Mächtigkeit der Menge nicht, denn die Mächtigkeit einer unendlichen Menge bleibt erhalten, wenn man eine endliche Menge addiert oder subtrahiert.

Im folgenden werden wir ein abgeschlossenes Intervall mit $[a, b]$, ein offenes Intervall, d. h. ein Intervall ohne Randpunkte, mit (a, b) bezeichnen. Wenn der linke Randpunkt nicht, der rechte dagegen wohl zum Intervall gehört, wird die Bezeichnung $(a, b]$ benutzt, entsprechend $[a, b)$. Die Zahlen a und b können auch unendliche Werte annehmen: $a = -\infty$ und $b = \infty$, d. h., die betrachteten Intervalle können sich unendlich weit nach links oder rechts erstrecken. Zum Beispiel schreibt man für das abgeschlossene Intervall, das beide unendlich fernen Elemente enthält, $[-\infty, \infty]$. Entsprechend kann auch eine Funktion $f(x)$ für $x = -\infty$ und für $x = \infty$ definiert werden. Man schreibt dann beispielsweise $f(-\infty)$. Die Stetigkeit für $x = -\infty$ ist gleichbedeutend mit der Bedingung $\lim_{x \rightarrow -\infty} f(x) = f(-\infty)$. Dasselbe gilt für $x \rightarrow \infty$. Man kann hierfür auch die üblichen Bezeichnungen $\lim_{x \rightarrow -\infty} f(x) = f(-\infty + 0)$ und $\lim_{x \rightarrow \infty} f(x) = f(\infty - 0)$ verwenden.

Es ist leicht zu zeigen [I, 43], daß eine im abgeschlossenen Intervall $[-\infty, \infty]$ endliche und stetige Funktion $f(x)$ in diesem Intervall auch gleichmäßig stetig ist.

2. Das Stieltjessche Integral und seine Haupteigenschaften. Wir erinnern an die Definition des Riemannschen Integrals, die wir in den Teilen I bis IV benutzt haben. Es sei $[a, b]$ ein endliches Intervall und $f(x)$ eine in diesem Intervall definierte beschränkte Funktion. Wir zerlegen das Intervall durch Teilpunkte $a = x_0 < x_1 < \dots < x_{n-1} < x_n = b$, wählen in jedem der Teilintervalle $[x_{k-1}, x_k]$ einen gewissen Punkt

ξ_k und bilden den Ausdruck

$$\sigma = \sum_{k=1}^n f(\xi_k) (x_k - x_{k-1}) . \quad (1)$$

Wenn bei beliebiger Verfeinerung der Zerlegung und bei beliebiger Wahl der Punkte ξ_k die Summe auf der rechten Seite von (1) einen endlichen Grenzwert A hat, dann sagt man, $f(x)$ sei in $[a, b]$ integrierbar, und nennt diesen Grenzwert das Integral von $f(x)$ längs $[a, b]$. Es sei δ die größte der Differenzen $x_k - x_{k-1}$. Die beliebige Verfeinerung des Intervalls $[a, b]$ in Teilintervalle ist gleichbedeutend mit $\delta \rightarrow 0$, und die Existenz eines endlichen Grenzwertes A für die Summe (1) ist gleichbedeutend mit der folgenden Aussage: Bei beliebig vorgegebenem positivem ε existiert ein solches positives η , daß

$$\left| A - \sum_{k=1}^n f(\xi_k) [(x_k - x_{k-1})] \right| \leq \varepsilon$$

gilt für $\delta \leq \eta$.

Ganz analog kann man einen allgemeineren Integralbegriff aufstellen, der zum ersten Male von dem holländischen Mathematiker STIELTJES im Jahre 1894 bei der Untersuchung von Kettenbrüchen eingeführt, später weiterentwickelt und auf viele Fragen der Mathematik und der exakten Naturwissenschaften angewendet wurde. In dem endlichen Intervall $[a, b]$ seien zwei Funktionen $f(x)$ und $g(x)$ gegeben, die in jedem Punkt dieses Intervalls endliche Werte annehmen. An Stelle der Summe (1) bilden wir die Summe

$$\sigma = \sum_{k=1}^n f(\xi_k) [g(x_k) - g(x_{k-1})] . \quad (2)$$

Diese Summe nennt man *Riemann-Stieltjessche Summe*. Wenn bei beliebiger Verfeinerung der Zerlegung und bei beliebiger Wahl der Punkte ξ_k diese Summe gegen einen endlichen Grenzwert strebt, sagt man, daß die Funktion $f(x)$ im Intervall $[a, b]$ bezüglich der Funktion $g(x)$ integrierbar ist, und schreibt

$$\int_a^b f(x) dg(x) = \lim \sum_{k=1}^n f(\xi_k) [g(x_k) - g(x_{k-1})] .$$

Beim Riemannschen Integral spielt x die Rolle von $g(x)$. Für das neue Integral gelten offenbar viele Eigenschaften, die analog den Eigenschaften des Riemannschen Integrals sind, und diese Eigenschaften kann man genauso wie die entsprechenden für das Riemannsche Integral beweisen. Wir geben diese Eigenschaften an und nehmen dabei an, daß alle in den Formeln auftretenden Integrale existieren:

$$\left. \begin{aligned} \int_a^b \sum_{k=1}^p a_k f_k dg(x) &= \sum_{k=1}^p a_k \int_a^b f_k(x) dg(x) , \\ \int_a^b f(x) d \sum_{k=1}^p a_k g_k(x) &= \sum_{k=1}^p a_k \int_a^b f(x) dg_k(x) , \\ \int_a^b f(x) dg(x) &= \int_a^c f(x) dg(x) + \int_c^b f(x) dg(x) \end{aligned} \right\} \quad (3)$$

(a_k ist eine Konstante).

Außerdem gilt offenbar

$$\int_a^b dg(x) = g(b) - g(a) . \quad (4)$$

In der ersten und zweiten Formel von (3) folgt aus der Existenz der rechts stehenden Integrale die Existenz des links stehenden Integrals.

Wir gehen ausführlicher auf die partielle Integration ein. Wir setzen voraus, daß das Integral der Funktion $g(x)$ bezüglich der Funktion $f(x)$ existiert, und beweisen, daß dann auch das Integral von $f(x)$ bezüglich $g(x)$ existiert. Dazu ordnen wir die Summe (2) nach den Werten der Funktion $g(x)$, wir sammeln also die Glieder mit gleichem $g(x)$:

$$\sigma = - \sum_{k=1}^{n-1} g(x_k) [f(\xi_{k+1}) - f(\xi_k)] + g(b) f(\xi_n) - g(a) f(\xi_1).$$

Addieren und subtrahieren wir die Differenz

$$[f(x) g(x)]_a^b = f(b) g(b) - f(a) g(a),$$

so erhalten wir

$$\sigma = [f(x) g(x)]_a^b - \left\{ \sum_{k=1}^{n-1} g(x_k) [f(\xi_{k+1}) - f(\xi_k)] + g(a) [f(\xi_1) - f(a)] + g(b) [f(b) - f(\xi_n)] \right\}. \quad (5)$$

In der geschweiften Klammer steht die Riemann-Stieltjessche Summe (2) für das Integral von $g(x)$ bezüglich $f(x)$. Nach Voraussetzung existiert das Integral von $g(x)$ bezüglich $f(x)$, d. h., bei beliebiger Verfeinerung der Zerlegung strebt die geschweifte Klammer gegen dieses Integral. Nach (5) hat also die Summe σ einen Grenzwert, d. h., es existiert das Integral von $f(x)$ bezüglich $g(x)$, und wir können die Formel für die partielle Integration folgendermaßen schreiben:

$$\int_a^b f(x) dg(x) = [f(x) g(x)]_a^b - \int_a^b g(x) df(x) \quad (6)$$

oder

$$\int_a^b f(x) dg(x) + \int_a^b g(x) df(x) = [f(x) g(x)]_a^b, \quad (7)$$

wobei aus der Existenz des einen Integrals in (7) die Existenz des anderen folgt.

Wir betrachten nun zwei Sonderfälle des Stieltjesschen Integrals. Wir nehmen an, daß das Intervall $[a, b]$ durch endlich viele Teilpunkte $a = c_0 < c_1 < \dots < c_{p-1} < c_p = b$ zerlegt wird und daß die Funktion $g(x)$ im Innern jedes der entstehenden Teilintervalle (c_{k-1}, c_k) den konstanten Wert g_k annimmt. Also macht die Funktion $g(x)$ in jedem Punkt c_k im Innern des Intervalls $[a, b]$ einen Sprung von der Größe $s_k = g_{k+1} - g_k$. Auch in den Randpunkten des Intervalls können Sprünge auftreten: ein Sprung von der Größe $s_0 = g_1 - g(a)$ im linken Randpunkt und ein Sprung von der Größe $s_p = g(b) - g_p$ im rechten Randpunkt. Wir nehmen weiter an, daß die Funktion $f(x)$ in allen Sprungstellen c_k und in den Randpunkten des Intervalls stetig ist. Die Punkte c_k seien mit Ausnahme von c_0 und c_p keine Teilpunkte des Intervalls. In der Summe (2) sind alle Summanden, bei denen x_{k-1} und x_k im Innern des gleichen Intervalls (c_{q-1}, c_q) liegen, gleich 0, weil in diesem Fall $g(x_{k-1}) = g(x_k)$ ist. Wenn das Intervall $[x_{k-1}, x_k]$ eine Sprungstelle c_q enthält, streben bei beliebiger Verfeinerung der Zerlegung $f(\xi_k)$ gegen $f(c_q)$ und $g(x_k) - g(x_{k-1})$ gegen s_q . Daher folgt unmittelbar, daß man als Grenzwert der Summe (2) die endliche Summe

$$\lim \sum_{k=1}^n f(\xi_k) [g(x_k) - g(x_{k-1})] = \sum_{q=0}^n f(c_q) s_q \quad (8)$$

erhält.

Ist der Punkt c_q ein Teilpunkt von $[a, b]$, so kann man die beiden angrenzenden Intervalle betrachten und zu dem gleichen Ergebnis kommen. Beim zweiten Sonderfall des Stieltjesschen Integrals nehmen wir an, daß $f(x)$ und $g(x)$ in $[a, b]$ stetig sind und daß $g(x)$ im Innern von $[a, b]$ eine Ableitung $g'(x)$ besitzt, die im Riemannschen Sinne integrierbar und folglich beschränkt ist. Wenn man auf die Differenz $g(x_k) - g(x_{k-1})$ die Lagrangesche Formel¹⁾ anwendet, erhält man für die Summe (2)

$$\sum_{k=1}^n f(\xi_k) [g(x_k) - g(x_{k-1})] = \sum_{k=1}^n f(\xi_k) g'(\xi'_k) (x_k - x_{k-1}), \quad (9)$$

wobei ξ'_k ein Wert im Innern von $[x_{k-1}, x_k]$ ist. Setzt man $f(\xi_k) = f(\xi'_k) + \varepsilon_k$, so strebt wegen der gleichmäßigen Stetigkeit von $f(x)$ in $[a, b]$ der größte der Werte $|\varepsilon_k|$ bei beliebiger Verfeinerung der Zerlegung gegen 0, d. h., bei beliebig vorgegebenem positivem ε existiert ein solches positives η , daß $|\varepsilon_k| < \varepsilon$ ist für $\delta < \eta$. Die Summe (9) kann man nun wie folgt schreiben:

$$\sum_{k=1}^n f(\xi_k) [g(x_k) - g(x_{k-1})] = \sum_{k=1}^n f(\xi'_k) g'(\xi'_k) (x_k - x_{k-1}) + \sum_{k=1}^n \varepsilon_k g'(\xi'_k) (x_k - x_{k-1}). \quad (9_1)$$

Das Produkt zweier im Riemannschen Sinne integrierbarer Funktionen ist wieder integrierbar [I, 117], und der erste Summand auf der rechten Seite der letzten Formel strebt bei beliebiger Verfeinerung der Zerlegung gegen das Riemannsche Integral des Produkts $f(x) g'(x)$. Wie man leicht sieht, strebt der zweite Summand auf der rechten Seite gegen 0. Die Funktion $g'(x)$ ist nach Voraussetzung beschränkt, d. h. $|g'(x)| < M$, wobei M eine bestimmte positive Zahl ist. Wie wir erwähnt haben, existiert bei vorgegebenem positivem ε ein solches positives η , daß $|\varepsilon_k| < \varepsilon$ ist für $\delta < \eta$, und es gilt die Abschätzung

$$\left| \sum_{k=1}^n \varepsilon_k g'(\xi'_k) (x_k - x_{k-1}) \right| \leq \sum_{k=1}^n \varepsilon M (x_k - x_{k-1}) = \varepsilon M (b - a).$$

Da ε beliebig gewählt war, folgt daraus, daß der zweite Summand auf der rechten Seite der Gleichung (9₁) gegen 0 strebt. Somit erhalten wir als Grenzwert

$$\int_a^b f(x) dg(x) = \int_a^b f(x) g'(x) dx, \quad (10)$$

d. h., unter den obigen Voraussetzungen artet das Stieltjessche Integral in ein gewöhnliches Riemannsches Integral aus. Im vorhergehenden Fall artete es in eine endliche Summe aus. Wie man leicht sieht, gilt die Formel (10) auch, wenn $f(x)$ nicht als stetig, sondern nur als im Riemannschen Sinne integrierbar vorausgesetzt wird. Im folgenden werden wir die Existenz solcher oben definierter Stieltjescher Integrale und gewisser allgemeinerer, noch zu definierender Integrale untersuchen. Eine wesentliche Voraussetzung bei allen diesen Integralen ist, daß die Funktion $g(x)$ in $[a, b]$ nicht abnimmt. Eine nicht abnehmende Funktion werden wir im weiteren oft (monoton) wachsend nennen.²⁾ Bei einer solchen Funktion stellen $g(b)$ den größten und $g(a)$ den kleinsten Wert dar.

Der folgende Abschnitt dient der Vorbereitung für die Untersuchung der Existenz des oben definierten Stieltjesschen Integrals und allgemeinerer Integrale, die wir im weiteren noch definieren werden.

¹⁾ In der deutschen Literatur *Mittelwertsatz der Differentialrechnung* genannt (Anm. d. Red.).

²⁾ Oft auch *im weiteren Sinne monoton wachsend* genannt (Anm. d. Red.).

3. Darboux'sche Summen. Bei der Untersuchung des Riemannschen Integrals führten wir die sogenannten *Darboux'schen Summen* ein. Analoge Summen spielen eine grundlegende Rolle bei allen noch einzuführenden Verallgemeinerungen des Integralbegriffs. In diesem Abschnitt bilden wir die Darboux'schen Summen und untersuchen ihre Eigenschaften für den Fall des Stieltjesschen Integrals. Alle Begriffe, die wir in diesem Abschnitt einführen, und die Behauptungen, die wir dafür beweisen werden, lassen sich mit einigen unwesentlichen Änderungen auch auf die allgemeineren Integralbegriffe übertragen. Wir werden oft auf die Ergebnisse dieses Abschnitts zurückkommen.

Wir erinnern zuerst an die Definition der oberen bzw. unteren Grenze einer Menge reeller Zahlen [I, 39]. Es sei jetzt eine gewisse Menge $\mathfrak{E}$ reeller Zahlen gegeben. Diese Menge sei nach oben beschränkt, d. h., es existiere eine solche Zahl L , daß alle Zahlen der Menge kleiner als L sind. Dann existiert eine Zahl M mit folgender Eigenschaft: Keine Zahl der Menge $\mathfrak{E}$ ist größer als M , aber bei einem beliebigen positiven ε gibt es Zahlen der Menge $\mathfrak{E}$, die größer als $M - \varepsilon$ sind. Diese Zahl M heißt *obere Grenze* der Menge $\mathfrak{E}$. Ist eine Menge $\mathfrak{E}$ nach unten beschränkt, d. h., sind alle Zahlen der Menge größer als eine gewisse Zahl, so hat die Menge ganz analog eine *untere Grenze* m mit der folgenden Eigenschaft: Keine Zahl der Menge $\mathfrak{E}$ ist kleiner als m , aber bei einem beliebigen positiven ε gibt es Zahlen der Menge $\mathfrak{E}$, die kleiner als $m + \varepsilon$ sind. Ist eine Menge nicht nach oben beschränkt, so sagt man, ihre obere Grenze sei $+\infty$ (oder kürzer ∞); ebenso sagt man, wenn eine Menge nicht nach unten beschränkt ist, ihre untere Grenze sei $-\infty$. Zur Bezeichnung der Grenzen wird die folgende Schreibweise benutzt:

$$m = \inf \mathfrak{E}, \quad M = \sup \mathfrak{E}.$$

Es seien $f(x)$ und $g(x)$ in einem endlichen oder unendlichen Intervall $[a, b]$ beschränkte Funktionen, $g(x)$ möge nicht abnehmen, und durch die Teilpunkte

$$a = x_0 < x_1 < x_2 < \dots < x_{n-1} < x_n = b$$

sei eine Zerlegung des Intervalls $[a, b]$ gegeben, die wir mit δ bezeichnen wollen. Ist das Intervall nach links nicht beschränkt, so gilt $a = -\infty$; ist das Intervall nach rechts nicht beschränkt, so gilt $b = \infty$. Weiter seien m_k die untere und M_k die obere Grenze der Werte von $f(x)$ im Intervall $[x_{k-1}, x_k]$. Wir bilden die Darboux-Stieltjesschen Summen

$$s_\delta = \sum_{k=1}^n m_k [g(x_k) - g(x_{k-1})], \quad S_\delta = \sum_{k=1}^n M_k [g(x_k) - g(x_{k-1})], \quad (11)$$

die der gewählten Zerlegung δ des Intervalls $[a, b]$ entsprechen. Da die Funktion $f(x)$ beschränkt ist, gilt für eine gewisse positive Zahl L die Abschätzung $|f(x)| \leq L$. Mit Rücksicht darauf, daß $g(x_k) - g(x_{k-1}) \geq 0$ ist, erhalten wir für die Summen (11) bei beliebiger Zerlegung δ die Abschätzungen

$$|s_\delta| \leq \sum_{k=1}^n L [g(x_k) - g(x_{k-1})] = L [g(b) - g(a)],$$

$$|S_\delta| \leq L [g(b) - g(a)].$$

Neben den Summen (11) bilden wir die Riemann-Stieltjessche Summe

$$\sigma_\delta = \sum_{k=1}^n f(\xi_k) [g(x_k) - g(x_{k-1})], \quad (12)$$

wobei ξ_k ein gewisser Wert im Intervall $[x_{k-1}, x_k]$ ist. Da $m_k \leq f(\xi_k) \leq M_k$ und $g(x_k) - g(x_{k-1}) \geq 0$ ist, gilt für jede Zerlegung δ die Ungleichung

$$s_\delta \leq \sigma_\delta \leq S_\delta. \quad (13)$$

Wir führen jetzt einige neue Begriffe ein. Eine Zerlegung δ' heißt *Fortsetzung der Zerlegung* δ , wenn alle Teilpunkte von δ auch Teilpunkte von δ' sind. Es seien δ_1 und δ_2 zwei beliebige Zerlegungen. Wir bilden eine neue Zerlegung aus den Teilpunkten von δ_1 und von δ_2 . Diese neue Zerlegung nennen wir das *Produkt der Zerlegungen* δ_1 und δ_2 und bezeichnen sie mit $\delta_1\delta_2$. Die Zerlegung $\delta_1\delta_2$ bildet offenbar eine *Fortsetzung der Zerlegung* δ_1 als auch der Zerlegung δ_2 . Auch für eine beliebige endliche Anzahl von Zerlegungen $\delta_1, \delta_2, \dots, \delta_n$ kann man den Begriff des Produkts $\delta_1\delta_2 \dots \delta_n$ einführen. Wir erwähnen noch, daß die Summen s_δ und S_δ nur von der Wahl der Zerlegung δ abhängen, die Summe s_δ dagegen hängt noch von der Wahl der Punkte ξ_k ab. Wir beweisen nun einige elementare Sätze.

Satz 1. *Ist die Zerlegung δ' eine Fortsetzung der Zerlegung δ , so ist $s_{\delta'} \geq s_\delta$ und $S_{\delta'} \leq S_\delta$.*

Wir beweisen als Beispiel die Ungleichung $s_{\delta'} \geq s_\delta$. Beim Übergang von δ zu δ' kann jedes zur Zerlegung δ gehörige Teilintervall durch die Teilpunkte

$$x_{k-1} = x_0^{(k)} < x_1^{(k)} < \dots < x_{p_k-1}^{(k)} < x_{p_k}^{(k)} = x_k$$

in eine endliche Anzahl von Teilen zerlegt werden, und an Stelle der Summanden $m_k [g(x_k) - g(x_{k-1})]$ der Summe s_δ erhalten wir die Summe

$$\sum_{s=1}^{p_k} m_s^{(k)} [g(x_s^{(k)}) - g(x_{s-1}^{(k)})],$$

worin $m_s^{(k)}$ die untere Grenze der Werte von $f(x)$ im Intervall $[x_{s-1}^{(k)}, x_s^{(k)}]$ ist. Offensichtlich ist $m_s^{(k)} \geq m_k$, und folglich gilt mit Rücksicht darauf, daß die Differenzen $g(x_s^{(k)}) - g(x_{s-1}^{(k)})$ nicht negativ sind,

$$\sum_{s=1}^{p_k} m_s^{(k)} [g(x_s^{(k)}) - g(x_{s-1}^{(k)})] \geq \sum_{s=1}^{p_k} m_k [g(x_s^{(k)}) - g(x_{s-1}^{(k)})] = m_k [g(x_k) - g(x_{k-1})],$$

womit die Behauptung bewiesen ist (vgl. [I, 112]).

Satz 2. *Sind δ_1 und δ_2 zwei beliebige Zerlegungen, so ist $s_{\delta_1} \leq S_{\delta_2}$.*

Beweis. Die Ungleichung $s_\delta \leq S_\delta$ für ein und dieselbe Zerlegung δ folgt unmittelbar aus $m_k \leq M_k$ und $g(x_k) - g(x_{k-1}) \geq 0$. Für die Zerlegung $\delta_1\delta_2$ gilt also $s_{\delta_1\delta_2} \leq S_{\delta_1\delta_2}$. Nach Satz 1 ist andererseits $s_{\delta_1} \leq s_{\delta_1\delta_2}$ und $S_{\delta_2} \leq S_{\delta_1\delta_2}$, woraus $s_{\delta_1} \leq S_{\delta_2}$ folgt.

Wir bezeichnen mit i die obere Grenze der Summen s_δ bei allen möglichen Zerlegungen δ und mit I die untere Grenze der Summen S_δ :

$$i = \sup s_\delta, \quad I = \inf S_\delta. \quad (14)$$

Aus der Definition der oberen und unteren Grenze und aus Satz 2 folgt für beliebige Zerlegungen δ_1 und δ_2 unmittelbar die Ungleichung $s_{\delta_1} \leq i \leq I \leq S_{\delta_2}$, also auch speziell

$$s_\delta \leq i \leq I \leq S_\delta. \quad (15)$$

Wir leiten nun eine notwendige und hinreichende Bedingung dafür ab, daß die untere Grenze i gleich der oberen Grenze I ist. Eine wesentliche Rolle spielt dabei die Differenz

$$S_\delta - s_\delta = \sum_{k=1}^n (M_k - m_k) [g(x_k) - g(x_{k-1})]. \quad (16)$$

Satz 3. Für die Gleichheit von i und I ist notwendig und hinreichend, daß eine Folge von Zerlegungen δ_n ($n=1, 2, \dots$) existiert, für die $S_{\delta_n} - s_{\delta_n} \rightarrow 0$ gilt.

Wir beweisen zuerst, daß diese Bedingung hinreichend ist. Wenn eine Folge von Zerlegungen δ_n existiert, für die $S_{\delta_n} - s_{\delta_n} \rightarrow 0$ gilt, und wenn wir auf diese Folge die Ungleichung (15) anwenden, erhalten wir $i = I$.

Wir beweisen jetzt die Notwendigkeit: Es sei $i = I = A$. Nach der Definition der oberen und unteren Grenze existiert eine Folge von Zerlegungen δ'_n , für die $s_{\delta'_n} \rightarrow A$ gilt, und eine Folge von Zerlegungen δ''_n , für die $S_{\delta''_n} \rightarrow A$ gilt. Wir nehmen die Folge der Zerlegungen $\delta_n = \delta'_n \delta''_n$. Nach Satz 1 gilt $s_{\delta_n} \cong s_{\delta'_n}$ und $S_{\delta_n} \cong S_{\delta''_n}$, wobei $s_{\delta'_n}, s_{\delta_n} \cong A$ und $S_{\delta''_n}, S_{\delta_n} \cong A$ ist. Also gilt um so mehr $s_{\delta_n} \rightarrow A$ und $S_{\delta_n} \rightarrow A$, folglich $S_{\delta_n} - s_{\delta_n} \rightarrow 0$, womit der Satz bewiesen ist.

Wir erwähnen, daß in den Zerlegungen δ_n die Teilintervalle nicht notwendig unbegrenzt kleiner werden müssen. Es kann z. B. vorkommen, daß alle Zerlegungen δ_n ein und derselben Zerlegung δ gleich sind. Aus (15) kann man unmittelbar schließen:

Folgerung. Gilt $S_{\delta_n} - s_{\delta_n} \rightarrow 0$, so ist $i = I$, $s_{\delta_n} \rightarrow i$ und $S_{\delta_n} \rightarrow i$.

Die in Satz 3 dargestellte notwendige und hinreichende Bedingung für die Gleichheit $i = I$ läßt sich durch die Summen σ_δ ausdrücken.

Satz 4. Die Differenz $S_{\delta_n} - s_{\delta_n}$ strebt genau dann gegen 0, wenn σ_{δ_n} bei beliebiger Wahl der Punkte $\xi_k^{(n)}$ einen endlichen Grenzwert hat. Wenn diese Bedingung erfüllt ist, ist der Grenzwert von σ_{δ_n} gleich i (oder $I = i$).

Beweis. Wir beweisen zuerst die Notwendigkeit. Aus $S_{\delta_n} - s_{\delta_n} \rightarrow 0$ folgt, wie wir sahen, $s_{\delta_n} \rightarrow i$ und $S_{\delta_n} \rightarrow i$; also folgt aus (13) $\sigma_{\delta_n} \rightarrow i$. Wir zeigen nun, daß die Bedingung hinreichend ist. Es gelte

$$\sigma_{\delta_n} = \sum_{k=1}^{p_n} f(\xi_k^{(n)}) [g(x_k^{(n)}) - g(x_{k-1}^{(n)})] \rightarrow A,$$

wobei $x_k^{(n)}$ Teilpunkte von δ_n und $\xi_k^{(n)}$ gewisse Punkte im Intervall $[x_{k-1}^{(n)}, x_k^{(n)}]$ sind. Wir bezeichnen noch mit $m_k^{(n)}$ und $M_k^{(n)}$ die untere und obere Grenze der Werte von $f(x)$ im Intervall $[x_{k-1}^{(n)}, x_k^{(n)}]$. Es sei ε eine beliebig vorgegebene positive Zahl. Auf Grund der Bedingung $\sigma_{\delta_n} \rightarrow A$ existiert ein solches N , daß

$$|A - \sigma_{\delta_n}| \leq \varepsilon \quad (17)$$

gilt für $n > N$ und für beliebig gewählte Punkte $\xi_k^{(n)}$. Nach der Definition der unteren Grenze können wir die Punkte $\xi_k^{(n)}$ so wählen, daß die Ungleichungen $0 \leq f(\xi_k^{(n)}) - m_k^{(n)} \leq \varepsilon$ erfüllt sind. Dann haben wir

$$\begin{aligned} 0 \leq \sigma_{\delta_n} - s_{\delta_n} &= \sum_{k=1}^{p_n} [f(\xi_k^{(n)}) - m_k^{(n)}] [g(x_k^{(n)}) - g(x_{k-1}^{(n)})] \\ &\leq \sum_{k=1}^{p_n} \varepsilon [g(x_k^{(n)}) - g(x_{k-1}^{(n)})] = \varepsilon [g(b) - g(a)] \end{aligned} \quad (18)$$

und erhalten also, wenn wir die Differenz $A - s_{\delta_n}$ in der Form $A - s_{\delta_n} = (A - \sigma_{\delta_n}) + (\sigma_{\delta_n} - s_{\delta_n})$ darstellen, mit (17) und (18) $|A - s_{\delta_n}| \leq |A - \sigma_{\delta_n}| + |\sigma_{\delta_n} - s_{\delta_n}| \leq \varepsilon [1 +$

$+g(b) - g(a)]$ für $n > N$, woraus, da ε beliebig war, $s_{\delta_n} \rightarrow A$ folgt. Auf genau die gleiche Weise kann man zeigen, daß $S_{\delta_n} \rightarrow A$ und folglich $S_{\delta_n} - s_{\delta_n} \rightarrow 0$ gilt; damit ist der Satz bewiesen.

Der Grenzwert A stimmt offenbar mit den Zahlen i und I überein, die im betrachteten Fall einander gleich sind. Aus Satz 4 und aus früheren Sätzen ergibt sich unmittelbar die nachstehende

Folgerung. Für die Gültigkeit der Gleichung $i = I$ ist notwendig und hinreichend, daß eine Folge von Zerlegungen δ_n existiert, für welche die Folge der σ_{δ_n} bei beliebiger Wahl der Punkte $\xi_k^{(n)}$ einen endlichen Grenzwert hat. Wenn die Bedingung erfüllt ist, ist dieser Grenzwert gleich i (oder $i = I$).

Satz 5. Wenn σ_{δ_n} für eine Folge von Zerlegungen δ_n einen endlichen Grenzwert hat und δ'_n eine Fortsetzung von δ_n ist, hat $\sigma_{\delta'_n}$ denselben Grenzwert.

Aus den Voraussetzungen von Satz 5 und aus Satz 4 folgt, daß $S_{\delta_n} - s_{\delta_n} \rightarrow 0$ strebt. Nach Satz 1 ist $s_{\delta'_n} \equiv s_{\delta_n}$ und $S_{\delta'_n} \equiv S_{\delta_n}$. Also gilt erst recht $S_{\delta_n} - s_{\delta'_n} \rightarrow 0$, d. h. $\sigma_{\delta'_n} \rightarrow i$, womit der Satz bewiesen ist.

Im Fall Riemannscher Integrale, d. h. für $g(x) = x$, erhielten wir früher [I, 112] für eine beliebige beschränkte Funktion $f(x)$ bei beliebig kleiner Zerlegung der Intervalle $s_{\delta_n} \rightarrow i$ und $S_{\delta_n} \rightarrow I$. Im Fall Riemannscher Integrale ist also die Beziehung $i = I$ gleichbedeutend damit, daß die Summe σ_δ bei beliebiger Verfeinerung der Intervallzerlegung einen endlichen Grenzwert hat. Dann ist dieser Grenzwert gleich i . Im allgemeinen Fall gilt das nicht mehr. Wenn σ_δ bei beliebiger Verfeinerung der Intervallzerlegung einen endlichen Grenzwert hat, gilt nach der Folgerung aus Satz 4 die Gleichheit $i = I$. Die Umkehrung läßt sich nicht beweisen. Aus $i = I$ folgt nur, daß eine Folge von Zerlegungen δ_n existiert, für die σ_{δ_n} einen endlichen Grenzwert hat. Man kann aber nicht zeigen, daß σ_δ für jede Folge von Zerlegungen einen endlichen Grenzwert hat, sobald die Länge der Teilintervalle gegen 0 strebt. In der obigen Definition des Stieltjesschen Integrals verlangten wir, daß σ_δ bei beliebiger Verfeinerung der Intervallteilung einen endlichen Grenzwert besitzt. Bei weiteren Verallgemeinerungen des Integralbegriffs ersetzen wir diese Forderung durch die schwächere Forderung, daß die Beziehung $i = I$ gilt. Außerdem erweitern wir unsere Möglichkeiten für die Zerlegung des Grundintegrationsintervalls in Teilintervalle, was bei den neuen Integraldefinitionen erklärt wird. Im folgenden Abschnitt kommen wir auf die in [2] definierten Stieltjesschen Integrale zurück und geben eine wichtige hinreichende Bedingung für ihre Existenz.

4. Das Stieltjessche Integral einer stetigen Funktion.

Satz 1. Wenn die Funktion $f(x)$ in einem endlichen Intervall $[a, b]$ stetig ist und die Funktion $g(x)$ nicht abnimmt und beschränkt ist, so existiert das Stieltjessche Integral von $f(x)$ bezüglich $g(x)$ in $[a, b]$.

Unter Benutzung der Ungleichungen (13) und (15) kann man

$$|i - \sigma_\delta| \leq S_\delta - s_\delta = \sum_{k=1}^n (M_k - m_k) [g(x_k) - g(x_{k-1})] \quad (19)$$

schreiben. Es sei ε eine gegebene positive Zahl. Wegen der gleichmäßigen Stetigkeit von $f(x)$ existiert für $[a, b]$ eine solche positive Zahl η , daß $0 \leq M_k - m_k \leq \varepsilon$ ($k =$

$= 1, 2, \dots, n$) gilt, wenn die größte der Differenzen $x_k - x_{k-1}$ nicht größer als η ist. Für die Ungleichung (19) erhält man dann $|i - \sigma_\delta| \leq \varepsilon [g(b) - g(a)]$, und folglich gilt bei beliebiger Verfeinerung der Zerlegung des Intervalls $\sigma_\delta \rightarrow i$. In der gleichen Weise kann man zeigen, daß σ_δ gegen I strebt und folglich $i = I$ gilt. Diese Gleichheit ergibt sich auch unmittelbar aus der Folgerung des Satzes 4 aus [3], da σ_δ bei beliebiger Verfeinerung der Zerlegung einen endlichen Grenzwert besitzt.

Die Verwendung eines unendlichen Integrationsintervalls bringt im Fall des Stieltjesschen Integrals nichts wesentlich Neues mit sich. Man muß nur erklären, was man unter beliebiger Verfeinerung der Intervallteilung eines unendlichen Intervalls versteht. Betrachtet man z. B. das Intervall $[-\infty, \infty]$, so sagt man, daß sich für eine Folge von Zerlegungen dieses Intervalls in endlich viele Teilintervalle die Intervallteilung beliebig verfeinert, wenn bei einem beliebig vorgegebenen positiven A die größte der Differenzen $x_k - x_{k-1}$ für diejenigen Intervalle $[x_{k-1}, x_k]$, die mit $[-A, A]$ gemeinsame Punkte haben, gegen 0 strebt. Wenn $\varphi(x)$ im Intervall $[-\infty, \infty]$ stetig ist und monoton wächst, d. h., wenn für $\beta > \alpha$ stets $\varphi(\beta) > \varphi(\alpha)$ ist, dann läßt sich durch eine Variablentransformation $t = \varphi(x)$ das Intervall $-\infty \leq x \leq +\infty$ in ein endliches Intervall $[a, b]$ verwandeln. Dabei ist $a = \varphi(-\infty)$ und $b = \varphi(\infty)$. Die Zerlegung und die beliebige Verfeinerung der Zerlegung für das Intervall $[-\infty, \infty]$ entsprechen dabei der gewöhnlichen Zerlegung und der beliebigen Verfeinerung der Zerlegung für das endliche Intervall $[a, b]$.

Wenn z. B. in dem abgeschlossenen Intervall $[-\infty, \infty]$ die Funktion $f(x)$ stetig ist und die Funktion $g(x)$ beschränkt ist und nicht abnimmt, existiert das Integral. Man kann sich etwa davon überzeugen, wenn man für x eine neue Veränderliche $t = \arctan x$ einführt. Setzt man $f(\tan t) = f_1(t)$ und $g(\tan t) = g_1(t)$, so kann man das Integral über das unendliche Integrationsintervall $[-\infty, \infty]$ durch ein Integral über das endliche Integrationsintervall $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ darstellen:

$$\int_{-\infty}^{\infty} f(x) dg(x) = \int_{-\frac{\pi}{2}}^{+\frac{\pi}{2}} f_1(t) dg_1(t).$$

In dem Intervall $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ ist dabei $f_1(t)$ stetig und $g_1(t)$ beschränkt und nicht abnehmend.

Wir geben nun eine sehr wichtige Änderung des eben behandelten grundlegenden Existenzsatzes für Stieltjessche Integrale an.

Satz 2. *Wenn die Funktion $f(x)$ beschränkt und im Innern des Integrationsintervalls stetig ist und die Funktion $g(x)$ nicht abnimmt und in den Randpunkten des Intervalls stetig ist, ist $f(x)$ bezüglich $g(x)$ integrierbar.*

Beweis. Wir untersuchen den Fall des unendlichen Integrationsintervalls $[-\infty, \infty]$. Wir schätzen die auf der rechten Seite der Formel (19) stehenden Summanden ab. Da $f(x)$ beschränkt ist, existiert eine positive Zahl L mit $|f(x)| \leq L$. Also ist auch $0 \leq M_k - m_k \leq 2L$. Diejenigen Summanden von (19), die solchen Intervallen $[x_{k-1}, x_k]$ entsprechen, die punktfremd zu dem Intervall $[-A, A]$ sind, ergeben eine Summe, die nicht größer als

$$2L [g(-A) - g(-\infty)] + 2L [g(\infty) - g(A)] \quad (20)$$

ist. Da $g(x)$ für $x = \pm \infty$ als stetig vorausgesetzt worden ist, kann man A so groß wählen, daß der Ausdruck (20) kleiner ist als ein beliebig vorgegebenes positives ε .

Wir legen auf diese Weise A fest und betrachten die anderen Summanden der Summe (19). Entweder liegen alle zugehörigen Intervalle $[x_{k-1}, x_k]$ vollständig in $[-A, A]$, oder die beiden Randintervalle können aus $[-A, A]$ hinausragen, wobei dann die Länge der hinausragenden Teile nicht größer als η ist, wenn η die größte der Differenzen $x_k - x_{k-1}$ ist für die Intervalle, die mit $[-A, A]$ gemeinsame Punkte haben. Bei beliebiger Verfeinerung der Intervallteilung strebt η gegen 0. Wir wählen die Anfangsstufe der Zerlegung schon so fein, daß auf jeden Fall $\eta < 1$ ist. Dadurch liegen alle betrachteten $[x_{k-1}, x_k]$ im Intervall $[-A-1, A+1]$, in dem die Funktion $f(x)$ gleichmäßig stetig ist. Folglich gilt für alle hinreichend kleinen Werte von δ die Abschätzung $0 \leq M_k - m_k \leq \varepsilon$, und damit gilt für die Summanden der Summe (19), deren zugehörige Intervalle $[x_{k-1}, x_k]$ mit $[-A, A]$ gemeinsame Punkte haben, die Abschätzung

$$0 \leq (M_k - m_k) [g(x_k) - g(x_{k-1})] \leq \varepsilon [g(x_k) - g(x_{k-1})],$$

und die Summe dieser Summanden ist nicht größer als

$$\varepsilon [g(A+1) - g(-A-1)].$$

Die Ungleichung (19) ergibt schließlich

$$|i - \sigma_\delta| \leq \varepsilon [1 + g(A+1) - g(-A-1)] \leq \varepsilon [1 + g(\infty) - g(-\infty)],$$

woraus, da ε beliebig war, $\sigma_\delta \rightarrow i$ folgt. Damit ist der Satz bewiesen.

Wir erwähnen einige zusätzliche Eigenschaften des Stieltjesschen Integrals im Fall einer stetigen Funktion $f(x)$ und einer zunehmenden (wachsenden) Funktion $g(x)$. Ist $|f(x)| \leq L$, so gilt die Abschätzung

$$\left| \int_a^b f(x) dg(x) \right| \leq L [g(b) - g(a)], \quad (21)$$

die sich durch Grenzübergang aus der offenbar geltenden Abschätzung für die Summe σ_δ ergibt. Offenbar gilt der erste Mittelwertsatz der Integralrechnung (vgl. [I, 95])

$$\int_a^b f(x) dg(x) = f(\xi) [g(b) - g(a)] \quad (\xi \text{ aus } [a, b]). \quad (21_1)$$

Wir nehmen jetzt an, daß eine Folge in $[a, b]$ stetiger Funktionen $f_n(x)$ in diesem Intervall gleichmäßig gegen die Grenzfunktion $f(x)$ strebt. Die Grenzfunktion ist ebenfalls in $[a, b]$ stetig und folglich bezüglich $g(x)$ integrierbar. Für ein beliebig vorgegebenes positives ε existiert wegen der gleichmäßigen Konvergenz der Folge der $f_n(x)$ ein solches N , daß $|f(x) - f_n(x)| \leq \varepsilon$ gilt für alle x in $[a, b]$, sobald $n > N$ ist. Mit (21) erhält man

$$\left| \int_a^b [f(x) - f_n(x)] dg(x) \right| \leq \varepsilon [g(b) - g(a)].$$

Da ε beliebig war, folgt daraus

$$\lim_{n \rightarrow \infty} \int_a^b f_n(x) dg(x) = \int_a^b f(x) dg(x). \quad (22)$$

Verwendet man die gleichen Abschätzungen wie beim Beweis des Satzes 2, so kann man leicht beweisen, daß die Gleichung (22) auch unter folgenden Voraussetzungen gültig bleibt: Die Funktionen $f_n(x)$ sind gleichmäßig beschränkt (d. h.,

mit einer von n unabhängigen positiven Zahl L gilt $|f_n(x)| \leq L$ für alle n und im Innern von $[a, b]$ stetig; sie streben in jedem im Innern von $[a, b]$ liegenden abgeschlossenen Intervall gleichmäßig gegen $f(x)$, und die nicht abnehmende Funktion $g(x)$ ist auf den Randpunkten des Intervalls $[a, b]$ stetig.

5. Das uneigentliche Stieltjessche Integral. Wenn $f(x)$ beschränkt und im Innern des Intervalls $[-\infty, \infty]$ stetig ist und $g(x)$ nicht abnimmt und in den Randpunkten stetig ist, so läßt sich, wie wir gesehen haben, das Integral von $f(x)$ bezüglich $g(x)$ im Intervall $[-\infty, \infty]$ in der üblichen Weise als Grenzwert endlicher Summen σ_s definieren. Wir nehmen jetzt an, daß die im Innern des Intervalls $[-\infty, \infty]$ stetige Funktion $f(x)$ nicht beschränkt ist und daß $g(x)$ wie bisher nicht abnimmt und beschränkt ist. Wir können für beliebige endliche a und b im Intervall $[a, b]$ das Integral von $f(x)$ bezüglich $g(x)$ bilden. Wenn bei $a \rightarrow -\infty, b \rightarrow \infty$ dieses Integral einen endlichen Grenzwert hat, so nehmen wir diesen Grenzwert als Wert des Integrals für das Intervall $(-\infty, \infty)$:

$$\int_{-\infty}^{\infty} f(x) dg(x) = \lim_{\substack{a \rightarrow -\infty \\ b \rightarrow \infty}} \int_a^b f(x) dg(x). \quad (23)$$

Wenn die Bedingungen erfüllt sind, die am Anfang dieses Abschnittes angeführt wurden, also das Integral für das Intervall $[-\infty, \infty]$ als Grenzwert der Summen σ_s existiert, dann läßt sich leicht zeigen, daß die Formel (23) gilt.

Wir nehmen an, daß die Integrale $\int_a^b |f(x)| dg(x)$ bei beliebiger Wahl von a und b beschränkt bleiben. Dann existiert das Integral

$$\int_{-\infty}^{\infty} |f(x)| dg(x) = \lim_{\substack{a \rightarrow -\infty \\ b \rightarrow \infty}} \int_a^b |f(x)| dg(x)$$

und offenbar auch das Integral (23) (vgl. [II, 85]), das man in diesem Fall *absolut konvergent* nennt.

Wir betrachten irgendeine Zerlegung des unendlichen Intervalls durch Teilpunkte x_k ($k = \dots, -3, -2, -1, 0, 1, 2, 3, \dots$):

$$\dots < x_{-2} < x_{-1} < x_0 < x_1 < x_2 < \dots \quad \left(\lim_{k \rightarrow -\infty} x_k = -\infty \quad \text{und} \quad \lim_{k \rightarrow \infty} x_k = \infty \right). \quad (24)$$

Es seien m_i und M_i der kleinste bzw. der größte Wert von $f(x)$ im Intervall $[x_{i-1}, x_i]$, und es sei $\omega_i = M_i - m_i$. Mit der Formel (21₁) aus [4] folgt

$$\left| \int_{x_{i-1}}^{x_i} f(x) dg(x) - f(\xi_i) [g(x_i) - g(x_{i-1})] \right| \leq \omega_i [g(x_i) - g(x_{i-1})] \quad (x_{i-1} \leq \xi_i \leq x_i)$$

und

$$\left| \int_{x_{-p}}^{x_q} f(x) dg(x) - \sum_{i=1-p}^q f(\xi_i) [g(x_i) - g(x_{i-1})] \right| \leq \sum_{i=1-p}^q \omega_i [g(x_i) - g(x_{i-1})]. \quad (25)$$

Wir nehmen an, daß die Menge der Zahlen ω_i ($i = 0, \pm 1, \pm 2, \dots$) eine endliche obere Grenze $\omega = \sup \omega_i$ hat. Wegen der Stetigkeit von $f(x)$ kann man insbesondere eine Zerlegung (24) des unendlichen Intervalls wählen, bei der ω kleiner als eine beliebig vorgegebene positive Zahl ist. Wir führen die folgenden Bezeichnungen ein:

$$A = \lim_{x \rightarrow -\infty} g(x), \quad B = \lim_{x \rightarrow \infty} g(x),$$

$$S_{p,q} = \sum_{t=1-p}^q f(\xi_t) [g(x_t) - g(x_{t-1})],$$

$$S'_{p,q} = \sum_{t=1-p}^q |f(\xi_t)| [g(x_t) - g(x_{t-1})].$$

Es sei ferner ω'_i die ω_i entsprechende Größe für $|f(x)|$ und

$$\omega' = \sup \omega'_i.$$

Offenbar gilt $\omega'_i \leq \omega_i$ und $\omega' \leq \omega$. Aus (25) folgt

$$\left| \int_{x-p}^{x_q} f(x) dg(x) - S_{p,q} \right| \leq \omega (B - A) \quad (26_1)$$

und analog

$$\left| \int_{x-p}^{x_q} |f(x)| dg(x) - S'_{p,q} \right| \leq \omega' (B - A); \quad (26_2)$$

daraus folgt

$$S'_{p,q} \leq \int_{x-p}^{x_q} |f(x)| dg(x) + \omega' (B - A) \quad (27)$$

und

$$\int_{x-p}^{x_q} |f(x)| dg(x) \leq S'_{p,q} + \omega' (B - A). \quad (28)$$

Wir beweisen nun einen Satz, der eine notwendige und hinreichende Bedingung für die absolute Konvergenz des Integrals (23) gibt.

Satz. *Das Integral (23) ist dann und nur dann absolut konvergent, wenn eine Zerlegung mit endlichem ω und zugehörige Zahlen ξ_i existieren, die der Ungleichung $x_{i-1} \leq \xi_i \leq x_i$ genügen und für die die Reihe*

$$\sum_{t=-\infty}^{\infty} f(\xi_t) [g(x_t) - g(x_{t-1})] \quad (29)$$

absolut konvergiert. Wenn diese Bedingung erfüllt ist, dann konvergiert die Reihe (29) für jede beliebige Zerlegung (24) mit endlichem ω und beliebig gewählten ξ_i aus dem Intervall $[x_{i-1}, x_i]$, und es ist

$$\int_{-\infty}^{\infty} f(x) dg(x) = \lim_{\omega \rightarrow 0} \sum_{t=-\infty}^{\infty} f(\xi_t) [g(x_t) - g(x_{t-1})]. \quad (30)$$

Beweis. Wir nehmen an, daß das Integral (23) absolut konvergiert. Dann gibt die Ungleichung (27) für eine beliebige Zerlegung mit endlichem ω

$$S'_{p,q} \leq \int_{-\infty}^{\infty} |f(x)| dg(x) + \omega' (B - A),$$

d. h., die Summe $S'_{p,q}$, die bei wachsendem p und q zunimmt, bleibt beschränkt, folglich ist die Reihe (29) bei beliebiger Zerlegung (24) mit endlichem ω und beliebig gewählten ξ_i absolut konvergent. Aus der Ungleichung (26₁) folgt unmittelbar (30). Wir nehmen nun umgekehrt an, daß die Reihe (29) bei einer gewissen Zerlegung (24) mit endlichem ω und bei einer gewissen Wahl der ξ_i absolut konvergiert.

Aus (28) ergibt sich unmittelbar

$$\int_{x-p}^{x_g} |f(x)| \, dg(x) \leq \sum_{i=-\infty}^{\infty} |f(\xi_i)| [g(x_i) - g(x_{i-1})] + \omega' (B - A).$$

Hieraus erkennt man, daß das Integral auf der linken Seite bei wachsendem p und q beschränkt bleibt, d. h., das Integral (23) konvergiert absolut. Die Reihe (29) ist dann, wie wir soeben festgestellt haben, bei beliebiger Zerlegung mit endlichem ω und bei beliebiger Wahl der ξ_i absolut konvergent, und es gilt (30).

Bemerkung. Wenn $f(x)$ im Innern von $[-\infty, \infty]$ gleichmäßig stetig ist und δ die größte der Differenzen $x_i - x_{i-1}$ ist, folgt aus der Voraussetzung $\delta \rightarrow 0$ sofort $\omega \rightarrow 0$, und wir können in (30) statt $\omega \rightarrow 0$ auch $\delta \rightarrow 0$ schreiben. Das gilt z. B. im Fall $f(x) = x$.

6. Die Sprungfunktion. Wir untersuchen jetzt die Eigenschaften der nicht abnehmenden Funktion $g(x)$. Da eine monotone beschränkte Veränderliche einen Grenzwert hat, hat die Funktion $g(x)$ in jedem inneren Punkt des Intervalls $[a, b]$ einen Grenzwert $g(x-0)$ von links und einen Grenzwert $g(x+0)$ von rechts. Im linken Randpunkt existiert ein Grenzwert $g(a+0)$ von rechts und im rechten Randpunkt ein Grenzwert $g(b-0)$ von links. Für $g(x-0) = g(x+0)$ ist x eine Stetigkeitsstelle von $g(x)$.

Analog dazu definiert man die Stetigkeit in den Randpunkten, wenn die Beziehungen $g(a+0) = g(a)$ und $g(b-0) = g(b)$ gelten. In den Unstetigkeitsstellen gilt $g(x+0) > g(x-0)$, und die positive Differenz $S_x = g(x+0) - g(x-0)$ heißt *Sprung von $g(x)$ im Punkt x* . Analog kann man den Sprung in den Randpunkten definieren.

Die Funktion $g(x)$ kann eine unendliche Menge von Unstetigkeitsstellen besitzen. Wir zeigen nun, daß *in diesem Fall die Menge der Unstetigkeitsstellen von $g(x)$ abzählbar ist*. Der gesamte Zuwachs der Funktion $g(x)$ im Intervall $[a, b]$ wird durch die positive Differenz $g(b) - g(a)$ ausgedrückt. Folglich ist die Anzahl der Unstetigkeitsstellen, in denen ein Sprung größer als 1 auftritt, nicht größer als der ganzzahlige Anteil von $g(b) - g(a)$, d. h., es gibt nur endlich viele solcher Unstetigkeitsstellen.

Ebenso ist die Anzahl der Unstetigkeitsstellen, in denen ein Sprung $> \frac{1}{2}$, aber ≤ 1 auftritt, nicht größer als der ganzzahlige Anteil von $2 [g(b) - g(a)]$, usw. Nun läßt sich leicht zeigen, daß die Unstetigkeitsstellen von $g(x)$ numeriert werden können. Zuerst numerieren wir in beliebiger Reihenfolge die endlich vielen Unstetigkeitsstellen, in denen Sprünge größer als 1 auftreten, dann numerieren wir die Punkte, in denen Sprünge größer als $\frac{1}{2}$, aber ≤ 1 auftreten usw.

Bei der Integration einer stetigen Funktion $f(x)$ können wir die Unstetigkeitsstellen von $g(x)$, die im Innern von $[a, b]$ liegen, für die Zerlegung des Integrationsintervalls nicht verwenden. Folglich spielen die Werte von $g(x)$ in diesen Punkten bei der Bildung des Integrals keine Rolle. Anders ist es in den Randpunkten des Intervalls, die offenbar zu den Teilpunkten gehören. Man kann nun z. B. annehmen, daß $g(x)$ in den Unstetigkeitsstellen von rechts stetig ist, d. h., daß $g(x) = g(x+0)$ gilt. Die durch diese Änderung aus der Funktion $g(x)$ entstehende Funktion werde mit $h(x)$ bezeichnet, es sei also $h(x) = g(x)$ in den Stetigkeitspunkten von $g(x)$ bzw. im rechten Randpunkt und $h(x) = g(x+0)$ in den Unstetigkeitsstellen. Auf die Größe des Integrals hat nur eine Änderung von $g(x)$ im linken Randpunkt des

Intervalls Einfluß, und es gilt

$$\int_a^b f(x) dh(x) = \int_a^b f(x) dg(x) - f(a) [g(a+0) - g(a)] .$$

Wir nehmen an, daß die Funktion $g(x)$ unstetig ist, und zerlegen sie in zwei Summanden, von denen der eine, $g_c(x)$, eine stetige nichtabnehmende Funktion ist und der andere, $g_d(x)$, die Summe der Sprünge von $g(x)$ im Intervall $[a, x]$ angibt. Dieser zweite Summand heißt *Sprungfunktion* für $g(x)$. Diese Funktion soll jetzt genauer untersucht werden.

Im Intervall $[a, b]$ sei eine endliche oder abzählbare Menge von Punkten c_k ($k = 1, 2, 3, \dots$) gegeben. Wir definieren zwei zunehmende Funktionen $\varphi_k(x)$ und $\psi_k(x)$ durch

$$\varphi_k(x) = \begin{cases} 0 & \text{für } x < c_k, \\ \alpha_k & \text{für } x \geq c_k, \end{cases} \quad \psi_k(x) = \begin{cases} 0 & \text{für } x \leq c_k, \\ \beta_k & \text{für } x > c_k. \end{cases}$$

Dabei seien α_k und β_k nichtnegative Konstante, für die die Reihen

$$\sum_{k=1}^{\infty} \alpha_k \quad \text{und} \quad \sum_{k=1}^{\infty} \beta_k \tag{31}$$

konvergieren. Ist eine Konstante α_k gleich 0, so ist die zugehörige Funktion $\varphi_k(x)$ identisch gleich 0. Dasselbe gilt für $\psi_k(x)$, wenn $\beta_k = 0$ ist. Durch die Verwendung dieser Funktionen erreichen wir in den folgenden Formeln eine symmetrische Schreibweise. Im Fall $c_k = a$ nehmen wir an, daß das zugehörige α_k gleich 0 ist, im Fall $c_k = b$ setzen wir das zugehörige β_k als 0 voraus. Aus der Konvergenz der Reihen (31) folgt unmittelbar, daß die Reihen

$$\varphi(x) = \sum_{k=1}^{\infty} \varphi_k(x), \quad \psi(x) = \sum_{k=1}^{\infty} \psi_k(x), \tag{32}$$

deren Glieder nichtnegative zunehmende Funktionen sind, für alle x und insbesondere in $[a, b]$ gleichmäßig konvergieren. Ist der Punkt x von sämtlichen Punkten c_k verschieden, so sind alle Glieder dieser beiden Reihen im Punkt x stetig, und wegen der gleichmäßigen Konvergenz sind folglich auch die Funktionen $\varphi(x)$ und $\psi(x)$ in allen von den c_k verschiedenen Punkten stetig. In einem Punkt $x = c_k$ macht der Summand $\varphi_k(x)$ einen Sprung von links von der Größe α_k und der Summand $\psi_k(x)$ einen Sprung von rechts von der Größe β_k , während die übrigen Summanden stetig sind. Wegen der gleichmäßigen Konvergenz ist auch die Summe der übrigen Summanden für $x = c_k$ stetig. Also hat auch die Funktion $\varphi(x)$ im Punkt $x = c_k$ einen Sprung von links von der Größe α_k und ist von rechts stetig; ebenso hat die Funktion $\psi(x)$ im Punkt $x = c_k$ einen Sprung von rechts von der Größe β_k und ist von links stetig. Dies gilt auch für den Fall, in dem die Menge der Punkte c_k endlich ist.

Jetzt nehmen wir an, daß $g(x)$ eine zunehmende Funktion sei, $x = c_k$ seien ihre Unstetigkeitsstellen und α_k und β_k ihre Sprünge von links bzw. von rechts in diesen Punkten, d. h. $\alpha_k = g(c_k) - g(c_k - 0)$ und $\beta_k = g(c_k + 0) - g(c_k)$. Die Differenz $g(b) - g(a)$ gibt den Gesamtzuwachs der Funktion $g(x)$ im Intervall $[a, b]$ an, und bei beliebigem n ist die Summe ihrer Sprünge $\gamma_k = \alpha_k + \beta_k$ in den ersten n Unstetigkeitsstellen $c_1, c_2, \dots, c_n$ nicht größer als die Differenz $g(b) - g(a)$. Somit ist die unendliche Reihe, die aus den Sprüngen γ_k der Funktion $g(x)$ gebildet wird, sicher konvergent. Dann konvergieren erst recht auch die Reihen, die aus den Sprüngen α_k von links und β_k von rechts gebildet werden. Wir bilden die Funktionen $\varphi(x)$ und $\psi(x)$ und setzen $g_d(x) = \varphi(x) + \psi(x)$. Die Größe $g_d(x)$ ist offenbar gleich der Summe, die aus den

Sprüngen der Funktion $g(x)$ in allen links vom Punkt x liegenden Sprungstellen und aus dem Sprung von links im Punkt x gebildet wird, wenn dieser Sprung existiert; die Differenz $g_a(\beta) - g_a(\alpha)$ ist gleich der Summe der Sprünge in den zwischen α und β liegenden Unstetigkeitsstellen, des Sprungs von rechts im Punkt α und des Sprungs von links im Punkt β . Die Differenz $g(\beta) - g(\alpha)$ gibt den Gesamtzuwachs der Funktion $g(x)$ an, wenn x von α bis β läuft, und die Differenz $g_a(\beta) - g_a(\alpha)$ gibt den Zuwachs von $g(x)$ an, der nur von den Sprüngen in den Unstetigkeitsstellen herrührt. Offenbar gilt also die Ungleichung

$$g(\beta) - g(\alpha) \geq g_a(\beta) - g_a(\alpha) \quad \text{für } \beta > \alpha.$$

Wir setzen nun $g_c(x) = g(x) - g_a(x)$. Wenn x eine Stetigkeitsstelle von $g(x)$ ist, so ist es auch eine Stetigkeitsstelle von $g_a(x)$ und somit auch von $g_c(x)$. Wir nehmen jetzt an, daß x gleich einem der c_k ist. In diesem Punkt macht die Funktion $g_a(x)$ von links und von rechts die gleichen Sprünge wie auch $g(x)$, und darum ist auch $g_c(x)$ stetig für $x = c_k$. Somit haben wir bewiesen, daß $g_c(x)$ stetig ist und nicht abnimmt. Die gesuchte Zerlegung läßt sich nun in der Gestalt

$$g(x) = g_a(x) + g_c(x) \quad (33)$$

schreiben. Diese Zerlegung kann man im Fall eines beliebigen abgeschlossenen oder nicht abgeschlossenen endlichen oder unendlichen Intervalls durchführen. Für eine beliebige stetige Funktion $f(x)$ kann man

$$\int_a^b f(x) dg(x) = \int_a^b f(x) dg_a(x) + \int_a^b f(x) dg_c(x) \quad (34)$$

schreiben.

Wir zeigen, daß das erste Integral auf der rechten Seite in Form der Summe

$$\int_a^b f(x) dg_a(x) = \sum_k f(c_k) \gamma_k \quad (35)$$

dargestellt werden kann, wenn c_k die Unstetigkeitsstellen von $g(x)$ und γ_k die Sprünge von $g(x)$ in diesen Punkten sind. Wir nehmen an, daß die Anzahl der Unstetigkeitsstellen unendlich ist. Setzen wir $\omega_k(x) = \varphi_k(x) + \psi_k(x)$, so können wir

$$g_a(x) = s_m(x) + r_m(x)$$

schreiben; dabei ist

$$s_m(x) = \sum_{k=1}^m \omega_k(x), \quad r_m(x) = \sum_{k=m+1}^{\infty} \omega_k(x).$$

Es gilt die Ungleichung $0 \leq r_m(x) \leq \gamma_{m+1} + \gamma_{m+2} + \dots$, und wegen der Konvergenz der Reihe, die aus den γ_k gebildet wird, können wir bei beliebig vorgegebenem positivem ε ein solches N wählen, daß für beliebige Werte x und $m > N$

$$0 \leq r_m(x) \leq \varepsilon \quad (36)$$

gilt. Wegen der Stetigkeit von $f(x)$ gilt ferner [2] $\int_a^b f(x) d\omega_k(x) = f(c_k) \gamma_k$ und folglich

$$\int_a^b f(x) ds_m(x) = \sum_{k=1}^m f(c_k) \gamma_k. \quad (37)$$

Die Funktion $f(x)$ ist beschränkt, d. h. $|f(x)| \leq L$, also gilt für die Summanden der letzten Summe die Abschätzung $|f(c_k) \gamma_k| \leq L \gamma_k$. Hieraus sieht man, daß die Reihe aus den Zahlen $f(c_k) \gamma_k$ absolut konvergent ist.

Für das Integral bezüglich der nicht abnehmenden Funktion $r_m(x)$ gilt wegen (36) die Abschätzung

$$\left| \int_a^b f(x) dr_m(x) \right| \leq L\varepsilon \quad (m > N).$$

Da ε beliebig gewählt war, folgt hieraus, daß die Differenz

$$\int_a^b f(x) dg_a(x) - \int_a^b f(x) ds_m(x) = \int_a^b f(x) dr_m(x)$$

bei wachsendem m gegen 0 strebt, d. h.

$$\int_a^b f(x) dg_a(x) = \lim_{m \rightarrow \infty} \int_a^b f(x) ds_m(x)$$

ist. Mit (37) ergibt sich hieraus

$$\int_a^b f(x) dg_a(x) = \sum_{k=1}^{\infty} f(c_k) \gamma_k. \quad (38)$$

7. Physikalische Interpretation. Wir geben nun eine physikalische Interpretation der Funktion $g(x)$ und des Stieltjesschen Integrals. Wir nehmen an, daß auf dem Intervall $[a, b]$ Materie verteilt ist. Es sei $g(x)$ die Masse, die sich im Intervall $[a, x]$ befindet, und $g(a)$ die Masse im Punkt $x=a$, wenn eine solche konzentrierte Masse existiert. Anderenfalls gilt $g(a)=0$. Die Differenz $g(d)-g(c)$ gibt die Masse an, die sich im Intervall $[c, d]$ befindet. Wenn die positive Zahl h gegen 0 strebt, schrumpft das Intervall $(x, x+h]$ zusammen, und ein beliebiger Punkt gehört bei hinreichend kleinem h nicht mehr zum Intervall $(x, x+h]$, weil der linke Randpunkt nicht zum Intervall gehört. Die Funktion $g(x)$ ist eine zunehmende Funktion (die Masse ist positiv), und nach dem oben Gesagten unterwirft man naturgemäß die Funktion $g(x)$, welche die Massenverteilung charakterisiert, der Bedingung $g(x+h)-g(x) \rightarrow 0$ oder $g(x)=g(x+0)$, d. h., man verlangt, daß $g(x)$ in allen Unstetigkeitsstellen außer in $x=b$ von rechts stetig ist. Es hat keinen Sinn, von Stetigkeit im rechten Randpunkt des Intervalls zu sprechen, weil die Funktion für $x > b$ nicht definiert ist. Im Innern des Intervalls existieren in den Unstetigkeitsstellen von $g(x)$ konzentrierte Massen, und die Größe der konzentrierten Masse wird durch die Differenz $g(x) - g(x-0)$ definiert. Dasselbe gilt auch für den rechten Randpunkt des Intervalls. Die Gesamtmasse im Intervall $[a, b]$ ist gleich $g(b)$. Alles Gesagte gilt für endliche und für unendliche Intervalle. Charakteristisch an diesen Überlegungen war die Tatsache, daß wir nicht von dem Begriff der Dichte der Massenverteilung Gebrauch machten. Der Massemittelpunkt der verteilten Materie wird durch

$$x_c = \frac{1}{g(b)} \int_a^b x dg(x)$$

definiert. Diese Formel gilt für ein endliches Intervall. Im Fall eines unendlichen Intervalls ist die Funktion $f(x)=x$, über die integriert wird, nicht mehr beschränkt, und man muß die Definition des uneigentlichen Integrals verwenden.

In der Wahrscheinlichkeitsrechnung stellt die Funktion $g(x)$ gewöhnlich die Wahrscheinlichkeit für die Verteilung einer zufälligen Größe dar, und zwar ist $g(x)$ gleich der Wahrscheinlichkeit dafür, daß die zufällige Größe im Intervall $(-\infty, x]$ liegt. Dabei ist wie oben $g(x)$ von rechts stetig. Der Begriff des Stieltjesschen Integrals einer stetigen Funktion kann, wie wir sehen werden, unmittelbar

auf den Fall ausgedehnt werden, daß $g(x)$ die Differenz zweier nicht abnehmender Funktionen ist: $g(x) = g_1(x) - g_2(x)$. In diesem Fall kann man leicht für $g(x)$ eine physikalische Interpretation geben. Wir nehmen an, daß im Intervall $(-\infty, \infty)$ positive und negative Ladungen verteilt sind. Die positive Gesamtladung im Intervall $(-\infty, x]$ wird dabei durch $g_1(x)$, die negative Gesamtladung im gleichen Intervall durch $g_2(x)$ dargestellt.

8. Funktionen von beschränkter Variation. Bisher haben wir angenommen, daß die Funktion $g(x)$, über die integriert wird, zunimmt. Wir gehen nun zu Integralen mit allgemeineren Funktionen $g(x)$ über und führen dazu eine gewisse Funktionenklasse ein, die für uns die grundlegende Funktionenklasse ist und zu der alle Funktionen $g(x)$, über die integriert wird, gehören müssen. In dem abgeschlossenen endlichen oder unendlichen Intervall $[a, b]$ existiere eine Funktion $g(x)$, die in jedem Punkt dieses Intervalls einen endlichen Wert annimmt. Es sei δ eine gewisse Zerlegung von $[a, b]$ in Teilintervalle: $a = x_0 < x_1 < \dots < x_{n-1} < x_n = b$. Wir bilden die Summe

$$t_\delta = \sum_{k=1}^n |g(x_k) - g(x_{k-1})|. \quad (39)$$

Definition. Wenn diese Summe bei allen möglichen Zerlegungen δ beschränkt bleibt, nennt man die Funktion $g(x)$ eine *Funktion von beschränkter Schwankung* oder eine *Funktion von beschränkter Variation* im Intervall $[a, b]$ und die obere Grenze der Summe (39) die *Totalvariation* oder einfach die *Variation* der Funktion $g(x)$ im Intervall $[a, b]$.

Wir werden die Variation mit $V_a^b(g)$ bezeichnen. Wir stellen nun einige einfache Eigenschaften der Summen t_δ und der Variation $V_a^b(g)$ zusammen. Wenn zwischen den Punkten x_k und x_{k-1} ein neuer Teilpunkt c eingeführt wird, so folgt aus

$$g(x_k) - g(x_{k-1}) = [g(x_k) - g(c)] + [g(c) - g(x_{k-1})]$$

unmittelbar

$$|g(x_k) - g(x_{k-1})| \leq |g(x_k) - g(c)| + |g(c) - g(x_{k-1})|,$$

d. h., bei *Hinzufügung neuer Teilpunkte nimmt die Summe t_δ nicht ab*. Wenn ferner die aus nichtnegativen Summanden bestehende Summe t_δ für das Intervall $[a, b]$ beschränkt bleibt, bleibt sie um so mehr noch für ein beliebiges Teilintervall $[\alpha, \beta]$ von $[a, b]$ beschränkt, d. h., *wenn die Funktion $g(x)$ im Intervall $[a, b]$ von beschränkter Variation ist, ist sie auch in einem beliebigen Teilintervall $[\alpha, \beta]$ von $[a, b]$ von beschränkter Variation, und es gilt $V_\alpha^\beta(g) \leq V_a^b(g)$.*

Das unzerlegte Intervall $[a, b]$ stellt auch eine mögliche Zerlegung δ dar; weil offenbar für eine beliebige Zerlegung δ die Ungleichung $t_\delta \leq V_a^b(g)$ gilt, gilt speziell

$$|g(b) - g(a)| \leq V_a^b(g). \quad (40)$$

Wenn $g(x)$ im Intervall $[a, b]$ monoton ist, so haben alle Differenzen $g(x_k) - g(x_{k-1})$ ein und dasselbe Vorzeichen, und die Summe t_δ ist bei beliebiger Zerlegung δ im Fall einer zunehmenden Funktion gleich $g(b) - g(a)$ und im Fall einer abnehmenden Funktion gleich $g(a) - g(b)$, d. h., *jede monotone Funktion ist eine Funktion von beschränkter Variation.*

Wir formulieren nun eine Reihe von Eigenschaften für Funktionen von beschränkter Variation.

Satz 1. Ist die Funktion $g(x)$ in $[a, b]$ von beschränkter Variation, so ist sie in diesem Intervall beschränkt.

Für einen beliebigen Punkt x in $[a, b]$ gilt nämlich $g(x) = g(a) + [g(x) - g(a)]$ und folglich $|g(x)| \leq |g(a)| + |g(x) - g(a)|$ oder wegen (40) die Abschätzung $|g(x)| \leq |g(a)| + V_a^x(g) \leq |g(a)| + V_a^b(g)$, womit die Beschränktheit von $g(x)$ bewiesen ist.

Satz 2. Sind $g(x)$ und $h(x)$ in $[a, b]$ Funktionen von beschränkter Variation und ist c eine Konstante, so sind auch $cg(x)$ und $g(x) + h(x)$ von beschränkter Variation.

Wir beweisen das für die Summe. Wir bilden t_δ für $g(x) + h(x)$:

$$\begin{aligned} t_\delta &= \sum_{k=1}^n |[g(x_k) + h(x_k)] - [g(x_{k-1}) + h(x_{k-1})]| \\ &\leq \sum_{k=1}^n |g(x_k) - g(x_{k-1})| + \sum_{k=1}^n |h(x_k) - h(x_{k-1})|. \end{aligned}$$

Die letzten beiden Summen sind beschränkt, weil nach Voraussetzung $g(x)$ und $h(x)$ von beschränkter Variation sind. Also ist erst recht t_δ bei allen Zerlegungen δ beschränkt, d. h., $g(x) + h(x)$ ist von beschränkter Variation.

Folgerung. Jede endliche Linearkombination

$$c_1 f_1(x) + c_2 f_2(x) + \dots + c_p f_p(x)$$

von Funktionen von beschränkter Variation ist eine Funktion von beschränkter Variation.

Satz 3. Sind $g(x)$ und $h(x)$ von beschränkter Variation, so ist auch ihr Produkt $g(x) h(x)$ von beschränkter Variation. Ist außerdem $|h(x)| \equiv m > 0$, so ist auch der Quotient $g(x) : h(x)$ von beschränkter Variation.

Beweis. Wir betrachten das Produkt und bilden das zugehörige t_δ :

$$t_\delta = \sum_{k=1}^n |g(x_k) h(x_k) - g(x_{k-1}) h(x_{k-1})|. \quad (41)$$

Da $g(x)$ und $h(x)$ beschränkt sind, gelten für eine gewisse positive Zahl L die Abschätzungen $|g(x)| \leq L$ und $|h(x)| \leq L$.

Offensichtlich gilt

$$g(x_k) h(x_k) - g(x_{k-1}) h(x_{k-1}) = g(x_k) [h(x_k) - h(x_{k-1})] + h(x_{k-1}) [g(x_k) - g(x_{k-1})];$$

hieraus folgt mit (41)

$$t_\delta \leq \sum_{k=1}^n |g(x_k)| |h(x_k) - h(x_{k-1})| + \sum_{k=1}^n |h(x_{k-1})| |g(x_k) - g(x_{k-1})|$$

oder

$$t_\delta \leq L \sum_{k=1}^n |h(x_k) - h(x_{k-1})| + L \sum_{k=1}^n |g(x_k) - g(x_{k-1})|.$$

Die in dieser Abschätzung auftretenden Summen sind aber beschränkt, weil nach Voraussetzung $g(x)$ und $h(x)$ von beschränkter Variation sind. Folglich ist auch die Summe t_δ bei allen Zerlegungen δ beschränkt, was zu beweisen war.

Satz 4. Gilt $a < c < b$ und ist die Funktion $g(x)$ in $[a, b]$ von beschränkter Variation, so ist sie auch in $[a, c]$ und in $[c, b]$ von beschränkter Variation, und umgekehrt, wenn die Funktion $g(x)$ in $[a, c]$ und in $[c, b]$ von beschränkter Variation ist, ist sie auch in

$[a, b]$ von beschränkter Variation. Dabei gilt

$$V_a^b(g) = V_a^c(g) + V_c^b(g). \quad (42)$$

Beweis. Wie wir vorhin erkannten, ist die Funktion $g(x)$ in $[a, c]$ und in $[c, b]$ von beschränkter Variation, wenn sie in $[a, b]$ von beschränkter Variation ist. Wir beweisen nun die umgekehrte Behauptung und die Formel (42). Wir bezeichnen mit t_δ die Summe (39) für $[a, b]$ und mit $t_{\delta_1}^{(1)}$ und $t_{\delta_2}^{(2)}$ die entsprechenden Summen für $[a, c]$ bzw. $[c, b]$. Ist der Punkt c ein Teilpunkt der Zerlegung δ , so wird die Zerlegung δ in die Zerlegung δ_1 des Intervalls $[a, c]$ und die Zerlegung δ_2 des Intervalls $[c, b]$ zerlegt, und wir haben $t_\delta = t_{\delta_1}^{(1)} + t_{\delta_2}^{(2)}$. Ist $g(x)$ in $[a, c]$ und in $[c, b]$ von beschränkter Variation, so ergibt dies $t_\delta \leq V_a^c(g) + V_c^b(g)$. Folglich bleibt die Summe t_δ beschränkt, wenn c einen Teilpunkt darstellt. Im anderen Fall bleibt sie erst recht beschränkt, weil durch Hinzufügung eines Teilpunktes die Summe t_δ nur vergrößert wird. Aus dieser Überlegung folgt, daß $g(x)$ in $[a, b]$ von beschränkter Variation ist und $V_a^b(g) \leq V_a^c(g) + V_c^b(g)$ gilt. Wir leiten nun die entgegengesetzte Ungleichung her und erhalten dann die Formel (42). Es sei ε eine gegebene positive Zahl. Nach der Definition der oberen Grenze können wir in $t_\delta = t_{\delta_1}^{(1)} + t_{\delta_2}^{(2)}$ zwei Zerlegungen δ_1 und δ_2 so auswählen, daß $t_{\delta_1}^{(1)} > V_a^c(g) - \varepsilon$ und $t_{\delta_2}^{(2)} > V_c^b(g) - \varepsilon$ gilt. Dann erhalten wir $t_\delta > V_a^c(g) + V_c^b(g) - 2\varepsilon$, woraus also $V_a^b(g) > V_a^c(g) + V_c^b(g) - 2\varepsilon$ oder, da ε beliebig ist, schließlich $V_a^b(g) \geq V_a^c(g) + V_c^b(g)$ folgt. Damit ist der Satz bewiesen.

Folgerung. Wir haben den Satz für den Fall bewiesen, daß das Intervall $[a, b]$ in zwei Teile zerlegt wird. Wenn wir den Satz wiederholt anwenden, so erhalten wir das analoge Ergebnis für den Fall, daß $[a, b]$ in eine endliche Anzahl von Teilintervallen zerlegt wird, d. h., wenn das Intervall $[a, b]$ durch $c_1, \dots, c_{n-1}$ in endlich viele Teilintervalle zerlegt wird und wenn die Funktion $g(x)$ im ganzen Intervall von beschränkter Variation ist, ist sie in jedem Teilintervall von beschränkter Variation und umgekehrt; weiterhin ist die Totalvariation im ganzen Intervall gleich der Summe der Totalvariationen in den einzelnen Teilintervallen. Diese Eigenschaft nennt man gewöhnlich *Additivität der Totalvariation*. Man kann sie in der Gestalt

$$V_a^b(g) = V_a^{c_1}(g) + V_{c_1}^{c_2}(g) + \dots + V_{c_{n-1}}^b(g) \quad (43)$$

schreiben.

Satz 5. Die Funktion $g(x)$ ist dann und nur dann von beschränkter Variation, wenn man sie als Differenz zweier zunehmender Funktionen darstellen kann.

Beweis. Offenbar ist die Bedingung hinreichend. Die zunehmenden Funktionen seien Funktionen beschränkter Variation; nach der Folgerung aus Satz 2 ist dann die Differenz ebenfalls eine Funktion von beschränkter Variation. Wir beweisen nun, daß die Bedingung notwendig ist, daß man also eine Funktion $g(x)$ von beschränkter Variation als Differenz zweier zunehmender Funktionen darstellen kann. Setzen wir

$$g_1(x) = \frac{1}{2} [V_a^x(g) + g(x)], \quad g_2(x) = \frac{1}{2} [V_a^x(g) - g(x)], \quad (44)$$

so folgt

$$g(x) = g_1(x) - g_2(x), \quad (45)$$

und es genügt zu zeigen, daß die Funktionen $g_1(x)$ und $g_2(x)$ zunehmen. Wir beweisen das für $g_1(x)$. Die Zahlen α und β mögen in $[a, b]$ liegen, und es sei $\alpha < \beta$.

Es gilt

$$g_1(\beta) - g_1(\alpha) = \frac{1}{2} [V_a^\beta(g) - V_a^\alpha(g) + g(\beta) - g(\alpha)]$$

oder wegen der Additivität der Totalvariation

$$g_1(\beta) - g_1(\alpha) = \frac{1}{2} [V_a^\beta(g) + g(\beta) - g(\alpha)] .$$

Wegen (40) gilt aber $V_a^\beta(g) \equiv |g(\beta) - g(\alpha)|$ und folglich $g_1(\beta) - g_1(\alpha) \equiv 0$.

Folgerung. Die zunehmenden Funktionen $g_1(x)$ und $g_2(x)$ können nur eine endliche oder eine abzählbare Menge von Unstetigkeitsstellen haben, und in jeder Unstetigkeitsstelle haben sie einen Grenzwert von links und von rechts. Dasselbe gilt folglich auch für die Funktion $g(x)$.

Satz 6. Ist die Funktion $g(x)$ in einem Punkt $x=c$ stetig, so ist auch die Funktion $V_a^x(g) = v(x)$ in diesem Punkt stetig und umgekehrt. Ist $g(x)$ von rechts (links) stetig, so ist auch $v(x)$ von rechts (links) stetig und umgekehrt.

Beweis. Wir nehmen $c < b$ an und betrachten beispielsweise die Stetigkeit von rechts. Es sei ε eine gegebene positive Zahl. Wir zerlegen $[c, b]$ durch die Punkte $c = x_0 < x_1 < \dots < x_{n-1} < x_n = b$ derart in Teile, daß die Ungleichung

$$\sum_{k=1}^n |g(x_k) - g(x_{k-1})| > V_c^b(g) - \varepsilon \quad (46)$$

gilt. Wenn neue Teilpunkte hinzugefügt werden, gilt diese Ungleichung erst recht. Wir können deshalb annehmen, daß der Punkt x_1 so nahe bei c liegt, daß $|g(x_1) - g(c)| < \varepsilon$ gilt. Dabei benutzen wir die rechtsseitige Stetigkeit von $g(x)$. Die Ungleichung (46) können wir in der Gestalt

$$|g(x_1) - g(c)| + \sum_{k=2}^n |g(x_k) - g(x_{k-1})| > V_c^b(g) - \varepsilon$$

schreiben, woraus sich wegen $|g(x_1) - g(c)| < \varepsilon$

$$\sum_{k=2}^n |g(x_k) - g(x_{k-1})| > V_c^b(g) - 2\varepsilon$$

ergibt.

Die links stehende Summe ist eine gewisse Summe t_δ für das Intervall $[x_1, b]$, und aus der letzten Ungleichung folgt

$$V_{x_1}^b(g) > V_c^b(g) - 2\varepsilon$$

oder, wegen der Additivität der Totalvariation, $V_{x_1}^c(g) < 2\varepsilon$, d. h. $v(x_1) - v(c) < 2\varepsilon$. Die Funktion $v(x)$ nimmt zu, und aus der letzten Ungleichung folgt $v(c+0) - v(c) < 2\varepsilon$. Da ε beliebig gewählt war, erhalten wir daraus $v(c+0) = v(c)$, d. h., $v(x) = V_a^x(g)$ ist im Punkt $x=c$ von rechts stetig. Nimmt man umgekehrt an, daß $v(x)$ von rechts stetig ist, so gilt wegen (40) $|g(c+h) - g(c)| \leq v(c+h) - v(c)$. Wenn die positive Zahl h gegen 0 strebt, strebt auch die rechte Seite gegen 0; folglich strebt auch die linke Seite gegen 0, womit bewiesen ist, daß $g(x)$ im Punkt $x=c$ von rechts stetig ist.

Folgerung. Ist $g(x)$ im Punkt $x=c$ stetig, so ist bewiesen, daß die durch (44) definierten Funktionen $g_1(x)$ und $g_2(x)$ auch im Punkt c stetig sind. Das gleiche gilt auch für die Stetigkeit von rechts oder links.

Satz 7. Ist $g(x)$ von beschränkter Variation und ist

$$g(x) = g_1^*(x) - g_2^*(x) \quad (47)$$

irgendeine von (45) verschiedene Darstellung von $g(x)$ als Differenz zweier zunehmender Funktionen, so gelten für beliebige, in $[a, b]$ liegende Punkte α und β ($\alpha < \beta$) die Ungleichungen

$$g_1(\beta) - g_1(\alpha) \leq g_1^*(\beta) - g_1^*(\alpha), \quad g_2(\beta) - g_2(\alpha) \leq g_2^*(\beta) - g_2^*(\alpha). \quad (48)$$

Beweis. Wir beweisen nur die erste Ungleichung, die in der Gestalt

$$\frac{1}{2} [V_\alpha^\beta(g) + g(\beta) - g(\alpha)] \leq g_1^*(\beta) - g_1^*(\alpha) \quad (49)$$

geschrieben werden kann, und zwar indirekt. Wir nehmen also an, daß

$$\frac{1}{2} [V_\alpha^\beta(g) + g(\beta) - g(\alpha)] > g_1^*(\beta) - g_1^*(\alpha) \quad (50)$$

gilt, und wählen eine Zerlegung δ des Intervalls $[\alpha, \beta]$, für welche sich die Summe t_δ so wenig von der Totalvariation $V_\alpha^\beta(g)$ unterscheidet, daß die Ungleichung (50) gültig bleibt, wenn wir in ihr die Totalvariation durch diese Summe ersetzen. Für eine gewisse Zerlegung $\alpha = x_0 < x_1 < \dots < x_{n-1} < x_n = \beta$ gilt also

$$\frac{1}{2} \left[\sum_{k=1}^n |g(x_k) - g(x_{k-1})| + g(\beta) - g(\alpha) \right] > g_1^*(\beta) - g_1^*(\alpha). \quad (51)$$

Andererseits können wir

$$g(\beta) - g(\alpha) = \sum_{k=1}^n [g(x_k) - g(x_{k-1})],$$

$$g_1^*(\beta) - g_1^*(\alpha) = \sum_{k=1}^n [g_1^*(x_k) - g_1^*(x_{k-1})]$$

schreiben; folglich nimmt die Ungleichung (51) die Gestalt

$$\frac{1}{2} \sum_{k=1}^n [|g(x_k) - g(x_{k-1})| + g(x_k) - g(x_{k-1})] > \sum_{k=1}^n [g_1^*(x_k) - g_1^*(x_{k-1})]$$

an. Damit die Ungleichung richtig ist, muß mindestens ein Summand auf der linken Seite größer sein als der entsprechende Summand auf der rechten Seite. Wir nehmen an, das sei für $k = p$ der Fall. Dann gilt also

$$\frac{1}{2} [|g(x_p) - g(x_{p-1})| + g(x_p) - g(x_{p-1})] > g_1^*(x_p) - g_1^*(x_{p-1}). \quad (52)$$

Für $g(x_p) - g(x_{p-1}) \leq 0$ ist diese Ungleichung nicht richtig, weil ihre linke Seite gleich 0, ihre rechte Seite aber nichtnegativ ist, da $g_1^*(x)$ nach Voraussetzung zunimmt. Es bleibt also der Fall $g(x_p) - g(x_{p-1}) > 0$ zu untersuchen. In diesem Fall kann die Ungleichung (52) in der Gestalt

$$g(x_p) - g(x_{p-1}) > g_1^*(x_p) - g_1^*(x_{p-1})$$

oder wegen (47) in der Gestalt

$$- [g_2^*(x_p) - g_2^*(x_{p-1})] > 0$$

geschrieben werden. Das ist aber nicht richtig, weil $g_2^*(x)$ nach Voraussetzung zunimmt. Die Annahme (50) führt also zu einem Widerspruch. Damit sind die Ungleichung (49) und der Satz bewiesen.

Die Darstellung (45) als Differenz der beiden durch (44) definierten zunehmenden Funktionen $g_1(x)$ und $g_2(x)$ heißt *kanonische Darstellung* einer Funktion von beschränkter Variation als Differenz zweier zunehmender Funktionen. Wie wir soeben bewiesen haben, nehmen die in der kanonischen Darstellung auftretenden Funktionen $g_1(x)$ und $g_2(x)$ nicht schneller zu als die entsprechenden Funktionen einer beliebigen anderen Darstellung. Wenn wir zu $g_1(x)$ und $g_2(x)$ ein und dieselbe Konstante c addieren, ändern sich weder ihre Differenz noch ihr Zuwachs in einem beliebigen Teilintervall $[\alpha, \beta]$ von $[a, b]$, und die erhaltene Darstellung von $g(x)$ als Differenz von $g_1(x) + c$ und $g_2(x) + c$ kann ebenfalls kanonisch genannt werden.

Bemerkung. Jede der beiden zunehmenden Funktionen $g_1(x)$ und $g_2(x)$ kann man in eine Sprungfunktion und einen stetigen Teil zerlegen:

$$g_1(x) = g_{1d}(x) + g_{1c}(x), \quad g_2(x) = g_{2d}(x) + g_{2c}(x).$$

Das führt zu der vollständig bestimmten Zerlegung der ursprünglichen Funktion $g(x)$ in eine Sprungfunktion und einen stetigen Teil:

$$g(x) = [g_{1d}(x) - g_{2d}(x)] + [g_{1c}(x) - g_{2c}(x)]. \quad (53)$$

9. Integrierbare Funktionen von beschränkter Variation. Ist in $[a, b]$ die Funktion $f(x)$ stetig und die Funktion $g(x)$ von beschränkter Variation, so können wir $g(x)$ als Differenz zweier zunehmender Funktionen darstellen; dann ist

$$\begin{aligned} & \sum_{k=1}^n f(\xi_k) [g(x_k) - g(x_{k-1})] \\ &= \sum_{k=1}^n f(\xi_k) [g_1(x_k) - g_1(x_{k-1})] - \sum_{k=1}^n f(\xi_k) [g_2(x_k) - g_2(x_{k-1})]. \end{aligned} \quad (54)$$

Die rechts stehenden Summen haben bei beliebiger Verfeinerung der Intervallteilung einen endlichen Grenzwert, also hat auch die Summe auf der linken Seite einen endlichen Grenzwert, d. h., *eine stetige Funktion ist bezüglich einer Funktion von beschränkter Variation integrierbar*. Der Grenzübergang in (54) ergibt

$$\int_a^b f(x) dg(x) = \int_a^b f(x) dg_1(x) - \int_a^b f(x) dg_2(x). \quad (55)$$

Wir weisen auf die Änderung hin, die sich für die Formulierung der Eigenschaften eines Stieltjesschen Integrals ergibt, wenn $g(x)$ eine Funktion von beschränkter Variation ist. Es gilt

$$\left| \sum_{k=1}^n f(\xi_k) [g(x_k) - g(x_{k-1})] \right| \leq L \sum_{k=1}^n |g(x_k) - g(x_{k-1})| \leq L V_a^b(g),$$

wenn $|f(x)| \leq L$ ist. Durch Grenzübergang erhält man

$$\left| \int_a^b f(x) dg(x) \right| \leq L V_a^b(g). \quad (56)$$

Diese Ungleichung steht an Stelle der Ungleichung (21) in [4]. Wir erinnern noch an die Formel (vgl. [2])

$$\int_a^b f(x) d \sum_{k=1}^p a_k g_k(x) = \sum_{k=1}^p a_k \int_a^b f(x) dg_k(x). \quad (57)$$

Sind die Funktionen $g_k(x)$ von beschränkter Variation, so ist auch ihre Linearkombination $a_1 g_1(x) + \dots + a_p g_p(x)$ eine Funktion von beschränkter Variation.

Wir betrachten das Stieltjessche Integral mit veränderlicher oberer Grenze

$$F(x) = \int_a^x f(t) \, dg(t) \quad (58)$$

für den Fall, daß $f(x)$ stetig und $g(x)$ von beschränkter Variation ist. Wir beweisen, daß $F(x)$ eine Funktion von beschränkter Variation ist. Wir bilden die Summe t_δ für $[a, x]$:

$$t_\delta = \sum_{k=1}^n \left| \int_{x_{k-1}}^{x_k} f(t) \, dg(t) \right|.$$

Mit der Ungleichung (56) erhalten wir

$$t_\delta \leq L \sum_{k=1}^n V_{x_{k-1}}^{x_k}(g) = L V_a^b(g),$$

woraus die Behauptung folgt. In den Stetigkeitspunkten von $g(x)$ ist die Funktion $V_a^x(g)$ stetig, und aus der Ungleichung

$$\left| \int_x^{x+h} f(x) \, dg(x) \right| \leq L V_x^{x+h}(g)$$

folgt, daß auch $F(x)$ in diesen Punkten stetig ist. Wir beweisen noch die Behauptung: *Wenn in $[a, b]$ die Funktionen $f(x)$ und $\varphi(x)$ stetig sind und $g(x)$ von beschränkter Variation ist, dann gilt*

$$\int_a^b \varphi(x) \, d \left[\int_a^x f(t) \, dg(t) \right] = \int_a^b \varphi(x) f(x) \, dg(x). \quad (59)$$

Es genügt, diese Gleichung für den Fall zu beweisen, daß $g(x)$ zunimmt. Wir bilden für das Integral auf der linken Seite der Formel (59) die Summe σ_δ :

$$\sigma_\delta = \sum_{k=1}^n \varphi(\xi_k) \int_{x_{k-1}}^{x_k} f(t) \, dg(t).$$

Nach dem Mittelwertsatz [4] folgt daraus

$$\sigma_\delta = \sum_{k=1}^n \varphi(\xi_k) f(\xi'_k) [g(x_k) - g(x_{k-1})]. \quad (60)$$

Die Punkte ξ_k und ξ'_k liegen in ein und demselben Intervall, und durch die gleiche Überlegung, die zu der Formel (9) in [2] führte, überzeugen wir uns, daß die Summe auf der rechten Seite der Gleichung (60) beim Grenzübergang das Integral auf der rechten Seite von (59) ergibt. Damit ist die Formel (59) bewiesen.

10. Existenz des Stieltjesschen Integrals. Bisher haben wir das Stieltjessche Integral einer stetigen Funktion $f(x)$ bezüglich einer Funktion $g(x)$ von beschränkter Variation betrachtet. Aus der Formel für die partielle Integration folgt [2], daß die Funktion $g(x)$ von beschränkter Variation bezüglich der stetigen Funktion $f(x)$ integrierbar ist. Wir nennen jetzt einige einfache Bedingungen, die die Existenz des Stieltjesschen Integrals in anderen Fällen liefern. Wir werden voraussetzen, daß $f(x)$ und $g(x)$ in dem endlichen Intervall $[a, b]$ beschränkt sind und $g(x)$ eine nicht abnehmende Funktion ist. Da wir im weiteren das Stieltjessche Integral nur für den Fall, daß $f(x)$ stetig ist, betrachten werden, geben wir hier nur einige Bedingungen ohne Beweis an.

I. Für die Existenz des Stieltjesschen Integrals ist notwendig, daß $f(x)$ in allen Sprungstellen von $g(x)$ stetig ist; ist diese Bedingung erfüllt, so ist $f(x)$ bezüglich der Sprungfunktion $g_d(x)$ integrierbar und das Integral von $f(x)$ bezüglich $g_d(x)$ durch (35) gegeben.

Wenn diese notwendige Bedingung der Integrierbarkeit erfüllt ist, ist die Frage der Integrierbarkeit von $f(x)$ bezüglich $g(x)$ gleichbedeutend mit der Frage der Integrierbarkeit von $f(x)$ bezüglich der stetigen, nicht abnehmenden Funktion $g_c(x)$. Ist etwa $f(x)$ eine Funktion von beschränkter Variation, so ist, wie oben angegeben wurde, $f(x)$ integrierbar. Wir geben jetzt eine notwendige und hinreichende Bedingung für die Integrierbarkeit von $f(x)$ bezüglich $g_c(x)$ an:

II. Eine Funktion $f(x)$ ist bezüglich $g_c(x)$ dann und nur dann integrierbar, wenn folgende Bedingung gilt: Bei beliebig vorgegebenem positivem ε kann man die Unstetigkeitsstellen von $f(x)$ mit einer endlichen oder abzählbaren Menge von Intervallen $[a_k, b_k]$ (die sich auch überdecken können) so überdecken, daß die Ungleichung

$$\sum_k [g_c(b_k) - g_c(a_k)] \leq \varepsilon \quad (61)$$

gilt.

Wir nehmen nun an, daß $g(x)$ eine Funktion von beschränkter Variation ist. Es seien die kanonische Zerlegung (45) und irgendeine andere Zerlegung (47) dieser Funktion gegeben. Wir bilden die Differenz $S_\delta - s_\delta$ für $g_1(x)$ und $g_1^*(x)$:

$$S_\delta - s_\delta = \sum_{k=1}^n (M_k - m_k) [g_1(x_k) - g_1(x_{k-1})] .$$

$$S_\delta^* - s_\delta^* = \sum_{k=1}^n (M_k - m_k) [g_1^*(x_k) - g_1^*(x_{k-1})] .$$

Wie wir oben sahen, ist

$$g_1(x_k) - g_1(x_{k-1}) \leq g_1^*(x_k) - g_1^*(x_{k-1}) .$$

Folglich gilt: Wenn bei beliebiger Verfeinerung der Intervallteilung $S_\delta^* - s_\delta^* \rightarrow 0$ gilt, strebt erst recht $S_\delta - s_\delta$ gegen 0; d. h., ist die Funktion $f(x)$ bezüglich $g_1^*(x)$ integrierbar, so ist sie es auch bezüglich $g_1(x)$. Ganz analog gilt: Ist die Funktion $f(x)$ bezüglich $g_2^*(x)$ integrierbar, so ist sie auch bezüglich $g_2(x)$ integrierbar. Aber aus der Integrierbarkeit bezüglich $g_k(x)$ folgt nicht die Integrierbarkeit bezüglich $g_k^*(x)$ ($k=1, 2$). Also muß man bei der Untersuchung, ob $f(x)$ bezüglich $g(x)$ integrierbar ist, die Integrierbarkeit bezüglich $g_1(x)$ und bezüglich $g_2(x)$ untersuchen. Wenn $f(x)$ bezüglich $g_1(x)$ und bezüglich $g_2(x)$ integrierbar ist, existiert auch das Integral von $f(x)$ bezüglich $g(x)$, und es gilt die Formel (55).

III. Die Integrierbarkeit der Funktion $f(x)$ bezüglich $g_1(x)$ und $g_2(x)$ ist gleichbedeutend mit ihrer Integrierbarkeit bezüglich der Totalvariation $V_a^x(g) = g_1(x) + g_2(x)$.

Diese Behauptung ist, wie auch die Behauptung I, sehr einfach zu beweisen. Etwas komplizierter ist der Beweis der Behauptung II.

11. Grenzübergang unter dem Stieltjesschen Integral. Wir beweisen in diesem und den nächsten Abschnitten einige Sätze über den Grenzübergang unter dem Stieltjesschen Integral. Einer dieser Sätze wurde schon früher behandelt. Er betrifft den Fall, daß die integrierbaren Funktionen gleichmäßig gegen eine Grenzfunktion $f(x)$ konvergieren. Es seien $f_n(x)$ im Intervall $[a, b]$ stetige Funktionen, die im ganzen Intervall gleichmäßig gegen eine Funktion $f(x)$ konvergieren, und $g(x)$ sei eine Funktion von beschränkter Variation in $[a, b]$. Auf Grund von [4] und Gleichung (55) gilt

$$\lim_{n \rightarrow \infty} \int_a^b f_n(x) dg(x) = \int_a^b f(x) dg(x) . \quad (63)$$

Wir beweisen nun einige einfache Verallgemeinerungen dieser Behauptung, wobei wir uns auf die Untersuchung eines unendlichen Intervalls beschränken.

Satz 1. Die Funktionen $f_n(x)$ seien im Innern des Intervalls $[-\infty, \infty]$ stetig und gleichmäßig beschränkt, d. h., es gibt eine von n unabhängige Konstante L , für die $|f_n(x)| \leq L$ ist.

Ferner mögen die $f_n(x)$ gleichmäßig in jedem endlichen Intervall gegen eine Grenzfunktion $f(x)$ konvergieren, und $g(x)$ sei im Intervall $[-\infty, \infty]$ von beschränkter Variation und in den Randpunkten dieses Intervalls stetig. Dann gilt im Intervall $[-\infty, \infty]$ die Formel (63).

Beweis. Die Funktion $f(x)$ ist im Innern des Intervalls $[-\infty, \infty]$ stetig und beschränkt, folglich ist sie bezüglich $g(x)$ integrierbar. Wenn wir beachten, daß die Funktion $g(x)$ und damit auch ihre Totalvariation in den Randpunkten des Intervalls stetig sind und $|f(x) - f_n(x)| \leq 2L$ ist, können wir unter Berücksichtigung von (56) zeigen: Für ein beliebig vorgegebenes positives ε existiert eine solche positive Zahl A , daß für beliebige n

$$\left| \int_A^\infty [f(x) - f_n(x)] dg(x) \right| \leq \varepsilon, \quad \left| \int_{-\infty}^{-A} [f(x) - f_n(x)] dg(x) \right| \leq \varepsilon$$

gilt.

Im Intervall $[-A, A]$ ist der Grenzübergang $f_n(x) \rightarrow f(x)$ gleichmäßig, folglich gilt nach dem oben erwähnten Satz für alle hinreichend großen n

$$\left| \int_{-A}^A [f(x) - f_n(x)] dg(x) \right| < \varepsilon$$

und damit

$$\left| \int_{-\infty}^\infty [f(x) - f_n(x)] dg(x) \right| \leq 3\varepsilon.$$

Da ε beliebig gewählt war, folgt hieraus die Behauptung.

Wir beweisen nun einen analogen Satz für den Fall, daß die Funktionen $f_n(x)$ im Intervall $[-\infty, \infty]$ nicht beschränkt sind und das Integral in diesem Intervall als uneigentliches Integral aufzufassen ist.

Satz 2. Die Funktionen $f_n(x)$ seien im Innern von $[-\infty, \infty]$ stetig und mögen gleichmäßig in jedem endlichen Intervall gegen eine Grenzfunktion $f(x)$ konvergieren, die uneigentlichen Integrale

$$\int_{-\infty}^\infty f_n(x) dg(x) = \lim_{\substack{\alpha \rightarrow -\infty \\ \beta \rightarrow \infty}} \int_\alpha^\beta f_n(x) dg(x) \quad (64)$$

mögen gleichmäßig in bezug auf n konvergieren, und die Funktion $g(x)$ sei in jedem endlichen Intervall von beschränkter Variation. Dann existiert das (uneigentliche) Integral von $f(x)$ bezüglich $g(x)$ im Intervall $[-\infty, \infty]$, und es gilt die Gleichung (63).

Beweis. Die Funktion $f(x)$ ist in jedem endlichen Intervall stetig und bezüglich $g(x)$ integrierbar. Wir beweisen, daß sie in dem unendlichen Intervall $[-\infty, \infty]$ integrierbar ist. Es sei ε eine vorgegebene positive Zahl. Da die Integrale (64) gleichmäßig in bezug auf n konvergieren, existiert eine solche positive Zahl A , daß für ein beliebiges Intervall $[B', B'']$ außerhalb von $[-A, A]$ und für jeden Index n

$$\left| \int_{B'}^{B''} f_n(x) dg(x) \right| \leq \varepsilon \quad (65)$$

gilt. Wir legen B' und B'' so fest, daß das Intervall $[B', B'']$ außerhalb von $[-A, A]$ liegt. Da die $f_n(x)$ im Intervall $[B', B'']$ gleichmäßig gegen $f(x)$ konvergieren, gilt für alle genügend großen Werte von n

$$\left| \int_{B'}^{B''} [f(x) - f_n(x)] dg(x) \right| \leq \varepsilon. \quad (66)$$

Unter Benutzung der Beziehung

$$\int_{B'}^{B''} f(x) dg(x) = \int_{B'}^{B''} f_n(x) dg(x) + \int_{B'}^{B''} [f(x) - f_n(x)] dg(x)$$

so erhalten wir auf Grund von (65) und (66)

$$\left| \int_{B'}^{B''} f(x) dg(x) \right| \leq 2\varepsilon.$$

Hieraus folgt, daß das Integral von $f(x)$ bezüglich $g(x)$ im Intervall $[-\infty, \infty]$ existiert. Für den Beweis von (63) genügt die Bemerkung, daß das Integral der Differenz $f(x) - f_n(x)$ für genügend weite Intervalle dem absoluten Betrag nach genügend klein ist und in endlichen Intervallen wegen der gleichmäßigen Konvergenz der $f_n(x)$ gegen $f(x)$ für alle hinreichend großen Werte von n genügend klein ist. Damit ist der Satz bewiesen.

Wir weisen darauf hin, daß im Fall eines endlichen Intervalls die gleichmäßige Konvergenz der $f_n(x)$ gegen $f(x)$ keine notwendige Bedingung für die Richtigkeit der Formel (63) darstellt. Man braucht nur anzunehmen, daß die stetigen Funktionen $f_n(x)$ gegen eine stetige Funktion $f(x)$ konvergieren und unabhängig von n beschränkt bleiben, d. h., es existiere eine positive Zahl L , so daß für alle n und alle x aus $[a, b]$ die Ungleichung $|f_n(x)| \leq L$ gilt. Diese Behauptung wird später [50] bewiesen werden.

12. Satz von Helly. Wir betrachten jetzt den Grenzübergang bei einer Folge von Integrationsfunktionen $g_n(x)$. Vorher untersuchen wir die Konvergenz der Funktionen von beschränkter Variation gegen eine Grenzfunktion. Im Intervall $[a, b]$ seien die $g_n(x)$ eine Folge von Funktionen von beschränkter Variation, wobei die Variation aller dieser Funktionen unter ein und derselben, von n unabhängigen Schranke liegt:

$$V_a^b(g_n) \leq L. \quad (67)$$

Wir nehmen an, daß die $g_n(x)$ in jedem Punkt des Intervalls $[a, b]$ gegen eine Grenzfunktion $g(x)$ streben, die endliche Werte annimmt. Wie man leicht sieht, ist auch $g(x)$ eine Funktion von beschränkter Variation. In der Tat gilt für die Summen $t_\delta^{(n)}$ der Funktionen $g_n(x)$

$$t_\delta^{(n)} = \sum_{k=1}^m |g_n(x_k) - g_n(x_{k-1})| \leq L;$$

hieraus erhält man durch Grenzübergang dieselbe Schranke für die Summe t_δ von $g(x)$:

$$t_\delta = \sum_{k=1}^m |g(x_k) - g(x_{k-1})| \leq L.$$

Aus dieser Abschätzung folgt, daß die Variation von $g(x)$ ebenfalls nicht größer als L ist. Wenn die Funktionen $g_n(x)$ nicht in allen Punkten des Intervalls $[a, b]$, sondern nur auf einer überall dichten Menge $\mathcal{S}$ von Punkten x_k ($k=1, 2, 3, \dots$) gegen die Funktion $g(x)$ konvergieren, kann man schon nicht mehr behaupten, daß $g(x)$ eine Funktion von beschränkter Variation ist. Im folgenden werden wir für diesen Fall voraussetzen, daß $g(x)$ eine Funktion von beschränkter Variation ist. Wir bemerken, daß eine Menge $\mathcal{S}$ von Punkten x_k *dicht in* $[a, b]$ heißt, wenn jedes Teilintervall von $[a, b]$ eine unendliche Menge von Punkten aus $\mathcal{S}$ enthält. Wir nehmen an, daß die $g_n(x)$ eine Folge zunehmender Funktionen bilden, die in jedem Punkt von $[a, b]$ gegen eine Grenzfunktion $g(x)$ konvergiert, die endliche Werte annimmt. Dann ist auch die Grenzfunktion zunehmend und folglich auch eine Funktion von beschränkter Variation.

Wir beweisen nun den folgenden Satz:

Satz 1. Wenn die im Intervall $[a, b]$ zunehmenden Funktionen $g_n(x)$ auf einer in $[a, b]$ dichten Punktmenge $\mathcal{S}$ gegen die Funktion $g(x)$ streben, konvergieren sie in jeder Stetigkeitsstelle von $g(x)$, die im Innern von $[a, b]$ liegt.

Beweis. Es sei x_0 eine Stetigkeitsstelle von $g(x)$, und x' und x'' seien zwei links bzw. rechts von x_0 liegende Punkte der Menge $\mathcal{S}$, also $x' < x_0 < x''$. Es gilt $g_n(x') \leq g_n(x_0) \leq g_n(x'')$ und folglich

$$g(x_0) - g_n(x'') \leq g(x_0) - g_n(x_0) \leq g(x_0) - g_n(x').$$

Diese Ungleichung können wir wie folgt schreiben:

$$\begin{aligned} [g(x_0) - g(x'')] + [g(x'') - g_n(x'')] &\leq g(x_0) - g_n(x_0) \\ &\leq [g(x_0) - g(x')] + [g(x') - g_n(x')] . \end{aligned} \quad (68)$$

Es sei ε eine vorgegebene positive Zahl. Wir können die Punkte x' und x'' der überall in $[a, b]$ dichten Menge $\mathfrak{S}$ so nahe an x_0 wählen, daß für den Stetigkeitspunkt x_0 von $g(x)$ die Ungleichungen $|g(x_0) - g(x'')| < \varepsilon$ und $|g(x_0) - g(x')| < \varepsilon$ gelten. Wenn wir x' und x'' auf diese Weise festlegen, gelten für alle genügend großen Werte von n die Ungleichungen $|g(x') - g_n(x')| < \varepsilon$ und $|g(x'') - g_n(x'')| < \varepsilon$, da die Funktionen $g_n(x)$ in x' und x'' gegen $g(x)$ konvergieren. Daraus und aus der Ungleichung (68) folgt unmittelbar $-2\varepsilon \leq g(x_0) - g_n(x_0) \leq +2\varepsilon$. Da ε beliebig gewählt war, erhält man daraus $g_n(x_0) \rightarrow g(x_0)$.

Wir formulieren nun den grundlegenden Satz für den Grenzübergang.

Satz 2 (HELLY). In $[a, b]$ sei die Funktion $f(x)$ stetig, und die Funktionen $g_n(x)$ seien von beschränkter Variation, wobei die Variationen $V_a^b(g_n)$ nicht größer als eine von n unabhängige Zahl L sind. Die $g_n(x)$ mögen in allen Punkten von $[a, b]$ gegen die Funktion $g(x)$ konvergieren. Dann gilt

$$\lim_{n \rightarrow \infty} \int_a^b f(x) dg_n(x) = \int_a^b f(x) dg(x) . \quad (69)$$

Beweis. Wie wir oben gesehen haben, ist $g(x)$ eine Funktion von beschränkter Variation, folglich ist $f(x)$ bezüglich $g(x)$ integrierbar. Wir zerlegen das Intervall in Teilintervalle durch $a = x_0 < x_1 < \dots < x_m = b$ und bilden die Formel

$$\int_a^b f(x) dg(x) = \sum_{k=1}^m \int_{x_{k-1}}^{x_k} f(x) dg(x) = \sum_{k=1}^m \int_{x_{k-1}}^{x_k} [f(x) - f(x_k)] dg(x) + \sum_{k=1}^m f(x_k) \int_{x_{k-1}}^{x_k} dg(x)$$

oder

$$\int_a^b f(x) dg(x) = \sum_{k=1}^m \int_{x_{k-1}}^{x_k} [f(x) - f(x_k)] dg(x) + \sum_{k=1}^m f(x_k) [g(x_k) - g(x_{k-1})] . \quad (70)$$

Es sei ε eine vorgegebene positive Zahl. Es läßt sich eine genügend feine Zerlegung von $[a, b]$ konstruieren, so daß für $f(x) - f(x_k)$ bei beliebigem k die Abschätzung $|f(x) - f(x_k)| \leq \varepsilon$ gilt. Das folgt unmittelbar aus der gleichmäßigen Stetigkeit von $f(x)$ in $[a, b]$. Bei einer solchen Zerlegung erhalten wir

$$\left| \int_{x_{k-1}}^{x_k} [f(x) - f(x_k)] dg(x) \right| \leq \varepsilon V_{x_{k-1}}^{x_k}(g)$$

und daraus

$$\left| \sum_{k=1}^m \int_{x_{k-1}}^{x_k} [f(x) - f(x_k)] dg(x) \right| \leq \varepsilon \sum_{k=1}^m V_{x_{k-1}}^{x_k}(g)$$

oder

$$\left| \sum_{k=1}^m \int_{x_{k-1}}^{x_k} [f(x) - f(x_k)] dg(x) \right| \leq \varepsilon V_a^b(g) \leq \varepsilon L .$$

Die Formel (70) können wir für die soeben konstruierte Zerlegung in der Gestalt

$$\int_a^b f(x) dg(x) = \theta \varepsilon L + \sum_{k=1}^m f(x_k) [g(x_k) - g(x_{k-1})]$$

schreiben; dabei ist $|\theta| \leq 1$. Durch eine analoge Überlegung erhalten wir mit $|\theta_n| \leq 1$

$$\int_a^b f(x) dg_n(x) = \theta_n \varepsilon L + \sum_{k=1}^m f(x_k) [g_n(x_k) - g_n(x_{k-1})] .$$

Durch Subtraktion ergibt sich

$$\int_a^b f(x) dg(x) - \int_a^b f(x) dg_n(x) = (\theta - \theta_n) \varepsilon L + \sum_{k=1}^m f(x_k) \{[g(x_k) - g_n(x_k)] - [g(x_{k-1}) - g_n(x_{k-1})]\}.$$

Die Punkte x_k sind festgelegt, und da in ihnen die $g_n(x)$ gegen $g(x)$ konvergieren, ist die in der letzten Gleichung auftretende Summe für alle genügend großen Werte von n dem absoluten Betrag nach kleiner als ε . Für solche Werte von n entsteht folglich aus der letzten Gleichung die Abschätzung

$$\left| \int_a^b f(x) dg(x) - \int_a^b f(x) dg_n(x) \right| \leq \varepsilon (2L + 1),$$

aus der wegen der Willkürlichkeit von ε die Beziehung (69) folgt.

Bemerkung. Wir nehmen an, daß die Funktionen $g_n(x)$ nicht überall in $[a, b]$, sondern nur auf einer in $[a, b]$ dichten Punktmenge (die auch die beiden Randpunkte des Intervalls enthält) gegen $g(x)$ konvergieren. Die Grenzfunktion $g(x)$ sei dabei eine Funktion von beschränkter Variation. Wenn wir dann als Teilpunkte x_n Punkte nehmen, die zu dieser Punktmenge gehören, auf denen die Funktionen $g_n(x)$ gegen $g(x)$ konvergieren, dann bleibt der soeben geführte Beweis richtig, und wir erhalten unmittelbar die Gleichung (69).

Wir leiten nun eine Verallgemeinerung dieses Satzes her, die in Analogie zu der Verallgemeinerung in [11] steht.

Satz 3. *Wir nehmen an, daß im Intervall $[-\infty, \infty]$ die Funktion $f(x)$ beschränkt und im Innern stetig ist, die Funktionen $g_n(x)$ und $g(x)$ zunehmen und in den Randpunkten des Intervalls stetig sind und die $g_n(x)$ auf einer dichten Punktmenge des Intervalls $[-\infty, \infty]$ (die speziell die beiden Randpunkte des Intervalls enthält) gegen $g(x)$ konvergieren. Dann gilt*

$$\lim_{n \rightarrow \infty} \int_{-\infty}^{\infty} f(x) dg_n(x) = \int_{-\infty}^{\infty} f(x) dg(x). \quad (71)$$

Beweis. Wir bemerken, daß in diesem Fall die Totalvariation einer Funktion $g_n(x)$ gleich der Differenz $g_n(\infty) - g_n(-\infty)$ ist, und aus den Voraussetzungen des Satzes folgt unmittelbar, daß alle diese Totalvariationen nicht größer sind als eine gewisse von n unabhängige Zahl L . Die Integrale von $f(x)$ bezüglich $g_n(x)$ und bezüglich $g(x)$ existieren. Wir schätzen die Differenz dieser Integrale ab, indem wir das Integrationsintervall in die drei Teile $[-\infty, a]$, $[a, b]$ und $[b, \infty]$ zerlegen (die Punkte a und b gehören dabei der Menge an, auf der die $g_n(x)$ gegen $g(x)$ streben):

$$\left| \int_{-\infty}^{\infty} f(x) dg(x) - \int_{-\infty}^{\infty} f(x) dg_n(x) \right| \leq \left| \int_{-\infty}^a f - \int_{-\infty}^a f \right| + \left| \int_a^b f - \int_a^b f \right| + \left| \int_b^{\infty} f - \int_b^{\infty} f \right|. \quad (72)$$

Die Funktion $f(x)$ ist beschränkt, d. h. $|f(x)| \leq L$. Für die erste Differenz gilt die Abschätzung

$$\left| \int_{-\infty}^a f(x) dg(x) - \int_{-\infty}^a f(x) dg_n(x) \right| \leq L \{[g(a) - g(-\infty)] + [g_n(a) - g_n(-\infty)]\},$$

die wir in der Gestalt

$$\left| \int_{-\infty}^a f - \int_{-\infty}^a f \right| \leq L \{2[g(a) - g(-\infty)] + [g_n(a) - g(a)] + [g(-\infty) - g_n(-\infty)]\}$$

schreiben können.

Da $g(x)$ für $x = -\infty$ stetig ist, können wir a so nahe an $-\infty$ wählen, daß die positive Differenz $g(a) - g(-\infty)$ kleiner ist als eine beliebig vorgegebene positive Zahl. Wir legen a in dieser Weise fest und bemerken weiterhin, daß die Differenzen $g_n(a) - g(a)$ und $g(-\infty) - g_n(-\infty)$ ebenfalls für alle genügend großen Werte von n dem absoluten Betrag

nach beliebig klein sind. Ganz analog kann man auch den dritten Summanden auf der rechten Seite von (72) untersuchen. Man kann also zu einem beliebig vorgegebenen positiven ε aus der oben erwähnten, überall in $[-\infty, \infty]$ dichten Menge solche Werte für a und b auswählen, daß der erste und der dritte Summand auf der rechten Seite von (72) für alle genügend großen n kleiner als ε sind. Wendet man auf das endliche Intervall $[a, b]$ den oben bewiesenen Satz an, so erhält man, daß auch der zweite Summand auf der rechten Seite von (72) für alle genügend großen Werte von n kleiner als ε ist. Folglich ist die linke Seite der Ungleichung (72) für alle genügend großen Werte von n kleiner als 3ε ; daraus folgt, da ε beliebig gewählt war, schließlich die Gleichung (71).

Wir kommen nun zur zweiten Verallgemeinerung des Satzes 2, welche uneigentliche Stieltjessche Integrale betrifft.

Satz 4. *Wir nehmen an, daß $f(x)$ im Innern des Intervalls $[-\infty, \infty]$ stetig und die $g_n(x)$ und $g(x)$ in diesem Intervall von beschränkter Variation sind, wobei die Variationen der $g_n(x)$ nicht größer sind als eine gewisse von n unabhängige Zahl. Ferner mögen die $g_n(x)$ auf einer überall in $[-\infty, \infty]$ dichten Menge gegen $g(x)$ konvergieren, und die uneigentlichen Integrale*

$$\int_{-\infty}^{\infty} f(x) dg_n(x) = \lim_{\substack{a \rightarrow -\infty \\ b \rightarrow \infty}} \int_a^b f(x) dg_n(x)$$

mögen gleichmäßig bezüglich n konvergieren. Dann ist $f(x)$ bezüglich $g(x)$ in $[-\infty, \infty]$ integrierbar, und es gilt (71).

Beweis. Es sei ε eine vorgegebene positive Zahl. Es existiert ein solches positives A , daß für ein beliebiges, außerhalb von $[-A, A]$ liegendes Intervall $[B', B'']$ die Abschätzung

$$\left| \int_{B'}^{B''} f(x) dg_n(x) \right| \leq \varepsilon \quad (73)$$

gilt. Wir wählen ein solches beliebiges Intervall $[B', B'']$ und schreiben die offenbar geltende Ungleichung

$$\left| \int_{B'}^{B''} f(x) dg(x) \right| \leq \left| \int_{B'}^{B''} f(x) dg_n(x) \right| + \left| \int_{B'}^{B''} f(x) dg(x) - \int_{B'}^{B''} f(x) dg_n(x) \right|$$

auf. Nach der Bemerkung zu Satz 2 kann man n so groß annehmen, daß der zweite Summand auf der rechten Seite kleiner als ε ist. Dann folgt aus (73) unmittelbar

$$\left| \int_{B'}^{B''} f(x) dg(x) \right| \leq 2\varepsilon.$$

Da ε beliebig gewählt werden kann, ergibt sich hieraus die Existenz des Integrals von $f(x)$ bezüglich $g(x)$ in $[-\infty, \infty]$. Die Formel (71) wird vollkommen analog zu dem Dargestellten durch eine Zerlegung des Intervalls $[-\infty, \infty]$ in drei Teile bewiesen.

13. Das Auswahlprinzip. Früher haben wir bereits das Auswahlprinzip für Mengen stetiger Funktionen untersucht [IV/1,15]. Jetzt beweisen wir den Satz, der das Auswahlprinzip für Funktionen von beschränkter Variation liefert.

Satz (HELLY). *Es sei $\mathfrak{S}$ eine Menge von Funktionen von beschränkter Variation in dem (endlichen oder unendlichen) Intervall $[a, b]$, wobei eine solche positive Zahl L existiert, daß für alle Funktionen $g(x)$, die in $\mathfrak{S}$ liegen, die Ungleichungen*

$$|g(x)| \leq L, \quad V_a^b(g) \leq L \quad (74)$$

gelten, d. h., alle Funktionen $g(x)$ seien dem absoluten Betrag nach beschränkt und ebenso ihre Variationen im Intervall $[a, b]$. Dann kann man aus einer beliebigen unendlichen Folge

$g_n(x)$ von Funktionen der Menge $\mathfrak{S}$ eine Teilfolge $g_{n_k}(x)$ auswählen, die in allen Punkten von $[a, b]$ gegen eine gewisse Funktion $g(x)$ von beschränkter Variation strebt.

Man braucht nur zu zeigen, daß man eine Folge $g_{n_k}(x)$ auswählen kann, die in allen Punkten von $[a, b]$ gegen eine Grenzfunktion strebt. Danach kann man mit Hilfe des in [12] Bewiesenen aus den Bedingungen des Satzes unmittelbar folgern, daß auch die Grenzfunktion eine Funktion von beschränkter Variation ist. Wir beweisen vorher ein Lemma.

Lemma. *Existiert eine Folge von Funktionen $h_n(x)$, die in dem Intervall $[a, b]$ wachsen und durch ein und dieselbe Zahl L beschränkt sind, so kann man aus ihr eine gewisse Teilfolge von Funktionen auswählen, die in allen Punkten von $[a, b]$ einen Grenzwert hat.*

Wir bilden die abzählbare Punktmenge x_k ($k=1, 2, \dots$), die in $[a, b]$ liegt und die aus dem linken Randpunkt $x=a$ und allen Punkten x mit rationalzahliger Abszisse besteht. Diese Punktmenge ist überall in $[a, b]$ dicht, und wir können eine Teilfolge $h_{n_k}(x)$ auswählen, die in allen Punkten x_k konvergiert [IV/1,5]. Auf diese Weise erhalten wir eine Grenzfunktion $h(x)$, die vorerst nur in den Punkten a und x_k definiert ist. Auf die übrigen Punkte des Intervalls $[a, b]$ wird sie in der folgenden Weise ausgedehnt. Ist x ein Punkt aus dem Intervall $[a, b]$, der nicht in der Punktmenge x_k liegt, so definieren wir $h(x)$ als obere Grenze der Werte von $h(x)$ für alle links von x liegenden Punkte x_k , d. h., wir setzen

$$h(x) = \sup_{x_k < x} h(x_k).$$

Diese Funktion $h(x)$ ist offenbar im Intervall $[a, b]$ wachsend und beschränkt. Sie kann nur eine endliche oder eine abzählbare Menge von Unstetigkeitsstellen $\xi_1, \xi_2, \dots$ haben. In allen Punkten x_k der überall im Intervall $[a, b]$ dichten Menge konvergiert die Folge $h_{n_k}(x)$ gegen $h(x)$. Nach Satz 1 in [12] findet auch in allen im Innern von $[a, b]$ liegenden Stetigkeitsstellen von $h(x)$ Konvergenz statt. Somit kann die Folge $h_{n_k}(x)$ nur in den Unstetigkeitsstellen ξ_k der Funktion $h(x)$ und im rechten Randpunkt des Intervalls nicht gegen $h(x)$ konvergieren. Wenden wir auf die Folge $h_{n_k}(x)$ noch einmal das Auswahlprinzip an, so ergibt sich, daß sie auch in diesen Punkten konvergiert [IV/1,15]. Damit ist das Lemma bewiesen.

Nun ist es leicht, den grundlegenden Satz 1 zu beweisen. Jede Funktion $g(x)$ von beschränkter Variation aus der Menge $\mathfrak{S}$ kann als Differenz zunehmender Funktionen dargestellt werden:

$$g(x) = \frac{1}{2} [V_a^x(g) + g(x)] - \frac{1}{2} [V_a^x(g) - g(x)], \quad (75)$$

wobei wegen (74) diese beiden zunehmenden Funktionen dem absoluten Betrag nach nicht größer als eine Zahl L sind. Durch Anwendung des Lemmas kann man zeigen, daß aus der Folge $g_n(x)$ eine solche Folge ausgewählt werden kann, für die der Minuend auf der rechten Seite von (75) in allen Punkten des Intervalls $[a, b]$ gegen eine Grenzfunktion konvergiert. Durch nochmalige Anwendung des Lemmas läßt sich zeigen, daß aus der erhaltenen Folge eine Teilfolge ausgewählt werden kann, für die auch der Subtrahend auf der rechten Seite von (75) in allen Punkten des Intervalls $[a, b]$ gegen eine Grenzfunktion strebt. Wir erhalten also eine solche Teilfolge $g_{n_k}(x)$, die in allen Punkten von $[a, b]$ gegen eine Grenzfunktion konvergiert. Damit ist der Satz bewiesen.

14. Der Raum der stetigen Funktionen. Wir betrachten die Menge aller in einem gegebenen endlichen Intervall $[a, b]$ reellwertigen und stetigen Funktionen und nennen diese Menge der Kürze halber den *Raum* C . Eine beliebige, in $[a, b]$ gegebene stetige Funktion heißt *Element* dieses Raumes oder, wie wir auch sagen werden, *Vektor* dieses Raumes. Voneinander verschiedene stetige Funktionen stellen voneinander verschiedene Elemente des Raumes C dar. Die in $[a, b]$ identisch verschwindende Funktion heißt *Null-*

element dieses Raumes. Wenn wir eine beliebige endliche Linearkombination $c_1 f_1(x) + c_2 f_2(x) + \dots + c_m f_m(x)$ von in $[a, b]$ stetigen Funktionen mit reellen Koeffizienten bilden, erhalten wir wieder eine in $[a, b]$ stetige reelle Funktion. Wenn man also Elemente des Raumes C mit reellen Zahlen multipliziert und addiert, erhält man wieder ein Element aus C . Diese Operation ordnet sich den gewöhnlichen Gesetzen der Algebra unter, z. B.

$$\begin{aligned} f_1(x) + f_2(x) &= f_2(x) + f_1(x), & c[f_1(x) + f_2(x)] &= cf_1(x) + cf_2(x), \\ (c_1 + c_2)f(x) &= c_1 f(x) + c_2 f(x), & c_1(c_2 f(x)) &= (c_1 c_2)f(x). \end{aligned}$$

Wir führen nun im Raum C den Begriff der Norm eines Elements, oder anders ausgedrückt, der Länge eines Vektors ein. Wir verstehen unter der Norm von $f(x)$ den größten Wert, den $|f(x)|$ in $[a, b]$ annimmt. Das Nullelement hat die Norm 0, jedes andere Element hat eine positive Norm. Wir bezeichnen die Norm eines Elementes $f(x)$ mit $\|f\|$. Schließlich führen wir im Raum C den Begriff der Konvergenz ein. Wir sagen, daß eine Folge von Elementen $f_n(x)$ aus C gegen ein Element $f(x)$ aus C konvergiert, wenn $\|f(x) - f_n(x)\| \rightarrow 0$ gilt. Diese Bedingung ist gleichbedeutend damit, daß der größte Wert von $|f(x) - f_n(x)|$ im Intervall $[a, b]$ gegen 0 strebt. Die letzte Bedingung besagt offenbar, daß die $f_n(x)$ gleichmäßig in $[a, b]$ gegen $f(x)$ konvergieren.

Wir führen nun die Begriffe Funktional und Operator im Raum C ein. Unter einem Funktional in C verstehen wir jede Vorschrift, mit deren Hilfe jedem Element $f(x)$ aus C eine bestimmte reelle Zahl zugeordnet wird. Für Funktionale werden oft die Bezeichnungen $\Phi[f(x)]$, $\Psi[f(x)]$ usw. benutzt. Der Begriff des Funktionals stellt eine Abart des gewöhnlichen Funktionsbegriffes dar. Als Argument treten im Fall des Funktionals Elemente aus C auf, und als Wert des Funktionals erscheint eine reelle Zahl. Ein Funktional $\Phi[f(x)]$ heißt *distributiv*, wenn es für jede endliche Linearkombination $c_1 f_1(x) + c_2 f_2(x) + \dots + c_m f_m(x)$ der Gleichung

$$\Phi[c_1 f_1(x) + c_2 f_2(x) + \dots + c_m f_m(x)] = c_1 \Phi[f_1(x)] + c_2 \Phi[f_2(x)] + \dots + c_m \Phi[f_m(x)] \quad (76)$$

genügt. Ein Funktional $\Phi[f(x)]$ heißt *beschränkt*, wenn eine solche positive Zahl N existiert, daß für jedes Element $f(x)$ aus C die Ungleichung

$$|\Phi[f(x)]| \leq N \|f(x)\| \quad (77)$$

erfüllt ist. Auf der linken Seite dieser Ungleichung steht der absolute Betrag der reellen Zahl $\Phi[f(x)]$, durch die der Wert des Funktionals in C für das Element $f(x)$ ausgedrückt wird, rechts steht dagegen das Produkt einer positiven Zahl N mit der Norm des Elementes $f(x)$, d. h. das Produkt von N mit dem größten Wert, den $|f(x)|$ im Intervall $[a, b]$ annimmt. Ein distributives und beschränktes Funktional heißt ein *lineares Funktional*. Man kann schließlich den Begriff der Stetigkeit eines Funktionals einführen, und zwar heißt ein Funktional $\Phi[f(x)]$ *stetig*, wenn die folgende Bedingung erfüllt ist: Konvergiert die Folge der $f_n(x)$ gleichmäßig in $[a, b]$ gegen $f(x)$, so konvergieren die Werte $\Phi[f_n(x)]$ gegen $\Phi[f(x)]$. Man sieht leicht, daß lineare Funktionale stetig sind. Unter Benutzung von (76) und (77) kann man nämlich

$$|\Phi[f(x)] - \Phi[f_n(x)]| = |\Phi[f(x) - f_n(x)]| \leq N \|f(x) - f_n(x)\|$$

schreiben, und da wegen der gleichmäßigen Konvergenz der $f_n(x)$ gegen $f(x)$ die Beziehung $\|f(x) - f_n(x)\| \rightarrow 0$ und folglich $\Phi[f(x)] - \Phi[f_n(x)] \rightarrow 0$ gilt, ergibt sich $\Phi[f_n(x)] \rightarrow \Phi[f(x)]$. Bei der Definition des linearen Funktionals könnte man auch Distributivität und Stetigkeit voraussetzen und daraus seine Beschränktheit herleiten, d. h., unter der Voraussetzung der Distributivität sind Beschränktheit und Stetigkeit eines Funktionals gleichbedeutend. Auf den Beweis dieser Tatsache, der keine Schwierigkeiten bietet, werden wir nicht eingehen.

Wir bringen ein Beispiel eines Funktionals. Es sei x_0 ein fester Punkt im Intervall $[a, b]$. Die Werte $f(x_0)$ der stetigen Funktionen in diesem Punkt stellen ein lineares Funktional in C dar. Das bestimmte Integral $\int_a^b f(x) dx$ ist ebenfalls ein lineares Funktional. Es sei $g(x)$ eine Funktion von beschränkter Variation in $[a, b]$. Für jedes Element

$f(x)$ aus C können wir das Stieltjessche Integral

$$\Phi[f(x)] = \int_a^b f(x) dg(x) \quad (78)$$

bilden. Es stellt ein lineares Funktional $\Phi[f(x)]$ dar. Seine Distributivität ergibt sich aus der Distributivität des Integrals bezüglich der Funktion $f(x)$, seine Beschränktheit folgt unmittelbar aus der Abschätzung

$$\left| \int_a^b f(x) dg(x) \right| \leq L V_a^b(g),$$

in der L der größte Wert von $|f(x)|$ im Intervall $[a, b]$ ist. Für das Funktional (78) kann also die Totalvariation $V_a^b(g)$ die Rolle der in der Ungleichung (77) auftretenden Zahl N spielen.

Wir wenden uns der Ungleichung (77) zu. Wenn sie für eine beliebig gewählte positive Zahl N gilt, gilt sie erst recht für größere Werte von N . Wir zeigen jetzt, daß ein kleinster Wert N existiert, für den die Ungleichung (77) gilt.

Wir bemerken zuerst: Gilt $\|f\| = 0$, so ist $f(x)$ in $[a, b]$ identisch gleich 0, d. h., $f(x)$ ist das Nullelement von C , wenn $\|f\| = 0$ ist. Für jedes andere Element ist $\|f\| > 0$. Das Nullelement von C kann man in Gestalt eines Produktes $f_0(x) = 0 \cdot f(x)$ darstellen, wobei $f(x)$ eine beliebige stetige Funktion ist. Aus (76) folgt $\Phi(f_0) = 0 \cdot \Phi(f) = 0$, d. h., jedes distributive Funktional hat für das Nullelement von C den Wert 0. Somit hat für das Nullelement die Ungleichung (77) die Form $0 \leq N \cdot 0$ für beliebig gewähltes N . Wir brauchen also (77) nur für $\|f\| > 0$ zu betrachten.

Unter Berücksichtigung von (76) können wir (77) in der Gestalt

$$\left| \Phi \left[\frac{f(x)}{\|f\|} \right] \right| \leq N \quad (79)$$

schreiben. Das Element $\varphi(x) = \frac{f(x)}{\|f\|}$ ist aber ein Element von C mit der Norm 1.

Wir bezeichnen mit n_Φ die obere Grenze der nichtnegativen Zahlen $|\Phi(\varphi)|$, wobei φ ein beliebiges Element von C mit der Norm 1 ist:

$$n_\Phi = \sup_{\|\varphi\|=1} |\Phi(\varphi)|. \quad (80)$$

Aus (79) folgt unmittelbar, daß n_Φ der kleinstmögliche Wert von N in Ungleichung (77) ist. Diese Zahl n_Φ heißt *Norm des Funktionals $\Phi(f)$* . Oft wird sie auch mit $\|\Phi\|$ bezeichnet. Es ist

$$|\Phi(f)| \leq n_\Phi \|f\|, \quad (81)$$

aber (77) kann nicht für alle $f(x)$ aus C gelten, wenn $N < n_\Phi$ genommen wird. Stets ist $n_\Phi \geq 0$. Im Fall $n_\Phi = 0$ folgt aus (81) die Beziehung $\Phi(f) = 0$ für jedes Element $f(x)$ aus C , d. h., das Funktional $\Phi(f)$ ist für jedes Element $f(x)$ gleich 0. Aus dem oben Gesagten folgt, daß die Norm des durch (78) dargestellten Funktionals nicht größer als $V_a^b(g)$ ist. Wir erinnern daran, daß wir in Teil IV den Raum F der stetigen Funktionen mit einer anderen Definition der Norm eines Elements betrachtet haben [IV/1, 33].

15. Die allgemeine Form der Funktionale in C . Wir beweisen jetzt einen wichtigen Satz von F. RIESZ, der besagt, daß jedes lineare Funktional $\Phi(f)$ in C durch die Formel (78) dargestellt werden kann, wobei $g(x)$ eine Funktion von beschränkter Variation ist, die dieses Funktional definiert. Wie wir oben gesehen haben, ist jedes Integral (78), wo $g(x)$ eine fest vorgegebene Funktion von beschränkter Variation ist, tatsächlich ein lineares Funktional in C . Somit kann man nach dem Beweis des Satzes von RIESZ sagen, daß das Integral (78), wo $g(x)$ eine Funktion von beschränkter Variation ist, die allgemeine Form des linearen Funktionals in C ist.

Die Totalvariation der Funktion $g_n(x)$ ist offenbar gleich der Summe über die absoluten Beträge der Sprünge von $g_n(x)$ in den Teilpunkten und in den Randpunkten des Intervalls. Wegen (85) gilt $V_a^b(g_n) \leq n_\Phi$. In genau derselben Weise folgt aus (85) und der Definition der Funktion $g_n(x)$ unmittelbar die Ungleichung $|g_n(x)| \leq n_\Phi$. Also ist Satz 1 aus [18] auf die Funktionenfolge $g_n(x)$ anwendbar, und wir können schließen, daß eine wachsende Folge ganzer positiver Zahlen n_k existiert, für die in allen Punkten von $[0, 1]$ die zugehörige Funktionenfolge $g_{n_k}(x)$ gegen eine Funktion $g(x)$ von beschränkter Variation konvergiert. Wir beweisen nun, daß diese Funktion $g(x)$ gerade die Funktion der rechten Seite von (78) ist. Wir bilden dazu das Stieltjessche Integral von $f(x)$ bezüglich $g_n(x)$. Es ist gleich der Summe über die Produkte der Funktionswerte von $f(x)$ in den Unstetigkeitsstellen von $g_n(x)$ mit den Sprunggrößen in diesen Punkten, d. h. (vgl. [6])

$$\int_0^1 f(x) dg_n(x) = \sum_{m=0}^n f\left(\frac{m}{n}\right) \Phi\left[\binom{n}{m} x^m (1-x)^{n-m}\right].$$

Auf Grund der Formeln (76) und (82) stellt die rechte Seite den Wert des Funktional $\Phi[f(x)]$ für $f(x) = P_n(x)$ dar, d. h.

$$\int_0^1 f(x) dg_n(x) = \Phi[P_n(x)].$$

Für $n = n_k$ erhalten wir

$$\int_0^1 f(x) dg_{n_k}(x) = \Phi[P_{n_k}(x)]. \quad (87)$$

Wenn n_k unbegrenzt wächst, streben die $P_{n_k}(x)$ gleichmäßig in $[0, 1]$ gegen $f(x)$, und wegen der Stetigkeit des Funktional $\Phi[f(x)]$ ergibt die Formel (87)

$$\Phi[f(x)] = \lim_{n_k \rightarrow \infty} \int_0^1 f(x) dg_{n_k}(x).$$

Auf die rechte Seite ist der Satz 2 aus [12] anwendbar, und wir kommen schließlich zu der Darstellung

$$\Phi[f(x)] = \int_0^1 f(x) dg(x). \quad (88)$$

Wir zeigen jetzt, daß $n_\Phi = V_0^1(g)$ ist. Aus (88) folgt unmittelbar, daß, wie wir in [14] sahen, $n_\Phi \leq V_0^1(g)$ ist. Andererseits folgt aus der oben bewiesenen Ungleichung $V_0^1(g_{n_k}) \leq n_\Phi$, daß auch $V_0^1(g) \leq n_\Phi$ ist. Aus diesen beiden Ungleichungen folgt aber $n_\Phi = V_0^1(g)$. Wir fragen jetzt nach der Eindeutigkeit des durch Formel (88) dargestellten Funktional $\Phi(f)$. Es sei $h(x)$ eine Funktion von beschränkter Variation, deren Werte auf einer endlichen oder abzählbar unendlichen Punktmenge $\mathcal{E}$ des Intervalls $[a, b]$ von den Werten von $g(x)$ verschieden sind. Es ist leicht zu zeigen, daß das Integral

$$\int_0^1 f(x) dh(x) \quad (89)$$

für eine beliebig gewählte stetige Funktion $f(x)$ gleich dem Integral (88) ist. Da nämlich die Menge der Punkte jedes in $[0, 1]$ liegenden Intervalls die Mächtigkeit des Kontinuums hat, bilden die nicht in $\mathcal{E}$ liegenden Punkte eine in $[0, 1]$ dichte Menge. Somit können wir beim Aufstellen der Riemann-Stieltjesschen Summen für das Integral (88)

$$\sum_{k=1}^n f(\xi_k) [g(x_k) - g(x_{k-1})]$$

und bei beliebiger Verfeinerung der Zerlegung Teilpunkte auswählen, die nicht in $\mathcal{E}$ liegen; hieraus folgt unmittelbar, daß die Integrale (88) und (89) übereinstimmen.

Wenn wir also $g(x)$ in einer endlichen oder abzählbar unendlichen Menge von Punkten so ändern, daß auch die neue Funktion $h(x)$ von beschränkter Variation ist, und das Integral (89) bilden, erhalten wir dasselbe lineare Funktional in C wie durch (88). Nach dem in [14] Gesagten gilt $n_\Phi \leq V_0^1(h)$, wobei auch das Zeichen $<$ auftreten kann. Konstruieren wir aber $g(x)$ so wie beim Beweis des Satzes, so erhalten wir $n_\Phi = V_a^b(g)$.

Wir stellen jetzt folgende allgemeine Frage: Für welche Funktionen $h(x)$ von beschränkter Variation definiert das Integral (89) dasselbe lineare Funktional in C wie das Integral (88)? Führen wir eine neue Funktion $\omega(x) = h(x) - g(x)$ von beschränkter Variation ein, so können wir die Frage folgendermaßen formulieren: Für welche Funktionen $\omega(x)$ von beschränkter Variation gilt

$$\int_0^1 f(x) d\omega(x) = 0 \quad (90)$$

für jede in $[0, 1]$ stetige Funktion $f(x)$? Wir beantworten diese Frage durch den folgenden

Satz 2. *Für eine beliebig gewählte stetige Funktion $f(x)$ gilt genau dann die Formel (90), wenn die Funktion $\omega(x)$ von beschränkter Variation ist und den folgenden Bedingungen genügt:*

a) *In jeder Stetigkeitsstelle $x = x_0$ von $\omega(x)$, die im Innern von $[0, 1]$ liegt, gilt die Beziehung $\omega(x_0) = \omega(0)$;*

b) $\omega(1) = \omega(0)$.

Beweis. Wir beweisen zuerst die Notwendigkeit. Wir wählen ein genügend großes n , so daß $x_0 > \frac{1}{n}$ ist und der Punkt $x_0 + \frac{1}{n}$ innerhalb von $[0, 1]$ liegt, und definieren eine stetige Funktion $f(x)$ folgendermaßen:

$$f(x) = \begin{cases} 1 & \text{für } 0 \leq x \leq x_0, \\ -nx + (1 + nx_0) & \text{für } x_0 < x < x_0 + \frac{1}{n}, \\ 0 & \text{für } x_0 + \frac{1}{n} \leq x \leq 1. \end{cases}$$

In dem mittleren Intervall $\left[x_0, x_0 + \frac{1}{n}\right]$ ist $f(x)$ eine lineare Funktion, die von 1 bis 0 abnimmt. Die Formel (90) liefert bei dieser Wahl von $f(x)$

$$\int_0^{x_0} d\omega(x) + \int_{x_0}^{x_0 + \frac{1}{n}} [-nx + (1 + nx_0)] d\omega(x) = 0$$

oder

$$\omega(x_0) - \omega(0) + \int_{x_0}^{x_0 + \frac{1}{n}} [-nx + (1 + nx_0)] d\omega(x) = 0. \quad (91)$$

Nun ist aber [9]

$$\left| \int_{x_0}^{x_0 + \frac{1}{n}} [-nx + (1 + nx_0)] d\omega(x) \right| \leq 1 \cdot V_{x_0}^{x_0 + \frac{1}{n}}(\omega).$$

Weil $\omega(x)$ nach Voraussetzung in $x = x_0$ stetig ist, folgt $V_{x_0}^{x_0 + \frac{1}{n}}(\omega) \rightarrow 0$ für $n \rightarrow \infty$, und die Formel (91) liefert in der Grenze $\omega(x_0) = \omega(0)$. Die zweite Bedingung $\omega(1) = \omega(0)$ erhält man aus (90), wenn man $f(x) = 1$ setzt.

Wir beweisen jetzt, daß die Bedingungen hinreichend sind. Weil die Stetigkeitspunkte einer Funktion von beschränkter Variation im Intervall $[0, 1]$ dicht liegen, brauchen wir bei der Bildung der Riemann-Stieltjesschen Summen nur diese Punkte zu benutzen. Da auf Grund der Bedingung $\omega(x_0) = \omega(0) = \omega(1)$ dabei alle Differenzen $\omega(x_k) - \omega(x_{k-1})$ gleich 0 werden, gilt also Formel (90) bei beliebiger Wahl der stetigen Funktion $f(x)$. Damit ist der Satz bewiesen.

Genauso können wir beweisen, daß das Integral (89) dann und nur dann dasselbe lineare Funktional in C liefert wie das Integral (88), wenn die Differenz $h(x) - g(x)$ in allen Stetigkeitspunkten und in $x=1$ gleich $h(0) - g(0)$ ist.

Wir wenden uns jetzt wieder der Formel (89) zu. Bekanntlich muß bei jeder Darstellung von $\Phi(f)$ durch ein Integral (89) die Ungleichung $n_\Phi \leq V_0^1(h)$ gelten. In (88) ist $n_\Phi = V_0^1(g)$. Natürlich kann man zu $g(x)$ einen beliebigen konstanten Summanden hinzufügen. Um diese Mehrdeutigkeit auszuschließen, wollen wir in (88) $g(0) = 0$ annehmen. Die Funktion $g(x)$ kann nur endlich viele oder abzählbar viele Unstetigkeitsstellen haben. Wir betrachten zunächst eine solche Unstetigkeitsstelle ξ , für die $g(\xi - 0) = g(\xi + 0)$, aber $g(\xi) \neq g(\xi - 0)$ ist (also eine sogenannte hebbare Unstetigkeit). Wenn wir $g(x)$ in dem einen Punkt $x = \xi$ abändern und $g(\xi) = g(\xi - 0)$ setzen, beseitigen wir die Unstetigkeit. Das Integral (88) ändert sich dabei nicht, aber $V_0^1(g)$ wird kleiner. Das kann aber nicht sein, da wir früher $V_0^1(g) = n_\Phi$ hatten, nach der Änderung jedoch $V_0^1(g) < n_\Phi$ erhalten würden. Das ist aber unmöglich. Somit folgt aus $V_0^1(g) = n_\Phi$, daß $g(x)$ keine hebbaren Unstetigkeitsstellen besitzt. Es bleiben noch die Unstetigkeiten mit $g(\xi - 0) \neq g(\xi + 0)$ zu untersuchen. Wenn wir zunächst einen Wert $g(\xi)$ auswählen, der in dem abgeschlossenen Intervall i mit den Randpunkten $g(\xi - 0)$ und $g(\xi + 0)$ liegt, und dann einen außerhalb dieses Intervalls, so wird $V_0^1(g)$ im zweiten Fall offenbar größer als im ersten. Somit folgt in den nicht hebbaren Unstetigkeitsstellen aus der Bedingung $V_0^1(g) = n_\Phi$, daß $g(\xi)$ in dem abgeschlossenen Intervall i liegt. Die Lage der Zahlen $g(\xi)$ in dem Intervall i hat keinen Einfluß auf die Totalvariation $V_a^b(g)$. Also ist unter der Bedingung $V_0^1(g) = n_\Phi$ die Definition von $g(\xi)$ in den Unstetigkeitsstellen ganz beliebig, sofern nur $g(\xi)$ in i liegt. Meist setzt man

$$g(\xi) = \frac{g(\xi - 0) + g(\xi + 0)}{2},$$

manchmal auch $g(\xi) = g(\xi + 0)$, dann hat man Stetigkeit von rechts, oder $g(\xi) = g(\xi - 0)$, dann hat man Stetigkeit von links.

Hierzu analog kann man den Raum C der in dem Intervall $[0, 1]$ stetigen komplexen Funktionen $\varphi(x) + i\psi(x)$ betrachten. Die Elemente dieses Raumes kann man addieren und mit komplexen Zahlen multiplizieren. Norm und lineares Funktional werden genauso wie im reellen Fall definiert, allerdings kann ein Funktional jetzt auch komplexe Werte annehmen. Es gilt auch der Satz über die allgemeine Form eines linearen Funktionals, wobei die komplexe Funktion von beschränkter Variation die Form $\lambda(x) + i\mu(x)$ hat; dabei sind λ und μ reelle Funktionen von beschränkter Variation.

16. Lineare Operatoren in C . Wir gehen über zur Definition eines Operators im Raum C . Unter einem Operator in C verstehen wir eine Vorschrift, mit deren Hilfe jedem Element $f(x)$ aus C ein bestimmtes ebenfalls in C liegendes Element $\varphi(x)$ zugeordnet wird.

Wir führen für Operatoren die Bezeichnung $F[f(x)]$ ein. Für eine beliebige Funktion $f(x)$ aus C definiert $F[f(x)]$ eine gewisse, ebenfalls in C liegende Funktion $\varphi(x)$. Die Distributivität eines Operators wird ebenso wie die Distributivität eines Funktionals, d. h. durch eine zu (76) analoge Formel definiert. Die Beschränktheit wird durch die folgende, zu (77) analoge Ungleichung definiert, in der nur auf der linken Seite an Stelle des absoluten Betrages die Norm zu schreiben ist, denn $F[f(x)]$ ist keine Zahl, sondern ein Element aus C :

$$\|F[f(x)]\| \leq N \|f(x)\|. \quad (77_1)$$

Ein distributiver und beschränkter Operator heißt *linear*. Ein solcher Operator ist sicher *stetig*, d. h., aus $f_n(x) \rightarrow f(x)$ folgt $F[f_n(x)] \rightarrow F[f(x)]$, wobei in beiden Fällen die

Konvergenz als gleichmäßige Konvergenz der entsprechenden Funktionenfolge im Intervall $[a, b]$ aufzufassen ist.

Wir geben ohne Beweis ein grundlegendes allgemeines Ergebnis an, das sich auf die allgemeine Form eines linearen Operators in C bezieht. Es sei $g(x, y)$ eine in dem abgeschlossenen zweidimensionalen Intervall $0 \leq x \leq 1, 0 \leq y \leq 1$ definierte Funktion. Bei beliebigen im Intervall $[0, 1]$ liegenden Werten von y sei $g(x, y)$ von beschränkter Variation bezüglich x im Intervall $[0, 1]$. Wenn wir diese Funktion $g(x, y)$ in die rechte Seite von (88) einführen, so erhalten wir als Ergebnis der Integration nicht mehr eine Zahl, sondern eine gewisse Funktion eines in dem Intervall $[0, 1]$ definierten Parameters y :

$$\varphi(y) = \int_0^1 f(x) \, dx g(x, y).$$

Diese Formel liefert dann und nur dann einen linearen Operator, wenn die Funktion $g(x, y)$ außer den oben angegebenen Eigenschaften noch die folgende Eigenschaft besitzt: Die durch die obige Formel definierte Funktion $\varphi(y)$ stellt bei beliebiger Wahl der in $[0, 1]$ stetigen Funktion $f(x)$ eine im Intervall $[0, 1]$ stetige Funktion dar. Ist y_0 ein beliebiger Wert des Intervalls $[0, 1]$ und y_n ($n=1, 2, 3, \dots$) eine Folge von Zahlen aus dem Intervall $[0, 1]$, die y_0 zum Grenzwert hat, so führt uns die Definition der Stetigkeit von $\varphi(y)$ im Punkt y_0 unmittelbar zu der folgenden notwendigen Bedingung, der die Funktion $g(x, y)$ genügen muß: Für jedes y_0 aus dem Intervall $[0, 1]$ und jede in diesem Intervall stetige Funktion $f(x)$ muß die Formel

$$\lim_{y_n \rightarrow y_0} \int_0^1 f(x) \, dx g(x, y_n) = \int_0^1 f(x) \, dx g(x, y_0) \quad (92)$$

gelten, in der die y_n eine beliebige Zahlenfolge aus dem Intervall $[0, 1]$ ist, die y_0 zum Grenzwert hat. Funktionen $g(x, y)$ mit dieser Eigenschaft heißen *schwach stetig bezüglich des Parameters y* . Ist $g(x, y)$ von beschränkter Variation bezüglich x und schwach stetig bezüglich y , so liefert die Formel (92) einen linearen Operator in C . Man kann auch den umgekehrten Satz beweisen: Jeder lineare Operator in C ist durch die Formel (92) darstellbar, in der $g(x, y)$ bezüglich x von beschränkter Variation und bezüglich y schwach stetig ist.¹⁾

Ist $K(x, y)$ eine in dem zweidimensionalen Intervall $0 \leq x \leq 1, 0 \leq y \leq 1$ stetige Funktion, so wird durch

$$\varphi(y) = \int_0^1 K(y, x) f(x) \, dx$$

offensichtlich ein linearer Operator in C definiert. Mit solchen Operatoren hatten wir in der Theorie der Integralgleichungen zu tun. Nicht jeder Operator in C kann jedoch in dieser Form dargestellt werden.

17. Intervallfunktionen. Bei den weiteren Verallgemeinerungen des Integralbegriffes ist es zweckmäßiger, an Stelle von Punktfunktionen *Intervallfunktionen* zu verwenden. Für alle Werte von x im Intervall $(-\infty, \infty)$ sei eine nicht abnehmende beschränkte Funktion $g(x)$ gegeben. Wir ordnen jedem halboffenen Intervall $\Delta = (\alpha, \beta]$ die nichtnegative Zahl $g(\beta+0) - g(\alpha+0)$ (die in diesem Intervall enthaltene Masse) zu. Auf diese Weise erhalten wir eine Funktion halboffener Intervalle, die wir mit $G(\Delta)$ bezeichnen:

$$G(\Delta) = g(\beta+0) - g(\alpha+0). \quad (93)$$

Zur Formulierung der Eigenschaften dieser Funktion führen wir einen neuen Be-

¹⁾ Der Beweis dieses Satzes und der Begriff „schwach stetig“ wird z. B. erläutert in dem Buch von W. I. GLIWENKO, Das Stieltjessche Integral (vgl. das Literaturverzeichnis).

griff ein. Wir nennen eine Folge $\Delta^{(1)}, \Delta^{(2)}, \dots$ halboffener Intervalle eine *verschwindende Folge*, wenn jedes Intervall $\Delta^{(k+1)}$ im Intervall $\Delta^{(k)}$ liegt und kein einziger Punkt in allen Intervallen enthalten ist. Wir untersuchen die Struktur einer verschwindenden Folge von Intervallen. Es sei $\Delta^{(k)} = (a_k, b_k]$. Nach unserer Definition gilt $a_k \leq a_{k+1}$, $b_k \geq b_{k+1}$ und $b_k - a_k \rightarrow 0$ bei unbegrenzt wachsendem k . Die monotonen Folgen a_k und b_k haben einen gemeinsamen Grenzwert c , für den bei beliebigem k die Ungleichung $a_k \leq c \leq b_k$ gilt. Da kein einziger Punkt in allen Intervallen $\Delta^{(k)}$ enthalten ist, ist speziell der Punkt c nicht in allen diesen Intervallen enthalten; für alle genügend großen Werte von k muß also der Punkt c mit dem linken Randpunkt des offenen Intervalls übereinstimmen, d. h., $\Delta^{(k)}$ ist für alle genügend großen Werte von k gleich $(c, b_k]$, und es gilt $b_k \rightarrow c$. Wegen $g(b_k + 0) \rightarrow g(c + 0)$ gilt für solche Werte von k

$$G(\Delta^{(k)}) = g(b_k + 0) - g(c + 0) \rightarrow 0.$$

Aus der Definition (93) und den soeben angestellten Überlegungen ergeben sich unmittelbar die folgenden drei Haupteigenschaften von $G(\Delta)$:

1. $G(\Delta)$ ist nicht negativ;
2. $G(\Delta)$ ist additiv, d. h., wenn das halboffene Intervall Δ in eine endliche Anzahl paarweise punktfremder halboffener Intervalle $\Delta_1, \Delta_2, \dots, \Delta_q$ zerlegt wird, was man gewöhnlich durch

$$\Delta = \Delta_1 + \Delta_2 + \dots + \Delta_q \quad (94)$$

ausdrückt, dann gilt

$$G(\Delta) = \sum_{k=1}^q G(\Delta_k); \quad (95)$$

3. die Funktion $G(\Delta)$ strebt für jede verschwindende Intervallfolge gegen 0.

Diese letzte Eigenschaft nennen wir die *Normalität* der Funktion $G(\Delta)$. Ihre physikalische Bedeutung ist klar. Wir betrachten jedoch nicht nur nach links offene und nach rechts abgeschlossene, sondern alle möglichen Intervalle $(\alpha, \beta]$, $[\alpha, \beta)$, $[\alpha, \beta]$, (α, β) und fassen dabei auch einen einzelnen Punkt α als Intervall $[\alpha, \alpha]$ auf. Gehen wir von einer nicht abnehmenden beschränkten Punktfunktion $g(x)$ aus, dann können wir für beliebige Intervalle die Intervallfunktion $G(\Delta)$ mit den drei angegebenen Eigenschaften bilden. Dazu genügt es, neben der Definition (93) noch die Definitionen

$$\left. \begin{aligned} G([\alpha, \beta)) &= g(\beta - 0) - g(\alpha - 0), \\ G([\alpha, \beta]) &= g(\beta + 0) - g(\alpha - 0), \\ G((\alpha, \beta)) &= g(\beta - 0) - g(\alpha + 0) \end{aligned} \right\} \quad (96)$$

einzuführen. Im Fall eines nur aus einem einzigen Punkt α bestehenden Intervalls $[\alpha, \alpha]$ gilt

$$G([\alpha, \alpha]) \equiv G([\alpha, \alpha]) = g(\alpha + 0) - g(\alpha - 0). \quad (97)$$

Ist die nicht abnehmende Funktion $g(x)$ nur in einem endlichen Intervall, z. B. in einem nach links offen und nach rechts abgeschlossenen Intervall $(a, b]$ definiert, so kann man sie durch die Definition $g(x) = g(a + 0)$ für $x \leq a$ und $g(x) = g(b)$ für $x > b$ für alle Werte von x fortsetzen.

Wir sind von einer nicht abnehmenden Punktfunktion $g(x)$ ausgegangen und zum Begriff der Intervallfunktion $G(\Delta)$ gelangt. Man kann umgekehrt auch von einer Intervallfunktion $G(\Delta)$ mit den angegebenen drei Eigenschaften ausgehen und eine

Punktfunktion $g(x)$ bilden, die nach der obigen Methode wieder zu $G(\Delta)$ führt. Man braucht dazu nur

$$g(x) = G((-\infty, x]) \quad (98)$$

zu setzen. Diese Funktion $g(x)$ ist offensichtlich von rechts stetig. Ist $G(\Delta)$ nur für Intervalle definiert, die in einem gewissen Intervall Δ_0 liegen, so kann man sie für alle Intervalle definieren, wenn man $G(\Delta) = G(\Delta \cdot \Delta_0)$ annimmt. Unter dem Produkt $\Delta \cdot \Delta_0$ zweier Intervalle Δ und Δ_0 verstehen wir dabei das Intervall, das aus den Punkten besteht, die gleichzeitig in Δ und Δ_0 liegen. Wenn es keine solchen Punkte gibt, d. h., wenn $\Delta \cdot \Delta_0$ die leere Menge ist, setzen wir natürlich $G(\Delta \cdot \Delta_0) = 0$. Wir bemerken noch, daß es auf die Werte von $G(\Delta)$ keinen Einfluß hat, wenn zu $g(x)$ eine beliebige Konstante addiert wird. Ist $g(x)$ eine Funktion von beschränkter Variation, so benutzen wir sie in der kanonischen Darstellung als Differenz zweier nicht abnehmender Funktionen, $g(x) = g_1(x) - g_2(x)$. Die Funktionen $g_1(x)$ und $g_2(x)$ führen uns zu den Intervallfunktionen $G_1(\Delta)$ und $G_2(\Delta)$ mit den angegebenen drei Eigenschaften, die Funktion $g(x)$ zu der Intervallfunktion $G(\Delta) = G_1(\Delta) - G_2(\Delta)$. Diese Funktion $G(\Delta)$ kann auch nach den Formeln (93), (96) und (97) unmittelbar aus der Funktion $g(x)$ hergeleitet werden. Sie ist additiv und normal, kann aber auch negative Werte annehmen.

Für das abgeschlossene Intervall $[-\infty, +\infty]$ müssen wir statt (98)

$$g(x) = c + G((-\infty, x]) \quad (99)$$

setzen, wobei c eine Konstante ist ($c = g(-\infty)$).

18. Das allgemeine Stieltjessche Integral. Wir gehen nun zur Verallgemeinerung des Stieltjesschen Integralbegriffes über. Wie wir bereits in [8] erwähnt haben, erhalten wir diese Verallgemeinerung, wenn wir nur fordern, daß die dort definierten Zahlen i und I übereinstimmen. Außerdem betrachten wir die Integration in Intervallen Δ beliebigen Typs und zerlegen auch Δ in Teilintervalle Δ_k beliebigen Typs, die weder im Innern noch auf dem Rand gemeinsame Punkte haben. Natürlich lassen wir auch einzelne Punkte als Teilintervalle zu.

Gegeben sei ein endliches oder unendliches Intervall Δ , und in ihm seien zwei beschränkte Funktionen $f(x)$ und $g(x)$ definiert; $g(x)$ sei dabei eine nicht abnehmende Funktion. Wir zerlegen Δ in beliebige punktfremde Teilintervalle Δ_k ($k = 1, 2, \dots, p$). Es seien m_k und M_k die untere bzw. die obere Grenze von $f(x)$ in Δ_k und ξ_k ein beliebiger Punkt aus Δ_k . Neben $g(x)$ führen wir die Intervallfunktion $G(\Delta_k)$ ein, die in [17] definiert wurde, und bilden die Summen

$$s_\delta = \sum_{k=1}^p m_k G(\Delta_k), \quad S_\delta = \sum_{k=1}^p M_k G(\Delta_k), \quad \sigma_\delta = \sum_{k=1}^p f(\xi_k) G(\Delta_k). \quad (100)$$

Wir verweisen auf einige Besonderheiten, die mit der Verwendung beliebiger Intervalltypen zusammenhängen. Ist der Punkt P mit der Abszisse x selbständiges Element einer Zerlegung von Δ , so sind in den Summen (100) die entsprechenden Summanden identisch und haben die Form $f(x)[g(x+0) - g(x-0)]$. Stetigkeitsstellen von $g(x)$ als selbständige Elemente einer Zerlegung einzuführen, ist nicht sinnvoll, und wir können im folgenden annehmen, daß die Punkte, die als selbständige Elemente einer Zerlegung auftreten, Unstetigkeitsstellen von $g(x)$ sind.

Sind Δ' und Δ'' zwei Intervalle, so ist ihr Produkt $\Delta'\Delta''$ die Menge der Punkte, die gleichzeitig zu Δ' und Δ'' gehören. Es ist entweder auch ein Intervall oder die leere Menge. Es seien δ' und δ'' zwei Zerlegungen von Δ . Das Produkt der Zerlegungen δ' und δ'' ist die Zerlegung $\delta'\delta''$, die aus allen möglichen Intervallen $\Delta'\Delta''$ besteht, wo Δ' zu δ' und Δ'' zu δ'' gehört. Die Intervalle $\Delta'\Delta''$ sind offensichtlich paarweise punktfremd, und ihre Summe ergibt das Ausgangsintervall Δ . Die Zerlegung δ' heißt Fortsetzung

der Zerlegung δ , wenn jedes Element von δ' ganz zu einem Element der Zerlegung δ gehört. Das Produkt $\delta'\delta''$ ist eine Fortsetzung sowohl von δ' als auch von δ'' . Ist das Intervall rechtsseitig abgeschlossen und $x=b$ der rechte Randpunkt von Δ , so gilt nach Definition von $G(\Delta)$

$$g(b+0)=g(b).$$

Analog gilt für den linken Randpunkt

$$g(a-0)=g(a).$$

Das bezieht sich auch auf den Fall, daß $b=\infty$ oder $a=-\infty$ ist. Wie schon in [8] bezeichnen wir mit i die obere Grenze der Summen s_δ und mit I die untere Grenze der Summen S_δ bei allen möglichen Zerlegungen. Für s_δ , S_δ , σ_δ , i und I gilt dann das in [8] Gesagte.

Definition 1. Man sagt, $f(x)$ sei bezüglich $g(x)$ (oder $G(\Delta)$) *integrierbar*, wenn $i=I$ gilt, und diese Größe i nimmt man als Wert des Integrals

$$\int_{\Delta} f(x) dg(x) = \int_{\Delta} f(x) G(d\Delta). \quad (101)$$

Das Integral (101) heißt *allgemeines Stieltjessches Integral*.

Zur Unterscheidung vom allgemeinen Integral nennen wir das in [2] definierte Integral einfach Stieltjessches Integral oder ursprüngliches Stieltjessches Integral. Im weiteren geben wir Existenzbedingungen und die Eigenschaften des allgemeinen Integrals an.

Satz 1. Für die Existenz des Integrals (101) ist notwendig und hinreichend, daß eine Folge δ_n ($n=1, 2, 3, \dots$) von Zerlegungen existiert, für die die Differenz $S_{\delta_n} - s_{\delta_n}$ gegen 0 strebt oder, was gleichbedeutend damit ist, daß σ_{δ_n} bei beliebiger Wahl der Punkte $\xi_k^{(n)}$ einen endlichen Grenzwert A hat. Dieser Grenzwert ist der Wert des Integrals. Es gilt dann $s_{\delta_n} \rightarrow A$ und $S_{\delta_n} \rightarrow A$.

Dieser Satz folgt unmittelbar aus Betrachtungen, die ganz analog zu [8] verlaufen. Dabei sind die Teilintervalle Δ_k jetzt punktfremd. Wir bemerken noch, daß die in Satz 1 genannte Folge δ_n von Zerlegungen nicht unbedingt eine Folge mit beliebiger Verkleinerung der Intervalle zu sein braucht. Existiert z. B. eine Zerlegung δ für die σ_δ nicht von der Wahl der Punkte ξ_k abhängt, so können wir alle δ_n mit δ zusammenfallen lassen.

Wir führen nun einen neuen Begriff ein.

Definition 2. Eine Folge von Zerlegungen δ_n heißt für die Funktion $g(x)$ (oder $G(\Delta)$) *regulär*, wenn die beiden folgenden Bedingungen erfüllt sind: 1. Jeder Unstetigkeitspunkt von $g(x)$ stellt ein selbständiges Element dar für alle Zerlegungen von einem gewissen Wert n an. 2. Die Teilintervalle der Zerlegungen δ_n werden bei wachsendem n beliebig klein. Im Fall eines unendlichen Intervalls Δ ist das so zu verstehen, wie es in [4] dargestellt wurde.

Satz 2. Wenn die δ_n eine reguläre Folge von Zerlegungen bilden, dann gilt $s_{\delta_n} \rightarrow i$ und $S_{\delta_n} \rightarrow I$ bei beliebiger Wahl der beschränkten Funktion $f(x)$.

Beweis. Es sei δ eine beliebige vorgegebene Zerlegung von Δ , und es gelte $\delta'_n = \delta\delta_n$. Wir nehmen n so groß, daß erstens alle Unstetigkeitspunkte von $g(x)$, die selbständige Elemente der Zerlegung δ sind, bereits Teilpunkte der Zerlegung δ_n darstellen und daß sich zweitens in jedem Teilintervall von δ_n nicht mehr als ein Punkt der Zerlegung δ befindet. Im weiteren ist zu beachten, daß die Anzahl dieser Punkte fest ist. Wir betrachten die Differenz $s_{\delta'_n} - s_{\delta_n}$. Die den Punkten der Zerlegung δ entsprechenden Summanden, die Unstetigkeitsstellen von $g(x)$ in den Summen $s_{\delta'_n}$ und s_{δ_n} sind, sind identisch, weil sie nach dem oben Gesagten selbständige Elemente der Zerlegung in δ_n und in δ'_n sind. Wir werden im weiteren nur von solchen Punkten der Zerlegung δ sprechen, in denen $g(x)$ stetig ist. Ihre Anzahl sei gleich q . Wenn ein Teilintervall der Zerlegung δ_n

solche Punkte der Zerlegung nicht im Innern enthält, dann sind die ihnen entsprechenden Summanden in s_{δ_n} und $s_{\delta'_n}$ identisch. Es bleiben nicht mehr als q Teilintervalle der Zerlegung δ_n , welche die erwähnten Punkte der Zerlegung δ im Innern enthalten. Es sei $x = x'$ ein Punkt der Zerlegung δ , der im Innern des Teilintervalls $\Delta^{(n)}$ der Zerlegung δ_n liegt. Beim Übergang von δ_n zu δ'_n wird ein Summand der Form $\lambda^{(n)} G(\Delta^{(n)})$ durch eine Summe der Form $\mu^{(n)} G(\Delta_1^{(n)}) + \nu^{(n)} G(\Delta_2^{(n)})$ ersetzt, wobei $\Delta^{(n)}$ den Punkt $x = x'$ im Innern enthält und $\Delta_1^{(n)}$ oder $\Delta_2^{(n)}$ diesen Punkt als Randpunkt haben (wozu x' gehört, ist nicht wichtig, weil $g(x)$ im Punkt $x = x'$ stetig ist). Mit $\lambda^{(n)}$, $\mu^{(n)}$ und $\nu^{(n)}$ bezeichnen wir die Zahlen, die dem absoluten Betrag nach die obere Grenze von $|f(x)|$ in Δ nicht übersteigen.

Weil in einer regulären Folge die Teilintervalle für wachsendes n beliebig klein werden, gilt das auch für $\Delta^{(n)}$, $\Delta_1^{(n)}$ und $\Delta_2^{(n)}$, und weil diese Intervalle im Innern oder auf dem Rand einen fest vorgegebenen Stetigkeitspunkt $x = x'$ der Funktion $g(x)$ enthalten, kann man behaupten, daß

$$\lambda^{(n)} G(\Delta^{(n)}) \rightarrow 0 \quad \text{und} \quad \mu^{(n)} G(\Delta_1^{(n)}) + \nu^{(n)} G(\Delta_2^{(n)}) \rightarrow 0$$

für $n \rightarrow \infty$ gilt.

Wenn wir noch beachten, daß die Anzahl der Punkte x' nicht größer als q ist, folgt

$$s_{\delta'_n} - s_{\delta_n} \rightarrow 0 \quad \text{für} \quad n \rightarrow \infty.$$

Andererseits gelten die Ungleichungen $s_{\delta_n} \leq s_{\delta'_n} \leq i$ und $s_{\delta} \leq s_{\delta'_n} \leq i$. Man kann die Zerlegung δ so wählen, daß s_{δ} beliebig nahe bei i liegt, d. h., bei beliebig vorgegebenem positivem ε kann man δ so wählen, daß $i - s_{\delta} < \varepsilon$ gilt. Aus der Ungleichung $s_{\delta} \leq s_{\delta'_n} \leq i$ folgt dann unmittelbar $i - s_{\delta'_n} < \varepsilon$, und das oben erhaltene Ergebnis $s_{\delta'_n} - s_{\delta_n} \rightarrow 0$ liefert schließlich für alle genügend großen n die Ungleichung $i - s_{\delta_n} < 2\varepsilon$; da ε willkürlich war, gilt also $s_{\delta_n} \rightarrow i$, was zu beweisen war. Ganz analog läßt sich zeigen, daß $S_{\delta_n} \rightarrow I$ gilt.

Ist $g(x)$ stetig, so ist für eine reguläre Folge schon die Tatsache charakteristisch, daß ihre Teilintervalle beliebig klein werden. Das gilt z. B. im Fall Riemannscher Integrale, d. h. für $g(x) = x$. Da x für unendliche Intervalle nicht beschränkt ist, konnten wir bei der Definition des eigentlichen Riemannschen Integrals nur beschränkte Intervalle heranziehen.

Satz 3. Wenn das Integral (101) existiert und gleich A ist, dann gilt für eine beliebige reguläre Folge von Zerlegungen die Beziehung $\sigma_{\delta_n} \rightarrow A$.

Aus der Voraussetzung des Satzes folgt sofort $i = I = A$. Nach Satz 2 gilt dann $s_{\delta_n} \rightarrow A$ und $S_{\delta_n} \rightarrow A$, und wegen der Ungleichung $s_{\delta_n} \leq \sigma_{\delta_n} \leq S_{\delta_n}$ konvergiert folglich auch σ_{δ_n} bei beliebiger Wahl der Punkte $\xi_k^{(n)}$ gegen A .

Folgerung. Aus den bewiesenen Sätzen ergibt sich unmittelbar: Wenn für irgend-eine reguläre Folge δ_n der Grenzwert von σ_{δ_n} existiert und gleich A ist, existiert er auch für jede andere reguläre Folge und ist gleich derselben Zahl A . Dann existiert auch das Integral und ist gleich A . Wenn σ_{δ_n} für eine gewisse reguläre Folge δ_n keinen endlichen Grenzwert hat, existiert das Integral (101) nicht. Somit wird bei der Verwendung der Summen σ_{δ} mit regulären Zerlegungsfolgen zugleich die Frage nach der Existenz des Stieltjesschen Grundintegrals gelöst.

Wir bemerken noch: Wenn δ_n eine reguläre Folge für $g(x)$ und δ'_n eine Fortsetzung von δ_n ist, so ist auch δ'_n eine reguläre Folge. Dies folgt unmittelbar aus der Definition der regulären Folgen.

19. Eigenschaften des (allgemeinen) Stieltjesschen Integrals. Wie wir gesehen haben, kann man das allgemeine Stieltjessche Integral als Grenzwert der Summen σ_{δ} für eine bestimmte Folge von Zerlegungen erhalten. Dabei gilt: Wenn die σ_{δ_n} einen endlichen

Grenzwert haben und δ'_n eine Fortsetzung von δ_n ist, dann haben die $\sigma_{\delta'_n}$ denselben Grenzwert. Wenn wir die Summen σ_{δ} benutzen, können wir also in derselben Weise die Eigenschaften des allgemeinen Stieltjesschen Integrals beweisen, wie das im Fall des Riemannschen und des ursprünglichen Stieltjesschen Integrals möglich war. Wir bringen diese Eigenschaften mit gewissen Ergänzungen. Im folgenden werden die Funktionen $f(x)$ und $g(x)$ im Integrationsintervall als beschränkt und die Funktion $g(x)$ außerdem als nicht abnehmend vorausgesetzt.

I. Wenn die c_k Konstanten sind, gilt

$$\int_{\Delta} \sum_{k=1}^p c_k f_k(x) dg(x) = \sum_{k=1}^p c_k \int_{\Delta} f_k(x) dg(x), \quad (102)$$

wobei aus der Existenz der rechts stehenden Integrale die Existenz des links stehenden Integrals folgt.

Zum Beweis braucht man nur irgendeine reguläre Folge δ_n für die Funktion $g(x)$ zu wählen.

II. Sind die $g_k(x)$ ($k=1, 2, \dots, p$) nicht abnehmende beschränkte Funktionen und die c_k positive Konstanten, so gilt

$$\int_{\Delta_0} f(x) d\left(\sum_{k=1}^p c_k g_k(x)\right) = \sum_{k=1}^p c_k \int_{\Delta_0} f(x) dg_k(x), \quad (103)$$

wobei aus der Existenz der rechts stehenden Integrale die Existenz des links stehenden Integrals folgt und umgekehrt.

Es sei $\delta_n^{(k)}$ eine reguläre Folge von Zerlegungen für $g_k(x)$. Die Folge $\delta_n = \delta_n^{(1)}, \delta_n^{(2)}, \dots, \delta_n^{(p)}$ ist eine reguläre Folge für alle Funktionen $g_k(x)$ ($k=1, 2, \dots, p$). Offensichtlich gilt die Beziehung

$$S_{\delta_n} - s_{\delta_n} = \sum_{k=1}^p c_k (S_{\delta_n^{(k)}} - s_{\delta_n^{(k)}}), \quad (104)$$

in der s_{δ_n} und S_{δ_n} zu dem Integral auf der linken Seite von (103), $s_{\delta_n^{(k)}}$ und $S_{\delta_n^{(k)}}$ zu den Integralen auf der rechten Seite von (103) gehören. Wenn die Integrale auf der rechten Seite von (103) existieren, dann gilt $S_{\delta_n^{(k)}} - s_{\delta_n^{(k)}} \rightarrow 0$ ($k=1, 2, \dots, p$) und folglich $S_{\delta_n} - s_{\delta_n} \rightarrow 0$, d. h., es existiert das Integral auf der linken Seite. Wir nehmen nun an, das letzte Integral existiere. Dann muß eine Folge δ_n von Zerlegungen existieren, für die $S_{\delta_n} - s_{\delta_n} \rightarrow 0$ gilt. Die Summanden auf der rechten Seite von (104) sind nicht negativ, folglich gilt auch $S_{\delta_n^{(k)}} - s_{\delta_n^{(k)}} \rightarrow 0$ für $k=1, 2, \dots, p$, d. h., die Integrale auf der rechten Seite von (103) existieren. Die Formel (103) selbst ergibt sich unmittelbar aus der analogen Eigenschaft für endliche Summen σ_{δ_n} durch Grenzübergang.

III. Wenn das Intervall Δ_0 in eine endliche Anzahl von punktfremden Intervallen $\Delta_1, \Delta_2, \dots, \Delta_m$ zerlegt wird, gilt

$$\int_{\Delta} f(x) dg(x) = \sum_{k=1}^m \int_{\Delta_k} f(x) dg(x), \quad (105)$$

wobei aus der Existenz der Integrale auf der rechten Seite die Existenz des Integrals auf der linken Seite folgt und umgekehrt.

Wir nehmen an, daß das Integral auf der linken Seite existiert. Nach Satz 1 gibt es eine Folge δ_n von Zerlegungen des Intervalls Δ , für die $S_{\delta_n} - s_{\delta_n} \rightarrow 0$ gilt. Wir bezeichnen mit δ die Zerlegungen von Δ in Teilintervalle Δ_k ($k=1, 2, \dots, m$) und setzen $\delta'_n = \delta_n \delta$. Offensichtlich gilt $S_{\delta'_n} - s_{\delta'_n} \rightarrow 0$ wegen $s_{\delta'_n} \geq s_{\delta_n}$ und $S_{\delta'_n} \leq S_{\delta_n}$. Die dem Ausdruck (16) entsprechende Summe, durch die $S_{\delta'_n} - s_{\delta'_n}$ dargestellt wird, kann in m nichtnegative Sum-

manden zerlegt werden, von denen jeder wiederum eine analoge Summe für ein gewisses Δ_k darstellt, und da die ganze Summe gegen 0 strebt, gilt dasselbe auch für die einzelnen Summanden, d. h., die Integrale auf der rechten Seite der Formel (105) existieren. Umgekehrt nehmen wir an, daß die rechts stehenden Integrale existieren. Für jedes von ihnen existiert eine Folge $\delta_n^{(k)}$ von Zerlegungen, für die die Differenz $S_{\delta_n^{(k)}} - s_{\delta_n^{(k)}}(k)$ gegen 0 strebt.

Das Produkt dieser Zerlegungen führt unmittelbar zu einer Folge von Zerlegungen des ganzen Intervalls Δ , bei der die analoge Differenz für das Integral auf der linken Seite ebenfalls gegen 0 strebt. Wir bemerken, daß für das ursprüngliche Stieltjessche Integral in der dritten Formel von (3) die Existenz der Integrale auf der rechten Seite nicht die Existenz des Integrals auf der linken Seite nach sich zieht.

IV. Gilt im Intervall Δ die Abschätzung $|f(x)| \leq L$, so gilt auch

$$\left| \int_{\Delta} f(x) dg(x) \right| \leq LG(\Delta), \quad (106)$$

wobei $G(\Delta)$ der volle Zuwachs der Funktion $g(x)$ im Intervall ist und wobei vorausgesetzt wird, daß das Integral existiert.

V. Wenn im Intervall Δ die Funktionen $f_p(x)$ für $p \rightarrow \infty$ gleichmäßig gegen $f(x)$ konvergieren und die Integrale von $f_p(x)$ bezüglich $g(x)$ existieren, dann existiert auch das Integral von $f(x)$ bezüglich $g(x)$, und es gilt

$$\lim_{p \rightarrow \infty} \int_{\Delta} f_p(x) dg(x) = \int_{\Delta} f(x) dg(x). \quad (107)$$

Es seien $\Delta_k^{(n)}$ ($k = 1, 2, \dots, t_n$) die Teilintervalle einer Zerlegung, die zu einer gewissen regulären Folge δ_n von Zerlegungen für die Funktion $g(x)$ gehört. Wir betrachten die Summen $\sigma_{\delta_n^{(p)}}$ und σ_{δ_n} für die Funktionen $f_p(x)$ und für die Funktion $f(x)$, die offenbar wegen der gleichmäßigen Konvergenz der $f_p(x)$ beschränkt ist:

$$\sigma_{\delta_n^{(p)}} = \sum_{k=1}^{t_n} f_p(\xi_k^{(n)}) G(\Delta_k^{(n)}), \quad \sigma_{\delta_n} = \sum_{k=1}^{t_n} f(\xi_k^{(n)}) G(\Delta_k^{(n)}). \quad (108)$$

Wir nehmen dabei in allen Summen dieselben Punkte $\xi_k^{(n)}$. Nun bilden wir die Differenz

$$\sigma_{\delta_n} - \sigma_{\delta_n^{(p)}} = \sum_{k=1}^{t_n} [f(\xi_k^{(n)}) - f_p(\xi_k^{(n)})] G(\Delta_k^{(n)}).$$

Bei beliebig vorgegebenem positivem ε existiert wegen der gleichmäßigen Konvergenz der $f_p(x)$ ein solches N , daß für $p > N$ und beliebiges x aus Δ die Ungleichung $|f(x) - f_p(x)| < \varepsilon$ gilt. Für die Differenz erhalten wir $\sigma_{\delta_n} - \sigma_{\delta_n^{(p)}}$ bei $p > N$ und bei beliebiger

Wahl der $\xi_k^{(n)}$ die Abschätzung $|\sigma_{\delta_n} - \sigma_{\delta_n^{(p)}}| < \varepsilon G(\Delta)$. Hieraus erkennt man, daß $\sigma_{\delta_n^{(p)}}$ für $p \rightarrow \infty$ gegen σ_{δ_n} strebt, und zwar gleichmäßig bezüglich n und bezüglich der Wahl der Punkte $\xi_k^{(n)}$. Die Funktionen $f_p(x)$ sind nach Voraussetzung bezüglich $g(x)$ integrierbar; folglich hat jede der Summen $\sigma_{\delta_n^{(p)}}$ bei unbegrenzt wachsendem n einen Grenzwert, den wir mit A_p bezeichnen. Dieser Grenzwert ist das Integral von $f_p(x)$ bezüglich $g(x)$. Wir zeigen nun, daß die Folge der Zahlen A_p einen Grenzwert hat. Nach (106) gilt

$$|A_p - A_q| \leq G(\Delta) \max |f_p(x) - f_q(x)|.$$

Nach Voraussetzung strebt die rechte Seite dieser Ungleichung für $p, q \rightarrow \infty$ gegen 0, und somit hat A_p für $p \rightarrow \infty$ einen Grenzwert, den wir mit A bezeichnen. Wir müssen nur noch beweisen, daß $\sigma_{\delta_n} \rightarrow A$ für $n \rightarrow \infty$ gilt. Wir schätzen die Differenz $A - \sigma_{\delta_n}$ ab, die wir dazu wie folgt schreiben:

$$A - \sigma_{\delta_n} = (A - A_p) + (A_p - \sigma_{\delta_n^{(p)}}) + (\sigma_{\delta_n^{(p)}} - \sigma_{\delta_n}). \quad (109)$$

Es sei $\varepsilon > 0$ gegeben. Zuerst wählen wir p so groß, daß für beliebige n und ξ_k

$$|A - A_p| \leq \varepsilon \quad \text{und} \quad |\sigma_{\delta_n}(p) - \sigma_{\delta_n}| \leq \varepsilon$$

ist. Weiter gilt für alle genügend großen n und das oben festgelegte p , daß $|A_p - \sigma_{\delta_n}(p)| < \varepsilon$ ist. Somit ist $|A - \sigma_{\delta_n}| < 3\varepsilon$; und hieraus folgt, da ε beliebig war, $\sigma_{\delta_n} \rightarrow A$.

VI. Wenn das Integral von $f(x)$ bezüglich $g(x)$ existiert, dann existiert auch das Integral von $|f(x)|$ bezüglich $g(x)$, und es gilt die Ungleichung

$$\left| \int_{\Delta_0} f(x) dg(x) \right| \leq \int_{\Delta_0} |f(x)| dg(x). \quad (110)$$

Die Größen m_k und M_k mögen die übliche Bedeutung für die Funktion $f(x)$ haben. Wenn beide Zahlen positiv sind, haben wir für $|f(x)|$ dieselbe untere bzw. obere Grenze wie für $f(x)$. Wenn beide Zahlen negativ sind, stellen die Zahlen $|M_k|$ und $|m_k|$ die untere bzw. obere Grenze für $|f(x)|$ dar, und die Differenz zwischen der oberen und der unteren Grenze bleibt dieselbe. Ist schließlich m_k negativ und M_k positiv, so stellt die größte der Zahlen $|m_k|$ und M_k die obere Grenze von $|f(x)|$ dar, und die untere Grenze ist eine nicht-negative Zahl. Folglich ist in allen Fällen die Differenz zwischen der oberen und der unteren Grenze für $|f(x)|$ nicht größer als für $f(x)$. Wenn daher bei einer gewissen Folge von Zerlegungen die Differenz $S_{\delta_n} - s_{\delta_n}$ für $f(x)$ gegen 0 strebt, dann strebt sie erst recht bei derselben Folge von Zerlegungen für $|f(x)|$ gegen 0, d. h., aus der Existenz des Integrals für $f(x)$ folgt die Existenz des Integrals für $|f(x)|$. Die Ungleichung (110) erhält man aus der analogen Ungleichung für die Summen durch Grenzübergang.

VII. Wenn die Funktionen $f_1(x)$ und $f_2(x)$ bezüglich $g(x)$ integrierbar sind, dann ist auch ihr Produkt $f_1(x) f_2(x)$ bezüglich $g(x)$ integrierbar.

Wir beweisen zuerst: Ist $f(x)$ bezüglich $g(x)$ integrierbar, so ist auch $f^2(x)$ bezüglich $g(x)$ integrierbar. Wir nehmen vorübergehend an, $f(x)$ sei positiv, und bilden die Summen σ_δ für $f(x)$ und $f^2(x)$:

$$\sigma_\delta = \sum_{k=1}^n (M_k - m_k) G(\Delta_k), \quad \sigma_{\delta'} = \sum_{k=1}^n (M_k^2 - m_k^2) G(\Delta_k) = \sum_{k=1}^n (M_k + m_k) (M_k - m_k) G(\Delta_k).$$

Wenn die erste Summe für eine gewisse Folge von Zerlegungen gegen 0 strebt, dann strebt wegen der Beschränktheit der Faktoren $M_k + m_k$ auch die zweite Summe für diese Folge von Zerlegungen gegen 0. Also folgt für eine positive Funktion $f(x)$ aus ihrer Integrierbarkeit die Integrierbarkeit ihres Quadrates $f^2(x)$. Ist die Funktion $f(x)$ nicht positiv, so existiert wegen ihrer Beschränktheit eine positive Konstante a derart, daß $f(x) + a$ positiv ist. Diese letzte Funktion ist nach Eigenschaft I integrierbar. Also ist nach dem soeben Bewiesenen auch die Funktion $[f(x) + a]^2 = f^2(x) + 2af(x) + a^2$ integrierbar. Hieraus folgt unmittelbar auch die Integrierbarkeit der Funktion $f^2(x)$, da man sie als Summe integrierbarer Funktionen darstellen kann: $f^2(x) = [f(x) + a]^2 - 2af(x) - a^2$. Um schließlich die Integrierbarkeit von $f_1(x) f_2(x)$ zu beweisen, genügt es, dieses Produkt wie folgt darzustellen:

$$f_1(x) f_2(x) = \frac{1}{2} [f_1(x) + f_2(x)]^2 - \frac{1}{2} f_1^2(x) - \frac{1}{2} f_2^2(x).$$

Die rechte Seite dieser Formel ist als Summe integrierbarer Funktionen integrierbar, also ist auch die linke Seite integrierbar.

20. Existenz des allgemeinen Stieltjesschen Integrals. Wir beweisen jetzt einige hinreichende Bedingungen für die Existenz des allgemeinen Stieltjesschen Integrals.

Satz. Jede beschränkte Funktion $f(x)$ ist bezüglich einer Sprungfunktion $g_a(x)$ im Sinne des allgemeinen Stieltjesschen Integrals integrierbar.

Wir nehmen zuerst an, die Menge $c_1, c_2, \dots, c_p$ der Unstetigkeitsstellen von $g(x)$ sei endlich. Ferner sei δ eine auf folgende Weise definierte Zerlegung des Grundintervalls:

Die Randpunkte des Intervalls, falls sie zum Integrationsgebiet gehören, und die Punkte $c_1, c_2, \dots, c_p$ bilden die selbständigen Elemente von δ . Die restlichen Elemente von δ sind die offenen Intervalle, die nach der Aussonderung der angegebenen Punkte übrigbleiben. In jedem dieser Intervalle nimmt $g_d(x)$ einen konstanten Wert an, und die dem Integral von $f(x)$ bezüglich $g_d(x)$ entsprechenden Summen sind offensichtlich identisch und gleich

$$s_\delta = S_\delta = \sum_{k=1}^p f(c_k) \gamma_k. \quad (111)$$

Also existiert das Integral von $f(x)$ bezüglich $g_d(x)$ und ist gleich der Summe (111). Wir nehmen nun an, daß die Menge der Unstetigkeitsstellen c_k der Funktion $g_d(x)$ unendlich ist; δ_n seien in derselben Weise wie soeben gebildete Zerlegungen des Grundintervalls, wobei als selbständige Elemente der Zerlegung die Randpunkte des Intervalls und die ersten n Unstetigkeitspunkte $c_1, c_2, \dots, c_n$ ausgewählt werden. Für die so konstruierte Folge von Zerlegungen δ_n bilden wir die Summen σ_{δ_n} . Die Summe derjenigen in σ_{δ_n} auftretenden Summanden, die von den als selbständigen Elementen der Zerlegung ausgesonderten Punkten herrühren, ist gleich $\sum_{k=1}^n f(c_k) \gamma_k$ und hängt nicht von den veränderlichen Punkten ξ_k ab. Wir betrachten eines der offenen Intervalle (α, β) , das zur Zerlegung δ_n gehört. Der entsprechende Summand der Summe σ_{δ_n} hat die Gestalt

$$f(\xi) [g(\beta - 0) - g(\alpha + 0)] \quad (\alpha < \xi < \beta).$$

Ist $|f(x)| \leq L$, so gilt

$$|f(\xi)| [g(\beta - 0) - g(\alpha + 0)] \leq L [g(\beta - 0) - g(\alpha + 0)].$$

Dabei stellt die Differenz $g(\beta - 0) - g(\alpha + 0)$ die Summe der Sprünge von $g(x)$ im Innern des Intervalls (α, β) dar. Folglich ist die Summe der zu σ_{δ} gehörenden und den offenen Intervallen der Zerlegung δ_n entsprechenden Summanden dem absoluten Betrag nach nicht größer als das Produkt von L und der Summe aller Sprünge der Funktion $g(x)$ mit Ausnahme der Sprünge in den ausgesonderten Punkten $c_1, c_2, \dots, c_n$. Diese Summe strebt bei unbegrenzt wachsendem n gegen 0, und die den ausgesonderten Punkten der Zerlegung entsprechende Summe ergibt für wachsendes n als Grenzwert die Summe der konvergenten Reihe

$$\sum_{k=1}^{\infty} f(c_k) \gamma_k. \quad (112)$$

Daraus folgt, daß das Integral von $f(x)$ bezüglich $g_d(x)$ existiert und durch die Reihe (112) ausgedrückt wird, was zu beweisen war.

Somit reduziert sich die Frage nach der Existenz des allgemeinen Stieltjesschen Integrals auf die Frage nach der Existenz des Integrals von $f(x)$ bezüglich der stetigen nicht abnehmenden Funktion $g_c(x)$. Für diese Funktion stellt jede Folge von Zerlegungen mit beliebig klein werdenden Teilintervallen eine reguläre Folge dar. Ist $f(x)$ in dem Intervall Δ eine monotone Funktion, so ist sie nach der Formel der partiellen Integration [2] in Δ bezüglich $g_c(x)$ integrierbar. Somit ist jede monotone Funktion bezüglich einer wachsenden beschränkten Funktion integrierbar.

Wir weisen noch auf die Klasse der Funktionen hin, die bezüglich jeder wachsenden beschränkten Funktion $g(x)$ integrierbar sind. Wir nennen eine Funktion $f(x)$ *stückweise konstant* im Intervall Δ , wenn dieses Intervall so in endlich viele paarweise punktfremde Intervalle $\Delta^{(1)}, \Delta^{(2)}, \dots, \Delta^{(m)}$ zerlegt werden kann, daß $f(x)$ in jedem Intervall $\Delta^{(k)}$ ($k = 1, 2, \dots, m$) einen konstanten Wert, etwa b_k , annimmt. Man erkennt leicht, daß eine *stückweise konstante Funktion bezüglich jeder wachsenden Funktion $g(x)$ integrierbar ist*. Denn ist δ eine Zerlegung von Δ in Intervalle $\Delta^{(k)}$ ($k = 1, 2, \dots, m$), so stimmen die Summen s_δ und S_δ überein:

$$s_\delta = S_\delta = \sum_{k=1}^m b_k G(\Delta^{(k)}). \quad (113)$$

Wenn wir diese Zerlegung δ wiederholen, dann können wir uns davon überzeugen, daß das Integral einer stückweise konstanten Funktion $f(x)$ bezüglich $g(x)$ existiert und gleich der Summe (113) ist. Durch Benutzung der Eigenschaft V aus [19] folgt: Wenn die Funktion $f(x)$ im Intervall Δ Grenzwert einer gleichmäßig konvergenten Folge stückweise konstanter Funktionen ist, dann ist $f(x)$ bezüglich jeder zunehmenden beschränkten Funktion $g(x)$ integrierbar. Wir bemerken noch, daß eine notwendige und hinreichende Bedingung für die Integrierbarkeit von $f(x)$ bezüglich $g_c(x)$ mit der in [10] gegebenen identisch ist. Auf den Beweis werden wir verzichten.

Bemerkung 1. Wir definierten in [8] eine Funktion $g(x)$ von beschränkter Variation in dem abgeschlossenen Intervall $[a, b]$. Ganz analog kann man diesen Begriff auch für Intervalle anderen Typs einführen. Wir nehmen z. B. das offene Intervall (a, b) . Eine Funktion $g(x)$ ist in (a, b) eine Funktion von beschränkter Variation, wenn sie in jedem abgeschlossenen Intervall $[c, x]$, das im Innern von (a, b) liegt, von beschränkter Variation ist und $V_c^x(g)$ für diese beliebigen Intervalle nicht größer als eine gewisse Zahl ist. Bei kleiner werdendem c nimmt der Wert von $V_c^x(g)$ nicht ab, und folglich hat $V_c^x(x)$ für $c \rightarrow a$ einen endlichen Grenzwert, den wir mit $V_{a+0}^x(g)$ bezeichnen. Wenn x gegen b strebt, hat dieser Wert auch einen endlichen Grenzwert, den wir die Totalvariation von $g(x)$ in (a, b) nennen. Die Formeln (44) definieren die Funktionen $g_1(x)$ und $g_2(x)$, die in der kanonischen Darstellung von $g(x)$ als Differenz nichtabnehmender beschränkter Funktionen, $g(x) = g_1(x) - g_2(x)$, auftreten. Das allgemeine Stieltjessche Integral läßt sich dann als Differenz der Integrale bezüglich dieser nichtabnehmenden Funktionen $g_1(x)$ und $g_2(x)$ definieren:

$$\int_{\Delta} f(x) dg(x) = \int_{\Delta} f(x) dg_1(x) - \int_{\Delta} f(x) dg_2(x) .$$

Die Existenz des Integrals von $f(x)$ bezüglich $g(x)$ ist gleichbedeutend mit der Existenz der Integrale von $f(x)$ bezüglich $g_1(x)$ und $g_2(x)$.

Auf Grund unserer bisherigen Überlegungen können wir behaupten, daß *im Sinne des allgemeinen Stieltjesschen Integrals jede Funktion von beschränkter Variation und jede Funktion, die Grenzwert einer gleichmäßig konvergenten Folge stückweise konstanter Funktionen ist, bezüglich jeder Funktion von beschränkter Variation integrierbar ist.*

Bemerkung 2. Bei der Integration über halboffene Intervalle $(a, b]$ wird manchmal noch ein weiterer Integralbegriff eingeführt, der sich von dem allgemeinen Stieltjesschen Integral nur dadurch unterscheidet, daß das Ausgangsintervall $(a, b]$ nicht in Intervalle beliebigen Typs zerlegt werden kann, sondern nur in punktfremde halboffene Intervalle vom Typ $(c, d]$. Weil dabei die Mengen der Zahlen s_i und S_i echte Teilmengen der Mengen dieser Zahlen für das allgemeine Integral sein können, kann die Zahl i kleiner, aber die Zahl I größer werden. Es kann also vorkommen, daß für das allgemeine Integral $i = I$, aber für das neu konstruierte Integral $i < I$ gilt, d. h., eine Funktion, die im Sinne des allgemeinen Integrals integrierbar ist, braucht bei dem neu definierten Integralbegriff nicht integrierbar zu sein.

21. Intervallfunktionen in der Ebene. Der Begriff der additiven Intervallfunktion und die Bildung des Stieltjesschen Integrals werden in der Ebene, im dreidimensionalen und überhaupt im mehrdimensionalen Raum komplizierter. Wir betrachten nun den ebenen Fall. Aus den angeführten Überlegungen erkennt man leicht, wie man sie für Räume mit einer größeren Zahl von Dimensionen ändern muß. Es sei eine Ebene mit den Achsen x und y gegeben. Wir nehmen an, daß auf der x -Achse ein gewisses Intervall Δ_x und auf der y -Achse ein gewisses Intervall Δ_y vorgegeben sind. Dabei fassen wir den Begriff des Intervalls jetzt in dem in [20] festgelegten allgemeinen Sinne auf. Diese Intervalle Δ_x und Δ_y definieren ein gewisses Intervall Δ in der Ebene, und zwar sagen wir, der Punkt (x, y) liege in Δ , wenn x in Δ_x und y in Δ_y liegt. Diese Intervalle in der Ebene können von verschiedenem Typ

sein. Es kann sich z. B. um ein durch die Ungleichungen $a \leq x \leq b$, $c \leq y \leq d$ definiertes abgeschlossenes Intervall, um ein durch $a \leq x < b$, $c \leq y < d$ definiertes halboffenes Intervall, um ein durch $a < x \leq b$, $c < y \leq d$ definiertes halboffenes Intervall, um ein durch $a \leq x < b$, $c < y \leq d$ definiertes halboffenes Intervall, um einen durch $a < x \leq b$, $y = c$ definierten Abschnitt auf einer Parallelen zur x -Achse oder um einen Punkt $x = a$, $y = c$ handeln usw. Die in diesen Ungleichungen auftretenden Zahlen a , b , c und d können endlich oder unendlich sein. Im folgenden werden für uns durch Ungleichungen der Gestalt $a < x \leq b$, $c < a \leq d$ definierte halboffene Intervalle eine große Bedeutung haben. Der Einfachheit halber werden wir im folgenden nur solche Intervalle *halboffen* nennen.

Wir nehmen an, daß für ein beliebiges, in einem bestimmten Grundintervall Δ_0 der Ebene liegendes Intervall Δ eine nichtnegative Funktion $G(\Delta)$ definiert ist, die additiv und normal ist. Anders ausgedrückt, wenn Δ die Summe $\Delta^{(1)} + \Delta^{(2)} + \dots + \Delta^{(m)}$ von punktfremden Intervallen ist, dann gilt

$$G(\Delta) = \sum_{k=1}^m G(\Delta^{(k)}),$$

und wenn $\Delta_1, \Delta_2, \dots$ eine verschwindende Intervallfolge ist, dann gilt $G(\Delta_n) \rightarrow 0$.

Wir leiten nun einige Eigenschaften dieser Funktionen her. Da die Funktion $G(\Delta)$ nichtnegativ und additiv ist, stellt $G(\Delta_0)$ den größten Wert von $G(\Delta)$ für alle in Δ_0 liegenden Δ dar. Den Wert $G(\Delta)$ bezeichnen wir in dem Fall, daß Δ ein einzelner Punkt P ist, einfach mit $G(P)$. Da $G(\Delta)$ nicht negativ ist, gilt dabei $G(P) \geq 0$. Es kann vorkommen, daß in allen Punkten P aus Δ_0 die Gleichung $G(P) = 0$ gilt. In diesem Fall heißt die Funktion $G(\Delta)$ in Δ_0 *stetig*. Wenn dagegen $G(P) > 0$ gilt, dann heißt der Punkt P *Unstetigkeitspunkt von $G(\Delta)$* . Man beweist leicht, daß *die Menge der Unstetigkeitspunkte endlich oder abzählbar ist* (vgl. [6]). Wir nehmen diejenigen Unstetigkeitspunkte, in denen $G(P) > 1$ gilt. Da $G(\Delta)$ additiv und nichtnegativ ist, ist die Anzahl dieser Punkte nicht größer als der ganzzahlige Anteil der Zahl $G(\Delta_0)$. Ähnlich ergibt sich, daß die Anzahl der Punkte P , in denen $G(P) > \frac{1}{2}$, aber ≤ 1 ist, nicht größer als der ganzzahlige Anteil von $2G(\Delta_0)$ ist, usw. Daraus folgt, genau wie in [6], daß die Menge der Unstetigkeitspunkte endlich oder abzählbar ist. Wenn sie abzählbar ist und $P_1, P_2, \dots$ die Folge von Unstetigkeitspunkten darstellt, dann konvergiert die Reihe, deren Glieder die positiven Zahlen $G(P_k)$ sind, und es gilt

$$\sum_{k=1}^{\infty} G(P_k) \leq G(\Delta_0). \quad (114)$$

Ist analog dazu l der zu Δ_0 gehörende Teil einer Geraden, die zu einer der Koordinatenachsen parallel verläuft, so heißt er eine *Unstetigkeitslinie*, wenn $G(l) > 0$ ist. Jede achsenparallele und durch einen Unstetigkeitspunkt gehende Gerade ergibt eine Unstetigkeitslinie. Es kann aber auch Unstetigkeitslinien geben, die keinen einzigen Unstetigkeitspunkt enthalten. In derselben Weise wie oben läßt sich zeigen: *Wenn Unstetigkeitslinien auftreten, dann ist ihre Anzahl endlich oder abzählbar*. Dabei nehmen wir jeweils den gesamten unzerlegten, in Δ_0 liegenden Abschnitt der achsenparallelen Gerade. Es möge nun eine Folge von Intervallen Δ_n ($n = 1, 2, \dots$) geben, für die Δ_{n+1} in Δ_n enthalten ist, und der Punkt P oder alle Punkte der Linie l mögen die einzigen Punkte darstellen, die allen Intervallen Δ_n gemeinsam sind. In diesem Fall heißt Δ_n ein *System gegen P oder l strebender ineinandergeschachtelter Intervalle* (l ist wie oben ein Abschnitt auf einer achsenparallelen Geraden). Wenn

man die Normalität der Funktion $G(\Delta)$ benutzt, läßt sich zeigen: *Ist Δ_n ein System ineinandergeschachtelter, gegen P oder l strebender Intervalle, so gilt $G(\Delta_n) \rightarrow G(P)$ bzw. $G(\Delta_n) \rightarrow G(l)$.* Wir beweisen das für den Fall des Punktes P und nehmen an, daß sich dieser Punkt im Innern aller Intervalle Δ_n befindet, die als offen vorausgesetzt werden. Durch P ziehen wir achsenparallele Geraden und zerlegen jedes Δ_n in die folgenden Teile: in den Punkt P , die vier in Δ_n liegenden Abschnitte auf diesen Geraden und die restlichen vier Intervalle. Falls sich die Δ_n auf den Punkt P zusammenziehen, stellen alle diese Elemente der Zerlegung mit Ausnahme des Punktes P ihrerseits verschwindende Intervallfolgen dar, und auf Grund seiner Normalität konvergiert $G(\Delta)$ für jede dieser Intervallfolgen gegen 0. Berücksichtigt man, daß die Funktion $G(\Delta)$ additiv ist, so sieht man leicht, daß $G(\Delta_n) \rightarrow G(P)$ gilt. Ganz analog lassen sich auch die anderen Fälle für den Punkt P und der Fall der Linie l untersuchen.¹⁾ Diese Definitionen erweisen sich als ganz naturgemäß, wenn $G(\Delta)$ als die Masse gedeutet wird, die sich im Intervall Δ befindet, wenn im Grundintervall Δ_0 Materie verteilt ist. Ist beispielsweise $G(P) > 0$, so befindet sich im Punkt P eine konzentrierte Masse $G(P)$. Analog dazu ist, wenn $G(l) > 0$ ist, die Masse $G(l)$ irgendwie längs der Linie l verteilt. Wir nehmen z. B. an, daß das halboffene Intervall Δ_0 ($0 < x \leq 2$, $0 < y \leq 2$) gegeben sei und in ihm auf dem Geradenstück l_0 ($x = 1$, $0 < y \leq 1$) Masse mit der linearen Dichte 1 verteilt sei. In diesem Fall ist $G(\Delta)$ gleich der Länge derjenigen Teilstrecke von l_0 , die in Δ liegt. In jedem Punkt P aus Δ_0 ist $G(P) = 0$. Ist Δ_n ein System gegen l_0 strebender ineinandergeschachtelter Intervalle, so streben ihre Flächeninhalte gegen 0, es ist aber $G(\Delta_n) = 1$ für alle n .

Wir erwähnen, daß die Funktion $G(\Delta)$, die für alle in Δ_0 liegenden Intervalle definiert wurde, leicht auf alle ebenen Intervalle erweitert werden kann, und zwar so, daß sie nichtnegativ, additiv und normal bleibt. Ist nämlich Δ ein beliebiges Intervall und $\Delta\Delta_0$ das Produkt der Intervalle Δ und Δ_0 , d. h. die Menge der Punkte, die gleichzeitig in Δ und Δ_0 liegen, dann ist dieses Produkt ein Intervall aus Δ_0 , und wir erhalten die gewünschte Erweiterung der Funktion $G(\Delta)$, wenn wir $G(\Delta) = G(\Delta\Delta_0)$ setzen. Diese Erweiterung der Funktion $G(\Delta)$ läßt sich leicht physikalisch deuten, wenn man $G(\Delta)$ als die Masse auffaßt, die im Intervall Δ enthalten ist, wenn die gesamte Masse im Intervall Δ_0 verteilt ist.

Im folgenden werden wir sehen: Ist eine nichtnegative additive normale Funktion nur in einem halboffenen Intervall gegeben, so kann sie in eindeutiger Weise nicht nur auf jedes Intervall, sondern auch auf bedeutend umfassendere Klassen von ebenen Punktmengen erweitert werden und behält dabei die angegebenen Eigenschaften. Ihr Wert in Punkten P oder auf Linien l kann dabei nach dem oben angegebenen Grenzübergang mit Hilfe eines Systems ineinandergeschachtelter Intervalle erhalten werden, wobei der Grenzwert nicht davon abhängt, welches System von ineinandergeschachtelten Intervallen gerade genommen wird.

Man kann leicht die Zerlegung der Funktion $G(\Delta)$ in die Sprungfunktion und den stetigen Teil durchführen. Die Funktion $G(\Delta)$ sei nichtnegativ, additiv, normal und für alle in Δ_0 liegenden Intervalle Δ definiert. Ihre Unstetigkeitsstellen bezeichnen wir mit P_k ($k = 1, 2, \dots$). Wir definieren die Sprungfunktion $G_d(\Delta)$ folgendermaßen: $G_d(\Delta)$ ist gleich der Summe der Werte $G(P_k)$ für die Punkte P_k , die in Δ liegen. Wenn die Menge dieser Punkte abzählbar ist, dann konvergiert die aus den Werten $G(P_k)$ gebildete Reihe. Die Differenz $G(\Delta) - G_d(\Delta)$ bezeichnen wir mit $G_c(\Delta)$. Diese Funktion besitzt keine Unstetigkeitsstellen mehr. Man erkennt leicht, daß die

¹⁾ An späterer Stelle beweisen wir die erwähnte Eigenschaft von $G(\Delta)$ nicht nur für Intervalle, sondern auch für viel allgemeinere Punktmengen.

beiden Funktionen $G_d(\Delta)$ und $G_c(\Delta)$ nichtnegativ, additiv und normal sind. Die Normalität folgt unmittelbar aus den Ungleichungen $0 \leq G_d(\Delta) \leq G(\Delta)$ und $0 \leq G_c(\Delta) \leq G(\Delta)$. Die erhaltene Zerlegung der Funktion $G(\Delta)$ lautet

$$G(\Delta) = G_d(\Delta) + G_c(\Delta).$$

22. Übergang zur Punktfunktion. Man kann, von einer Punktfunktion $g(x, y)$ ausgehend, die Intervallfunktion $G(\Delta)$ bilden. Wir nehmen an, es sei die Punktfunktion $g(x, y)$ gegeben, die für alle x im Intervall $\Delta_x^{(0)}$ der x -Achse und alle y im Intervall $\Delta_y^{(0)}$ der y -Achse definiert ist, wobei $\Delta_x^{(0)}$ und $\Delta_y^{(0)}$ das Grundintervall Δ_0 in der Ebene definieren. Weiterhin nehmen wir an, daß die Funktion $g(x, y)$ als Funktion jeweils einer Veränderlichen bei beliebigem festem Wert der anderen Veränderlichen nicht abnimmt und daß außerdem für beliebige x und y und beliebige positive h und k

$$g(x+h, y+k) - g(x+h, y) - g(x, y+k) + g(x, y) \geq 0 \quad (115)$$

gilt. Dann existieren die Grenzwerte $g(x, y \pm 0)$ und $g(x \pm 0, y)$, und es gilt

$$a) \quad g(x_2, y+0) \geq g(x_1, y+0), \quad g(x_2, y-0) \geq g(x_1, y-0) \quad \text{für } x_2 > x_1$$

und

$$b) \quad g(x+0, y_2) \geq g(x+0, y_1), \quad g(x-0, y_2) \geq g(x-0, y_1) \quad \text{für } y_2 > y_1.$$

Wegen der Monotonie, die durch diese Ungleichungen ausgedrückt wird, können wir bezüglich beider Veränderlicher zur Grenze übergehen. Wir bilden die Grenzwerte

$$A = \lim_{k \rightarrow +0} \lim_{h \rightarrow +0} g(x+h, y+k) \quad \text{und} \quad B = \lim_{h \rightarrow +0} \lim_{k \rightarrow +0} g(x+h, y+k)$$

und beweisen, daß sie übereinstimmen: Es gilt $g(x+h, y+k) \geq g(x+h, y+0)$, und wir erhalten, wenn wir zuerst bezüglich h und dann bezüglich k zur Grenze übergehen, $A \geq B$; in der gleichen Weise kann man $B \geq A$ beweisen, also ist $A = B$. Die Größe $A = B$ wird mit $g(x+0, y+0)$ bezeichnet. Ganz analog kann man

$$\lim_{k \rightarrow +0} \lim_{h \rightarrow +0} g(x-h, y-k) = \lim_{h \rightarrow +0} \lim_{k \rightarrow +0} g(x-h, y-k)$$

beweisen. Diesen Grenzwert bezeichnet man mit $g(x-0, y-0)$. Bisher wurde nur benutzt, daß die Funktion $g(x, y)$ bezüglich jeder der beiden Veränderlichen nicht abnimmt. Wir verwenden jetzt auch die Bedingung (115). Wir bilden die Grenzwerte

$$A_1 = \lim_{h \rightarrow +0} \lim_{k \rightarrow +0} g(x+h, y-k)$$

und

$$B_1 = \lim_{k \rightarrow +0} \lim_{h \rightarrow +0} g(x+h, y-k)$$

und beweisen, daß $A_1 = B_1$ ist. Wenn wir $0 < h_1 < h$ und $0 < k_1 < k$ annehmen, können wir wegen (115)

$$g(x+h, y-k_1) - g(x+h, y-k) - g(x+h_1, y-k_1) + g(x+h_1, y-k) \geq 0$$

schreiben. Lassen wir zuerst h_1 und dann k_1 gegen 0 streben, so erhalten wir

$$g(x+h, y-0) - g(x+h, y-k) - A_1 + g(x+0, y-k) \geq 0.$$

Lassen wir nun hierin zuerst h und dann k gegen 0 streben, so haben wir $B_1 - A_1 \geq 0$, d. h. $B_1 \geq A_1$. Ganz analog kann man $A_1 \geq B_1$ zeigen, folglich gilt $A_1 = B_1$. Die

Größe $A_1 = B_1$ bezeichnen wir mit $g(x+0, y-0)$. Ganz analog gilt

$$\lim_{k \rightarrow +0} \lim_{h \rightarrow +0} g(x-h, y+k) = \lim_{h \rightarrow +0} \lim_{k \rightarrow +0} g(x-h, y+k).$$

Diesen letzten Grenzwert bezeichnen wir mit $g(x-0, y+0)$. Man kann leicht zeigen, daß sich in allen vier betrachteten Fällen derselbe Grenzwert ergibt, wenn h und k in beliebiger Weise gleichzeitig gegen $+0$ streben. Zum Beispiel kann man folgendes beweisen: Bei beliebig vorgegebenem positivem ε existiert ein solches positives η , daß für $0 < h < \eta$ und $0 < k < \eta$ die Abschätzung

$$|g(x+h, y-k) - g(x+0, y-0)| \leq \varepsilon \quad \text{für} \quad 0 < h < \eta \quad \text{und} \quad 0 < k < \eta$$

gilt.

Wenn die Bedingung (115) erfüllt ist, haben folglich die Ausdrücke $g(x \pm 0, y \pm 0)$ einen wohldefinierten Sinn. Mit Hilfe von $g(x, y)$ wird eine nichtnegative, additive und normale Funktion $G(\Delta)$ folgendermaßen gebildet: Es seien Δ_x und Δ_y Intervalle auf den Achsen, die das Intervall Δ in der Ebene definieren; a und b seien die Randpunkte des Intervalls Δ_x , ebenso c und d die Randpunkte des Intervalls Δ_y . Dann sei $G(\Delta)$ gleich dem Ausdruck

$$g(b \pm 0, d \pm 0) - g(b \pm 0, c \pm 0) - g(a \pm 0, d \pm 0) + g(a \pm 0, c \pm 0).$$

Dabei wird das Pluszeichen genommen, wenn das Intervall an dem betreffenden Ende rechts abgeschlossen oder links offen ist, das Minuszeichen, wenn das Intervall an dem betreffenden Ende rechts offen oder links abgeschlossen ist. So gilt z. B. im Fall des abgeschlossenen Intervalls $a \leq x \leq b$, $c \leq y \leq d$

$$G(\Delta) = g(b+0, d+0) - g(b+0, c-0) - g(a-0, d+0) + g(a-0, c-0),$$

im Fall des halboffenen Intervalls $a < x \leq b$, $c < y \leq d$

$$G(\Delta) = g(b+0, d+0) - g(b+0, c+0) - g(a+0, d+0) + g(a+0, c+0)$$

und im Fall des Punktes $x=a$, $y=c$

$$G(\Delta) = g(a+0, c+0) - g(a+0, c-0) - g(a-0, c+0) + g(a-0, c-0).$$

Ist umgekehrt eine Funktion $G(\Delta)$ gegeben, so kann leicht eine Punktfunktion $g(x, y)$ konstruiert werden, mit deren Hilfe sich in der eben skizzierten Weise $G(\Delta)$ ergibt. Wir nehmen z. B. an, daß $G(\Delta)$ in der ganzen offenen Ebene definiert ist. Dann kann man $g(x, y)$ gleich dem Wert von $G(\Delta)$ für das Intervall $\Delta_{x,y}$ annehmen, das durch die folgenden Intervalle auf den Achsen definiert wird: $-\infty < x' \leq x$; $-\infty < y' \leq y$. Die so gebildete Funktion $g(x, y)$ ist bezüglich x und y von rechts stetig. Wir erwähnen, daß sich die linke Seite von (115) nicht ändert, wenn zu $g(x, y)$ ein beliebiges Polynom ersten Grades in x und y hinzugefügt wird. Ist ferner $g(x, y)$ eine stetige Funktion, so hat die ihr zugeordnete Funktion $G(\Delta)$ weder Unstetigkeitspunkte noch Unstetigkeitslinien, und umgekehrt. Wenn im Punkt (x, y) die Funktion $g(x, y)$ zwar einen Unstetigkeitspunkt hat, aber

$$g(x+0, y+0) - g(x-0, y+0) - g(x+0, y-0) + g(x-0, y-0) = 0$$

gilt, dann ist dieser Punkt kein Unstetigkeitspunkt für $G(\Delta)$.

Als Beispiel betrachten wir die schon erwähnte Massenverteilung, und zwar nehmen wir an, daß die Masse mit einer linearen Dichte 1 auf der Strecke $x=1$, $0 \leq y \leq 1$ verteilt ist. Es gilt also

$$g(x, y) = \begin{cases} 0 & \text{für } x < 1 \quad \text{oder} \quad y < 0, \\ y & \text{für } x = 1 \quad \text{und} \quad 0 \leq y \leq 1, \\ 0 & \text{für } x \geq 1 \quad \text{und} \quad y > 1. \end{cases}$$

Ganz analog kann man Intervalle im dreidimensionalen x, y, z -Raum betrachten. Ein solches Intervall wird durch drei Intervalle auf den Achsen definiert. Außer Punkten und achsenparallelen Geraden muß man hier auch Ebenen betrachten, die den Koordinatenebenen parallel sind. Im übrigen stehen alle Überlegungen und Ergebnisse in Analogie zum zweidimensionalen Fall.

23. Das Stieltjessche Integral in der Ebene. Der Begriff des Stieltjesschen Integrals läßt sich leicht auf den Fall der Ebene erweitern. Dabei entsteht die Aufgabe, ein zweidimensionales Intervall zu zerlegen. Ist Δ ein beliebiges, durch die Intervalle Δ_x und Δ_y auf den Achsen definiertes Intervall in der Ebene, so verstehen wir unter einer *Zerlegung* von Δ eine solche Einteilung von Δ in Teilintervalle, die sich durch eine Zerlegung von Δ_x und Δ_y in Teilintervalle ergibt. Jedes Teilintervall von Δ wird durch ein Teilintervall von Δ_x und durch ein Teilintervall von Δ_y definiert. Abb. 1 zeigt eine solche Zerlegung eines halboffenen Intervalls in sechs Teilintervalle.

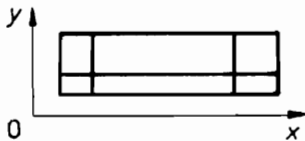


Abb. 1

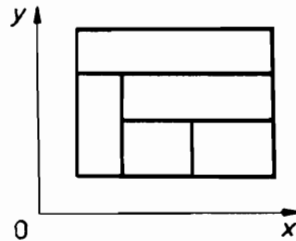


Abb. 2

Abb. 2 zeigt eine Einteilung ganz anderer Art. Bei ihr enden die Teilungslinien im Innern des Intervalls Δ , man kann sie also nicht in der angegebenen Weise durch Zerlegung der Intervalle Δ_x und Δ_y erhalten. Man kann aber leicht von einer Einteilung der zweiten Art zu einer Zerlegung der ersten Art kommen, wenn man alle Teilungslinien fortsetzt; wir können daher im folgenden nur Zerlegungen der ersten Art verwenden, was keine wesentliche Rolle spielt. Die Zerlegung δ' heißt Fortsetzung der Zerlegung δ , wenn die zu δ' gehörenden Zerlegungen von Δ_x und Δ_y Fortsetzungen der zu δ gehörenden Zerlegungen von Δ_x und Δ_y darstellen. Sind δ_1 und δ_2 zwei Zerlegungen von Δ und $\delta_x^{(1)}, \delta_x^{(2)}$ bzw. $\delta_y^{(1)}, \delta_y^{(2)}$ die zugehörigen Zerlegungen von Δ_x und Δ_y , so versteht man unter dem Produkt $\delta_1 \delta_2$ die Zerlegung, die durch die Zerlegungen $\delta_x^{(1)} \delta_x^{(2)}$ und $\delta_y^{(1)} \delta_y^{(2)}$ der Intervalle Δ_x und Δ_y definiert ist. Die Zerlegung $\delta_1 \delta_2$ ist offenbar eine Fortsetzung von δ_1 und von δ_2 . Wir bemerken noch: Wenn Δ' ein Intervall ist, das in Δ enthalten ist, so gibt es auch eine Zerlegung von Δ , bei der Δ' eines der Teilintervalle bildet.

Man kann nun leicht den dem Begriff des Stieltjesschen Integrals für den Fall der Ebene, des dreidimensionalen oder allgemein des n -dimensionalen Raumes analogen Begriff bilden. Wir beschränken uns auf den Fall der Ebene. In den anderen Fällen verläuft die Begriffsbildung ganz entsprechend.

Es sei in der Ebene ein gewisses endliches Intervall Δ_0 gegeben und in ihm eine gleichmäßig stetige und folglich beschränkte Punktfunktion $f(P)$ und eine nicht-negative, additive und normale Intervallfunktion $G(\Delta)$ definiert; δ sei eine bestimmte Zerlegung von Δ_0 in paarweise punktfremde Intervalle $\Delta_1, \Delta_2, \dots, \Delta_n$. Wir wählen

in jedem Δ_k einen bestimmten Punkt P_k und bilden die Summe

$$\sigma_\delta = \sum_{k=1}^n f(P_k) G(\Delta_k). \quad (116)$$

Ganz analog zu [4] läßt sich beweisen, daß diese Summe einen festen Grenzwert hat, wenn der größte Durchmesser der Bereiche Δ_k gegen 0 strebt. Dieser Grenzwert heißt das Integral von $f(P)$ bezüglich $G(\Delta)$:

$$\int_{\Delta_0} f(P) G(d\Delta) = \lim \sum_{k=1}^n f(P_k) G(\Delta_k). \quad (117)$$

Ist $f(x, y)$ in dem abgeschlossenen Intervall Δ_0 ($a_1 \leq x \leq b_1$, $a_2 \leq y \leq b_2$) stetig und besitzt $g(x, y)$ die oben angegebenen Eigenschaften, so kann man das Stieltjessche Integral als Grenzwert der Summen

$$\sigma_\delta = \sum_{k=1}^p \sum_{l=1}^q f(\xi_k, \eta_l) [g(x_k, y_l) - g(x_{k-1}, y_l) - g(x_k, y_{l-1}) + g(x_{k-1}, y_{l-1})] \quad (118)$$

bei beliebig kleiner werdenden Intervallen definieren, wobei

$$\begin{aligned} a_1 &= x_0 < x_1 < \dots < x_{p-1} < x_p = b_1, \\ a_2 &= y_0 < y_1 < \dots < y_{q-1} < y_q = b_2, \\ x_{k-1} &\leq \xi_k \leq x_k, \quad y_{l-1} \leq \eta_l \leq y_l \end{aligned} \quad (119)$$

ist.

Wie auch in [4] kann man zeigen, daß das Stieltjessche Integral unter der Voraussetzung existiert, daß $f(P)$ beschränkt und im Innern von Δ_0 stetig ist, wenn $G(\Delta)$ die folgende zusätzliche Bedingung erfüllt:

Gegeben seien abgeschlossene Intervalle $\Delta^{(n)}$ im Innern von Δ_0 , die mit n größer werden und derart gegen Δ_0 streben, daß jeder innere Punkt von Δ_0 in allen $\Delta^{(n)}$ von einem genügend großen n ab enthalten ist. Wir fordern dann, daß $G(\Delta^{(n)}) \rightarrow G(\Delta_0)$ gilt. Diese Eigenschaft steht in Analogie zur Stetigkeit von $g(x)$ in den Randpunkten des Intervalls, die wir in [4] forderten.

Das Stieltjessche Integral kann auch für die gesamte Ebene $-\infty < x < \infty$, $-\infty < y < \infty$ definiert werden, die wir mit Q bezeichnen. $G(\Delta)$ sei eine nichtnegative, additive und normale Funktion, die für alle endlichen und unendlichen in Q liegenden Intervalle definiert ist. $\Delta^{(n)}$ sei eine Folge von Intervallen, die nach allen Seiten unbegrenzt wachsen, z. B. sei $\Delta^{(n)}$ gegeben durch $-n \leq x \leq n$, $-n \leq y \leq n$. Aus der Normalität von $G(\Delta)$ folgt, daß $G(\Delta^{(n)}) \rightarrow G(Q)$ für $n \rightarrow \infty$ gilt. Ist $f(P)$ in Q stetig und beschränkt, so existiert das Stieltjessche Integral (117). Dabei muß die Folge von Zerlegungen so beschaffen sein, daß bei beliebig festgelegtem n der größte Durchmesser der Intervalle, die mit $\Delta^{(n)}$ gemeinsame Punkte haben, gegen 0 strebt.

Als Integrationsgebiet kann an Stelle des Intervalls Δ_0 ein Bereich S auftreten, der die Summe endlich vieler Intervalle ist. Wir können beliebig feine Zerlegungen in Teilintervalle durchführen, die Summen (116) bilden und zur Grenze übergehen. Das Integral über S wird zurückgeführt auf eine endliche Summe von Integralen über die Intervalle, in die S zerlegt wird. Offenbar hängt das Integral über S nicht davon ab, wie die Intervallzerlegung von S gewählt wird.

Die Eigenschaften des Stieltjesschen Doppelintegrals sind den früher behandelten Eigenschaften des einfachen Integrals völlig analog.

24. Funktionen von beschränkter Variation in der Ebene. Das Studium von Funktionen von beschränkter Variation in der Ebene verläuft in vielem analog dem

Vorigen. Die Formulierung ist etwas verschieden, weil wir für die Darstellung nicht die Terminologie der Punktfunktionen, sondern die der Intervallfunktionen verwenden. Es sei $G(\Delta)$ eine additive und normale Intervallfunktion, die für alle Intervalle (in der allgemeinen Bedeutung dieses Begriffs) definiert ist, die in einem bestimmten Grundintervall Δ_0 liegen. Dabei wird nicht vorausgesetzt, daß diese Funktion $G(\Delta)$ nicht negativ ist, und $\Delta_1, \Delta_2, \dots, \Delta_n$ sei eine gewisse Zerlegung δ des Intervalls Δ_0 in Teilintervalle. Wir bilden die Summe

$$t_\delta = \sum_{k=1}^n |G(\Delta_k)|. \quad (120)$$

Definition. Wenn für alle möglichen Zerlegungen δ die Menge der Werte t_δ beschränkt bleibt, nennt man die Funktion $G(\Delta)$ eine Funktion von beschränkter Variation im Intervall Δ_0 , die obere Grenze dieser Summen t_δ nennt man Totalvariation oder einfach Variation der Funktion $G(\Delta)$ im Intervall Δ_0 . Wir bezeichnen sie mit $V_{\Delta_0}(G)$.

Die Eigenschaften der Summen t_δ und der Totalvariation stehen in vollständiger Analogie zu den in [8] angeführten Eigenschaften, und wir werden die meisten dieser Eigenschaften ohne Beweis angeben.

Ist δ' eine Fortsetzung der Zerlegung δ , so gilt $t_{\delta'} \geq t_\delta$. Wenn die Funktion $G(\Delta)$ in Δ_0 von beschränkter Variation ist, dann ist sie auch in einem beliebigen, in Δ_0 liegenden Intervall Δ' von beschränkter Variation, und es gilt $V_{\Delta'}(G) \leq V_{\Delta_0}(G)$. Jede nichtnegative oder nichtpositive Funktion $G(\Delta)$ ist eine Funktion von beschränkter Variation. Liegt das Intervall Δ' in Δ_0 , so gilt

$$|G(\Delta')| \leq V_{\Delta_0}(G), \quad (121)$$

und eine Funktion $G(\Delta)$, die in Δ_0 von beschränkter Variation ist, ist in allen im Intervall Δ_0 liegenden Intervallen Δ (dem absoluten Betrag nach) beschränkt. Jede Linearkombination von Funktionen von beschränkter Variation ist eine Funktion von beschränkter Variation. Es gilt Satz 3 aus [8] über Produkt und Quotient. Die Totalvariation $V_\Delta(G)$ ist eine in Δ_0 definierte nichtnegative Funktion. Durch Wiederholung des Beweises zu Satz 4 in [8] läßt sich zeigen, daß $V(\Delta) = V_\Delta(G)$ additiv ist. Wir beweisen jetzt, daß die Funktion $V(\Delta)$ in Δ_0 normal ist. Es sei Δ_m ($m = 1, 2, 3, \dots$) eine verschwindende Folge von Intervallen. Wir haben zu zeigen, daß $V(\Delta_m) \rightarrow 0$ gilt. Es sei ε eine vorgegebene positive Zahl. Wir wählen eine Zerlegung $\delta: \Delta^{(1)}, \Delta^{(2)}, \dots, \Delta^{(p)}$ des Intervalls Δ_0 , für die

$$\sum_{k=1}^p |G(\Delta^{(k)})| \geq V(\Delta_0) - \varepsilon \quad (122)$$

gilt.

Für ein beliebiges k aus der Reihe der Zahlen $k = 1, 2, \dots, p$ ist das Produkt $\Delta_m \Delta^{(k)}$ ($m = 1, 2, \dots$) eine verschwindende Folge. Wegen der Normalität von $G(\Delta)$ kann man einen Wert $m = m_0$ festlegen derart, daß für $m > m_0$

$$|G(\Delta_m \Delta^{(k)})| \leq \frac{\varepsilon}{p} \quad (k = 1, 2, \dots, p) \quad (123)$$

gilt. Wir legen ein beliebiges $m > m_0$ fest. Jedes Intervall $\Delta^{(k)}$ wird nun noch weiterhin zerlegt, und zwar derart, daß $\Delta_m \Delta^{(k)}$ als Teilintervall auftritt. Es seien $\Delta_s^{(k)}$ ($s = 1, 2, \dots, n_k$) die übrigen Teilintervalle bei dieser Zerlegung von $\Delta^{(k)}$. Man erhält also eine bestimmte Zerlegung von Δ_0 , die eine Fortsetzung der Zerlegung δ dar-

stellt, für die also erst recht die Ungleichung (122), d. h.

$$\sum_{k=1}^p |G(\Delta_m \Delta^{(k)})| + \sum_{k=1}^p \sum_{s=1}^{n_k} |G(\Delta_s^{(k)})| \equiv V(\Delta_0) - \varepsilon$$

gilt.

Wegen der Additivität von $V(\Delta)$ und der Ungleichung $|G(\Delta_s^{(k)})| \leq V(\Delta_s^{(k)})$ ergibt die letzte Ungleichung für $m > m_0$

$$\sum_{k=1}^p |G(\Delta_m \Delta^{(k)})| + \sum_{k=1}^p \sum_{s=1}^{n_k} V(\Delta_s^{(k)}) \equiv V(\Delta_m) + \sum_{k=1}^p \sum_{s=1}^{n_k} V(\Delta_s^{(k)}) - \varepsilon$$

oder

$$V(\Delta_m) \leq \sum_{k=1}^p |G(\Delta_m \Delta^{(k)})| + \varepsilon.$$

Unter Berücksichtigung von (123) erhält man

$$V(\Delta_m) \leq \sum_{k=1}^p \frac{\varepsilon}{p} + \varepsilon = 2\varepsilon \quad \text{für } m > m_0,$$

woraus, da ε beliebig war, schließlich $V(\Delta_m) \rightarrow 0$ folgt. Also ist $V(\Delta)$ eine nichtnegative, additive und normale Intervallfunktion in Δ_0 . Wir definieren weiter zwei nichtnegative, additive und normale Funktionen durch

$$G_1(\Delta) = \frac{1}{2} [V(\Delta) + G(\Delta)], \quad G_2(\Delta) = \frac{1}{2} [V(\Delta) - G(\Delta)] \quad (124)$$

und erhalten damit eine kanonische Darstellung der Funktion $G(\Delta)$ von beschränkter Variation als Differenz zweier nichtnegativer additiver und normaler Funktionen,

$$G(\Delta) = G_1(\Delta) - G_2(\Delta). \quad (125)$$

Ist eine beliebige andere analoge Darstellung

$$G(\Delta) = G_1^*(\Delta) - G_2^*(\Delta) \quad (126)$$

gegeben, so gilt für jedes Δ aus Δ_0

$$G_1(\Delta) \leq G_1^*(\Delta) \quad \text{und} \quad G_2(\Delta) \leq G_2^*(\Delta).$$

Ist $G(\Delta)$ umgekehrt die Differenz zweier nichtnegativer additiver und normaler Funktionen, so ist $G(\Delta)$ eine Funktion von beschränkter Variation.

Wenn das Intervall Δ aus dem Punkt P besteht, dann stimmt $V(P)$ mit $G(P)$ überein. Ist $G(P) = 0$, so ist auch $V(P) = 0$, d. h., Stetigkeit von $G(\Delta)$ in einem bestimmten Punkt zieht Stetigkeit von $V(\Delta)$ in demselben Punkt nach sich. Für achsenparallele Strecken verhält sich die Sache etwas komplizierter. Wir deuten $G(\Delta)$ als Menge der Ladungen, die im Intervall Δ verteilt sind. Sind z. B. auf einer achsenparallelen Strecke l mit einer bestimmten linearen Dichte elektrische Ladungen verteilt, deren Gesamtsumme gleich 0 ist, so ist $G(l) = 0$. Andererseits gilt $V(l) > 0$, weil $V(l)$ die Gesamtsumme der Ladungen angibt, wobei alle Ladungen mit positivem Vorzeichen genommen werden.

Man könnte die additive und normale Funktion $G(\Delta)$ nur für halboffene Intervalle definieren und nur Zerlegungen in halboffene Intervalle durchführen. Dann könnten wir alle oben angegebenen Überlegungen anstellen und würden zur Formel (125) gelangen. Die nichtnegativen, additiven und normalen Funktionen $G_1(\Delta)$ und $G_2(\Delta)$ für halboffene Intervalle ließen sich auf alle möglichen Intervalle ausdehnen;

dadurch lieferte uns die Formel (126) die Ausdehnung der gegebenen Funktion $G(\Delta)$ auf alle möglichen Intervalle, und diese Funktion wäre additiv, normal und von beschränkter Variation für alle möglichen Intervalle.

Benutzen wir die kanonische Darstellung (125), so können wir das Integral bezüglich $G(\Delta)$ in der gewöhnlichen Weise definieren:

$$\int_{\Delta^{(0)}} f(P) G(d\Delta) = \int_{\Delta^{(0)}} f(P) G_1(d\Delta) - \int_{\Delta^{(0)}} f(P) G_2(d\Delta). \quad (127)$$

Wenn wir das abgeschlossene Intervall $\Delta^{(0)}$ ($a_1 \leq x \leq b_1$, $a_2 \leq y \leq b_2$) haben, können wir unter Benutzung von (115) eine Funktion $g(x, y)$ von beschränkter Variation definieren, und ihre Totalvariation erhalten wir, wie auch in [8], aus den Summen

$$t_g = \sum_{k=1}^p \sum_{l=1}^q |g(x_k, y_l) - g(x_{k-1}, y_l) - g(x_k, y_{l-1}) + g(x_{k-1}, y_{l-1})|. \quad (128)$$

25. Der Raum der stetigen Funktionen mehrerer Veränderlicher. Wir nehmen an, daß in einem endlichen abgeschlossenen Intervall Δ_0 ($a_1 \leq x \leq b_1$, $a_2 \leq y \leq b_2$) eine stetige Funktion $f(x, y)$ gegeben ist. Durch die lineare Transformation $x' = ax + b$, $y' = cy + d$ ($a \neq 0$ und $c \neq 0$) können wir Δ_0 in das Intervall $0 \leq x' \leq 1$, $0 \leq y' \leq 1$ überführen, und die stetige Funktion bleibt bei dieser Transformation auch stetig. Damit können wir annehmen, daß das Ausgangsintervall Δ_0 bereits das Intervall $0 \leq x \leq 1$, $0 \leq y \leq 1$ ist. Wir bilden für $f(x, y)$ die Polynome von S. N. BERNSTEIN [II, 168]:

$$P_{m,n}(x, y) = \sum_{k=0}^m \sum_{l=0}^n f\left(\frac{k}{m}, \frac{l}{n}\right) \binom{m}{k} \binom{n}{l} x^k y^l (1-x)^{m-k} (1-y)^{n-l}. \quad (129)$$

Wenn wir die gleichmäßige Stetigkeit von $f(x, y)$ in Δ_0 benutzen, können wir wie in [II, 168] zeigen, daß $P_{m,n}(x, y)$ bei unbegrenzt wachsenden n und m in Δ_0 gleichmäßig gegen $f(x, y)$ strebt. Wir nehmen nun an, daß die Funktion $f(x, y)$ auf einer beschränkten abgeschlossenen Menge F stetig ist. Wenn wir die lineare Transformation (siehe oben) benutzen, können wir annehmen, daß F zu dem oben definierten Intervall Δ_0 gehört. Weiter können wir $f(x, y)$ auf das ganze Intervall so fortsetzen, daß die Funktion in Δ_0 stetig ist und der größte Wert $|f(x, y)|$ in Δ_0 gleich dem größten Wert $|f(x, y)|$ in F ist. Für die so fortgesetzte Funktion $f(x, y)$ können wir die Folge der Polynome $P_{m,n}(x, y)$ so konstruieren, daß $P_{m,n}(x, y)$ auf Δ_0 und somit auch auf F gleichmäßig gegen $f(x, y)$ strebt.

Wir betrachten jetzt analog dazu, wie auch schon in [14], den Raum C der in Δ_0 stetigen Funktionen mit der Norm $\|f\| = \max |f(x, y)|$ in Δ_0 . Wie auch in [15] kann man zeigen, daß die allgemeine Form des linearen Funktional in C ein Stieltjesches Integral ist:

$$\Phi(f) = \int_{\Delta_0} f(x, y) G(d\Delta), \quad (130)$$

wobei $G(\Delta)$ eine Funktion von beschränkter Variation in Δ_0 ist, die das Funktional $\Phi(f)$ charakterisiert.

Für die Definition einer Funktion von beschränkter Variation kann man die Summen (128) benutzen, und das Integral kann bei unbeschränkt kleiner werdenden Intervallen als Grenzwert der Summen (118) definiert werden.¹⁾

26. Das Fourier-Stieltjesche Integral. Wir betrachten die durch ein Fourier-Stieltjesches Integral dargestellte Funktion

$$\varphi(t) = \int_{-\infty}^{+\infty} e^{itx} dG(x) \quad (-\infty < t < \infty); \quad (131)$$

¹⁾ Ausführungen über den Raum der stetigen Funktionen auf beschränkten abgeschlossenen Mengen finden sich bei J. RADON, Über lineare Funktionaltransformationen und Funktionalgleichungen, Sitzungsber. Akad. Wiss. Wien 128 (1919), 1083–1121, und bei F. RIESZ und B. SZ. NAGY, Vorlesungen über Funktionalanalysis (vgl. das Literaturverzeichnis).

dabei sei $g(x)$ eine nicht abnehmende, beschränkte und für $x = \pm \infty$ stetige Funktion, d. h., es gelte

$$g(-\infty) = \lim_{x \rightarrow -\infty} g(x), \quad g(\infty) = \lim_{x \rightarrow \infty} g(x).$$

Das Integral (131) existiert offenbar, da die Funktion e^{ix} stetig und beschränkt ist [4]. Wir leiten einige einfache Eigenschaften der Funktion $\varphi(t)$ her. Es gilt

$$|\varphi(t)| \leq \int_{-\infty}^{\infty} |e^{itx}| dg(x) = \int_{-\infty}^{\infty} dg(x) = g(\infty) - g(-\infty) = \varphi(0)$$

oder $|\varphi(t)| \leq \varphi(0)$, d. h., $\varphi(t)$ ist eine beschränkte Funktion.

Aus der Formel (131) folgt unmittelbar auch die Identität

$$\varphi(-t) = \overline{\varphi(t)}. \quad (132)$$

Wir beweisen noch, daß die Funktion $\varphi(t)$ im Intervall $(-\infty, \infty)$ gleichmäßig stetig ist. Wir schätzen dazu den absoluten Betrag der Differenz $\varphi(t+h) - \varphi(t)$ ab:

$$|\varphi(t+h) - \varphi(t)| \leq \int_{-\infty}^{\infty} |e^{ixt}| |(e^{ixh} - 1)| dg(x) = 2 \int_{-\infty}^{\infty} \left| \sin \frac{hx}{2} \right| dg(x). \quad (133)$$

Zuerst legen wir n so groß fest, daß

$$[g(-n) - g(-\infty)] \leq \varepsilon, \quad [g(\infty) - g(n)] \leq \varepsilon$$

gilt. Ferner wählen wir ein nicht von t abhängiges festes η so, daß für $|h| \leq \eta$ im Intervall $-n \leq x \leq n$ die Ungleichung

$$2 \left| \sin \frac{hx}{2} \right| \leq \varepsilon$$

gilt. Wegen (133) erhalten wir dann

$$|\varphi(t+h) - \varphi(t)| \leq [2 + g(n) - g(-n)] \varepsilon \leq [2 + g(\infty) - g(-\infty)] \varepsilon,$$

woraus wegen der Willkürlichkeit von ε und der Unabhängigkeit der Größe η von t schließlich die gleichmäßige Stetigkeit von $\varphi(t)$ folgt. Wir zeigen noch eine andere Eigenschaft der Funktion $\varphi(t)$. Wir nehmen m beliebige reelle Zahlen $t_1, t_2, \dots, t_m$ und bilden die hermitesche Form

$$\sum_{p,q=1}^m \varphi(t_p - t_q) \xi_p \overline{\xi_q} \quad (134)$$

in den Variablen ξ_s . Unter Berücksichtigung von (131) können wir

$$\varphi(t_p - t_q) \xi_p \overline{\xi_q} = \int_{-\infty}^{\infty} e^{it_p x} \xi_p \overline{e^{it_q x} \xi_q} dg(x)$$

schreiben und erhalten für die hermitesche Form (134)

$$\sum_{p,q=1}^m \varphi(t_p - t_q) \xi_p \overline{\xi_q} = \int_{-\infty}^{\infty} \left| \sum_{s=1}^m e^{it_s x} \xi_s \right|^2 dg(x).$$

Hieraus folgt unmittelbar, daß für beliebiges m und für beliebig gewählte Werte t_s der Veränderlichen t die hermitesche Form (34) nichtnegativ ist, d. h., daß

$$\sum_{p,q=1}^m \varphi(t_p - t_q) \xi_p \overline{\xi_q} \geq 0 \quad (135)$$

gilt.

Wir führen nun einen neuen Begriff ein. Die Funktion $\varphi(t)$ heißt *positiv definit*, wenn sie im Intervall $(-\infty, \infty)$ stetig und beschränkt ist, der Identität (132) genügt und wenn außerdem die hermitesche Form (134) bei beliebigem m und bei beliebig gewählten

Punkten t_s nicht negativ ist. Aus den vorigen Überlegungen folgt: Wenn die Funktion $\varphi(t)$ mit Hilfe einer nicht abnehmenden beschränkten Funktion $g(x)$ durch das Fourier-Stieltjessche Integral (131) dargestellt wird, dann ist sie positiv definit. Auch das Umgekehrte gilt: *Jede positiv definite Funktion $\varphi(t)$ ist durch ein Integral (131) mit Hilfe einer nicht abnehmenden beschränkten Funktion $g(x)$ darstellbar.*¹⁾ Wir betrachten nochmals das Integral (131) und stellen die Funktion $g(x)$ als eine Summe dar, $g(x) = g_d(x) + g_c(x)$, in der $g_d(x)$ die Sprungfunktion und $g_c(x)$ den stetigen Teil der Funktion $g(x)$ bilden. Die Funktion $\varphi(t)$ ergibt sich dann in Gestalt der Summe

$$\varphi(t) = \varphi_1(t) + \varphi_2(t)$$

mit

$$\varphi_1(t) = \int_{-\infty}^{\infty} e^{itx} dg_d(x), \quad \varphi_2(t) = \int_{-\infty}^{\infty} e^{itx} dg_c(x). \quad (136)$$

Es seien x_k ($k=1, 2, \dots$) die Abszissen der Sprungstellen der Funktion $g(x)$, und es sei $a_k = g(x_k+0) - g(x_k-0)$, wobei offenbar $a_k > 0$ gilt und die Reihe aus den a_k konvergiert. Es gilt folgende Entwicklung von $\varphi_1(t)$ in eine gleichmäßig konvergente Reihe:

$$\varphi_1(t) = \sum_k a_k e^{ix_k t}. \quad (137)$$

Wenn die Anzahl der Punkte x_k endlich ist, dann ist auch diese Summe endlich.

Wir führen jetzt den sogenannten *Mittelwert* einer beliebigen, stetigen und im Intervall $(-\infty, \infty)$ definierten Funktion $F(t)$ ein:

$$M\{F(t)\} = \lim_{\omega \rightarrow +\infty} \frac{1}{2\omega} \int_{-\omega}^{\omega} F(t) dt. \quad (138)$$

Dabei wird vorausgesetzt, daß der rechts stehende Grenzwert existiert. Es läßt sich zeigen, daß für die Funktion $\varphi_2(t)$, die mit Hilfe der stetigen Funktion $g_c(x)$ durch die Formel (136) definiert wurde, die Beziehung

$$M\{\varphi_2(t)\} = 0 \quad (139)$$

gilt. Denn wenn wir die Integration unter dem Integralzeichen durchführen, erhalten wir

$$\frac{1}{2\omega} \int_{-\omega}^{\omega} \varphi_2(t) dt = \int_{-\infty}^{\infty} \frac{\sin \omega x}{\omega x} dg_c(x).$$

Zerlegen wir das Integrationsintervall $[-\infty, \infty]$ in die drei Teile $[-\infty, -a]$, $[-a, a]$ und $[a, \infty]$, so erhalten wir durch eine elementare Abschätzung der Integrale

$$\left| \frac{1}{2\omega} \int_{-\omega}^{\omega} \varphi_2(t) dt \right| \leq \frac{1}{a\omega} [g_c(-a) - g_c(-\infty)] + \frac{1}{a\omega} [g_c(\infty) - g_c(a)] + [g_c(a) - g_c(-a)].$$

Es sei ε eine vorgegebene positive Zahl. Wegen der Stetigkeit von $g_c(x)$ im Punkt $x=0$ können wir a so klein wählen, daß $g_c(a) - g_c(-a) \leq \frac{\varepsilon}{2}$ gilt. Bei festgelegtem a streben die ersten beiden Summanden auf der rechten Seite der letzten Ungleichung für $\omega \rightarrow \infty$ gegen 0, also gilt für alle genügend großen ω

$$\left| \frac{1}{\omega} \int_{-\omega}^{\omega} \varphi_2(t) dt \right| \leq \varepsilon.$$

¹⁾ Vgl. S. BOCHNER, Vorlesungen über Fouriersche Integrale, Akademische Verlagsgesellschaft, Leipzig 1932, S. 74, oder S. BOCHNER, Monotone Funktionen, Stieltjessche Integrale und harmonische Analyse, Math. Ann. 108 (1933), 378–410.

Da ε beliebig war, folgt daraus schließlich (139). Daß die Integration nach t unter dem Integralzeichen ausgeführt werden kann, läßt sich leicht verifizieren.

Wir betrachten nun die Funktion $\varphi_2(t) e^{-i\lambda t}$, die wir in der Gestalt

$$\varphi_2(t) e^{-i\lambda t} = \int_{-\infty}^{+\infty} e^{i(x-\lambda)t} dg_c(x)$$

oder, nach Transformation der Integrationsvariablen,

$$\varphi_2(t) e^{-i\lambda t} = \int_{-\infty}^{+\infty} e^{itx} dg_c(x + \lambda)$$

schreiben können. Wegen der Stetigkeit der Funktion $g_c(x)$ gilt für jedes reelle λ

$$M\{\varphi_2(t) e^{-i\lambda t}\} = 0.$$

Wenn man die gleichmäßige Konvergenz der Reihe (137) im Intervall $(-\infty, \infty)$ und die Beziehung $M\{e^{i\lambda t}\} = 0$ für $\lambda \neq 0$ benutzt, erhält man

$$M\{\varphi_1(t) e^{-i\lambda t}\} = a_k$$

für $\lambda = x_k$ und

$$M\{\varphi_1(t) e^{-i\lambda t}\} = 0,$$

wenn λ mit keinem der x_k übereinstimmt. Unsere Überlegungen führen uns zu folgendem allgemeinem Ergebnis: *Ist $\varphi(t)$ mit Hilfe einer nicht abnehmenden beschränkten Funktion $g(x)$ in der Gestalt des Integrals (131) darstellbar, so gilt $M\{\varphi(t) e^{-i\lambda t}\} = g(\lambda + 0) - g(\lambda - 0)$, und die rechte Seite ist gleich 0, wenn λ keine Unstetigkeitsstelle von $g(x)$ ist.* Der Mittelwert des Produktes $\varphi(t) e^{-i\lambda t}$ ist eine Verallgemeinerung des Begriffs der Fourierkoeffizienten bei periodischen Funktionen. Wir kommen noch im Zusammenhang mit der Verallgemeinerung der Vollständigkeitsrelation¹⁾ [29] auf die verallgemeinerten Fourierkoeffizienten zurück.

Im folgenden Abschnitt leiten wir die Umkehrformel zu dem Integral (131) her, d. h. eine Formel, die $g(x)$ durch $\varphi(t)$ ausdrückt.

27. Die Umkehrformel. In Teil II haben wir den zweiten Mittelwertsatz, die Entwickelbarkeit einer Funktion in eine Fourierreihe und die Fouriersche Integralformel unter der Voraussetzung bewiesen, daß die untersuchte Funktion den Dirichletschen Bedingungen genügt. Verfolgt man den Beweisgang, so kann man sich leicht davon überzeugen, daß alle Beweise gültig bleiben, wenn man die Dirichletsche Bedingung durch die Forderung ersetzt, daß die Funktion von beschränkter Variation ist. Für das Fouriersche Integral gilt also das folgende Ergebnis: Wenn die Funktion $f(x)$ in jedem endlichen Intervall von beschränkter Variation und über das unendliche Intervall im Riemannschen Sinne absolut integrierbar ist, dann folgt aus

$$\varphi(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f(x) e^{itx} dx \quad (140)$$

die Umkehrformel

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \varphi(t) e^{-itx} dt; \quad (141)$$

das letzte Integral ist dabei im Sinne des Hauptwertes zu verstehen, und auf der linken Seite ist $f(x)$ in einem Unstetigkeitspunkt durch $\frac{1}{2} [f(x-0) + f(x+0)]$ zu ersetzen.

¹⁾ Parsevalsche Gleichung (Anm. d. Red.).

Dieses Ergebnis wird manchmal in etwas anderer Form geschrieben, und zwar: Aus

$$\varphi(t) = \int_{-\infty}^{\infty} f(x) e^{itx} dx \quad (142)$$

folgt nach [II, 173] (Integral im Sinne des Hauptwertes!)

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \varphi(t) e^{-itx} dt = \frac{1}{2\pi} \lim_{N \rightarrow \infty} \int_{-N}^N \varphi(t) e^{-itx} dt. \quad (143)$$

Unsere Aufgabe besteht nun darin, die Umkehrformel für das Integral (131) zu bilden. Man kann leicht vorausbestimmen, welches Aussehen diese Formel haben muß. Wir wenden das folgende heuristische Verfahren an, das natürlich keine Beweiskraft besitzt. Wir ersetzen im Integral (131) $dg(x)$ durch $g'(x) dx$, wobei $g'(x)$ die Ableitung von $g(x)$ sei. Auf diese Weise erhalten wir

$$\varphi(t) = \int_{-\infty}^{\infty} g'(x) e^{itx} dx,$$

und die Umkehrformel für das gewöhnliche Fouriersche Integral liefert

$$g'(x) = \frac{1}{2\pi} \lim_{N \rightarrow \infty} \int_{-N}^N \varphi(t) e^{-itx} dt.$$

Wenn wir beide Seiten dieser Formel nach x integrieren, etwa von $x=0$ bis zu einem bestimmten Wert x , und auf der rechten Seite die Integration unter dem Integralzeichen durchführen, erhalten wir schließlich die Umkehrformel

$$g(x) - g(0) = \frac{1}{2\pi} \lim_{N \rightarrow \infty} \int_{-N}^N \varphi(t) \frac{1 - e^{-itx}}{it} dt. \quad (144)$$

Auf der linken Seite tritt der Wert von $g(x)$ in dem festen Punkt $x=0$ auf, weil die Funktion $g(x)$ nur bis auf eine positive Konstante bestimmt ist.

Wir haben nun die Formel (144) zu beweisen, legen dazu auf irgendeine Weise einen Wert x der Veränderlichen fest und betrachten die Funktion

$$h(y) = g(y+x) - g(y) \quad (145)$$

als Funktion der Veränderlichen y . Als Differenz zweier zunehmender Funktionen ist sie in jedem endlichen Intervall eine Funktion von beschränkter Variation. Wir beweisen noch, daß sie im unendlichen Intervall absolut integrierbar ist. Wir nehmen $x > 0$ an. Für $x < 0$ verläuft der Beweis ebenso. Wenn wir berücksichtigen, daß die Funktion $g(y)$ wächst, dann können wir

$$\int_p^q |h(y)| dy = \int_p^q [g(y+x) - g(y)] dy = \int_p^q g(y+x) dy - \int_p^q g(y) dy \quad (q > p) \quad (146)$$

oder, wenn wir im ersten Integral der rechten Seite die Integrationsveränderliche $z = y+x$ einführen,

$$\int_p^q |h(y)| dy = \int_{p+x}^{q+x} g(z) dz - \int_p^q g(z) dz = \int_q^{q+x} g(z) dz - \int_p^{p+x} g(z) dz$$

schreiben. Wenn wir berücksichtigen, daß im Intervall $[p, p+x]$ die Ungleichung $g(p) \leq g(z)$ und im Intervall $[q, q+x]$ die Ungleichung $g(q+x) \geq g(z)$ gilt, erhalten wir

$$\int_p^q |h(y)| dy \leq [g(q+x) - g(p)] x.$$

Hieraus erkennt man, daß für genügend große Werte von p und für beliebige $q > p$ das Integral (146) beliebig klein ist, womit die absolute Integrierbarkeit der Funktion $h(y)$ in einem unendlichen Intervall bewiesen ist. Also ist die Formel (143) auf (145) anwendbar:

$$g(y+x) - g(y) = \frac{1}{2\pi} \lim_{N \rightarrow \infty} \int_{-N}^N e^{-ity} \left[\int_{-\infty}^{\infty} [g(z+x) - g(z)] e^{itz} dz \right] dt$$

oder, für $y = 0$,

$$g(x) - g(0) = \frac{1}{2\pi} \lim_{N \rightarrow \infty} \int_{-N}^N \left[\int_{-\infty}^{\infty} [g(z+x) - g(z)] e^{itz} dz \right] dt. \quad (147)$$

Wir betrachten das in der Klammer stehende Integral I und formen es durch partielle Integration um:

$$I = \frac{1}{it} \int_{-\infty}^{\infty} [g(z+x) - g(z)] d e^{itz} = - \frac{1}{it} \int_{-\infty}^{\infty} e^{itz} d [g(z+x) - g(z)].$$

Durch Anwendung der Gleichung (131) erhalten wir

$$I = \frac{1}{it} \varphi(t) - \frac{1}{it} \int_{-\infty}^{\infty} e^{itz} dg(z+x).$$

Wenn wir in diesem Integral die Integrationsveränderliche $u = z + x$ einführen, so kommen wir schließlich zu der Gleichung

$$\int_{-\infty}^{\infty} [g(z+x) - g(z)] e^{itz} dz = \frac{1 - e^{-itx}}{it} \varphi(t). \quad (148)$$

Durch Einsetzen dieser Formel in (147) erhalten wir die Umkehrformel (144).

Bemerkung. In der Gleichung (143), die die Umkehrung der gewöhnlichen Fouriertransformation darstellt, muß man $f(x)$ auf der linken Seite in den Unstetigkeitsstellen dieser Funktion bekanntlich [II, 155] gleich dem arithmetischen Mittel der Grenzwerte von links und von rechts wählen. Folglich gilt dasselbe sowohl für die linke Seite von (147) als auch für die linke Seite der Umkehrformel (144). Die Umkehrformel (144) für das Integral (131) ist auch in dem Fall richtig, daß $g(x)$ eine Funktion von beschränkter Variation ist. Man kann sich davon überzeugen, wenn man $g(x)$ als Differenz zweier zunehmender Funktionen darstellt, das Integral (131) in zwei Integrale zerlegt und auf jedes dieser beiden Integrale die soeben bewiesene Umkehrformel anwendet.

28. Der Faltungssatz. Die Umkehrformel für das gewöhnliche Fouriersche Integral steht bekanntlich [IV/1, 51, 52] in unmittelbarem Zusammenhang mit der Umkehrformel für das Laplacesche Integral. Für dieses letzte Integral galt der für Anwendungen wichtige Faltungssatz. Wir beweisen den analogen Satz für das Fourier-Stieltjessche Integral. Es seien $g_1(x)$ und $g_2(x)$ zwei Funktionen mit den oben für $g(x)$ angegebenen Eigenschaften. Dann existiert das allgemeine Stieltjessche Integral

$$g_3(t) = \int_{-\infty}^{\infty} g_2(t-x) dg_1(x). \quad (149)$$

Man sieht leicht, daß $g_3(x)$ die oben angegebenen Eigenschaften besitzt, und wenn $g_1(x)$ und $g_2(x)$ stetig sind, ist auch $g_3(x)$ stetig.

Wir bilden für die Funktionen $g_k(x)$ die Fourier-Stieltjessche Transformation

$$\varphi_k(t) = \int_{-\infty}^{\infty} e^{itz} dg_k(x). \quad (150)$$

Unter dem Faltungssatz versteht man nun die Behauptung, daß *diese transformierten Funktionen der einfachen Beziehung*

$$\psi_3(t) = \varphi_1(t) \psi_2(t) \quad (151)$$

genügen.

Wenn wir auf die Funktion $\psi_3(t)$ die Gleichung (148) anwenden, erhalten wir

$$\frac{1 - e^{-iut}}{it} \psi_3(t) = \int_{-\infty}^{\infty} [g_3(z+u) - g_3(z)] e^{itz} dz.$$

Durch Einsetzen des Ausdrucks (149) für $g_3(t)$ folgt

$$\frac{1 - e^{-iut}}{it} \psi_3(t) = \int_{-\infty}^{\infty} \left\{ \int_{-\infty}^{\infty} [g_2(z+u-x) - g_2(z-x)] dg_1(x) \right\} e^{itz} dz.$$

Wir vertauschen nun die Reihenfolge der Integration. Daß diese Umformung erlaubt ist, folgt aus einem allgemeinen Satz über die Vertauschung der Reihenfolge der Integration, den wir später beweisen werden.

Durch die Vertauschung ergibt sich

$$\frac{1 - e^{-iut}}{it} \psi_3(t) = \int_{-\infty}^{\infty} \left\{ \int_{-\infty}^{\infty} [g_2(z+u-x) - g_2(z-x)] e^{itz} dz \right\} dg_1(x).$$

In dem Integral in der geschweiften Klammer führen wir an Stelle von z die neue Integrationsveränderliche $y = z - x$ ein. Dann bekommt die letzte Formel die Gestalt

$$\frac{1 - e^{-iut}}{it} \psi_3(t) = \int_{-\infty}^{\infty} \left\{ \int_{-\infty}^{\infty} [g_2(y+u) - g_2(y)] e^{ity} dy \right\} e^{itx} dg_1(x).$$

Das Integral in der geschweiften Klammer läßt sich nach Gleichung (148) durch $\psi_2(t)$ ausdrücken. Das führt zu der Formel

$$\frac{1 - e^{-iut}}{it} \psi_3(t) = \frac{1 - e^{-iut}}{it} \psi_2(t) \int_{-\infty}^{\infty} e^{itx} dg_1(x),$$

die wegen (150) mit (151) übereinstimmt.

Aus dem Faltungssatz folgt unmittelbar: Wenn die Funktionen $\psi_1(t)$ und $\psi_2(t)$ als Integrale vom Typ (131) darstellbar sind, d. h., wenn sie zur Klasse P der positiv definiten Funktionen gehören, dann gilt dasselbe auch von ihrem Produkt. Es ist klar, daß dann auch die Linearkombination $c_1\psi_1(t) + c_2\psi_2(t)$ mit positiven Koeffizienten zur Klasse P gehört. Weiterhin gehört mit $\psi(t)$ auch $\overline{\psi(t)}$ zur Klasse P ; denn wenn $\psi(t)$ durch $g_0(x)$ definiert wird, gilt für $\overline{\psi(t)}$ mit $g_1(x) = -g_0(-x)$

$$\overline{\psi(t)} = \int_{-\infty}^{\infty} e^{-itx} dg_0(x) = \int_{-\infty}^{\infty} e^{itx} dg_1(x).$$

Folglich gehört auch $|\psi(t)|^2$ zu P und wird durch die Funktion

$$h(x) = \int_{-\infty}^{\infty} g_0(x-t) dg_1(t)$$

definiert. Ist $g_0(x)$ stetig, so ist auch $h(x)$ stetig, und wegen des in [26] Bewiesenen gilt $M\{|\psi(t)|^2\} = 0$. Wir kehren nun zu der durch (137) definierten Funktion $\varphi_1(t)$ zurück; die a_k seien wie oben die den Werten $x = x_k$ entsprechenden verallgemeinerten Fourierkoeffizienten. Wenn wir beachten, daß die positiven Zahlen a_k eine konvergente Reihe

bilden, und die Formel (137) berücksichtigen, erhalten wir

$$\sum_k |\alpha_k|^2 = M\{|\varphi_1(t)|^2\}. \quad (152)$$

Andererseits gilt nach dem oben Bewiesenen

$$M\{|\varphi_2(t)|^2\} = 0. \quad (153)$$

Ferner haben wir

$$|\varphi(t)|^2 = |\varphi_1(t) + \varphi_2(t)|^2 = |\varphi_1(t)|^2 + |\varphi_2(t)|^2 + \varphi_1(t) \overline{\varphi_2(t)} + \overline{\varphi_1(t)} \varphi_2(t).$$

Nach der Schwarzschen Ungleichung¹⁾ ist

$$\left| \frac{1}{\omega} \int_{-\omega}^{\omega} \varphi_1(t) \overline{\varphi_2(t)} dt \right|^2 \leq \frac{1}{\omega} \int_{-\omega}^{\omega} |\varphi_1(t)|^2 dt \cdot \frac{1}{\omega} \int_{-\omega}^{\omega} |\varphi_2(t)|^2 dt;$$

unter Berücksichtigung der Formeln (152) und (153) ergibt sich, daß die rechte Seite dieser Ungleichung bei unbegrenzt wachsendem ω gegen 0 strebt. Also gilt $M\{\varphi_1(t) \overline{\varphi_2(t)}\} = 0$. In der gleichen Weise folgt $M\{\varphi_1(t) \varphi_2(t)\} = 0$. Aus den Formeln (152) und (153) erhalten wir also für eine beliebige Funktion der Klasse P die Vollständigkeitsrelation

$$\sum_k \alpha_k^2 = M\{|\varphi|^2\}. \quad (154)$$

29. Das Cauchy-Stieltjessche Integral. Wir betrachten das Cauchysche Integral längs der gesamten reellen Achse,

$$\omega(z) = \int_{-\infty}^{\infty} \frac{\psi(x)}{x-z} dx. \quad (155)$$

Unter bestimmten Voraussetzungen über $\psi(x)$ existiert dieses Integral, und die Funktion $\omega(z)$ der komplexen Veränderlichen z ist eine reguläre Funktion sowohl in der oberen als auch in der unteren Halbebene. In diesen Halbebenen sind die angeführten regulären Funktionen voneinander verschiedene analytische Funktionen. Wir wissen, daß $\psi(x)$ durch den Sprung von $\omega(z)$ auf der reellen Achse ausgedrückt werden kann, und zwar gilt die Formel [IV/1, 67]

$$\psi(x) = \lim_{\tau \rightarrow +0} \frac{1}{2\pi i} [\omega(x + \tau i) - \omega(x - \tau i)].$$

Wir nehmen an, daß $g(x)$ in dem unendlichen Intervall $[-\infty, \infty]$ eine Funktion von beschränkter Variation ist, und bilden für komplexe Werte von z das Cauchy-Stieltjessche Integral

$$\omega(z) = \int_{-\infty}^{\infty} \frac{1}{x-z} dg(x). \quad (156)$$

Die Funktion $\frac{1}{x-z}$ unter dem Integralzeichen ist auf der ganzen reellen Achse stetig und strebt für $x \rightarrow \pm \infty$ gegen 0. Also können wir das Integral (156) als Stieltjessches Integral [4] auffassen. Wir beweisen nun für das Integral (156) die Umkehrformel

$$g(x) - g(0) = \lim_{\tau \rightarrow +0} \frac{1}{2\pi i} \int_0^x [\omega(\sigma + \tau i) - \omega(\sigma - \tau i)] d\sigma, \quad (157)$$

¹⁾ Im Original Bunjakowski-Schwarzsche Ungleichung (Anm. d. Red.).

wobei für die Unstetigkeitsstellen der Funktion $g(x)$ auf der linken Seite die Hälfte der Summe der Grenzwerte von links und von rechts zu nehmen ist. Diese Formel hätten wir genau in der gleichen Weise wie oben bei der Umkehrung der Fourier-Stieltjesschen Integrale heuristisch ansetzen können.

Bevor wir die Gleichung (157) beweisen, betrachten wir das Poissonsche Integral für den Fall der Halbebene. Wir setzen $z = \sigma + \tau i$ und zerlegen den Cauchyschen Kern in Real- und Imaginärteil:

$$\frac{1}{x-z} = \frac{x-\sigma}{(x-\sigma)^2 + \tau^2} + \frac{\tau}{(x-\sigma)^2 + \tau^2} i.$$

Wenn wir im Integral (155) den Imaginärteil betrachten und den Faktor $1/\pi$ hinzufügen, gelangen wir zum Poissonschen Integral für die Halbebene:

$$F(\sigma, \tau) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{\tau}{(x-\sigma)^2 + \tau^2} \psi(x) dx. \quad (158)$$

Es ist offenbar sowohl in der oberen als auch in der unteren Halbebene eine harmonische Funktion.

Wir beweisen über dieses Integral folgendes: *Ist $\psi(x)$ eine in jedem endlichen Intervall beschränkte und im Riemannschen Sinne integrierbare Funktion (z. B. eine Funktion $g(x)$ von beschränkter Variation), so strebt das (offenbar existierende) Integral (158), wenn τ von der Seite positiver Werte gegen 0 strebt, in den Stetigkeitsstellen von $\psi(x)$ gegen $\psi(\sigma)$ und in den Unstetigkeitsstellen erster Art gegen die Hälfte der Summe der Grenzwerte von links und von rechts. Dabei ist die Konvergenz in jedem im Innern eines Stetigkeitsintervalls von $\psi(x)$ liegenden abgeschlossenen Intervall der Veränderlichen τ gleichmäßig bezüglich σ .* Zum Beweis erinnern wir zunächst an die offenbar geltende Beziehung

$$\frac{2}{\pi} \int_0^{\infty} \frac{\tau}{x^2 + \tau^2} dx = 1. \quad (159)$$

Im Integral (158) führen wir an Stelle von x die neue Integrationsveränderliche $y = x - \sigma$ ein:

$$F(\sigma, \tau) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{\tau}{y^2 + \tau^2} \psi(y + \sigma) dy.$$

Wir zerlegen das Integrationsintervall in die beiden Intervalle $[-\infty, 0]$ und $[0, \infty]$ und führen im ersten der entstehenden Integrale die neue Integrationsveränderliche $y_1 = -y$ ein. Auf diese Weise erhalten wir

$$F(\sigma, \tau) = \frac{2}{\pi} \int_0^{\infty} \frac{\tau}{x^2 + \tau^2} \frac{\psi(\sigma + x) + \psi(\sigma - x)}{2} dx. \quad (160)$$

Wir nehmen an, daß σ ein Stetigkeitspunkt oder ein Unstetigkeitspunkt erster Art von $\psi(x)$ ist. Wenn man beide Seiten der Formel (159) mit $\frac{1}{2} [\psi(\sigma + 0) + \psi(\sigma - 0)]$ multipliziert und das Ergebnis von (160) subtrahiert, erhält man

$$F(\sigma, \tau) - \frac{\psi(\sigma + 0) + \psi(\sigma - 0)}{2} = \frac{2}{\pi} \int_0^{\infty} \frac{\tau}{x^2 + \tau^2} w(x) dx \quad (161)$$

mit

$$w(x) = \frac{\psi(\sigma+x) - \psi(\sigma+0)}{2} + \frac{\psi(\sigma-x) - \psi(\sigma-0)}{2}. \quad (162)$$

Es sei ε eine gegebene positive Zahl. Dann existiert eine solche positive Zahl η , daß für $0 < x \leq \eta$ die Abschätzung $|w(x)| < \varepsilon$ gilt. Wenn sich σ in einem abgeschlossenen Intervall befindet, das im Innern eines Stetigkeitsintervalls von $\psi(x)$ liegt, dann wird die Zahl η wegen der gleichmäßigen Stetigkeit von $\psi(x)$ nur durch die Zahl ε bestimmt und hängt nicht von σ ab. Im Integral (161) zerlegen wir das Integrationsintervall in die beiden Intervalle $[0, \eta]$ und $[\eta, \infty]$. Für das erste Integrationsintervall gilt

$$\left| \frac{2}{\pi} \int_0^\eta \frac{\tau}{x^2 + \tau^2} w(x) dx \right| \leq \varepsilon \cdot \frac{2}{\pi} \int_0^\eta \frac{\tau}{x^2 + \tau^2} dx < \varepsilon \cdot \frac{2}{\pi} \int_0^\infty \frac{\tau}{x^2 + \tau^2} dx = \varepsilon.$$

Für die Abschätzung des Integrals bezüglich des zweiten Intervalls erwähnen wir, daß die Funktion $w(x)$ beschränkt ist, was aus der Beschränktheit der Funktion $\psi(x)$ folgt; es gilt also $|w(x)| \leq L$, wobei L eine gewisse positive Zahl ist. Wir erhalten damit

$$\left| \frac{2}{\pi} \int_\eta^\infty \frac{\tau}{x^2 + \tau^2} w(x) dx \right| \leq L \frac{2}{\pi} \int_\eta^\infty \frac{\tau}{x^2 + \tau^2} dx = \frac{2L}{\pi} \left(\frac{\pi}{2} - \arctan \frac{\eta}{\tau} \right),$$

und für die Differenz auf der linken Seite der Formel (161) ergibt sich die Abschätzung

$$\left| F(\sigma, \tau) - \frac{\psi(\sigma+0) + \psi(\sigma-0)}{2} \right| \leq \varepsilon + \frac{2L}{\pi} \left(\frac{\pi}{2} - \arctan \frac{\eta}{\tau} \right).$$

Die Differenz im zweiten Summanden strebt offensichtlich gegen 0, wenn die positive Zahl τ gegen 0 strebt; für alle genügend nahe bei 0 liegenden τ ist also dieser zweite Summand kleiner als ε , und wir erhalten für diese τ die Abschätzung

$$\left| F(\sigma, \tau) - \frac{\psi(\sigma+0) + \psi(\sigma-0)}{2} \right| \leq 2\varepsilon,$$

aus der wegen der willkürlichen Wahl von ε die Richtigkeit der obigen Behauptung folgt. Daß τ beliebig nahe bei 0 liegt, wurde durch den Wert von η bedingt; dabei hängt η für die oben erwähnten Stetigkeitsintervalle von $\psi(x)$ nicht von σ ab. Daraus folgt die gleichmäßige Konvergenz gegen einen Grenzwert in diesen Stetigkeitsintervallen.

Wir beweisen jetzt die Umkehrformel (157). Wir bilden die Funktion

$$F_1(\sigma, \tau) = \frac{1}{2\pi i} [\omega(\sigma + \tau i) - \omega(\sigma - \tau i)] = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{\tau}{(x - \sigma)^2 + \tau^2} dg(x).$$

Durch partielle Integration ergibt sich

$$F_1(\sigma, \tau) = -\frac{1}{\pi} \int_{-\infty}^{\infty} g(x) \frac{\partial}{\partial x} \left(\frac{\tau}{(x - \sigma)^2 + \tau^2} \right) dx = \frac{1}{\pi} \int_{-\infty}^{\infty} g(x) \frac{\partial}{\partial \sigma} \left(\frac{\tau}{(x - \sigma)^2 + \tau^2} \right) dx.$$

Wegen der Beschränktheit von $g(x)$ konvergiert dieses uneigentliche Integral gleichmäßig bezüglich σ , das in einem beliebigen endlichen Intervall liegt. Integriert man beide Seiten der letzten Formel im Intervall $[0, x_0]$ nach σ und führt rechts die

Integration unter dem Integralzeichen durch, so erhält man

$$\frac{1}{2\pi i} \int_0^{x_0} [\omega(\sigma + \tau i) - \omega(\sigma - \tau i)] d\sigma = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{\tau}{(x - x_0)^2 + \tau^2} g(x) dx - \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{\tau}{x^2 + \tau^2} g(x) dx.$$

Die rechts stehenden Integrale sind Poissonsche Integrale, und wenn man den oben bewiesenen Satz anwendet, erhält man die Umkehrformel (157). Diese wurde zuerst von STIELTJES angegeben und wird meist *Stieltjessche Formel* genannt. Es sei noch erwähnt, daß für das Integral (156) die Werte der Funktion $g(x)$ in den Randpunkten des Intervalls $[-\infty, \infty]$ unwesentlich sind, weil die Funktion unter dem Integralzeichen für $x \rightarrow \pm \infty$ gegen 0 strebt.

II. Mengenfunktionen und das Lebesguesche Integral

§ 1. Mengenfunktionen und Maßtheorie

30. Mengenoperationen. Wir kommen jetzt zu einem allgemeinen Integralbegriff, bei dem das Integrationsintervall nicht in einzelne Teilintervalle, sondern in Punktmengen allgemeineren Typs zerlegt wird. Auch das Integrationsgebiet wird oft kein Intervall sein, sondern eine Punktmenge allgemeineren Typs. In § 1 werden wir uns gerade mit derartigen Mengen beschäftigen sowie mit Funktionen, die auf solchen Mengen allgemeineren Typs definiert sind. Zuerst behandeln wir die grundlegenden Begriffe und Sätze, die nicht nur Punktmengen, sondern auch Mengen mit beliebigen Elementen betreffen. Für solche allgemeinen Mengen führen wir einige Grundbegriffe und Bezeichnungen ein, die wir dann im folgenden vielfach benutzen werden. Hauptsächlich werden wir mit Punktmengen rechnen, d. h. mit Mengen, deren Elemente Punkte auf der Geraden, auf der Ebene oder in dem mehrdimensionalen Raum sind.

Gehört ein Element x zu einer Menge A , so schreibt man dafür $x \in A$. Gehört x nicht zur Menge A , so schreibt man $x \notin A$. Wenn alle Elemente, die in der Menge A liegen, auch in der Menge B liegen, sagt man, A ist eine *Teil-* oder *Untermenge* von B und schreibt $A \subset B$ oder $B \supset A$. Wenn die Mengen A und B ein und dieselben Elemente enthalten, schreibt man $A = B$. Wenn alle Elemente, die in A liegen, auch in B liegen, aber unter den Elementen von B solche sind, die nicht in A liegen, dann sagt man, um dies besonders hervorzuheben, A ist eine *eigentliche* oder *echte Untermenge* von B . Gelten $A \subset B$ und $B \subset C$, so folgt hieraus $A \subset C$. Es seien

$$A_1, A_2, A_3, \dots \quad (1)$$

endlich oder abzählbar unendlich viele Mengen. Unter der *Summe*¹⁾ der Mengen (1) versteht man die Menge $\mathcal{E}_1$ der Elemente, die in wenigstens einer der Mengen A_n enthalten sind. Die Summe von Mengen bezeichnet man gewöhnlich mit

$$\mathcal{E}_1 = A_1 + A_2 + \dots \quad \text{oder} \quad \mathcal{E}_1 = \sum_n A_n.$$

Unter dem *Produkt*²⁾ der Mengen (1) versteht man die Menge $\mathcal{E}_2$ der Elemente, die in allen Mengen A_n enthalten sind. Das Produkt von Mengen bezeichnet man gewöhnlich mit

$$\mathcal{E}_2 = A_1 A_2 \dots \quad \text{oder} \quad \mathcal{E}_2 = \prod_n A_n.$$

¹⁾ Die Menge $\mathcal{E}_1$ wird in der Literatur oft auch *Vereinigung(-smenge)* genannt. Man schreibt auch $\mathcal{E}_1 = A_1 \cup A_2 \cup \dots$ oder $\mathcal{E}_1 = \bigcup_n A_n$ (Anm. d. Red.).

²⁾ In der Literatur nennt man die Menge $\mathcal{E}_2$ auch den *Durchschnitt* der Mengen (1) und schreibt $\mathcal{E}_2 = A_1 \cap A_2 \cap \dots$ oder $\mathcal{E}_2 = \bigcap_n A_n$ (Anm. d. Red.).

Das Produkt von Mengen braucht kein einziges Element zu enthalten. Die Menge, die kein Element enthält, nennt man die *leere Menge*, und wir bezeichnen sie mit dem Symbol $\emptyset$. Wenn also die Mengen A und B kein Element gemeinsam haben¹⁾, ist $AB = \emptyset$. Für die Summe und das Produkt von Mengen gilt offenbar das Kommutativ- und das Assoziativgesetz. Es ist

$$\left. \begin{aligned} A + B &= B + A, & AB &= BA, \\ A + (B + D) &= (A + B) + D, & A(BD) &= (AB)D. \end{aligned} \right\} \quad (2)$$

Es gilt auch das Distributivgesetz, d. h., es ist

$$B \sum_n A_n = \sum_n BA_n. \quad (3)$$

Diese Formel werden wir jetzt beweisen. Dazu müssen wir zeigen, daß jedes Element x aus der Menge, die auf der linken Seite der Formel (3) steht, auch in der rechtsstehenden Menge enthalten ist und umgekehrt. Ist x ein Element der linksstehenden Menge, so gehört x gleichzeitig zu B und zur Summe der Mengen A_n , d. h. wenigstens zu einer der Mengen A_n , etwa $x \in A_k$. Es ist also $x \in B$ und $x \in A_k$ und somit $x \in BA_k$, und daraus folgt, daß x auch in der Menge liegt, die auf der rechten Seite der Formel (3) steht. Liegt umgekehrt x in dieser Menge, so gehört es wenigstens zu einer der Produktmengen BA_n , also etwa $x \in BA_k$; somit ist $x \in B$ und $x \in A_k$. Hieraus folgt, daß $x \in B$ ist und daß x zu der Summe der Mengen A_n gehört, d. h., x gehört zu der Menge, die auf der linken Seite der Formel (3) steht. Damit ist die Formel (3) bewiesen.

Für $B \subset A$ ist offensichtlich $A + B = A$. Hieraus und aus (2) und (3) folgt unmittelbar

$$(A + B)(A + D) = A + BD. \quad (4)$$

Wir definieren jetzt die Differenz. Unter der *Differenz* $A - B$ von Mengen versteht man die Menge aller derjenigen Elemente der Menge A , die nicht Elemente der Menge B sind. Für $A \subset B$ ist $A - B$ die leere Menge.

Bemerkung. Bei der Definition der Differenz $A - B$ wurde nicht $B \subset A$ vorausgesetzt.

Ist $B \subset A$, so gilt offenbar

$$A = B + (A - B).$$

Allgemein gilt

$$A + B = B + (A - B). \quad (5)$$

Wir führen nun einige Formeln auf, die im folgenden nützlich sind. Sie sind leicht zu beweisen. Ist $A - B = \mathcal{E}_2$ und $B - A = \mathcal{E}_1$, so ist $A + \mathcal{E}_1 = B + \mathcal{E}_2$. Für $A_n \subset B_n$ gilt

$$\sum_n A_n \subset \sum_n B_n$$

und

$$\sum_n B_n - \sum_n A_n \subset \sum_n (B_n - A_n). \quad (6)$$

Wir geben noch einige Formeln an, die mit dem Begriff der Differenz zusammenhängen:

$$A - (B - D) \subset (A - B) + D, \quad (7)$$

¹⁾ Hierfür sagt man auch, daß die Mengen A und B *durchschnittsfremd*, *disjunkt* oder *elementfremd* sind (Anm. d. Red.).

$$AB = A - (A - B) ; \quad (7_1)$$

$$(A_1 - A_2) - (B_1 - B_2) \subset (A_1 - B_1) + (B_2 - A_2) , \quad (8)$$

$$A + B = (A - B) + (B - A) + AB . \quad (8_1)$$

Wir definieren jetzt die Begriffe der monotonen Mengenfolge und des Grenzwertes. Gilt für eine unendliche Mengenfolge $\mathfrak{E}_1, \mathfrak{E}_2, \dots$

$$\mathfrak{E}_1 \subset \mathfrak{E}_2 \subset \mathfrak{E}_3 \subset \dots , \quad (9)$$

so sagt man, die Mengenfolge sei *wachsend*. Eine *fallende* Mengenfolge wird durch die Bedingung

$$\mathfrak{E}_1 \supset \mathfrak{E}_2 \supset \mathfrak{E}_3 \supset \dots \quad (10)$$

definiert.

Unter dem *Grenzwert einer wachsenden Folge* von Mengen $\mathfrak{E}_n$ versteht man die Menge $\mathfrak{E}$ der Elemente, die in wenigstens einer der Mengen $\mathfrak{E}_n$ enthalten sind. Dafür schreibt man

$$\mathfrak{E} = \lim_{n \rightarrow \infty} \mathfrak{E}_n .$$

Bemerkung. Wenn im Fall der Bedingung (9) das Element x in einer Menge $\mathfrak{E}_k$ liegt, dann liegt es auch in allen Mengen $\mathfrak{E}_n$ mit $n > k$. Für eine wachsende Mengenfolge gilt offensichtlich

$$\mathfrak{E} = \lim_{n \rightarrow \infty} \mathfrak{E}_n = \sum_{n=1}^{\infty} \mathfrak{E}_n \quad (11)$$

oder – in anderer Schreibweise

$$\mathfrak{E} = \lim_{n \rightarrow \infty} \mathfrak{E}_n = \mathfrak{E}_1 + \sum_{k=1}^{\infty} (\mathfrak{E}_{k+1} - \mathfrak{E}_k) . \quad (12)$$

In der letzten Formel sind die Mengen, die auf der rechten Seite als Summanden auftreten, paarweise elementfremd. Unter dem *Grenzwert einer fallenden Folge* von Mengen $\mathfrak{E}_n$ versteht man die Menge $\mathfrak{E}$ der Elemente, die in allen $\mathfrak{E}_n$ enthalten sind. Dann schreibt man

$$\mathfrak{E} = \lim_{n \rightarrow \infty} \mathfrak{E}_n = \prod_{n=1}^{\infty} \mathfrak{E}_n \quad (13)$$

oder auch

$$\mathfrak{E}_1 = \mathfrak{E} + \sum_{k=1}^{\infty} (\mathfrak{E}_k - \mathfrak{E}_{k+1}) . \quad (14)$$

In dieser Formel sind die rechtsstehenden Summanden wieder paarweise elementfremd. Wir haben den Begriff des Grenzwertes nur für monotone Mengenfolgen definiert. Man kann diesen jedoch auch allgemein definieren. Wir werden den Grenzwertbegriff hier nicht verallgemeinern, da wir ihn im folgenden nicht benötigen.

Jetzt führen wir noch einen Begriff speziell für Punktmengen ein. Es sei $\mathfrak{E}$ eine Punktmenge auf der Ebene. Man nennt die Menge, die aus allen Punkten der Ebene besteht, die nicht zu $\mathfrak{E}$ gehören, *Komplementärmenge* oder *Komplement* von $\mathfrak{E}$ und bezeichnet sie mit $\mathcal{C}\mathfrak{E}$. Wir betrachten jetzt einige Formeln für Komplementärmengen. Wenn man von einer Komplementärmenge wieder das Komplement bildet, erhält man die ursprüngliche Menge, $\mathcal{C}\mathcal{C}\mathfrak{E} = \mathfrak{E}$. Für $\mathfrak{E}_1 \subset \mathfrak{E}_2$ gilt $\mathcal{C}\mathfrak{E}_1 \supset \mathcal{C}\mathfrak{E}_2$.

Folgende Formeln kann man leicht beweisen:

$$\prod_n \mathcal{E}_n = \mathcal{C} \sum_n \mathcal{C}\mathcal{E}_n, \quad (15)$$

$$\sum_n \mathcal{E}_n = \mathcal{C} \prod_n \mathcal{C}\mathcal{E}_n, \quad (16)$$

$$\mathcal{C} \prod_n \mathcal{E}_n = \sum_n \mathcal{C}\mathcal{E}_n, \quad (17)$$

$$\mathcal{C} \sum_n \mathcal{E}_n = \prod_n \mathcal{C}\mathcal{E}_n, \quad (18)$$

$$\mathcal{C}\mathcal{E}_1 - \mathcal{C}\mathcal{E}_2 = \mathcal{E}_2 - \mathcal{E}_1, \quad (19)$$

$$\mathcal{E}_1 - \mathcal{E}_2 = \mathcal{E}_1 \cdot \mathcal{C}\mathcal{E}_2. \quad (20)$$

Der Begriff der Komplementärmenge kann auch in bezug auf eine Gerade oder einen beliebigen mehrdimensionalen Raum eingeführt werden, wenn man diese Gebilde als Grundmengen ansieht. Oft führt man den Begriff der Komplementärmenge in bezug auf eine beliebige Menge A ein: Wenn alle Punkte der Menge $\mathcal{E}$ in A liegen, dann ist die Komplementärmenge von $\mathcal{E}$ in bezug auf A die Differenz $A - \mathcal{E}$. Wir werden den Begriff der Komplementärmenge nur in bezug auf den ganzen Raum, d. h. nur in bezug auf eine Gerade oder auf eine Ebene usw., benutzen.

31. Punktmengen. Wir behandeln in diesem Abschnitt einige Begriffe und Resultate, die speziell die Punktmengen betreffen.

In der Theorie der mehrfachen Riemannschen Integrale (Teil II) führten wir Überlegungen durch, welche Punktmengen auf der Ebene oder im beliebigen n -dimensionalen Raum betreffen. Wir werden das in Teil II Gesagte nochmals mit wesentlichen Ergänzungen wiederholen. Wir werden Punktmengen auf der x, y -Ebene betrachten. Man kann das dann leicht für den Fall einer Geraden oder eines beliebigen n -dimensionalen Raumes verallgemeinern.

Wir betrachten eine Ebene im kartesischen x, y -Koordinatensystem und auf dieser Ebene eine Punktmenge A . Die Menge A heißt *beschränkt*, wenn der Abstand aller Punkte von A vom Punkt $(0, 0)$ kleiner als eine positive Zahl N ist, d. h., wenn $x^2 + y^2 < N^2$ für alle Punkte der Menge A gilt. Man nennt den abgeschlossenen Kreis mit dem Mittelpunkt P und dem Radius ε eine ε -Umgebung des Punktes $P(a, b)$ d. h., die Menge der Punkte (x, y) , die der Bedingung

$$(x - a)^2 + (y - b)^2 \leq \varepsilon^2$$

genügen, heißt ε Umgebung. Ein Punkt P heißt *Häufungspunkt für die Menge* $\mathcal{E}$, wenn in jeder Umgebung des Punktes P unendlich viele Punkte von $\mathcal{E}$ liegen. Der Punkt P selbst kann entweder zu $\mathcal{E}$ gehören oder nicht. Wenn alle Häufungspunkte für eine Menge $\mathcal{E}$ zu $\mathcal{E}$ gehören, heißt $\mathcal{E}$ *abgeschlossen*. Ein Punkt P , der zur Menge $\mathcal{E}$ gehört, heißt ein *innerer Punkt* von $\mathcal{E}$, wenn alle Punkte einer ε -Umgebung des Punktes P in $\mathcal{E}$ liegen. Eine Menge $\mathcal{E}$ heißt *offen*, wenn sie nur aus inneren Punkten besteht. Abgeschlossene Mengen werden gewöhnlich mit dem Buchstaben F und offene Mengen mit dem Buchstaben O bezeichnet.¹⁾ Eine Menge, die keinen Punkt enthält, heißt *leere Menge*. In den folgenden Sätzen können die Mengen, von denen die Rede ist, auch leer sein; die leere Menge ist als gleichzeitig offen und abgeschlossen anzusehen. Unter dem *Rand* einer offenen Menge O versteht man die Menge $\mathcal{L}$

¹⁾ Vom französischen Wort „fermé“, d. h. „abgeschlossen“, bzw. „ouvert“, d. h. „offen“.

der Punkte P , die folgende Eigenschaften besitzen: Der Punkt P liegt nicht in O , aber in jeder ε -Umgebung von P liegen Punkte, die zu O gehören. Da O aus inneren Punkten besteht, kann man behaupten, daß in jeder ε -Umgebung von P unendlich viele Punkte aus O liegen. Daher kann man den Rand l einer offenen Menge O als die Menge der Häufungspunkte von O , die nicht in O liegen, definieren. Es ist leicht zu zeigen, daß *der Rand einer offenen Menge abgeschlossen ist* (vgl. [II, 91]).

Es sei $\mathcal{E}$ eine beliebige Menge. Nimmt man zu einer Menge alle ihre Häufungspunkte hinzu, so erhält man eine Menge, die mit $\bar{\mathcal{E}}$ bezeichnet wird. $\bar{\mathcal{E}}$ heißt *abgeschlossene Hülle* von $\mathcal{E}$. Ist $\mathcal{E}$ eine abgeschlossene Menge, so ist $\bar{\mathcal{E}} = \mathcal{E}$.¹⁾ Wir zeigen jetzt, daß $\bar{\mathcal{E}}$ eine abgeschlossene Menge ist. Es sei P ein Häufungspunkt für $\bar{\mathcal{E}}$, d. h., es gibt eine unendliche Folge verschiedener Punkte P_n ($n=1, 2, \dots$) aus $\bar{\mathcal{E}}$ mit $P_n \rightarrow P$. Wenn von den P_n unendlich viele Punkte zu $\mathcal{E}$ gehören, ist P Häufungspunkt für $\mathcal{E}$ und gehört daher (infolge der Hüllenbildung) zu $\bar{\mathcal{E}}$. Wir nehmen jetzt an, daß alle Punkte der Folge P_n von einem gewissen n ab nicht in $\mathcal{E}$ liegen. Nach Voraussetzung gehören sie zu $\bar{\mathcal{E}}$, und folglich sind sie Häufungspunkte für $\mathcal{E}$. In jeder $\frac{\varepsilon}{2}$ -Umgebung des Punktes P befinden sich unendlich viele Punkte P_n , und in jeder $\frac{\varepsilon}{2}$ -Umgebung jedes Punktes P_n sind unendlich viele Punkte von $\mathcal{E}$ enthalten. Hieraus folgt unmittelbar, daß in jeder ε -Umgebung des Punktes P unendlich viele Punkte aus $\mathcal{E}$ liegen, d. h., P ist Häufungspunkt für $\mathcal{E}$, gehört also zu $\bar{\mathcal{E}}$. Somit haben wir gezeigt, daß *die Menge $\bar{\mathcal{E}}$, die man als abgeschlossene Hülle einer beliebigen gegebenen Menge $\mathcal{E}$ erhält, sicher abgeschlossen ist*.

Bemerkung. *Die ganze Ebene ist gleichzeitig abgeschlossen und offen.* Der unendlich ferne Punkt wird nicht zur Ebene gerechnet. *Jede endliche Menge von Punkten ist abgeschlossen*; denn sie besitzt keinen Häufungspunkt.

Wir führen jetzt den Begriff des Abstandes zwischen zwei Mengen ein. Unter dem *Abstand* zweier Mengen $\mathcal{E}_1$ und $\mathcal{E}_2$ versteht man die untere Grenze der Abstände aller Punkte von $\mathcal{E}_1$ zu den Punkten von $\mathcal{E}_2$. Wenn die Mengen einen gemeinsamen Punkt haben, ist der Abstand zwischen ihnen gleich 0. Jedoch kann der Abstand zwischen zwei Mengen auch gleich 0 sein, wenn die Mengen keinen gemeinsamen Punkt haben. Es kann nämlich unter den Punkten zweier Mengen, die keinen gemeinsamen Punkt haben, beliebig benachbarte Punkte geben. Das ist nicht möglich, wenn die beiden gegebenen Mengen beschränkt und abgeschlossen sind. In Teil II wurde folgender Satz bewiesen: *Sind $\mathcal{E}_1$ und $\mathcal{E}_2$ beschränkte und abgeschlossene Mengen, so ist der Abstand d zwischen ihnen positiv, und man findet wenigstens ein solches Paar von Punkten P aus $\mathcal{E}_1$ und Q aus $\mathcal{E}_2$, daß*

$$PQ = d$$

gilt. Aus dem Beweis dieses Satzes folgt unmittelbar, daß *er auch gilt, wenn nur eine der gegebenen abgeschlossenen Mengen beschränkt ist. Insbesondere ist der Abstand jedes gegebenen Punktes einer offenen Menge vom Rand dieser Menge positiv.*

32. Eigenschaften abgeschlossener und offener Mengen. Wir zeigen jetzt einige spezielle Eigenschaften der abgeschlossenen und der offenen Mengen.

¹⁾ Eine Menge, die keinen Häufungspunkt besitzt, heißt *diskret* (Ann. d. Red.).

Satz 1. *Die Summe endlich oder abzählbar unendlich vieler offener Mengen ist wieder eine offene Menge. Das Produkt endlich vieler offener Mengen ist eine offene Menge.*

Beweis. Wir betrachten die Summe endlich oder abzählbar unendlich vieler offener Mengen

$$\mathcal{G} = \sum_n O_n.$$

Ist $P \in \mathcal{G}$, so gehört P zu wenigstens einem der O_n . Es sei $P \in O_k$. Da O_k eine offene Menge ist, gehört eine ε -Umgebung von P zu O_k . Diese ε -Umgebung von P gehört also auch zur Summe $\mathcal{G}$, und hieraus folgt, daß $\mathcal{G}$ eine offene Menge ist. Wir betrachten jetzt das endliche Produkt

$$\mathcal{G} = \prod_{n=1}^m O_n.$$

P gehöre zu $\mathcal{G}$. Wie oben zeigen wir, daß auch eine gewisse ε -Umgebung von P in $\mathcal{G}$ liegt. Wenn P in $\mathcal{G}$ liegt, dann liegt P in allen O_k ($k=1, 2, \dots, m$). Weil die O_k offene Mengen sind, entspricht für jedes O_k dem Punkt P eine ε_k -Umgebung, die in O_k liegt. Ist ε gleich dem Minimum der (endlich vielen!) ε_k ($k=1, 2, \dots, m$), so gehört die ε -Umgebung des Punktes P zu jedem O_k und folglich auch zu $\mathcal{G}$.

Bemerkung. *Das Produkt abzählbar unendlich vieler offener Mengen braucht nicht offen zu sein.*

Satz 2. *Die Menge CF ist offen, die Menge CO ist abgeschlossen.*

Beweis. Wir beweisen die erste Behauptung. P liege in CF . Wir müssen zeigen, daß eine ε -Umgebung von P in CF liegt. Würden in jeder ε -Umgebung von P Punkte aus F liegen, so würde der Punkt P , der nach Voraussetzung nicht zu F gehört, Häufungspunkt für F sein, und da F abgeschlossen ist, müßte er in F liegen. Das ist aber ein Widerspruch zur Voraussetzung.

Satz 3. *Das Produkt endlich oder abzählbar unendlich vieler abgeschlossener Mengen ist abgeschlossen. Die Summe endlich vieler abgeschlossener Mengen ist abgeschlossen.*

Beweis. Wir zeigen nur, daß die Menge $\mathcal{G} = \prod_n F_n$ abgeschlossen ist. Geht man zu den Komplementärmengen über, so kann man $C\mathcal{G} = \sum_n CF_n$ schreiben [30]. Nach Satz 2 sind die CF_n offene Mengen, und nach Satz 1 ist die Menge $C\mathcal{G}$ ebenfalls offen, also ist die Komplementärmenge $\mathcal{G}$ abgeschlossen.

Bemerkung. *Die Summe abzählbar unendlich vieler abgeschlossener Mengen braucht nicht abgeschlossen zu sein.*

Satz 4. *Die Menge $O - F$ ist offen, die Menge $F - O$ ist abgeschlossen.*

Beweis. Folgende Beziehungen sind leicht nachzuprüfen:

$$O - F = O \cdot CF, \quad F - O = F \cdot CO.$$

Hieraus und auf Grund des vorigen Satzes folgt Satz 4.

Man sagt, daß die Menge $\mathcal{G}$ von einem System M von Mengen *überdeckt* wird, wenn jeder Punkt von $\mathcal{G}$ in wenigstens einer Menge aus dem Mengensystem M liegt.

Satz 5 (BOREL).¹⁾ *Wird eine abgeschlossene beschränkte Menge F von einem unendlichen System α offener Mengen O überdeckt, so kann man aus diesem unendlichen System endlich viele offene Mengen auswählen, die auch schon F überdecken.*

Beweis. Der Beweis wird indirekt geführt. Wir nehmen an, daß endlich viele offene Mengen aus dem System α die Menge F nicht überdecken können. Da F eine beschränkte Menge ist, liegen alle Punkte von F in irgendeinem endlichen zweidimensionalen Intervall Δ_0 ($a \leq x \leq b$, $c \leq y \leq d$). Wir teilen dieses abgeschlossene Intervall Δ_0 in vier gleiche Teile, indem wir die Intervalle $[a, b]$ und $[c, d]$ halbieren. Jedes der erhaltenen vier Intervalle sei abgeschlossen. Jene Punkte von F , die in eines der vier abgeschlossenen Intervalle fallen, stellen auf Grund des Satzes 3 selbst eine abgeschlossene Menge dar, und wenigstens eine dieser abgeschlossenen Mengen kann nicht von endlich vielen offenen Mengen des Systems α überdeckt werden. Wir betrachten das der oben beschriebenen vier abgeschlossenen Intervalle, in dem dies der Fall ist. Dieses Intervall teilen wir wieder in vier Teile und verfahren so wie oben. Auf diese Weise erhalten wir ein System ineinander geschachtelter Intervalle $\Delta_0, \Delta_1, \Delta_2, \dots$, in dem jedes folgende den vierten Teil des vorhergehenden darstellt. Somit haben wir folgenden Sachverhalt: Die Menge der Punkte von F , die in Δ_k (k beliebig) liegen, kann nicht von endlich vielen offenen Mengen des Systems α überdeckt werden. Bei unbeschränkt wachsendem k ziehen sich die Intervalle Δ_k auf einen Punkt P zusammen, der in allen Intervallen Δ_k liegt. Da Δ_k für beliebiges k eine unendliche Menge von Punkten aus F enthält, ist der Punkt P Häufungspunkt für F und liegt damit auch in F , weil F abgeschlossen ist. Der Punkt P wird von einer offenen Menge O' , die zum System α gehört, überdeckt. Eine gewisse ε -Umgebung des Punktes P gehört ebenfalls zu O' . Für hinreichend große Werte k liegen die Intervalle Δ_k im Innern der ε -Umgebung von P . Diese Δ_k werden von der einen offenen Menge O' des Systems α ganz überdeckt, und das ist ein Widerspruch dazu, daß die Punkte von F , die in Δ_k liegen, für beliebiges k nicht von endlich vielen offenen Mengen des Systems α überdeckt werden können. Damit ist der Satz bewiesen.

Satz 6. *Jede offene Menge kann als Summe abzählbar unendlich vieler halboffener und paarweise punktfremder Intervalle dargestellt werden.*

Beweis. Wir hatten vereinbart, unter einem halboffenen Intervall in der Ebene ein durch die Ungleichungen $a < x \leq b$, $c < y \leq d$ gegebenes endliches Intervall zu verstehen. Wir tragen auf der Ebene ein achsenparalleles quadratisches Gitter der Seitenlänge 1 ein. Die Menge dieser Quadrate ist abzählbar. Wir wählen von diesen Quadraten jene aus, deren sämtliche Punkte zu der gegebenen offenen Menge O gehören. Die Anzahl dieser Quadrate kann endlich oder abzählbar sein, sie kann auch 0 sein. Jedes der übrigbleibenden Quadrate teilen wir in vier gleiche Quadrate ein, und von diesen erhaltenen Quadraten wählen wir wieder die, deren sämtliche Punkte in O liegen. Jedes der übrigbleibenden Quadrate teilen wir wieder in vier gleiche Teile und wählen dann wieder die Quadrate aus, deren sämtliche Punkte in O liegen usw. Wir zeigen, daß jeder Punkt P der Menge O in einem dieser Quadrate liegt, dessen sämtliche Punkte zu O gehören. Mit d bezeichnen wir den positiven Abstand von P bis zum Rand von O . Wenn wir zu Quadraten gelangen, deren Diagonale kleiner als d ist, kann man offenbar behaupten, daß der Punkt P in einem Quadrat liegt, dessen sämtliche Punkte in O liegen. Nehmen wir die ausge-

¹⁾ Dieser Satz ist auch unter dem Namen *Heine-Borel-Lebesguescher Überdeckungssatz* bekannt. — Anm. d. Red.

wählten Quadrate als halboffen an, so sind sie paarweise punktfremd, und der Satz ist damit bewiesen.

Die Anzahl der ausgewählten Quadrate muß notwendig abzählbar sein; denn eine endliche Summe von halboffenen Intervallen ist offensichtlich keine offene Menge. Bezeichnen wir mit Δ_n jene halboffenen Quadrate, die wir als Ergebnis bei der obigen Konstruktion erhalten, so können wir

$$O = \sum_{n=1}^{\infty} \Delta_n \quad (21)$$

schreiben.

Im eindimensionalen Fall, d. h. für eine gerade Linie, ist folgende Behauptung richtig: Jede offene Menge auf einer Geraden kann als Summe endlich oder abzählbar unendlich vieler offener und paarweise punktfremder Intervalle dargestellt werden. Wir werden diese Behauptung im weiteren nicht benutzen und daher jetzt auch nicht beweisen.

Alles, was in den letzten beiden Paragraphen behandelt wurde, ist auch für Punktmengen auf einer Geraden, im dreidimensionalen Raum und im beliebigen n -dimensionalen Raum gültig. Der Unterschied besteht nur in der Definition der ε -Umgebung und des Intervalls. Im dreidimensionalen Raum ist die ε -Umgebung des Punktes P eine Kugel um P mit dem Radius ε , und das Intervall ist ein rechteckiges Parallelepiped, dessen Kanten parallel zu den Achsen sind. Ein halboffenes Intervall wird dann durch die Ungleichungen $a_1 < x \leq b_1$, $a_2 < y \leq b_2$, $a_3 < z \leq b_3$ definiert. Für eine Gerade wird die ε -Umgebung des Punktes x_0 durch die Ungleichung $x_0 - \varepsilon \leq x \leq x_0 + \varepsilon$ definiert.

33. Elementarbereiche. Im folgenden werden die *endlichen halboffenen Intervalle*, die wir kurz Intervalle nennen werden, eine grundlegende Rolle spielen. Es sei eine nichtnegative, additive und normale Intervallfunktion $G(\Delta)$ gegeben. Wir wollen diese Funktion auf eine umfangreiche Klasse von Punktmengen ausdehnen; sie soll aber alle ihre oben angegebenen Eigenschaften behalten. Die Summe endlich vieler paarweise punktfremder Intervalle Δ_k ($k=1, 2, \dots, m$) wollen wir einen *Elementarbereich* nennen. Wenn wir einen solchen Elementarbereich mit R bezeichnen, können wir

$$R = \sum_{k=1}^m \Delta_k \quad (22)$$

schreiben. Schließlich können wir diese Elementarbereiche auf beliebige Weise in paarweise punktfremde Intervalle zerlegen:

$$R = \sum_{k=1}^{m'} \Delta'_k. \quad (23)$$

Es ist leicht zu sehen, daß für zwei beliebige Zerlegungen die Beziehung

$$\sum_{k=1}^m G(\Delta_k) = \sum_{k=1}^{m'} G(\Delta'_k) \quad (24)$$

gilt. Um die Richtigkeit dieser Behauptung zu beweisen, ist es auf Grund der Additivität von $G(\Delta)$ hinreichend, eine neue Zerlegung von R durchzuführen, die das Produkt der Zerlegungen (22) und (23) ist. Dabei sieht man, daß die linke und die rechte Seite der Formel (24) die Summe der Werte von $G(\Delta)$ für die Intervalle der neuen Zerlegung von R ist. Um die linke Seite der Formel (24) zu erhalten,

genügt es, die Summanden dieser letzten Summe, die solchen Teilintervallen entsprechen, die alle in ein und demselben Δ_k liegen, zusammenzufassen, und um die rechte Seite von (24) zu erhalten, muß man die Summanden zusammenfassen, die den Teilintervallen entsprechen, welche in ein und demselben Δ_k liegen. Wenn man also den Elementarbereich R irgendwie in paarweise punktfremde Teilintervalle zerlegt, ist die Summe der Werte der Funktion $G(\Delta)$ für die erhaltenen Teilintervalle ein wohldefinierter Wert, d. h., er hängt nicht von der Art der Zerlegung ab. Wir können also diese Summe als Wert der Funktion $G(R)$ für den Elementarbereich R nehmen, d. h., es ist

$$G(R) = \sum_{k=1}^m G(\Delta_k) \quad (25)$$

für jede Zerlegung von R in endlich viele punktfremde Intervalle. Somit haben wir die Definition der Funktion $G(\Delta)$ ganz einfach auf Elementarbereiche erweitert. Unter Benutzung von (21) ließe sich ganz analog $G(\Delta)$ für alle offenen Mengen definieren. Aber wir gehen jetzt einen anderen Weg. Dabei werden die offenen Mengen eine grundlegende Rolle spielen. In diesem Paragraphen behandeln wir noch einige einfache Eigenschaften der Intervalle und Elementarbereiche.

Bemerkung. Ist $R_1 \subset R_2$, so gilt $G(R_1) \leq G(R_2)$. Dies folgt unmittelbar aus der Nichtnegativität von $G(\Delta)$, wenn eine solche Zerlegung von R_2 auf dem Intervall benutzt wird, bei der die Teilintervalle, die gemeinsame Punkte mit R_1 haben, ganz zu R_1 gehören. Es seien δ_k ($k = 1, \dots, p$) Intervalle, die auch gemeinsame Punkte haben können. Ziehen wir diejenigen Geraden, auf denen die Seiten von δ_k liegen, so zerlegen wir die Summe der Intervalle δ_k in Teilintervalle, wobei diese Teilintervalle folgende Eigenschaft besitzen: Haben zwei von ihnen einen gemeinsamen Punkt, so stimmen sie überein. Fassen wir aufeinanderliegende Intervalle als ein einziges Intervall auf, so erhalten wir einen Elementarbereich R_0 , der die Summe der Intervalle darstellt:

$$R_0 = \sum_{k=1}^p \delta_k,$$

wobei

$$G(R_0) \leq \sum_{k=1}^p G(\delta_k) \quad (26)$$

ist. Das Zeichen $<$ steht dann, wenn nur für eines dieser aufeinanderliegenden Intervalle der Wert der Funktion $G(\Delta)$ positiv ist.

Wir führen jetzt einen neuen Begriff ein, der im weiteren benutzt werden wird. Es sei Δ ($a < x \leq b$, $c < y \leq d$) ein beliebiges Intervall und α eine positive Zahl. Wir nennen ein Intervall, das durch die Ungleichungen $a + \alpha < x \leq b$, $c + \alpha < y \leq d$ definiert ist, eine α -Kontraktion des Intervalls Δ und bezeichnen es mit ${}^{(\alpha)}\Delta$. Ein Intervall, das durch die Ungleichungen $a < x \leq b + \alpha$, $c < y \leq d + \alpha$ definiert ist, nennen wir eine α -Dilatation des Intervalls Δ und bezeichnen es mit $\Delta^{(\alpha)}$. Die Differenzen $\Delta - {}^{(\alpha)}\Delta = {}^{(\alpha)}R$ und $\Delta^{(\alpha)} - \Delta = R^{(\alpha)}$ sind Elementarbereiche. Infolge der Nichtnegativität von $G(\Delta)$ ist

$$G({}^{(\alpha)}R) = G(\Delta) - G({}^{(\alpha)}\Delta) \geq 0$$

und

$$G(R^{(\alpha)}) = G(\Delta^{(\alpha)}) - G(\Delta) \geq 0,$$

und daraus, daß die Funktion $G(\Delta)$ normal ist, folgt unmittelbar

$$\lim_{\alpha \rightarrow +0} G({}^{(\alpha)}\Delta) = \lim_{\alpha \rightarrow +0} G(\Delta^{(\alpha)}) = G(\Delta). \quad (26_1)$$

Wir beweisen jetzt ein Lemma, das im folgenden gebraucht wird.

Lemma. *Wird der Elementarbereich R von endlich oder abzählbar unendlich vielen Intervallen δ'_k überdeckt (diese können auch gemeinsame Punkte haben), so ist*

$$\sum_k G(\delta'_k) \cong G(R).^1 \quad (27)$$

Beweis. Die Behauptung dieses Lemmas ist anschaulich klar. Exakt beweisen können wir es, wenn wir Satz 5 aus [32] benutzen. Es sei ε eine gegebene positive Zahl. Wir zerlegen R in endlich viele paarweise punktfremde Intervalle Δ_k ($k = 1, \dots, m$) und unterwerfen jedes Teilintervall Δ_k einer α -Kontraktion, wobei die positive Zahl α so klein gehalten wird, daß die Summe der Werte von $G(\Delta)$ für die kontrahierten Intervalle höchstens gleich $G(R) - \varepsilon$ wird. Bezeichnen wir mit R_α den Elementarbereich, der gleich der Summe der kontrahierten Intervalle ist, so können wir

$$G(R_\alpha) \cong G(R) - \varepsilon \quad (28)$$

schreiben. Jedes der Intervalle δ'_k , das in der Überdeckung von R liegt, erfahre eine α_k -Dilatation, wobei die positiven Zahlen α_k so klein gewählt werden, daß

$$G(\delta_k^{(\alpha_k)}) \cong G(\delta'_k) + \frac{\varepsilon}{2k} \quad (29)$$

ist.

Die kontrahierten Intervalle, deren Summe gleich R_α ist, machen wir zu abgeschlossenen Intervallen, d. h., wir schließen jedes dieser Intervalle ab. Die Summe dieser abgeschlossenen Intervalle (dies sind endlich viele) ist wieder eine abgeschlossene Menge F , wobei $F \subset R$ gilt. Wenn wir von $\delta_k^{(\alpha_k)}$ den Rand ausschließen, d. h. zwei Seiten und eine Ecke, dann bleibt ein offenes Intervall $\delta_k^{(\alpha_k)}$ übrig. Wenn wir die Dilatation der Intervalle δ'_k in Betracht ziehen, können wir behaupten, daß die offenen Intervalle $\delta_k^{(\alpha_k)}$ die oben erwähnten abgeschlossenen Intervalle überdecken, d. h., sie überdecken die beschränkte abgeschlossene Menge F . Es sei z. B. die Anzahl der Intervalle δ'_k unendlich. Nach Satz 5 aus [32] genügen schon endlich viele Intervalle $\delta_k^{(\alpha_k)}$ ($k = 1, \dots, q$), um F zu überdecken, und diese überdecken dann auch R_α .

Die Summe der Intervalle $\delta_k^{(\alpha_k)}$ ($k = 1, \dots, q$) ist ein Elementarbereich R' , wobei $R_\alpha \subset R'$ ist; folglich gilt $G(R_\alpha) \cong G(R')$. Wegen (26) und (29) ist

$$G(R') \cong \sum_{k=1}^q G(\delta_k^{(\alpha_k)}) \cong \sum_{k=1}^q G(\delta'_k) + \varepsilon \sum_{k=1}^q \frac{1}{2k};$$

hieraus folgt unmittelbar

$$G(R_\alpha) \cong G(R') \cong \sum_{k=1}^q G(\delta'_k) + \varepsilon.$$

Also gilt auch

$$G(R_\alpha) \cong \sum_{k=1}^{\infty} G(\delta'_k) + \varepsilon.$$

¹⁾ $\sum_k$ bedeutet $\sum_{k=1}^m$, wobei m unendlich oder endlich ($m \geq 1$) sein kann.

Vergleichen wir dies mit (28), so erhalten wir

$$G(R) - \varepsilon \leq \sum_{k=1}^{\infty} G(\delta'_k) + \varepsilon, \quad \text{d. h.} \quad \sum_{k=1}^{\infty} G(\delta'_k) \geq G(R) - 2\varepsilon.$$

Die Summe hängt nicht von ε ab, und da ε beliebig gewählt wurde, erhalten wir die Ungleichung (27).

Bemerkung. Die Summanden der links stehenden Summe von (27) sind nicht-negativ und endlich, die Reihe selbst kann aber divergieren.

Im folgenden werden wir Summen betrachten, die unendlich viele nichtnegative Summanden besitzen. Ist einer der Summanden der Summe gleich ∞ , so muß die ganze Summe gleich ∞ gesetzt werden. Aber wie wir schon erwähnt haben, können auch alle Summanden endlich und die Summe trotzdem ∞ sein, d. h., die Reihe kann divergieren.

34. Das äußere Maß und seine Eigenschaften. Wenn wir die Funktionen $G(\mathcal{A})$ benutzen, können wir jetzt zu einer beliebigen Punktmenge $\mathcal{E}$ der Ebene eine nicht-negative Zahl bilden, die man das äußere Maß dieser Menge nennt.

Definition. Es seien endlich oder abzählbar unendlich viele Intervalle $\mathcal{A}_n$ ($n=1, 2, \dots$) gegeben, die die Mengen $\mathcal{E}$ überdecken. Das *äußere Maß* von $\mathcal{E}$ ist die untere Grenze der Summe

$$\sum_n G(\mathcal{A}_n) \tag{30}$$

bei allen möglichen Überdeckungen der Menge $\mathcal{E}$ durch Intervalle $\mathcal{A}_n$.

Wir werden das äußere Maß mit $|\mathcal{E}|_G$ bezeichnen, wobei sich der untere Index auf die Funktion $G(\mathcal{A})$ bezieht, die die Grundlage für die Definition des äußeren Maßes bildete. Somit gilt für jede Überdeckung

$$\sum_n G(\mathcal{A}_n) \geq |\mathcal{E}|_G, \quad |\mathcal{E}|_G = \inf \sum_n G(\mathcal{A}_n). \tag{31}$$

Wenn für eine beliebige Überdeckung die Summen (30) gleich ∞ sind, ist auch das äußere Maß gleich ∞ . Bei beschränkten Mengen ist das äußere Maß immer endlich; denn eine solche Menge kann durch ein Intervall $\mathcal{A}_0$ überdeckt werden, und nach Voraussetzung ist $G(\mathcal{A}_0)$ endlich. Wir bemerken, daß eine unbeschränkte Menge $\mathcal{E}$ nicht von endlich vielen Intervallen überdeckt wird, denn wir hatten jedes der Intervalle als endlich angenommen. Jedoch kann das äußere Maß einer unbeschränkten Menge auch endlich sein. Wir beweisen jetzt einige Sätze für das äußere Maß.

Satz 1. *Gilt $\mathcal{E}' \subset \mathcal{E}''$, so ist $|\mathcal{E}'|_G \leq |\mathcal{E}''|_G$.*

Beweis. Jede Überdeckung von $\mathcal{E}''$ ist zugleich auch Überdeckung von $\mathcal{E}'$. Daher kann die untere Grenze der Summen (30) für $\mathcal{E}'$ kleiner sein als für $\mathcal{E}''$. Auf keinen Fall kann sie größer sein als für $\mathcal{E}''$, was zu beweisen war.

Satz 2. *Für jeden Elementarbereich R ist das äußere Maß gleich $G(R)$, d. h., es ist $|R|_G = G(R)$.*

Beweis. Wenn wir R in beliebiger Weise in Teilintervalle $\mathcal{A}_k$ zerlegen, überdecken diese natürlich R . Auf Grund von (25) und der Definition des äußeren Maßes als untere Grenze der Summen (30) haben wir somit $|R|_G \leq G(R)$ bei beliebiger Zerlegung von R .

Wir beweisen jetzt die entgegengesetzte Ungleichung. Wenn die Intervalle Δ'_n den Elementarbereich R überdecken, dann ist nach dem Lemma des vorigen Paragraphen $\sum_n G(\Delta'_n) \cong G(R)$. Hieraus folgt unmittelbar $|R|_G \cong G(R)$. Aus den beiden Ungleichungen ergibt sich die Gleichheit $|R|_G = G(R)$.

Satz 3. *Für endlich oder abzählbar unendlich viele Summanden ist das äußere Maß der Summe von Mengen höchstens gleich der Summe der äußeren Maße der Einzelmengen, d. h.*

$$\left| \sum_n \mathcal{E}_n \right|_G \leq \sum_n |\mathcal{E}_n|_G. \quad (32)$$

Beweis. Im folgenden werden wir eine Überdeckung einer Menge oft mit S und dabei die Summe (30) für diese Überdeckung mit $\sigma(S)$ bezeichnen. Es sei eine positive Zahl ε gegeben. Nach der Definition der unteren Grenze existiert eine Überdeckung S_n der Menge $\mathcal{E}_n$ derart, daß $\sigma(S_n) \leq |\mathcal{E}_n|_G + \frac{\varepsilon}{2^n}$ gilt. Wir nehmen dann die Intervalle, die in allen S_n ($n = 1, 2, \dots$) enthalten sind. Diese bilden eine Überdeckung S für $\sum_n \mathcal{E}_n$, und für diese Überdeckung gilt

$$\sigma(S) = \sum_n \sigma(S_n) \leq \sum_n |\mathcal{E}_n|_G + \varepsilon \sum_n \frac{1}{2^n} \leq \sum_n |\mathcal{E}_n|_G + \varepsilon.$$

Nach Definition der unteren Grenze ist

$$\left| \sum_n \mathcal{E}_n \right|_G \leq \sum_n |\mathcal{E}_n|_G + \varepsilon,$$

und da ε beliebig ist, erhalten wir die Ungleichung (32).

Wir bemerken, daß die rechts stehende Summe in Ungleichung (32) und sogar einzelne Summanden dieser Summe gleich ∞ sein können. Da die Summanden positiv sind, spielt die Ordnung der Summanden keine Rolle. Bei der Definition des äußeren Maßes erweiterten wir die Funktion $G(\Delta)$ auf alle möglichen Punktmengen der Ebene, aber dabei blieb im allgemeinen die Additivität dieser Funktion nicht erhalten. In der Tat kann man zeigen, daß in Formel (32) das Zeichen $<$ stehen kann, selbst wenn die Mengen $\mathcal{E}_n$ paarweise punktfremd sind.

Im folgenden werden wir einige Klassen von Mengen angeben, für die das äußere Maß additiv bleibt. Vorher werden wir noch einen Satz beweisen, der sich auf das äußere Maß bezieht.

Satz 4. *Jede Menge $\mathcal{E}$ kann von einer offenen Menge O überdeckt werden, deren äußeres Maß sich beliebig wenig vom äußeren Maß von $\mathcal{E}$ unterscheidet, d. h., ist $\mathcal{E}$ eine beliebige Menge und ε eine beliebig gegebene positive Zahl, so existiert eine offene Menge O mit $\mathcal{E} \subset O$ und $|O|_G \leq |\mathcal{E}|_G + \varepsilon$.*

Beweis. Ist $|\mathcal{E}|_G = \infty$, so ist die Ungleichung $|O|_G \leq |\mathcal{E}|_G + \varepsilon$ für jede Überdeckung der Menge $\mathcal{E}$ durch eine offene Menge erfüllt. Jetzt sei $|\mathcal{E}|_G$ endlich und ε eine gegebene positive Zahl. Wir wählen eine Überdeckung von $\mathcal{E}$ durch Intervalle Δ_n so, daß die Ungleichung

$$\sum_n G(\Delta_n) \leq |\mathcal{E}|_G + \frac{\varepsilon}{2} \quad (33)$$

gilt. Jedes der Intervalle Δ_n unterziehen wir einer α_n -Dilatation, wobei die positiven Zahlen α_n so gewählt seien, daß

$$G(\Delta_n^{(\alpha_n)}) \leq G(\Delta_n) + \frac{\varepsilon}{2^{n+1}} \quad (34)$$

gilt. Wenn wir den Rand von $\Delta_n^{(\alpha_n)}$, d. h. zwei Seiten und eine Ecke, ausschließen, ergibt die Summe der erhaltenen offenen Intervalle $\Delta_n^{(\alpha_n)}$ eine offene Menge O , die durch die Intervalle $\Delta_n^{(\alpha_n)}$ überdeckt wird. Nach der Definition der unteren Grenze ist

$$|O|_G \leq \sum_n G(\Delta_n^{(\alpha_n)}) .$$

Benutzen wir (34), so können wir weiter

$$|O|_G \leq \sum_n G(\Delta_n) + \frac{\varepsilon}{2}$$

schreiben, und wenn wir (33) berücksichtigen, erhalten wir schließlich

$$|O|_G \leq |\mathcal{E}|_G + \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = |\mathcal{E}|_G + \varepsilon .$$

35. Meßbare Mengen. Wir behandeln jetzt die Klasse von Mengen, die man meßbare Mengen nennt und für die im weiteren die Additivität des äußeren Maßes bewiesen wird. Für diese meßbaren Mengen nennt man das äußere Maß einfach Maß.

Definition. Eine Menge $\mathcal{E}$ heißt *meßbar*, wenn sie von einer offenen Menge O so überdeckt werden kann, daß das äußere Maß der Differenz $O - \mathcal{E}$ beliebig klein wird, d. h., die Menge $\mathcal{E}$ heißt *meßbar*, wenn zu jeder beliebig gegebenen positiven Zahl ε eine offene Menge O existiert mit $\mathcal{E} \subset O$ und $|O - \mathcal{E}|_G \leq \varepsilon$. Das äußere Maß einer meßbaren Menge nennt man einfach *Maß* dieser Menge.

Die Forderung, die in der Definition der meßbaren Menge steht, ist stärker als die Eigenschaft, die in Satz 4 formuliert wurde. Diese Eigenschaft gilt nämlich für alle Mengen, während bei bestimmter Wahl der Funktion $G(\Delta)$ Mengen existieren, die nicht meßbar sind, d. h., auf die die oben gegebene Definition nicht zutrifft. Wir können das Maß einer meßbaren Menge wieder mit $|\mathcal{E}|_G$ bezeichnen, da für die meßbaren Mengen, das Maß per definitionem mit dem äußeren Maß übereinstimmt. Wir werden zeigen, daß jedes Intervall meßbar ist. Sein äußeres Maß ist nach Satz 1 aus [34] gleich $G(\Delta)$. Im weiteren werden wir sehen, daß auch jeder Elementarbereich meßbar ist. Daher kann man das Maß einer beliebigen meßbaren Menge $\mathcal{E}$ einfach mit $G(\mathcal{E})$ bezeichnen. Wir verabreden noch, das äußere Maß (oder einfach das Maß) der leeren Menge gleich 0 zu setzen. Dies stimmt mit der oben gegebenen Definition überein. Wir definieren noch: Eine Menge $\mathcal{E}$ heißt *Menge vom Maß 0 bezüglich $G(\Delta)$* oder einfach *Menge vom Maß 0* (da $G(\Delta)$ als fest angenommen wird), wenn $|\mathcal{E}|_G = 0$ ist. Aus dieser Definition folgt, daß jede Untermenge einer Menge vom Maß 0 wieder eine Menge vom Maß 0 ist.

Wir beweisen jetzt eine Reihe von Eigenschaften der meßbaren Mengen. Diese Eigenschaften sind für alle weiteren Darlegungen grundlegend.

Satz 1. *Jede offene Menge ist meßbar.*

Ist $\mathcal{E}$ eine offene Menge, so genügt es, um die Meßbarkeit nachzuweisen, die Menge O zu nehmen, die mit $\mathcal{E}$ übereinstimmt. Dabei ist

$$|O - \mathcal{E}|_G = 0 .$$

Satz 2. *Jedes Intervall Δ ist eine meßbare Menge, und sein Maß ist gleich $G(\Delta)$.*

Beweis. Es sei Δ ein beliebiges Intervall. Unterziehen wir es einer α -Dilatation, so erhalten wir ein Intervall $\Delta^{(\alpha)}$. Die Differenz $\Delta^{(\alpha)} - \Delta$ ist ein Elementarbereich,

und da die Funktion $G(\Delta)$ normal ist, erhalten wir

$$G(\Delta^{(\alpha)} - \Delta) = G(\Delta^{(\alpha)}) - G(\Delta) \rightarrow 0$$

für $\alpha \rightarrow +0$, d. h., für jedes positive ε existiert ein α , für das $G(\Delta^{(\alpha)}) - G(\Delta) \leq \varepsilon$ ist. Es sei O das offene Intervall von $\Delta^{(\alpha)}$, d. h. das Intervall, das man aus $\Delta^{(\alpha)}$ erhält, wenn man den Rand von $\Delta^{(\alpha)}$ fortläßt. Auf Grund des Prozesses der Dilatation ist $\Delta \subset O$. Um die Meßbarkeit von Δ zu beweisen, muß man zeigen, daß die Ungleichung $|O - \Delta|_G \leq \varepsilon$ gilt. Es ist $O \subset \Delta^{(\alpha)}$, und nach Satz 1 aus [34] ist $|O - \Delta|_G \leq |\Delta^{(\alpha)} - \Delta|_G$. Nun ist $\Delta^{(\alpha)} - \Delta$ ein Elementarbereich, nach Satz 2 aus [34] gilt $|\Delta^{(\alpha)} - \Delta|_G = G(\Delta^{(\alpha)} - \Delta) \leq \varepsilon$, also $|O - \Delta|_G \leq \varepsilon$, was zu beweisen war. Das Maß von Δ , das gleich dem äußeren Maß von Δ ist, stimmt mit $G(\Delta)$ überein, denn Δ ist ein Spezialfall eines Elementarbereichs.

Satz 3. *Der Begriff der Menge vom Maß 0, d. h. der Mengen mit dem äußeren Maß 0, ist gleichbedeutend mit dem Begriff einer meßbaren Menge, deren Maß gleich 0 ist.*

Beweis. Ist $|\mathcal{E}|_G = 0$, so existiert nach Satz 4 aus [34] eine offene Menge O derart, daß für beliebiges $\varepsilon > 0$ die Beziehungen $\mathcal{E} \subset O$ und $|O|_G \leq \varepsilon$ gelten. Hieraus folgt nach Satz 1 aus [34], daß $|O - \mathcal{E}|_G \leq \varepsilon$ ist, d. h., daß $\mathcal{E}$ meßbar ist. Das Maß von $\mathcal{E}$ ist gleich dem äußeren Maß, d. h. gleich 0. Ist umgekehrt $\mathcal{E}$ eine meßbare Menge und ihr Maß gleich 0, so ist nach Definition des Maßes auch das äußere Maß von $\mathcal{E}$ gleich 0. Damit ist Satz 3 bewiesen.

Satz 4. *Die Summe endlich oder abzählbar unendlich vieler meßbarer Mengen ist eine meßbare Menge.*

Beweis. Es seien $\mathcal{E}_n$ meßbare Mengen, $\mathcal{E}$ ihre Summe, $\mathcal{E} = \sum_n \mathcal{E}_n$, und ε sei eine gegebene positive Zahl. Nach Definition der meßbaren Mengen existieren offene Mengen O_n , so daß $\mathcal{E}_n \subset O_n$ und $|O_n - \mathcal{E}_n|_G < \frac{\varepsilon}{2^n}$ gilt. Die Summe der offenen Mengen O_n ist eine offene Menge O , wobei $\mathcal{E} \subset O$ ist, und nach (6) aus [30] gilt

$$O - \mathcal{E} \subset \sum_n (O_n - \mathcal{E}_n).$$

Benutzen wir die Sätze 1 und 3 aus [34], so erhalten wir

$$|O - \mathcal{E}|_G \leq \left| \sum_n (O_n - \mathcal{E}_n) \right|_G \leq \sum_n |O_n - \mathcal{E}_n|_G$$

oder, da $|O_n - \mathcal{E}_n|_G \leq \frac{\varepsilon}{2^n}$ ist, $|O - \mathcal{E}|_G < \varepsilon$. Damit ist die Meßbarkeit von $\mathcal{E}$ bewiesen.

Bevor wir die Meßbarkeit der abgeschlossenen Mengen beweisen, beweisen wir noch das folgende

Lemma. *Ist der Abstand der beiden Mengen $\mathcal{E}_1$ und $\mathcal{E}_2$ eine positive Zahl, so ist*

$$|\mathcal{E}_1 + \mathcal{E}_2|_G = |\mathcal{E}_1|_G + |\mathcal{E}_2|_G.$$

Beweis. Es sei d der positive Abstand der Mengen $\mathcal{E}_1$ und $\mathcal{E}_2$. Für ein beliebig vorgegebenes positives ε existiert eine Überdeckung S der Menge $\mathcal{E}_1 + \mathcal{E}_2$ derart, daß

$$\sigma(S) \leq |\mathcal{E}_1 + \mathcal{E}_2|_G + \varepsilon \quad (35)$$

gilt. Jedes der Intervalle der Überdeckung S zerlegen wir so in endlich viele Intervalle, daß ihre Diagonalen kleiner als d sind. Alle Intervalle, die in S liegen, teilen wir in drei Klassen ein: Zur ersten Klasse gehören die Intervalle, die nur Punkte

von $\mathfrak{E}_1$ überdecken, zur zweiten jene, die nur Punkte von $\mathfrak{E}_2$ überdecken, und schließlich zur dritten jene, die weder Punkte aus $\mathfrak{E}_1$ noch Punkte aus $\mathfrak{E}_2$ überdecken. Intervalle, die sowohl Punkte aus $\mathfrak{E}_1$ als auch Punkte aus $\mathfrak{E}_2$ enthalten, gibt es nicht. Die Intervalle der dritten Klasse können wir einfach aus der Zerlegung von S entfernen. Dabei kann die Summe $\sigma(S)$ nur kleiner werden, und die Ungleichung (35) bleibt gültig. Also kann man annehmen, daß sich die Überdeckung S in die Überdeckungen S_1 und S_2 zerlegen läßt, wobei die Intervalle aus S_1 die Menge $\mathfrak{E}_1$ überdecken und elementfremd zu $\mathfrak{E}_2$ sind und die Intervalle aus S_2 die Menge $\mathfrak{E}_2$ überdecken und elementfremd zu $\mathfrak{E}_1$ sind. Es gilt $\sigma(S) = \sigma(S_1) + \sigma(S_2)$ und nach (35)

$$\sigma(S_1) + \sigma(S_2) \leq |\mathfrak{E}_1 + \mathfrak{E}_2|_G + \varepsilon. \quad (36)$$

Nach der Definition der unteren Grenze erhält man $|\mathfrak{E}_1|_G \leq \sigma(S_1)$ und $|\mathfrak{E}_2|_G \leq \sigma(S_2)$, und die Ungleichung (36) führt auf $|\mathfrak{E}_1|_G + |\mathfrak{E}_2|_G \leq |\mathfrak{E}_1 + \mathfrak{E}_2|_G + \varepsilon$. Hieraus folgt, da ε beliebig war, $|\mathfrak{E}_1|_G + |\mathfrak{E}_2|_G \leq |\mathfrak{E}_1 + \mathfrak{E}_2|_G$. Andererseits folgt aus Satz 3 aus [34] $|\mathfrak{E}_1 + \mathfrak{E}_2|_G \leq |\mathfrak{E}_1|_G + |\mathfrak{E}_2|_G$. Also ist $|\mathfrak{E}_1 + \mathfrak{E}_2|_G = |\mathfrak{E}_1|_G + |\mathfrak{E}_2|_G$. Damit ist das Lemma bewiesen.

Folgerung. Sind F_1 und F_2 zwei abgeschlossene punktfremde Mengen, von denen wenigstens eine beschränkt ist, so ist $|F_1 + F_2|_G = |F_1|_G + |F_2|_G$. Sind F_k ($k = 1, 2, \dots, m$) abgeschlossene beschränkte paarweise punktfremde Mengen, so ist

$$\left| \sum_{k=1}^m F_k \right|_G = \sum_{k=1}^m |F_k|_G.$$

Um diese Folgerung zu beweisen, genügt es, das in [32] bezüglich des Abstandes abgeschlossener und paarweise punktfremder Mengen Gesagte zu benutzen.

Satz 5. Abgeschlossene Mengen sind meßbar.

Beweis. Zuerst nehmen wir an, daß F eine beliebige beschränkte abgeschlossene Menge ist, und ε sei eine gegebene positive Zahl. Nach Satz 4 aus [34] existiert eine offene Menge O mit $F \subset O$ und $|O|_G \leq |F|_G + \varepsilon$. Wir zeigen jetzt, daß diese offene Menge O auch der Ungleichung

$$|O - F|_G \leq \varepsilon \quad (37)$$

genügt, die in der Definition der Meßbarkeit vorkommt. Nach Satz 4 aus [32] ist $O - F$ eine offene Menge, folglich kann sie auf Grund von Satz 6 aus [32] als Summe von abzählbar unendlich vielen paarweise punktfremden Intervallen Δ_n dargestellt werden:

$$O - F = \sum_{n=1}^{\infty} \Delta_n. \quad (38)$$

Wir betrachten die Summe der ersten m Summanden (m fest und eine ganze positive Zahl) von (38), wobei für jedes Intervall, das in der Summe steht, eine α -Kontraktion durchgeführt wird (α ist eine beliebig gewählte positive Zahl). Auf diese Weise erhalten wir einen Elementarbereich

$$R = \sum_{n=1}^m {}^{(\alpha)}\Delta_n.$$

Wenn wir jedes Intervall ${}^{(\alpha)}\Delta_n$ ($n = 1, 2, \dots, m$) abschließen, dann liefert diese Summe eine abgeschlossene Menge, die mit der abgeschlossenen Hülle $\bar{R}$ des Elementarbereichs R übereinstimmt. Jedes abgeschlossene Intervall ${}^{(\alpha)}\Delta_n$ wird durch ein entsprechendes Intervall Δ_n der Summe (38) überdeckt. Somit hat die abge-

geschlossene Menge R keine gemeinsamen Punkte mit F , und daher ist der Abstand zwischen ihnen positiv. Also ist der Abstand zwischen R und F erst recht positiv, und nach dem Lemma erhalten wir

$$\left| \sum_{n=1}^m {}^{(\alpha)}\Delta_n + F \right|_G = \left| \sum_{n=1}^m {}^{(\alpha)}\Delta_n \right|_G + |F|_G.$$

Nach (38) ist aber $\sum_{n=1}^m {}^{(\alpha)}\Delta_n + F \subset O$ und folglich

$$\left| \sum_{n=1}^m {}^{(\alpha)}\Delta_n \right|_G + |F|_G \leq |O|_G.$$

Benutzen wir die Ungleichung $|O|_G \leq |F|_G + \varepsilon$, so erhalten wir hieraus

$$\left| \sum_{n=1}^m {}^{(\alpha)}\Delta_n \right|_G + |F|_G \leq |F|_G + \varepsilon.$$

Da F nach Voraussetzung eine beschränkte Menge ist, ist folglich $|F|_G$ endlich. Die letzte Ungleichung führt auf

$$\left| \sum_{n=1}^m {}^{(\alpha)}\Delta_n \right|_G \leq \varepsilon.$$

Es ist aber

$$\left| \sum_{n=1}^m {}^{(\alpha)}\Delta_n \right|_G = G(R) = \sum_{n=1}^m G({}^{(\alpha)}\Delta_n) = \sum_{n=1}^m |{}^{(\alpha)}\Delta_n|_G,$$

also

$$\sum_{n=1}^m |{}^{(\alpha)}\Delta_n|_G \leq \varepsilon.$$

Lassen wir zuerst α gegen 0 und darauf m gegen ∞ streben, so erhalten wir $\sum_{n=1}^{\infty} |\Delta_n|_G \leq \varepsilon$. Schließlich folgt aus Formel (38) sowie Satz 3 aus [34]

$$|O - F|_G \leq \sum_{n=1}^{\infty} |\Delta_n|_G \leq \varepsilon,$$

d. h. die Ungleichung (37). Jetzt nehmen wir an, daß die abgeschlossene Menge F unbeschränkt ist. Es sei γ_n der abgeschlossene Kreis mit dem Ursprung als Mittelpunkt und dem Radius n . Mit den abgeschlossenen beschränkten Mengen $F_n = F \cdot \gamma_n$ können wir

$$F = \sum_{n=1}^{\infty} F_n$$

schreiben, und die Meßbarkeit von F erhalten wir unmittelbar aus Satz 4. Damit ist Satz 5 bewiesen.

Satz 6. Ist $\mathcal{E}$ eine meßbare Menge, so ist auch die Komplementärmenge $\mathcal{C}\mathcal{E}$ meßbar.

Beweis. Infolge der Meßbarkeit von $\mathcal{E}$ existieren offene Mengen O_n mit $\mathcal{E} \subset O_n$ und $|O_n - \mathcal{E}|_G < 1/n$. Wir bilden die abgeschlossenen Mengen $F_n = \mathcal{C}O_n$. Da $\mathcal{E} \subset O_n$ ist, folgt $F_n \subset \mathcal{C}\mathcal{E}$, und nach (19) aus [30] gilt $\mathcal{C}\mathcal{E} - F_n = O_n - \mathcal{E}$. Ersetzen wir die F_n durch die Summe über alle F_n , so erhalten wir

$$\mathcal{C}\mathcal{E} - \sum_{n=1}^{\infty} F_n \subset \mathcal{C}\mathcal{E} - F_n, \quad \text{d. h.} \quad \mathcal{C}\mathcal{E} - \sum_{n=1}^{\infty} F_n \subset O_n - \mathcal{E},$$

und wegen $|O_n - \mathcal{E}|_G < 1/n$ ist

$$\left| C\mathcal{E} - \sum_{n=1}^{\infty} F_n \right|_G < \frac{1}{n}.$$

Die linke Seite dieser Ungleichung hängt nicht von n ab, und wenn n gegen ∞ strebt, erhalten wir

$$\left| C\mathcal{E} - \sum_{n=1}^{\infty} F_n \right|_G = 0.$$

Hieraus folgt, daß die links stehende Differenz eine Menge $\mathcal{E}_0$ vom Maß 0 ist. Somit kann man $C\mathcal{E}$ als Summe von meßbaren Mengen schreiben:

$$C\mathcal{E} = \mathcal{E}_0 + \sum_{n=1}^{\infty} F_n.$$

Hieraus und aus Satz 4 folgt die Meßbarkeit von $C\mathcal{E}$.

Gemäß unserer Definition haben wir die Meßbarkeit einer Menge mit Hilfe der offenen Mengen nachgewiesen. Im folgenden Satz werden wir zeigen, daß man die Meßbarkeit einer Menge auch analog durch abgeschlossene Mengen feststellen kann.

Satz 7. *Für die Meßbarkeit von $\mathcal{E}$ ist notwendig und hinreichend, daß zu jeder beliebigen positiven Zahl ε eine abgeschlossene Menge F existiert, für welche $F \subset \mathcal{E}$ und $|\mathcal{E} - F|_G < \varepsilon$ gilt.*

Beweis. Die Meßbarkeit von $\mathcal{E}$ ist gleichbedeutend mit der Meßbarkeit von $C(\mathcal{E})$. Dafür ist jedoch notwendig und hinreichend, daß für ein beliebiges positives ε eine offene Menge O existiert, für welche $C\mathcal{E} \subset O$ und $|O - C\mathcal{E}|_G \leq \varepsilon$ gilt. Wenn wir $F = CO$ setzen und beachten, daß nach (19) aus [30]

$$O - C\mathcal{E} = \mathcal{E} - CO = \mathcal{E} - F \quad \text{und} \quad F \subset \mathcal{E}$$

gilt, erhalten wir die Behauptung des Satzes.

Satz 8. *Das Produkt endlich oder abzählbar unendlich vieler meßbarer Mengen ist wieder eine meßbare Menge. Die Differenz zweier meßbarer Mengen ist wieder eine meßbare Menge.*

Beweis. Wenn die Mengen $\mathcal{E}_n$ meßbar sind, ergibt sich die Meßbarkeit ihres Produkts unmittelbar aus der Formel

$$\prod_n \mathcal{E}_n = C \sum_n C\mathcal{E}_n$$

(vgl. [30]) und aus den Sätzen 6 und 4. Sind A und B meßbar, so folgt die Meßbarkeit ihrer Differenz unmittelbar aus der Formel

$$A - B = A \cdot CB$$

(vgl. [30]) und der soeben bewiesenen Meßbarkeit des Produktes.

Satz 9. *Das Maß der Summe endlich oder abzählbar unendlich vieler meßbarer paarweise punktfremder Mengen ist gleich der Summe der Maße der einzelnen Summanden.*

Beweis. Es seien $\mathcal{E}_n$ meßbare paarweise punktfremde Mengen. Die Meßbarkeit ihrer Summe folgt aus Satz 4. Zuerst nehmen wir an, daß jede der Mengen $\mathcal{E}_n$ beschränkt ist. Nach Satz 7 existieren für ein beliebig gegebenes positives ε abgeschlossene Mengen F_n derart, daß $F_n \subset \mathcal{E}_n$ und $|\mathcal{E}_n - F_n|_G \leq \frac{\varepsilon}{2^n}$ gilt, wobei die Men-

gen F_n offenbar beschränkte, paarweise punktfremde Mengen sind. Aus der Formel $\mathcal{E}_n = F_n + (\mathcal{E}_n - F_n)$ folgt unmittelbar

$$|\mathcal{E}_n|_G \leq |F_n|_G + \frac{\varepsilon}{2^n}.$$

Andererseits gilt, wenn wir die ersten m Mengen F_n betrachten, $\sum_{n=1}^m F_n \subset \sum_n \mathcal{E}_n$ und folglich

$$\left| \sum_{n=1}^m F_n \right|_G \leq \left| \sum_n \mathcal{E}_n \right|_G.$$

Auf die endliche Summe abgeschlossener paarweise punktfremder Mengen F_n können wir das oben bewiesene Lemma anwenden, und somit erhalten wir, wenn wir noch die Ungleichung $|F_n|_G \geq |\mathcal{E}_n|_G - \frac{\varepsilon}{2^n}$ benutzen,

$$\left| \sum_n \mathcal{E}_n \right|_G \geq \sum_{n=1}^m |F_n|_G \geq \sum_{n=1}^m |\mathcal{E}_n|_G - \sum_{n=1}^m \frac{\varepsilon}{2^n}.$$

Wir betrachten jetzt den komplizierteren Fall, daß die Anzahl der Mengen $\mathcal{E}_n$ unendlich ist. Wenn in der letzten Ungleichung m gegen ∞ strebt, erhalten wir

$$\left| \sum_{n=1}^{\infty} \mathcal{E}_n \right|_G \geq \sum_{n=1}^{\infty} |\mathcal{E}_n|_G - \varepsilon$$

oder, da ε beliebig war,

$$\left| \sum_{n=1}^{\infty} \mathcal{E}_n \right|_G \geq \sum_{n=1}^{\infty} |\mathcal{E}_n|_G.$$

Vergleichen wir diese Ungleichung mit (32), so erhalten wir

$$\left| \sum_{n=1}^{\infty} \mathcal{E}_n \right|_G = \sum_{n=1}^{\infty} |\mathcal{E}_n|_G, \quad (39)$$

was wir beweisen wollten.

Wenn wir die Meßbarkeit der $\mathcal{E}_n$ und ihrer Summe berücksichtigen, können wir die letzte Formel in der Gestalt

$$G(\mathcal{E}_1 + \mathcal{E}_2 + \mathcal{E}_3 + \dots) = G(\mathcal{E}_1) + G(\mathcal{E}_2) + G(\mathcal{E}_3) + \dots \quad (40)$$

schreiben.

Wir betrachten jetzt den Fall, daß unter den Mengen $\mathcal{E}_n$ auch unbeschränkte Mengen sind. Es sei γ_n der abgeschlossene Kreis mit dem Ursprung als Mittelpunkt und dem Radius n . Wir betrachten die Mengen

$$\mathcal{E}_n^{(1)} = \mathcal{E}_n \gamma_1, \quad \mathcal{E}_n^{(2)} = \mathcal{E}_n (\gamma_2 - \gamma_1), \quad \mathcal{E}_n^{(3)} = \mathcal{E}_n (\gamma_3 - \gamma_2), \dots$$

Jede von ihnen ist beschränkt, und alle sind sie meßbar, da die abgeschlossene Menge γ_1 und die Differenz der abgeschlossenen Mengen $\gamma_k - \gamma_{k-1}$ meßbare Mengen sind und das Produkt meßbarer Mengen wieder meßbar ist. Wir können jede der Mengen $\mathcal{E}_n$ als Summe beschränkter meßbarer und paarweise punktfremder Mengen darstellen:

$$\mathcal{E}_n = \sum_{k=1}^{\infty} \mathcal{E}_n^{(k)},$$

und nach dem oben Bewiesenen ist

$$|\mathfrak{E}_n|_G = \sum_{k=1}^{\infty} |\mathfrak{E}_n^{(k)}|_G. \quad (41)$$

Die Summe $\mathfrak{E}$ der Mengen $\mathfrak{E}_n$ können wir als Doppelsumme der beschränkten, paarweise punktfremden Mengen $\mathfrak{E}_n^{(k)}$ darstellen, wobei einige der Mengen $\mathfrak{E}_n^{(k)}$ auch leer sein können:

$$\mathfrak{E} = \sum_{n=1}^{\infty} \sum_{k=1}^{\infty} \mathfrak{E}_n^{(k)}.$$

Nach dem oben Bewiesenen gilt

$$|\mathfrak{E}|_G = \sum_{n=1}^{\infty} \sum_{k=1}^{\infty} |\mathfrak{E}_n^{(k)}|_G.$$

Die Reihenfolge der Summation ist, da die Summanden nicht negativ sind, unwesentlich [I, 142]. Wir werden zuerst über k summieren und dann über n . Berücksichtigen wir die Formel (41), so kommen wir wieder auf (39), und damit ist der Satz vollständig bewiesen.

Bemerkung. Lassen wir die Voraussetzung, daß die meßbaren Mengen $\mathfrak{E}_n$ paarweise punktfremd sind, fallen, so werden wir für ihre nach Satz 4 meßbare Summe statt (40) die Ungleichung

$$G(\mathfrak{E}) \leq \sum_{n=1}^{\infty} G(\mathfrak{E}_n) \quad (40_1)$$

erhalten, die unmittelbar aus Satz 3 aus [34] und der Tatsache folgt, daß für meßbare Mengen das äußere Maß einfach mit dem Maß übereinstimmt. Wenn für alle Mengen $\mathfrak{E}_n$ das Maß gleich 0 ist, folgt aus (40) $G(\mathfrak{E}) \leq 0$. Aber das Maß kann nicht negativ sein, und daher ist $G(\mathfrak{E}) = 0$, d. h., die Summe endlich oder abzählbar unendlich vieler Mengen vom Maß 0 ist eine Menge vom Maß 0.

Alle oben bewiesenen Sätze gelten auch für meßbare Mengen unendlichen Maßes. In den nächsten Sätzen machen wir aus diesem Anlaß eine Einschränkung.

Satz 10. Sind A und B meßbar, gilt $B \subset A$ und hat B endliches Maß, so ist

$$G(A - B) = G(A) - G(B). \quad (42)$$

Beweis. Die Differenz $A - B = D$ ist nach Satz 8 meßbar. Es gilt $A = B + D$ wobei B und D punktfremd sind. Nach Satz 9 ist $G(A) = G(B) + G(D)$, und subtrahieren wir von beiden Seiten die endliche Zahl $G(B)$, so erhalten wir (42).

Satz 11. Wenn die $\mathfrak{E}_n$ ($n = 1, 2, \dots$) eine nicht fallende Folge meßbarer Mengen bilden, ist die Grenzmenge $\mathfrak{E}$ meßbar, und es gilt

$$G(\mathfrak{E}) = \lim_{n \rightarrow \infty} G(\mathfrak{E}_n). \quad (43)$$

Beweis. Die Meßbarkeit von $\mathfrak{E}$ folgt unmittelbar aus der Formel

$$\mathfrak{E} = \lim_{n \rightarrow \infty} \mathfrak{E}_n = \mathfrak{E}_1 + (\mathfrak{E}_2 - \mathfrak{E}_1) + (\mathfrak{E}_3 - \mathfrak{E}_2) + \dots \quad (44)$$

Die auf der rechten Seite stehenden Summanden sind punktfremd, und wenn alle $\mathfrak{E}_n$ endliches Maß haben, ist

$$G(\mathfrak{E}) = G(\mathfrak{E}_1) + [G(\mathfrak{E}_2) - G(\mathfrak{E}_1)] + [G(\mathfrak{E}_3) - G(\mathfrak{E}_2)] + \dots$$

Auf der rechten Seite ist die Summe der ersten n Summanden gleich $G(\mathfrak{E}_n)$, d. h., aus der letzten Formel folgt (43). Wenn einige der $\mathfrak{E}_n$ unendliches Maß haben, hat die Grenzmenge ebenfalls unendliches Maß, und die Formel (43) ist evident. In dieser Formel ist der Wert ∞ sowohl für $G(\mathfrak{E}_n)$ als auch für $G(\mathfrak{E})$ zulässig.

Satz 12. *Wenn die $\mathfrak{E}_n$ ($n = 1, 2, \dots$) eine nicht wachsende Folge von Mengen endlichen Maßes bilden, ist die Grenzmenge meßbar, und es gilt die Formel (43).*

Beweis. Wir stellen $\mathfrak{E}_1$ als Summe von paarweise punktfremden Mengen dar:

$$\mathfrak{E}_1 = \mathfrak{E} + (\mathfrak{E}_1 - \mathfrak{E}_2) + (\mathfrak{E}_2 - \mathfrak{E}_3) + (\mathfrak{E}_3 - \mathfrak{E}_4) + \dots \quad (45)$$

Die Meßbarkeit von $\mathfrak{E}$ folgt aus den Sätzen 4 und 10. Wenn wir auf (45) die Sätze 9 und 10 anwenden, erhalten wir

$$G(\mathfrak{E}_1) = G(\mathfrak{E}) + [G(\mathfrak{E}_1) - G(\mathfrak{E}_2)] + [G(\mathfrak{E}_2) - G(\mathfrak{E}_3)] + [G(\mathfrak{E}_3) - G(\mathfrak{E}_4)] + \dots,$$

d. h.

$$G(\mathfrak{E}_1) = G(\mathfrak{E}) + G(\mathfrak{E}_1) - \lim_{n \rightarrow \infty} G(\mathfrak{E}_n).$$

Hieraus folgt (43).

Bemerkung. Die Meßbarkeit der Grenzmenge $\mathfrak{E}$ ergibt sich aus (45) auch ohne Voraussetzung der Endlichkeit des Maßes der $\mathfrak{E}_n$.

36. Meßbare Mengen (Fortsetzung). Wir ziehen jetzt eine Reihe von Folgerungen aus den bewiesenen Sätzen über meßbare Mengen. Ein Elementarbereich R ist als Summe endlich vieler Intervalle eine meßbare Menge, und das Maß von R , das mit dem äußeren Maß von R übereinstimmt, kann durch Formel (25) ausgedrückt werden, wobei Δ_k ($k = 1, 2, \dots, m$) irgendwelche elementfremden Zerlegungen von R auf dem Intervall sind. Mit L_G bezeichnen wir die Familie der meßbaren Mengen, wobei der Index G auf diejenige Funktion $G(\Delta)$ hinweist, welche der Konstruktion dieser Familie zugrunde liegt. Wir erweitern die Funktion $G(\Delta)$ auf alle Mengen $\mathfrak{E}$, die zu L_G gehören, wobei die erhaltene Funktion $G(\mathfrak{E})$ nichtnegativ und nicht nur für endlich viele, sondern nach Satz 9 aus [35] auch für abzählbar unendlich viele Summanden paarweise elementfremder Mengen additiv ist. Es sei eine verschwindende Folge von Mengen $\mathfrak{E}_n$ gegeben, die zu L_G gehören und endliches Maß haben, d. h. $\mathfrak{E}_1 \supset \mathfrak{E}_2 \supset \mathfrak{E}_3 \supset \dots$, deren Grenzmenge $\mathfrak{E}$ also die leere Menge ist. Aus Satz 12 von [35] folgt unmittelbar $G(\mathfrak{E}_n) \rightarrow 0$, d. h., die Funktion $G(\mathfrak{E})$ ist nicht nur nichtnegativ und additiv, sondern auch normal für die Familie L_G . Wenn hervorgehoben werden soll, daß sie nicht nur für endlich viele, sondern auch für abzählbar unendlich viele Mengen, die in L_G liegen, additiv ist, nennen wir sie *voll additiv*. Die Familie L_G enthält auch unbeschränkte Mengen. Einige von diesen können endliches Maß haben, und für andere kann das Maß gleich ∞ sein. Schließlich braucht nicht jede unbeschränkte Menge meßbar zu sein. Oft betrachtet man bei der Konstruktion einer Familie von meßbaren Mengen nur beschränkte Mengen oder nur Mengen, die in einem bestimmten endlichen Intervall liegen. Im vorhergehenden machten wir jedoch diese Einschränkung nicht. Wir bemerken noch, daß die Ausgangsfunktion $G(\Delta)$ für alle endlichen Intervalle definiert war. Wenn $G(\Delta)$ nur für die Intervalle Δ definiert ist, die in einem Intervall Δ_0 liegen, kann man sie natürlich auf alle Intervalle Δ ausdehnen, indem man die Formel $G(\Delta) = G(\Delta \cdot \Delta_0)$ benutzt, wobei beachtet werden muß, daß das Produkt von Intervallen wieder ein Intervall ist.

Die Familie L_G hängt von der Wahl der Ausgangsfunktion $G(\Delta)$ ab. Aber bei beliebiger Wahl dieser Funktion enthält sie in jedem Fall alle Intervalle, die Elementarbereiche, die offenen und die abgeschlossenen Mengen. Im folgenden werden

wir jene Mengen, die bei beliebiger Wahl von $G(\Delta)$ in L_G liegen, vollständiger charakterisieren. Wir werden die Mengenfunktion als Masse deuten. Die Vorgabe der ursprünglichen Funktion $G(\Delta)$ läßt sich auf die Vorgabe einer Masse auf einem beliebigen Intervall Δ zurückführen, wobei die üblichen Bedingungen erfüllt sind, d. h., $G(\Delta)$ nichtnegativ, additiv und normal ist. Eine Punktmenge $\mathcal{E}$ ist meßbar, wenn es sinnvoll ist, von einer Masse, mit der $\mathcal{E}$ belegt ist, zu sprechen, und $G(\mathcal{E})$ ist diese Masse.

Man kann ein einfaches Beispiel dafür angeben, wann die Menge L_G alle Punkt-mengen der Ebene enthält. Es sei eine Masse 1 gegeben, die im Punkt P konzentriert ist. Dabei sei $G(\Delta) = 1$, wenn das Intervall Δ den Punkt P enthält, und $G(\Delta) = 0$, wenn Δ den Punkt P nicht enthält. Man kann leicht nachprüfen, daß die Familie L_G für diese Funktion $G(\Delta)$ alle Mengen enthält, wobei $G(\mathcal{E}) = 1$ ist, wenn P in $\mathcal{E}$ enthalten ist, und $G(\mathcal{E}) = 0$, wenn $\mathcal{E}$ den Punkt P nicht enthält.

Wir betrachten den wichtigen Spezialfall, daß $G(\Delta)$ gleich dem Flächeninhalt des Intervalls Δ ist. Die Familie L_G werde in diesem Fall einfach mit L bezeichnet. Hier haben wir eine Erweiterung des Begriffs des Flächeninhalts für eine große Mengenfamilie L . Dieser Spezialfall wurde zuerst von dem französischen Mathematiker LEBESGUE behandelt. Die Funktion $G(\mathcal{E})$ wird in diesem Fall mit $m(\mathcal{E})$ bezeichnet. Die Mengenfamilie L nennt man gewöhnlich die *Familie der Lebesgue-meßbaren¹⁾ Mengen*. Bei solchen Mengen hat es Sinn, von ihrem Flächeninhalt zu sprechen. Ist $\mathcal{E}$ eine endliche oder abzählbare Menge von Punkten, so ist $m(\mathcal{E}) = 0$. Wenn $\mathcal{E}$ ein Geradenabschnitt oder die ganze Gerade ist, ist ebenfalls $m(\mathcal{E}) = 0$. Wenn wir ein und dasselbe Intervall Δ als halboffen, offen oder abgeschlossen betrachten, hat $m(\Delta)$ in jedem Fall ein und denselben Wert. Hat eine meßbare Menge innere Punkte, so ist $m(\mathcal{E}) > 0$. Man kann zeigen, daß beschränkte offene Mengen existieren, für deren Rand l (l ist abgeschlossen, also meßbar) $m(l) > 0$ gilt. Für eine offene Menge ist $m(O)$ die Summe der Flächeninhalte derjenigen Intervalle, die in Formel (21) auftreten. Diese Summe hängt hierbei nicht davon ab, wie O als Summe von Intervallen dargestellt wurde. Ist F eine beschränkte abgeschlossene Menge, so können wir sie mit einem offenen Intervall Δ_0 überdecken und $m(F)$ als Differenz der Maße zweier offener Mengen definieren, und zwar

$$m(F) = m(\Delta_0) - m(\Delta_0 - F).$$

Die ganze Konstruktion der Familie L_G kann genau so wie oben auch in einem beliebigen endlichdimensionalen Raum durchgeführt werden. Im Spezialfall des dreidimensionalen Raumes ist L die Familie der Mengen, die ein bestimmtes „Volumen“ haben; im eindimensionalen Fall haben die Mengen der Familie eine bestimmte „Länge“. Man spricht aber auch in Räumen beliebiger endlicher Dimension einfach von dem Maß der Menge, wenn diese Menge in L enthalten ist.

37. Meßbarkeitskriterien. Man kann verschiedene Definitionen der meßbaren Mengen geben, die gleichbedeutend der oben gegebenen Definition sind. Wir werden einige dieser Definitionen erwähnen, wobei wir zuerst nur beschränkte Mengen betrachten.

Satz 1. *Die beschränkte Menge $\mathcal{E}$ ist genau dann in der Familie L_G enthalten, wenn es für ein beliebig gegebenes positives ε einen Elementarbereich R mit*

$$\mathcal{E} + e_1 = R + e_2 \tag{46}$$

¹⁾ Häufig wird auch zur Abkürzung (L) -meßbar geschrieben (Anm. d. Red.)

gibt, wobei für die Mengen e_1 und e_2 die Ungleichungen

$$|e_1|_G \leq \varepsilon \quad \text{und} \quad |e_2|_G \leq \varepsilon \quad (47)$$

gelten.

Beweis. Wir beweisen zunächst die Notwendigkeit. Es sei $\mathcal{E}$ in L_G enthalten. Dann existiert eine offene Menge O mit $\mathcal{E} \subset O$ und $|O - \mathcal{E}|_G \leq \varepsilon$. Setzen wir $O - \mathcal{E} = e_1$, so ist $O = \mathcal{E} + e_1$, und für e_1 gilt die Ungleichung (47). Andererseits ist nach Satz 6 aus [32] O der Grenzwert einer wachsenden Folge von Elementarbereichen R_n , wobei R_n die Summe der ersten n Summanden auf der rechten Seite von (21) ist. Nach Satz 11 aus [35] ist $G(O) = \lim_{n \rightarrow \infty} G(R_n)$, und folglich können wir so große Werte $n = m$ nehmen, daß, wenn man $R = R_m$ setzt, $O = R + e_2$ mit $|e_2|_G \leq \varepsilon$ ist. Wenn wir die beiden für O erhaltenen Ausdrücke gleichsetzen, erhalten wir Formel (46), wobei für e_1 und e_2 die Ungleichungen (47) erfüllt sind. Wir beweisen jetzt, daß die Bedingung auch hinreichend ist. Es seien für ein beliebig gegebenes ε die Formel (46) und die Ungleichungen (47) erfüllt. Infolge der Meßbarkeit von R existiert eine offene Menge O_1 mit $R \subset O_1$ und $|O_1 - R|_G \leq \varepsilon$. Andererseits existiert nach Satz 4 aus [34] eine offene Menge O_2 mit $e_2 \subset O_2$ und $|O_2|_G \leq |e_2|_G + \varepsilon$ oder nach (47) $|O_2|_G \leq 2\varepsilon$. Die offene Menge $O = O_1 + O_2$ überdeckt $\mathcal{E} + e_1$, und es ist $O - \mathcal{E} \subset [O - (\mathcal{E} + e_1)] + e_1$ oder nach Formel (6) aus [30]

$$O - \mathcal{E} \subset [(O_1 + O_2) - (R + e_2)] + e_1 \subset (O_1 - R) + (O_2 - e_2) + e_1.$$

Beachten wir, daß $|O_1 - R|_G \leq \varepsilon$, $|O_2 - e_2|_G \leq |O_2|_G \leq 2\varepsilon$ und (47) gilt, so erhalten wir hieraus $|O - \mathcal{E}|_G \leq 4\varepsilon$. Das liefert, da ε beliebig gewählt war, die Meßbarkeit von $\mathcal{E}$.

Satz 2. Die beschränkte Menge $\mathcal{E}$ liegt genau dann in L_G , wenn für ein beliebig gegebenes positives ε ein Elementarbereich R mit

$$|\mathcal{E} - R|_G \leq \varepsilon \quad \text{und} \quad |R - \mathcal{E}|_G \leq \varepsilon \quad (48)$$

existiert.

Beweis. Wir beweisen zuerst die Notwendigkeit. Es sei $\mathcal{E}$ in L_G enthalten. Es existiert ein Elementarbereich R , für den die Beziehungen (46) und (47) gelten. Die Ungleichungen (48) folgen aus den Relationen $\mathcal{E} - R \subset e_2$, $R - \mathcal{E} \subset e_1$. Jetzt beweisen wir, daß die Bedingung auch hinreichend ist. Für ein beliebig gegebenes ε existiert ein Elementarbereich R , für den die Ungleichungen (48) gelten. Wenn wir $\mathcal{E} - R = e_2$ und $R - \mathcal{E} = e_1$ setzen, erhalten wir $\mathcal{E} + e_1 = R + e_2$, wobei e_1 und e_2 die Ungleichungen (47) befriedigen, und die Meßbarkeit von $\mathcal{E}$ folgt jetzt unmittelbar aus Satz 1.

Satz 3. Eine Menge $\mathcal{E}$ (die auch unbeschränkt sein kann) liegt genau dann in L_G , wenn für ein beliebig gegebenes positives ε eine offene Menge O und eine abgeschlossene Menge F mit

$$F \subset \mathcal{E} \subset O \quad \text{und} \quad |O - F|_G \leq \varepsilon \quad (49)$$

existieren.

Beweis. Ist $\mathcal{E}$ meßbar, so existieren per definitionem und nach Satz 7 aus [35] Mengen F und O mit

$$F \subset \mathcal{E} \subset O, \quad |\mathcal{E} - F|_G \leq \frac{\varepsilon}{2} \quad \text{und} \quad |O - \mathcal{E}|_G \leq \frac{\varepsilon}{2}.$$

Ferner ist $O - F = (O - \mathcal{E}) + (\mathcal{E} - F)$, und hieraus folgt (49). Umgekehrt sei jetzt (49) gegeben. Dabei gilt auch $|O - \mathcal{E}|_G \leq \varepsilon$, und folglich ist $\mathcal{E}$ nach Definition meßbar.

Wir geben schließlich noch ein Meßbarkeitskriterium an, ohne es jedoch zu beweisen: Für die Meßbarkeit von $\mathcal{E}$ ist notwendig und hinreichend, daß für eine beliebig gewählte Menge A die folgende Formel gilt:

$$|A|_{\mathcal{G}} = |A \cdot \mathcal{E}|_{\mathcal{G}} + |A - \mathcal{E}|_{\mathcal{G}}. \quad (50)$$

38. Mengenkörper. Wir führen jetzt einen neuen Begriff ein, der sich auf Familien von Punktmengen bezieht, wobei wir unter einer Mengenfamilie eine Gesamtheit von Mengen (Mengen von Mengen) verstehen. Man nennt eine Mengenfamilie einen *Mengenkörper*, wenn die beiden folgenden Eigenschaften gelten:

1. Wenn die Mengen $\mathcal{E}_1$ und $\mathcal{E}_2$ in der Familie enthalten sind, ist auch ihre Differenz $\mathcal{E}_1 - \mathcal{E}_2$ in dieser Familie enthalten.

2. Wenn $\mathcal{E}_1$ und $\mathcal{E}_2$ elementfremd und in der Familie enthalten sind, ist auch ihre Summe $\mathcal{E}_1 + \mathcal{E}_2$ in dieser Familie enthalten.

Aus dieser Definition kann man unmittelbar folgern: *Die leere Menge als Differenz von zwei gleichen Mengen, die dem Mengenkörper angehören, liegt bestimmt in jedem Mengenkörper.* Weiter folgt aus den Formeln

$$\mathcal{E}_1 \mathcal{E}_2 = \mathcal{E}_1 - (\mathcal{E}_1 - \mathcal{E}_2), \quad \mathcal{E}_1 + \mathcal{E}_2 = \mathcal{E}_1 + (\mathcal{E}_2 - \mathcal{E}_1)$$

(vgl. [31]) unmittelbar, daß das Produkt zweier Mengen, die zu einem Körper gehören, wieder in diesem Körper liegt und daß die Summe zweier Mengen, die in einem Körper liegen, wieder in diesem Körper liegt; dies gilt auch in dem Fall, daß die Mengen gemeinsame Punkte haben. Diese Behauptung gilt offenbar auch für eine beliebige endliche Anzahl von Faktoren und Summanden, d. h., *Summe und Produkt endlich vieler Mengen, die zu einem Körper gehören, gehören auch wieder zu diesem Körper.*

Wir verschärfen jetzt noch die Forderung 2 der Definition des Mengenkörpers, und zwar fordern wir, daß zu diesem Körper die Summe abzählbar vieler paarweise punktfremder Mengen gehört, die zu dem Körper gehören. Solche Mengenkörper heißen *abgeschlossene Mengenkörper*. Wir können also sagen: Eine Mengenfamilie heißt *abgeschlossener Mengenkörper*, wenn die beiden folgenden Eigenschaften gelten:

1. Wenn die Mengen $\mathcal{E}_1$ und $\mathcal{E}_2$ zu der Familie gehören, gehört auch ihre Differenz $\mathcal{E}_1 - \mathcal{E}_2$ zu dieser Familie.

2. Wenn die Familie endlich oder abzählbar unendlich viele paarweise punktfremde Mengen $\mathcal{E}_n$ enthält, enthält sie auch ihre Summe.

Wie oben können wir uns auch davon überzeugen, daß *ein abgeschlossener Mengenkörper alle endlichen Summen und Produkte zu ihm gehörender Mengen enthält.* Wir zeigen, daß *ein abgeschlossener Mengenkörper die Summen und Produkte abzählbar unendlich vieler zu ihm gehörender Mengen enthält.* Um uns davon zu überzeugen, schreiben wir die beiden folgenden Formeln auf:

$$\sum_{n=1}^{\infty} \mathcal{E}_n = \mathcal{E}_1 + (\mathcal{E}_2 - \mathcal{E}_1) + [\mathcal{E}_3 - (\mathcal{E}_1 + \mathcal{E}_2)] + [\mathcal{E}_4 - (\mathcal{E}_1 + \mathcal{E}_2 + \mathcal{E}_3)] + \dots, \quad (51)$$

$$\prod_{n=1}^{\infty} \mathcal{E}_n = \mathcal{E}_1 - [(\mathcal{E}_1 - \mathcal{E}_2) + (\mathcal{E}_1 - \mathcal{E}_3) + (\mathcal{E}_1 - \mathcal{E}_4) + \dots]. \quad (52)$$

Der Beweis dieser Formeln ist nicht schwer. Es genügt nachzuprüfen, daß jedes Element (jeder Punkt) einer Menge, welche links in den Formeln steht, auch in der rechts stehenden Menge enthalten ist, und umgekehrt. Es sei $\mathcal{E}_n$ in dem abge-

geschlossenen Mengenkörper T enthalten. Dann sind die Summanden, die auf der rechten Seite von (51) stehen, paarweise punktfremd und in T enthalten. Folglich ist dann nach Definition des abgeschlossenen Körpers T auch die Summe der Mengen $\mathfrak{E}_n$ in T enthalten. Die Summanden, die in eckigen Klammern auf der rechten Seite von (52) stehen, sind in T enthalten, daher ist auch ihre Summe in T enthalten. Folglich ist auch die ganze rechte Seite, d. h. das Produkt der Mengen $\mathfrak{E}_n$, in T enthalten, was zu beweisen war.

Aus den in [37] bewiesenen Sätzen folgt unmittelbar, daß L_G ein abgeschlossener Mengenkörper ist. Wir betrachten die Intervallfunktion $G(\Delta)$, die wir auf den abgeschlossenen Körper L_G erweitert haben. Die Familie der Intervalle selbst ist kein Körper, weil schon die Differenz zweier Intervalle kein Intervall mehr zu sein braucht. Die Familie der Elementarbereiche R ist dagegen ein Körper, jedoch kein abgeschlossener Körper. Unser Verfahren, die Funktion $G(\Delta)$ fortzusetzen, bestand darin, daß wir $G(\Delta)$ zuerst auf den Körper der Elementarbereiche erweiterten und dann auf den abgeschlossenen Körper L_G . Dabei erwies sich die Funktion $G(\mathfrak{E})$ in dem in [36] angegebenen Sinn als nichtnegativ, volladditiv und normal in L_G .

Wir erklären jetzt den Zusammenhang der Begriffe „normal“ und „additiv“ bei Mengenfunktionen. Gegeben sei ein beliebiger Körper T , der nicht abgeschlossen zu sein braucht. Eine Funktion $G(\mathfrak{E})$, die für alle Mengen aus T definiert ist, heißt *volladditiv* in T , wenn folgende Bedingung gilt: Ist die Menge $\mathfrak{E}$, die zu T gehört, die Summe von endlich oder abzählbar unendlich vielen paarweise punktfremden Mengen $\mathfrak{E}_n$, die ebenfalls zu T gehören, so gilt

$$G(\mathfrak{E}_1 + \mathfrak{E}_2 + \dots) = G(\mathfrak{E}_1) + G(\mathfrak{E}_2) + \dots$$

Wir hatten den Begriff einer volladditiven Funktion schon früher erwähnt. Zwischen den Begriffen der Volladditivität und Normalität von Funktionen besteht ein unmittelbarer Zusammenhang, der durch den folgenden Satz ausgedrückt wird:

Satz. Eine Funktion $G(\mathfrak{E})$, die in einem beliebigen Körper T definiert ist und nur endliche Werte annimmt, ist genau dann additiv und normal, wenn sie volladditiv ist.

Beweis. Aus der Additivität folgt unmittelbar, wenn $A \subset B$ ist,

$$G(B - A) = G(B) - G(A).$$

Die Funktion $G(\mathfrak{E})$ sei additiv und normal, und wir wollen beweisen, daß sie volladditiv ist. Wir nehmen an, daß $\mathfrak{E}$ die Summe von abzählbar unendlich vielen paarweise punktfremden Mengen $\mathfrak{E}_n$ ($n = 1, 2, \dots$) ist. Dann können wir

$$\mathfrak{E} = \mathfrak{E}_1 + \dots + \mathfrak{E}_n + [\mathfrak{E} - (\mathfrak{E}_1 + \mathfrak{E}_2 + \dots + \mathfrak{E}_n)]$$

und auf Grund der Additivität der Funktion

$$G(\mathfrak{E}) = G(\mathfrak{E}_1) + \dots + G(\mathfrak{E}_n) + G[\mathfrak{E} - (\mathfrak{E}_1 + \dots + \mathfrak{E}_n)] \quad (53)$$

schreiben. Gehen wir in (53) zur Grenze über und berücksichtigen die Normalität, so erhalten wir, da $\mathfrak{E} - (\mathfrak{E}_1 + \dots + \mathfrak{E}_n)$ eine verschwindende Folge ist,

$$G(\mathfrak{E}) = \lim_{n \rightarrow \infty} [G(\mathfrak{E}_1) + \dots + G(\mathfrak{E}_n)] = \sum_{n=1}^{\infty} G(\mathfrak{E}_n).$$

Damit ist bewiesen, daß $G(\mathfrak{E})$ volladditiv ist. Wir nehmen jetzt umgekehrt an, $G(\mathfrak{E})$ sei volladditiv, und beweisen, daß $G(\mathfrak{E})$ normal ist. Es sei $\mathfrak{E}'_1 \supset \mathfrak{E}'_2 \supset \dots$ eine beliebige verschwindende Folge. Wir müssen also zeigen, daß $G(\mathfrak{E}'_n) \rightarrow 0$ gilt. Es ist

$$\mathfrak{E}'_1 = (\mathfrak{E}'_1 - \mathfrak{E}'_2) + (\mathfrak{E}'_2 - \mathfrak{E}'_3) + \dots + (\mathfrak{E}'_{n-1} - \mathfrak{E}'_n) + \mathfrak{E}'_n,$$

wobei die einzelnen Summanden paarweise punktfremd sind, und hieraus erhalten wir infolge der Additivität von $G(\mathcal{E})$

$$G(\mathcal{E}'_n) = G(\mathcal{E}'_1) - [G(\mathcal{E}'_1 - \mathcal{E}'_2) + G(\mathcal{E}'_2 - \mathcal{E}'_3) + \dots + G(\mathcal{E}'_{n-1} - \mathcal{E}'_n)] . \quad (54)$$

Andererseits folgt aus der Formel

$$\mathcal{E}'_1 = \sum_{k=1}^{\infty} (\mathcal{E}'_k - \mathcal{E}'_{k+1})$$

und der Volladditivität von $G(\mathcal{E})$

$$G(\mathcal{E}'_1) = \sum_{k=1}^{\infty} G(\mathcal{E}'_k - \mathcal{E}'_{k+1}) = \lim_{n \rightarrow \infty} \sum_{k=1}^n G(\mathcal{E}'_k - \mathcal{E}'_{k+1}) ,$$

und wenn wir (54) berücksichtigen, erhalten wir $G(\mathcal{E}'_n) \rightarrow 0$, was zu beweisen war.

Oben setzten wir eine nichtnegative, additive und normale Intervallfunktion $G(\Delta)$ auf den abgeschlossenen Körper L_G fort, wobei sich die auf diese Weise erhaltene Funktion $G(\mathcal{E})$ als volladditiv herausstellte. Man kann beweisen, daß keine andere Fortsetzung von $G(\Delta)$ auf L_G unter der Bedingung der Volladditivität möglich ist.

39. Unabhängigkeit von der Wahl des Koordinatensystems. Wir machen noch einige Bemerkungen zur Unabhängigkeit des Maßes von der Wahl des Koordinatensystems. Die ursprüngliche Funktion $G(\Delta)$ war in einem halboffenen Rechteck definiert, dessen Seiten parallel zur x - und zur y -Achse waren.

Der Körper L_G enthält die halboffenen Rechtecke der Ebene mit beliebig gerichteten Seiten, weil jedes dieser Rechtecke die Differenz des abgeschlossenen Rechtecks und der abgeschlossenen Punktmenge ist, die aus zwei Seiten und drei Ecken des Rechtecks besteht. Somit ist die Funktion $G(\mathcal{E})$ insbesondere auf allen halboffenen Rechtecken Δ' definiert, deren Seiten parallel zu den Achsen eines kartesischen x', y' -Koordinatensystems sind, wobei die Funktion $G(\Delta')$ in diesem Rechteck additiv und normal ist. Wenn wir, indem wir ein neues x', y' -System wählen, von der Funktion $G(\Delta')$ ausgehen und sie so fortsetzen, wie oben gezeigt wurde, erhalten wir einen Körper $L_{G'}$. Es ist leicht zu zeigen, daß dieser Körper mit dem Körper L_G übereinstimmt und daß wir bei der neuen Fortsetzung, von $G(\Delta')$ ausgehend, in allen Intervallen jene Werte $G(\mathcal{E})$ erhalten, die wir auch bei der früheren Fortsetzung erhielten, die wir, ausgehend von $G(\Delta)$, hergeleitet haben. Das ergibt sich daraus, daß jede offene Menge O als Summe der paarweise punktfremden Intervalle Δ_k oder als Summe der paarweise punktfremden Intervalle Δ'_k dargestellt werden kann, wobei

$$G(O) = \sum_{k=1}^{\infty} G(\Delta_k) = \sum_{k=1}^{\infty} G(\Delta'_k)$$

gilt.

Nach Satz 4 aus [34] folgt hieraus unmittelbar, daß jede Menge in den beiden Koordinatensystemen dasselbe äußere Maß hat. Aus der Definition der Meßbarkeit folgt weiter unmittelbar, daß eine Menge in beiden Koordinatensystemen gleichzeitig meßbar ist, d. h., der Körper $L_{G'}$ stimmt mit dem Körper L_G überein. Daß das Maß auf beiden Systemen übereinstimmt, folgt sofort daraus, daß nach dem oben angegebenen Satz $G(\mathcal{E})$ die untere Grenze der Maße der offenen Mengen ist, die $\mathcal{E}$ überdecken. Wir bemerken noch: Wenn $G(\Delta)$ der gewöhnliche Flächeninhalt des Rechtecks Δ ist, d. h. das Lebesguesche Maß $m(\Delta)$, dann ist $G(\Delta')$ der gewöhnliche Flächeninhalt von Δ' (vgl. [II, 96]).

40. Der Körper B . Wie wir schon erwähnten, hängt der abgeschlossene Körper L_G von der Wahl der Funktion $G(A)$ ab. Wir geben jetzt einen abgeschlossenen Körper mit der Eigenschaft an, daß die Mengen, die in ihm liegen, zu jedem abgeschlossenen Körper L_G und insbesondere zu L gehören. Wir betrachten alle abgeschlossenen Körper T , von denen jeder einzelne alle abgeschlossenen Intervalle enthält, und bilden die Mengenfamilie B , die aus den Mengen besteht, die zu allen oben definierten abgeschlossenen Körpern T gehören. Wie man leicht sieht, ist die Mengenfamilie B ebenfalls ein abgeschlossener Körper. Gehören nämlich $\mathfrak{E}_1$ und $\mathfrak{E}_2$ zu B , so gehören sie zu allen abgeschlossenen Körpern T ; daher liegt die Differenz $\mathfrak{E}_1 - \mathfrak{E}_2$ ebenfalls in allen T , also auch in der Familie B . Analog kann man auch den zweiten Teil der Definition des abgeschlossenen Körpers verifizieren. *Somit ist der abgeschlossene Körper B ein Durchschnitt aller abgeschlossenen Körper, die alle möglichen abgeschlossenen Intervalle enthalten.* Dieser abgeschlossene Körper B ist in jedem L_G enthalten, weil L_G auch alle abgeschlossenen Intervalle enthält. Wie wir in [32] gesehen haben, kann jede offene Menge als Summe abzählbar vieler abgeschlossener Intervalle dargestellt werden; *somit enthält der abgeschlossene Körper B alle offenen Mengen.* Jede abgeschlossene Menge F ist das Komplement zu einer offenen Menge O , d. h., man kann sie als Differenz der ganzen Ebene (offene Menge) und der Menge O darstellen; *daher enthält der Körper B auch alle abgeschlossenen Mengen.* Der Mengenkörper B wurde schon vor LEBESGUE von dem französischen Mathematiker BOREL betrachtet. Die Mengen, die zum Körper B gehören, nennt man manchmal auch *B -meßbare* oder *Borel-meßbare Mengen*.

Den Körper B kann man auch noch auf andere Weise definieren: Wir sagen, die Menge $\mathfrak{E}$ gehört zum Körper B , wenn wir sie aus abgeschlossenen Intervallen durch die beiden folgenden Operationen, die endlich oder abzählbar oft angewendet werden, erhalten können: 1. Bildung der Summe endlich oder abzählbar unendlich vieler schon konstruierter Mengen; 2. Bildung des Produkts endlich oder abzählbar unendlich vieler schon konstruierter Mengen. Diese Definition verlangt einige Erläuterungen, auf die wir hier jedoch nicht eingehen werden. Wir werden auch nicht beweisen, daß die neue Definition des Körpers B der früheren gleichbedeutend ist. Dagegen beweisen wir jetzt abschließend zwei einfache Sätze.

Satz 1. *Ist $\mathfrak{E}$ eine beliebige Menge aus L_G , so existieren zwei Mengen $\mathfrak{E}_1$ und $\mathfrak{E}_2$ aus dem Körper B , für die*

$$\mathfrak{E}_1 \subset \mathfrak{E} \subset \mathfrak{E}_2 \quad \text{und} \quad G(\mathfrak{E}_1) = G(\mathfrak{E}_2) = G(\mathfrak{E}) \quad (55)$$

gilt.

Beweis. Wir wissen, daß für eine Menge $\mathfrak{E}$, die zu L_G gehört, solche abgeschlossenen Mengen F_n und offenen Mengen O_n existieren, daß

$$F_n \subset \mathfrak{E} \subset O_n, \quad G(\mathfrak{E} - F_n) \leq \frac{1}{n}, \quad G(O_n - \mathfrak{E}) \leq \frac{1}{n} \quad (56)$$

ist. Die Mengen F_n und O_n liegen im Körper B . Folglich kann man auf Grund der Definition des abgeschlossenen Körpers behaupten, daß der Körper B die Summe der Mengen F_n und das Produkt der Mengen O_n enthält:

$$\mathfrak{E}_1 = \sum_{n=1}^{\infty} F_n, \quad \mathfrak{E}_2 = \prod_{n=1}^{\infty} O_n. \quad (57)$$

Beachten wir $F_n \subset \mathfrak{E} \subset O_n$, so können wir $\mathfrak{E}_1 \subset \mathfrak{E} \subset \mathfrak{E}_2$ behaupten. Außerdem ist $\mathfrak{E} - \mathfrak{E}_1 \subset \mathfrak{E} - F_n$ und $\mathfrak{E}_2 - \mathfrak{E} \subset O_n - \mathfrak{E}$ und folglich $G(\mathfrak{E} - \mathfrak{E}_1) \leq 1/n$ und $G(\mathfrak{E}_2 - \mathfrak{E}) \leq 1/n$

für beliebiges n . Die linke Seite hängt nicht von n ab, und deshalb folgt aus diesen Ungleichungen die Gleichheit (55); damit ist der Satz bewiesen.

Man kann ihn auch auf folgende Weise formulieren: *Jede Menge $\mathcal{S}$ aus L_G kann zwischen Mengen aus B eingeschlossen werden, die dasselbe Maß wie die gegebene Menge $\mathcal{S}$ haben.*

Wir wählen jetzt aus dem Körper B eine Mengenfamilie aus, die wir im weiteren benutzen werden.

Definition. Eine Menge $\mathcal{S}$ heißt G_δ -Menge, wenn sie eine offene Menge oder das Produkt abzählbar unendlich vieler offener Mengen ist.

Zuerst bemerken wir, daß, wie wir auch schon in [32] erwähnt haben, das Produkt abzählbar unendlich vieler offener Mengen keine offene Menge zu sein braucht. Aus der Definition der G_δ -Mengen folgt unmittelbar, daß das Produkt endlich oder abzählbar unendlich vieler G_δ -Mengen auch eine G_δ -Menge ist. Wir zeigen, daß jedes endliche abgeschlossene Intervall $\Delta [a \leq x \leq b, c \leq y \leq d]$ eine G_δ -Menge ist. In der Tat können wir es als Produkt offener Intervalle $\Delta_n (a - \varepsilon_n < x < b + \varepsilon_n, c - \varepsilon_n < y < d + \varepsilon_n)$ darstellen, wobei ε_n eine Folge positiver Zahlen ist, die gegen 0 strebt. Es ist nicht schwer zu zeigen, daß jede abgeschlossene Menge eine G_δ -Menge ist. Im weiteren werden wir dies nicht gebrauchen. Aus dem Beweis des obigen Satzes kann man unmittelbar folgende Behauptung herleiten:

Satz 2. *Jede meßbare Menge $\mathcal{S}$ kann so von einer Menge H vom Typ G_δ überdeckt werden, daß $G(\mathcal{S}) = G(H)$ ist.*

Bemerkung. *Liegt die Menge $\mathcal{S}$ in einem abgeschlossenen Intervall Δ , so kann die Überdeckungsmenge H so gewählt werden, daß auch sie in Δ liegt.* Ist nämlich H eine G_δ -Menge, die $\mathcal{S}$ überdeckt und der Bedingung $G(\mathcal{S}) = G(H)$ genügt, so ist die Menge $H' = H \Delta$ ebenfalls eine G_δ -Menge, die $\mathcal{S}$ überdeckt und die der Bedingung $G(\mathcal{S}) = G(H')$ genügt, wobei H' offenbar in Δ liegt.

41. Der Fall einer Veränderlichen. Im Fall einer Veränderlichen wird die Maßtheorie einfacher. Eine nichtnegative, additive und normale Funktion halboffener Intervalle kann, wie wir wissen, mit Hilfe einer nichtfallenden Punktfunktion $g(x)$ dargestellt werden:

$$G(\Delta) = G((a, b]) = g(b+0) - g(a+0).$$

Geht man von $G(\Delta)$ aus, so erhält man, wie schon oben bewiesen wurde, die Mengenfunktion $G(\mathcal{S})$, die für alle in L_G liegenden Mengen definiert ist. Den Wert von $G(\mathcal{S})$ für eine Menge $\mathcal{S}$, die in L_G liegt, nennt man die *Variation von $g(x)$ auf der Menge $\mathcal{S}$* . Ist $g(x) = x$, so erhalten wir die Lebesgue-meßbaren Mengen, und $G(\mathcal{S})$ ist die Verallgemeinerung des Längenbegriffs auf diese Mengen. Ist $g(x)$ nur in einem Intervall definiert, so läßt sich diese Funktion, wie wir früher bewiesen haben, auf die ganze Achse fortsetzen.

Wir führen jetzt an Stelle von x eine neue Variable t durch die Formel

$$t = g(x) \tag{58}$$

ein, wobei man diese Variablensubstitution so zu verstehen hat: Ist in einem Punkt x die Funktion $g(x)$ stetig, so ist der entsprechende Wert für t durch (58) gegeben. Ist x eine Unstetigkeitsstelle, so ordnen wir dem Punkt x das abgeschlossene Intervall $[g(x-0), g(x+0)]$ der Veränderlichen t zu. Dabei geht das entsprechende halboffene Intervall $(a, b]$ der Veränderlichen x in das halboffene Intervall

$(g(a+0), g(b+0)]$ der Veränderlichen t über, wobei im Fall $g(b+0) = g(a+0)$ dieses halboffene Intervall in einen Punkt ausartet.

Sind e_x Mengen auf der x -Achse und e_t die entsprechenden Mengen auf der t -Achse, so kann man zeigen, daß das äußere Maß von e_x bezüglich $g(x)$ gleich dem äußeren Maß von e_t im Lebesgueschen Sinne ist, d. h. unter der Bedingung, daß man den Begriff der Länge eines halboffenen Intervalls zugrunde legt. Ein Elementarbereich ist im Fall einer Veränderlichen die Summe endlich vieler halboffener, paarweise punktfremder Intervalle, und wenn wir Satz 1 aus [35] benutzen würden, könnten wir leicht die folgende Behauptung beweisen: Ist e_x eine bezüglich $g(x)$ meßbare Menge, so ist auch e_t meßbar, und zwar im Lebesgueschen Sinne, und das Maß von e_x bezüglich $g(x)$ ist gleich dem Lebesgueschen Maß der Menge e_t .

§ 2. Meßbare Funktionen

42. Die Definition der meßbaren Funktionen. In diesem und dem folgenden Abschnitt werden wir eine Funktionenklasse konstruieren und die Eigenschaften dieser Funktionen untersuchen. Im weiteren werden wir auf Grund dieser Funktionenklasse einen allgemeinen Integralbegriff definieren. Bei der Darlegung werden wir annehmen, daß die Funktion $G(A)$, die der Maßtheorie zugrunde liegt, beliebig, aber fest ist, d. h., wir werden einen bestimmten Körper L_G betrachten. Das kann z. B. der Körper L der Lebesgue-meßbaren Mengen sein. Es sei auf der meßbaren Menge $\mathcal{E}$ eine Punktfunktion $f(P)$ gegeben, die nur reelle Werte annimmt. Diese Werte können endlich oder unendlich sein, d. h., die Funktion $f(P)$ kann außer allen endlichen Werten auch die Werte ∞ und $-\infty$ annehmen. Wir führen folgende Bezeichnung ein: Wir bezeichnen mit $\mathcal{E}[f > a]$ die Menge der Punkte aus $\mathcal{E}$, für die $f(P) > a$ ist. Analog bezeichnet $\mathcal{E}[f \leq a]$ die Menge der Punkte aus $\mathcal{E}$, für die $f(P) \leq a$ ist. Sind $f(P)$ und $g(P)$ zwei Funktionen, so bezeichnet man mit $\mathcal{E}[f = g]$ die Menge der Punkte von $\mathcal{E}$, für die $f(P) = g(P)$ ist, usw.

Definition. Die Funktion $f(P)$, die auf einer meßbaren Menge $\mathcal{E}$ gegeben ist, heißt *meßbar*, wenn für jedes reelle a die Mengen

$$\mathcal{E}[f \geq a], \quad \mathcal{E}[f < a], \quad \mathcal{E}[f > a], \quad \mathcal{E}[f \leq a] \quad (1)$$

meßbar sind.

Zuerst werden wir den folgenden Satz beweisen:

Satz 1. *Für die Meßbarkeit der Mengen (1) bei beliebigem a ist hinreichend, daß eine dieser Mengen für beliebiges a meßbar ist.*

Beweis. Die Mengen $\mathcal{E}[f \geq a]$ und $\mathcal{E}[f < a]$ sind komplementär, und wenn eine von ihnen für beliebiges a meßbar ist, ist es auch die andere. Dasselbe gilt für die Meßbarkeit der dritten und der vierten Menge von (1). Wir beweisen z. B., daß aus der Meßbarkeit der dritten Menge für beliebiges a auch die Meßbarkeit der übrigen Mengen folgt. In der Tat folgt aus der Meßbarkeit der dritten Menge die der vierten Menge und die der Menge $\mathcal{E}[f \geq a]$, die in der Form

$$\mathcal{E}[f \geq a] = \bigcap_{n=1}^{\infty} \mathcal{E}\left[f > a - \frac{1}{n}\right]$$

dargestellt werden kann, und damit die der zweiten Menge.

Wir bemerken noch, daß die Mengen $\mathfrak{E} [f = \infty]$ und $\mathfrak{E} [f = -\infty]$ in der Form

$$\mathfrak{E} [f = \infty] = \prod_{n=1}^{\infty} \mathfrak{E} [f > n], \quad \mathfrak{E} [f = -\infty] = \prod_{n=1}^{\infty} \mathfrak{E} [f < -n]$$

dargestellt werden können.

Wir bemerken weiter, daß die Meßbarkeit von (1) nur für rationale Werte a bewiesen werden muß. Jede irrationale Zahl a kann nämlich als Grenzwert einer fallenden Folge von rationalen Zahlen a_n dargestellt werden, und die Meßbarkeit von $\mathfrak{E} [f > a]$ folgt unmittelbar aus

$$\mathfrak{E} [f > a] = \sum_{n=1}^{\infty} \mathfrak{E} [f > a_n].$$

Wir erklären nun eine Reihe einfacher Eigenschaften der meßbaren Funktionen, die unmittelbar aus der oben gegebenen Definition folgen.

Satz 2. *Ist eine Funktion $f(P)$ auf $\mathfrak{E}$ meßbar, so ist sie auch auf einer beliebig gegebenen Teilmenge $\mathfrak{E}'$ von $\mathfrak{E}$ meßbar. Ist die Funktion $f(P)$ auf endlich oder abzählbar unendlich vielen paarweise punktfremden Mengen $\mathfrak{E}_n$ meßbar, so ist sie auch auf der Summe $\mathfrak{E}$ aller $\mathfrak{E}_n$ meßbar.*

Die Behauptung dieses Satzes folgt unmittelbar aus den Formeln

$$\mathfrak{E}' [f > a] = \mathfrak{E} [f > a] \cdot \mathfrak{E}', \quad \mathfrak{E} [f > a] = \sum_n \mathfrak{E}_n [f > a].$$

Satz 3. *Ist $\mathfrak{E}$ eine Menge vom Maß 0, so ist jede Funktion $f(P)$ auf dieser Menge meßbar.*

In der Tat ist für jedes a die Menge $\mathfrak{E} [f > a]$ eine Teilmenge von $\mathfrak{E}$, die das Maß 0 hat, und folglich hat auch die Menge $\mathfrak{E} [f > a]$ das Maß 0, d. h. ist meßbar.

Definition. Zwei Funktionen $f(P)$ und $g(P)$, die auf der Menge $\mathfrak{E}$ definiert sind, heißen *äquivalent auf dieser Menge* oder einfach *äquivalent*, wenn die Menge $\mathfrak{E} [f \neq g]$ das Maß 0 hat.

Wir beweisen für äquivalente Funktionen den folgenden

Satz 4. *Wenn $f(P)$ und $g(P)$ äquivalente Funktionen auf der meßbaren Menge $\mathfrak{E}$ sind und eine von ihnen meßbar ist, ist auch die andere Funktion meßbar.*

Nach der Voraussetzung des Satzes ist die Menge $\mathfrak{E} [f \neq g] = A$ eine Menge vom Maß 0. Auf der meßbaren Menge $\mathfrak{E}' = \mathfrak{E} - A$ gilt $f(P) = g(P)$. Aus der Meßbarkeit von $f(P)$ auf $\mathfrak{E}$ folgt auch die Meßbarkeit von $f(P)$ auf $\mathfrak{E}'$ und daraus die Meßbarkeit von $g(P)$ auf $\mathfrak{E}'$. Nach Satz 3 ist die Funktion $g(P)$ auf der Menge A meßbar. Somit ist $g(P)$ nach Satz 2 auf der Menge $\mathfrak{E} = \mathfrak{E}' + A$ meßbar, und der Satz ist bewiesen.

Man beweist leicht: Ist f_1 äquivalent zu g_1 , ferner f_2 äquivalent zu g_2 , so ist $f_1 + f_2$ äquivalent zu $g_1 + g_2$, ferner $f_1 f_2$ äquivalent zu $g_1 g_2$ und schließlich $f_1 : f_2$ äquivalent zu $g_1 : g_2$, wenn die betreffenden Operationen fast überall Sinn haben (zu „fast überall“ siehe S. 115).

Wenn zwei stetige Funktionen im Sinne des Lebesgueschen Maßes auf einem Intervall oder in der ganzen Ebene äquivalent sind, so sieht man leicht, daß ihre Werte in allen Punkten übereinstimmen. Würde nämlich z. B. in einem Punkt $f(P_0) - g(P_0) > 0$ gelten, so wäre diese Ungleichung infolge der Stetigkeit der Funktionen in einer hinreichend kleinen ε -Umgebung δ des Punktes P_0 gültig, wobei $m(\delta) > 0$ wäre, und das widerspricht der Definition der Äquivalenz von Funktionen.

Wir geben jetzt einige einfache Beispiele meßbarer Funktionen an. Wir nehmen an, daß $f(P)$ in dem endlichen abgeschlossenen Intervall Δ_0 stetig ist. Wir betrachten die Menge $\Delta_0 [f(P) \geq a]$ für beliebiges a und zeigen, daß sie abgeschlossen ist. Hieraus folgt dann ihre Meßbarkeit, somit ist $f(P)$ eine meßbare Funktion. Ist P_n ($n=1, 2, \dots$) eine Folge von Punkten, die als Grenzwert den Punkt P hat, und ist $f(P_n) \geq a$, so ist infolge der Stetigkeit der Funktion auch $f(P) \geq a$, d. h., die Menge $\Delta_0 [f(P) \geq a]$ ist abgeschlossen. Genauso zeigt man, daß $f(P)$ meßbar ist, wenn $f(P)$ auf der ganzen Ebene stetig ist. Ist nämlich Δ_0 ein beliebiges abgeschlossenes Intervall, so ist, wie wir schon zeigten, die Menge $\Delta_0 [f(P) \geq a]$ meßbar. Bei Vergrößerung von Δ_0 ist auch die Grenzmenge meßbar. Diese Grenzmenge ist die Menge aller derjenigen Punkte der Ebene, für die $f(P) \geq a$ gilt.

Wir nehmen jetzt an, daß $f(P)$ eine Sprungstelle P_0 hat. Wir überdecken sie durch eine Folge offener Intervalle Δ_n ($n=1, 2, \dots$), die sich auf P_0 zusammenziehen. Außerhalb Δ_n ist $f(P)$ stetig, und die Menge e_n derjenigen Punkte P , für die $f(P) \geq a$ ist, ist abgeschlossen. Wächst n , so nehmen die Mengen e_n nicht ab und streben gegen eine meßbare Menge e . Zu dieser Menge muß noch P_0 hinzugefügt werden, wenn $f(P_0) \geq a$ ist, und wir erhalten auf diese Weise die Menge aller Punkte, in denen $f(P) \geq a$ ist, wobei diese Menge nach dem vorher Gesagten meßbar ist. Dieselbe Schlußweise kann man anwenden, wenn endlich viele Unstetigkeitsstellen gegeben sind, d. h., *eine Funktion mit endlich vielen Unstetigkeitsstellen ist meßbar.*

Wir geben den folgenden Satz ohne Beweis an: Wenn $f(P)$ auf dem abgeschlossenen Intervall Δ_0 endliche Werte annimmt und die Menge ihrer Unstetigkeitsstellen das Maß 0 hat, ist $f(P)$ auf Δ_0 meßbar. Diese Meßbarkeitsbedingung ist aber nur hinreichend, aber keineswegs notwendig, wie folgendes einfache Beispiel zeigt, bei dem jeder Punkt der Menge $\mathcal{E}$ Unstetigkeitsstelle, aber die Funktion meßbar ist. Im Intervall $[0, 1]$ sei die Funktion $f(x)$ folgendermaßen definiert: $f(x)=0$, wenn x eine rationale Zahl ist, und $f(x)=1$, wenn x irrational ist. Wir betrachten die Meßbarkeit nach LEBESGUE, d. h. den Fall, daß $G(\Delta)$ die Länge des Intervalls ist. Dabei ist das Maß jedes Punktes gleich 0. Die rationalen Punkte des Intervalls $[0, 1]$ bilden eine abzählbare Menge, und da das Maß volladditiv ist, hat die Menge der rationalen Punkte ebenfalls das Maß 0. Unsere Funktion $f(x)$ ist nur auf der Menge der rationalen Punkte, die das Maß 0 hat, von der Funktion, die auf dem ganzen Intervall identisch 1 ist, verschieden, d. h., $f(x)$ ist der Funktion äquivalent, die identisch 1 ist, und nach Satz 4 ist $f(x)$ meßbar. Man sieht leicht, daß jeder Punkt x_0 des Intervalls $[0, 1]$ eine Unstetigkeitsstelle von $f(x)$ ist. In der Tat befinden sich in jeder ε -Umgebung von x_0 rationale und irrationale x -Werte, d. h., in jeder ε -Umgebung von $x=x_0$ nimmt $f(x)$ die Werte 0 und 1 an, also ist x_0 eine Sprungstelle. In [45] werden wir den Zusammenhang der Begriffe „Meßbarkeit“ und „Stetigkeit“ zeigen.

Wir betrachten noch die auf einer meßbaren Menge $\mathcal{E}$ *stückweise konstanten* Funktionen, d. h. solche Funktionen $f(x)$, die auf $\mathcal{E}$ endlich oder abzählbar unendlich viele Werte c_k ($k=1, 2, \dots$) annehmen. Wenn jede Menge $\mathcal{E}_k$, auf der $f(P)=c_k$ ist, meßbar ist, folgt aus der Meßbarkeitsdefinition unmittelbar, daß *die stückweise konstante Funktion $f(P)$ auf $\mathcal{E}$ meßbar ist.* Wir führen dazu ein Beispiel an. Die Funktion $f(P)$ sei auf der meßbaren Menge $\mathcal{E}$ meßbar und auf der Komplementärmenge $C\mathcal{E}$ gleich 0. Diese Funktion ist auf $\mathcal{E}$ und auf $C\mathcal{E}$ meßbar, d. h., nach Satz 2 ist sie auf der ganzen Ebene meßbar.

Wir betrachten noch den Fall einer Veränderlichen. Es sei $g(x)$ die nichtabnehmende Funktion, die dem Maß zugrunde gelegt ist, und $f(x)$ sei eine in diesem Sinne

meßbare Funktion. In diesem Fall sagt man manchmal, daß $f(x)$ bezüglich $g(x)$ meßbar ist; ist $g(x) = x$, so sagt man einfach, daß $f(x)$ meßbar ist.

43. Eigenschaften der meßbaren Funktionen. Wir behandeln jetzt einige weitere Eigenschaften der meßbaren Funktionen.

Satz 1. Ist $f(P)$ eine meßbare Funktion, so ist auch $|f(P)|$ eine meßbare Funktion.

Die Behauptung dieses Satzes folgt unmittelbar aus

$$\mathfrak{E}[|f| > a] = \mathfrak{E}[f > a] + \mathfrak{E}[f < -a].$$

Satz 2. Ist $f(P)$ eine meßbare Funktion und c eine endliche und von 0 verschiedene Konstante, so sind $c + f(P)$ und $cf(P)$ meßbare Funktionen.

Die erste Behauptung folgt unmittelbar aus

$$\mathfrak{E}[c + f(P) > a] = \mathfrak{E}[f(P) > a - c],$$

die zweite aus

$$\mathfrak{E}[cf(P) > a] = \mathfrak{E}\left[f(P) > \frac{a}{c}\right] \quad \text{für } c > 0,$$

$$\mathfrak{E}[cf(P) > a] = \mathfrak{E}\left[f(P) < \frac{a}{c}\right] \quad \text{für } c < 0.$$

Satz 3. Sind $f(P)$ und $g(P)$ meßbare Funktionen, so ist die Menge $\mathfrak{E}[f > g]$ meßbar.

Numerieren wir alle rationalen Zahlen $r_1, r_2, \dots$, so folgt die Meßbarkeit der in dem Satz erwähnten Menge unmittelbar aus

$$\mathfrak{E}[f > g] = \sum_{k=1}^{\infty} \mathfrak{E}[f > r_k] \mathfrak{E}[g < r_k].$$

Satz 4. Sind $f(P)$ und $g(P)$ meßbare Funktionen, die nur endliche Werte annehmen, so sind die Funktionen $f - g$, $f + g$, fg und $\frac{f}{g}$ (für $g \neq 0$) meßbar.

Beweis. Die Meßbarkeit der Differenz $f - g$ folgt unmittelbar aus der Formel

$$\mathfrak{E}[f - g > a] = \mathfrak{E}[f > a + g]$$

und aus Satz 2 und 3. Die Meßbarkeit der Summe folgt aus $f + g = f - (-g)$ und aus Satz 2 für $c = -1$. Die Meßbarkeit von f^2 , wenn f selbst meßbar ist, folgt aus der Formel

$$\mathfrak{E}[f^2 > a] = \mathfrak{E}[f > \sqrt{a}] + \mathfrak{E}[f < -\sqrt{a}] \quad (a \geq 0),$$

die Meßbarkeit des Produkts fg aus der Formel

$$fg = \frac{1}{4} [(f+g)^2 - (f-g)^2].$$

Wir beweisen jetzt die Meßbarkeit der Funktion $\frac{1}{g}$ unter der Bedingung, daß g ungleich 0 ist. Die Behauptung folgt unmittelbar aus den Formeln

$$\mathfrak{E}\left[\frac{1}{g} > a\right] = \mathfrak{E}[g > 0] \mathfrak{E}\left[g < \frac{1}{a}\right] \quad \text{für } a > 0,$$

$$\mathfrak{E} \left[\frac{1}{g} > a \right] = \mathfrak{E} [g > 0] + \mathfrak{E} \left[g < \frac{1}{a} \right] \quad \text{für } a < 0,$$

$$\mathfrak{E} \left[\frac{1}{g} > a \right] = \mathfrak{E} [g > 0] \quad \text{für } a = 0.$$

Schließlich folgt aus $\frac{f}{g} = f \frac{1}{g}$ die Meßbarkeit des Quotienten.

Bei diesem Satz ist die zusätzliche Voraussetzung, daß die Funktionen in allen Punkten von $\mathfrak{E}$ endliche Werte haben, notwendig. Sonst könnten die Operationen ihren Sinn verlieren. Ist z. B. in einem Punkt $f = \infty$ und $g = -\infty$, so können wir in diesem Punkt nichts über die Summe aussagen. Tritt diese Unbestimmtheit bei der Durchführung der Operationen nicht auf, so kann man auch für $f(P)$ und $g(P)$ unendliche Werte zulassen. Als Beispiel beweisen wir jetzt folgenden

Satz 5. *Sind $f(P)$ und $g(P)$ meßbare Funktionen, die endliche Werte und den Wert ∞ annehmen können, so ist auch die Funktion $f + g$ meßbar.*

Es sei A eine Menge, auf der wenigstens eine der Funktionen gleich ∞ ist. Diese Menge ist infolge der Meßbarkeit von f und g meßbar, und auf der Menge A hat die Summe $f + g$ den konstanten Wert ∞ , ist also meßbar. Auf der Menge $\mathfrak{E}' = \mathfrak{E} - A$ haben die Funktionen f und g endliche Werte, und nach Satz 4 ist die Summe $f + g$ auch auf $\mathfrak{E}'$ meßbar. Daraus folgt, daß sie auch auf $\mathfrak{E} = \mathfrak{E}' + A$ meßbar ist, was zu beweisen war.

44. Der Grenzwert meßbarer Funktionen. In diesem Abschnitt untersuchen wir den Grenzübergang bei meßbaren Funktionen. Als grundlegendes Ergebnis werden wir erhalten, daß der Grenzübergang bei meßbaren Funktionen wieder zu meßbaren Funktionen führt. Zuerst erklären wir einige Tatsachen, die mit dem Begriff des Grenzwerts zusammenhängen. Gegeben sei eine Folge reeller Zahlen

$$a_1, a_2, a_3, \dots, \quad (2)$$

wobei unter diesen Zahlen auch Zahlen mit dem Wert ∞ oder $-\infty$ auftreten können. Wir werden unter s_n die untere Grenze der Zahlenmenge $[a_n, a_{n+1}, \dots]$ und unter t_n die obere Grenze dieser Menge verstehen, d. h.

$$s_n = \inf [a_n, a_{n+1}, \dots], \quad t_n = \sup [a_n, a_{n+1}, \dots]. \quad (3)$$

Bei wachsendem n wird diese Zahlenmenge keine zusätzlichen Elemente bekommen; folglich nimmt s_n nicht ab, und t_n wächst nicht. Somit haben die monotonen Folgen s_n und t_n bei unbegrenzt wachsendem n endliche oder unendliche Grenzwerte

$$\lim_{n \rightarrow \infty} s_n = S, \quad \lim_{n \rightarrow \infty} t_n = T, \quad (4)$$

wobei wir infolge der Monotonie

$$S = \sup s_n, \quad T = \inf t_n \quad (5)$$

erhalten; außerdem folgt aus $s_n \leq t_n$

$$S \leq T.$$

Wir bemerken dabei, daß wir als Grenzwert der Folge $\infty, \infty, \dots$ den Wert ∞ ansehen und analog für $-\infty$. Der Wert S heißt *unterer Limes* (*Limes inferior*) der Folge (2) und der Wert T *oberer Limes* (*Limes superior*) dieser Folge.

Oft benutzt man hierfür folgende Schreibweise :

$$S = \varliminf a_n \quad \text{oder} \quad S = \liminf_{n \rightarrow \infty} a_n,$$

$$T = \varlimsup a_n \quad \text{oder} \quad T = \limsup_{n \rightarrow \infty} a_n.$$

Wir beweisen jetzt den folgenden

Hilfssatz. *Für die Existenz des (endlichen oder unendlichen) Grenzwerts der Folge (2) ist notwendig und hinreichend, daß $S = T$ ist; wenn diese Bedingung erfüllt ist, ist dieser Grenzwert gleich S .*

Wir beweisen zuerst, daß dies hinreichend ist. Es gilt $s_n \leq a_k \leq t_n$ für $k \geq n$, und wenn die Grenzwerte s_n und t_n übereinstimmen, d. h. $S = T$ ist, strebt offenbar a_n gegen S . Wir beweisen jetzt die Notwendigkeit. Die Folge (2) habe einen endlichen Grenzwert σ . Dabei liegen für hinreichend großes n alle Zahlen a_n in dem Intervall $(\sigma - \varepsilon, \sigma + \varepsilon)$, wobei ε eine beliebig kleine vorgegebene positive Zahl ist. Daher liegen in diesem Intervall auch alle s_n und t_n für hinreichend großes n . Hieraus folgt, da ε beliebig war, daß $s_n \rightarrow \sigma$ und $t_n \rightarrow \sigma$, also $S = T = \sigma$ ist. Ganz analog kann man das für den Fall eines unendlichen Grenzwerts der Folge (2) beweisen.

Wir beweisen jetzt einige Eigenschaften für Folgen meßbarer Funktionen.

Satz 1. *Ist $f_n(P)$ eine Folge meßbarer Funktionen, so sind die untere und die obere Grenze der Werte von $f_n(P)$ in einem beliebigen Punkt P der Menge $\mathfrak{E}$ meßbare Funktionen, d. h., die Funktionen*

$$\varphi(P) = \inf_n f_n(P) \quad \text{und} \quad \psi(P) = \sup_n f_n(P) \quad (6)$$

sind meßbar.

Wir beweisen nur die Meßbarkeit der Funktion $\varphi(P)$. Gilt in einem Punkt $\varphi(P) < a$, so ist wenigstens einer der Werte $f_n(P)$ kleiner als a , und umgekehrt, wenn wenigstens einer der Werte $f_n(P)$ kleiner als a ist, ist $\varphi(P) < a$. Somit folgt

$$\mathfrak{E}[\varphi(P) < a] = \sum_{n=1}^{\infty} \mathfrak{E}[f_n(P) < a],$$

und hieraus ergibt sich infolge der Meßbarkeit der Funktionen $f_n(P)$, daß auch die Funktion $\varphi(P)$ meßbar ist.

Satz 2. *Ist die Folge der meßbaren Funktionen $f_n(P)$ in jedem Punkt P der Menge $\mathfrak{E}$ monoton wachsend (oder monoton fallend), so ist die Grenzfunktion $f(P)$ meßbar.*

Der Satz folgt unmittelbar aus dem vorhergehenden; denn für eine monoton wachsende Funktionenfolge stimmt die Grenzfunktion mit der oberen Grenze $\psi(P)$ überein und für eine monoton fallende Folge mit der unteren Grenze $\varphi(P)$.

Satz 3. *Für eine Folge meßbarer Funktionen $f_n(P)$ sind auch der untere Limes $S(P)$ und der obere Limes $T(P)$ dieser Folge meßbare Funktionen.*

Wir führen die Funktionen

$$s_n(P) = \inf [f_n(P), f_{n+1}(P), \dots], \quad t_n(P) = \sup [f_n(P), f_{n+1}(P), \dots]$$

ein. Nach Satz 1 sind sie für beliebiges n meßbar. Die Funktionen $S(P)$ und $T(P)$ sind die Grenzfunktionen der monotonen Folgen $s_n(P)$ und $t_n(P)$, und daraus folgt nach Satz 2, daß auch sie meßbare Funktionen sind.

Satz 4. Für eine Folge meßbarer Funktionen $f_n(P)$, die in jedem Punkt P einer Menge $\mathcal{E}$ konvergiert, ist die Grenzfunktion $f(P)$ meßbar.

Die Meßbarkeit von $f(P)$ folgt unmittelbar aus Satz 3, denn wenn in jedem Punkt der Grenzwert $f(P)$ existiert, stimmt er mit $S(P)$ und $T(P)$ überein. Der letzte Satz ist für die weiteren Betrachtungen grundlegend, und wir werden ihn nun etwas verallgemeinern.

Man sagt, eine Eigenschaft gilt *fast überall* auf $\mathcal{E}$, wenn sie in allen Punkten von $\mathcal{E}$ mit Ausnahme einer Punktmenge vom Maß 0 gilt.

Satz 5. Bilden die auf $\mathcal{E}$ meßbaren Funktionen $f_n(P)$ eine Folge, die fast überall auf $\mathcal{E}$ konvergiert, so ist die Grenzfunktion $f(P)$ meßbar.

Wir bemerken, daß die Grenzfunktion $f(P)$ auf einer Teilmenge A von $\mathcal{E}$ nicht definiert zu sein braucht, wobei A vom Maß 0 ist. Wir definieren $f(P)$ auf A beliebig. Die Folge der $f_n(P)$ konvergiert in allen Punkten der meßbaren Menge $\mathcal{E}' = \mathcal{E} - A$, und nach Satz 4 ist $f(P)$ auf $\mathcal{E}'$ meßbar. Außerdem ist sie auch auf A nach Satz 3 aus [42] meßbar. Folglich ist $f(P)$ auf der Menge $\mathcal{E} = \mathcal{E}' + A$ meßbar, und der Satz ist bewiesen.

Wir führen nun einen neuen Konvergenzbegriff für Funktionenfolgen ein.

Definition. Es seien $f_n(P)$ und $f(P)$ auf $\mathcal{E}$ meßbare Funktionen, die endliche Werte haben. Man sagt, die Folge der $f_n(P)$ *konvergiere auf $\mathcal{E}$ dem Maß nach gegen $f(P)$* , wenn für beliebig vorgegebenes positives ε das Maß $G(\mathcal{E}_n)$ der Menge $\mathcal{E}_n$ der Punkte, für die die Ungleichung $|f(P) - f_n(P)| \geq \varepsilon$ gilt, für unbeschränkt wachsendes n gegen 0 strebt.

In den beiden folgenden Sätzen werden wir den Zusammenhang zwischen der „Konvergenz fast überall“ und der „Konvergenz dem Maß nach“ erklären.

Satz 6. Es sei $\mathcal{E}$ eine meßbare Menge mit endlichem Maß und $f_n(P)$ eine Folge auf $\mathcal{E}$ meßbarer Funktionen, die fast überall auf $\mathcal{E}$ endliche Werte annehmen und fast überall auf $\mathcal{E}$ gegen eine Funktion $f(P)$ konvergieren, die gleichfalls fast überall auf $\mathcal{E}$ endliche Werte annimmt. Dann konvergiert die Folge der $f_n(P)$ auf $\mathcal{E}$ dem Maß nach gegen $f(P)$.

Beweis. Es sei ε eine gegebene positive Zahl, und wir führen die Punktmenge

$$\mathcal{E}_n = \mathcal{E} [|f(P) - f_n(P)| \geq \varepsilon]$$

ein. Wir müssen beweisen, daß $G(\mathcal{E}_n)$ gegen 0 strebt. Wir führen die Punkt Mengen, auf denen $f(P)$ und $f_n(P)$ unendliche Werte annehmen, und die Menge ein, auf der $f_n(P)$ nicht gegen $f(P)$ strebt:

$$A = \mathcal{E} [|f(P)| = +\infty], \quad A_n = \mathcal{E} [|f_n(P)| = +\infty],$$

$$B = \mathcal{E} [f_n(P) \text{ nicht } \rightarrow f(P)].$$

Nach der Voraussetzung des Satzes haben diese Mengen das Maß 0. Das gleiche gilt für ihre Summe [36]

$$C = A + \sum_{n=1}^{\infty} A_n + B,$$

also ist $G(C) = 0$. Wenn P_0 nicht in C liegt, so haben $f_n(P_0)$ und $f(P_0)$ endliche Werte, und es gilt $f_n(P_0) \rightarrow f(P_0)$. Wir führen nun die Mengen

$$R_n = \sum_{k=n}^{\infty} \mathcal{E}_k \quad \text{und} \quad S = \prod_{n=1}^{\infty} R_n \quad (7)$$

ein. Die Folge R_n ($n = 1, 2, \dots$) ist eine nichtwachsende Folge von Mengen endlichen Maßes, da $\mathcal{E}$ endliches Maß hat, und S ist die Grenzmenge der R_n , so daß

$$G(R_n) \rightarrow G(S) \quad (8)$$

gilt. Wir zeigen jetzt, daß $S \subset C$ ist, d. h., wir zeigen, daß P_0 nicht zu S gehört, wenn P_0 nicht in C liegt. In der Tat, wenn P_0 nicht zu C gehört, sind $f_n(P_0)$ und $f(P_0)$ endlich und $f_n(P_0) \rightarrow f(P_0)$, d. h., es existiert ein N , so daß $|f(P_0) - f_n(P_0)| < \varepsilon$ für $n > N$ gilt. Hieraus folgt, daß P_0 für $n > N$ nicht zu $\mathcal{E}_n$ gehört, d. h., P_0 gehört nicht zu R_n für $n > N$, und daraus folgt weiter, daß P_0 nicht zu S gehört. Also gilt $S \subset C$. Da $G(C) = 0$ ist, ist auch $G(S) = 0$, und nach (8) gilt $G(R_n) \rightarrow 0$. Da mnach der ersten Formel (7) $\mathcal{E}_n \subset R_n$ ist, folgt erst recht $G(\mathcal{E}_n) \rightarrow 0$, was zu beweisen war.

Bemerkung. Wir weisen noch darauf hin, daß wir die Menge C zu allen $\mathcal{E}_n$ hinzufügen können. Da $G(C) = 0$ ist, haben wir auch nach einer solchen Vereinigung $G(\mathcal{E}_n) \rightarrow 0$, und in allen Punkten der Menge $\mathcal{E} - \mathcal{E}_n$ gilt die Ungleichung $|f(P) - f_n(P)| < \varepsilon$.

Aus der „Konvergenz dem Maß nach“ folgt nicht die „Konvergenz fast überall“, aber es gilt der folgende

Satz 7. *Es sei $\mathcal{E}$ eine meßbare Menge endlichen Maßes, $f_n(P)$ und $f(P)$ seien auf $\mathcal{E}$ meßbare endliche Funktionen, wobei $f_n(P)$ dem Maß nach auf $\mathcal{E}$ gegen $f(P)$ strebt. Dann existiert eine Teilfolge $f_{n_k}(P)$, die fast überall auf $\mathcal{E}$ gegen $f(P)$ strebt.*

Wir wählen eine Folge von positiven Zahlen δ_k ($k = 1, 2, \dots$) mit $\delta_k \rightarrow 0$ für $k \rightarrow \infty$ und eine Folge solcher positiven Zahlen ε_k , daß die Reihe $\varepsilon_1 + \varepsilon_2 + \dots$ konvergiert. Infolge der Konvergenz dem Maß nach existiert eine solche unbegrenzt wachsende Folge von Indizes n_k , daß für die Mengen $\mathcal{E}_k = \mathcal{E} [|f(P) - f_{n_k}(P)| \geq \delta_k]$ die Ungleichung $G(\mathcal{E}_k) \leq \varepsilon_k$ erfüllt ist. Wir führen die Mengen

$$R_n = \sum_{k=n}^{\infty} \mathcal{E}_k, \quad S = \prod_{n=1}^{\infty} R_n$$

ein. Es ist leicht zu zeigen, daß $G(S) = 0$ ist. In der Tat ist

$$G(R_n) \leq \sum_{k=n}^{\infty} G(\mathcal{E}_k) \leq \sum_{k=n}^{\infty} \varepsilon_k;$$

die letzte Summe strebt infolge der Konvergenz der Reihe $\varepsilon_1 + \varepsilon_2 + \dots$ gegen 0 für $n \rightarrow \infty$. Wir zeigen jetzt, daß $f_{n_k}(P) \rightarrow f(P)$ auf der Menge $\mathcal{E} - S$ gilt. Wegen $G(S) = 0$ würde damit dann der Satz bewiesen sein. Es sei $P_0 \in \mathcal{E} - S$, also $P_0 \notin S$. Hieraus folgt, daß P_0 für alle hinreichend großen k nicht zu R_k gehört, und folglich gehört P_0 für alle hinreichend großen k nicht zu $\mathcal{E}_k$, d. h., es existiert ein solches N , für das $P_0 \notin \mathcal{E}_k$ für $k \geq N$ gilt. Erinnern wir uns der Definition von $\mathcal{E}_k$, so erhalten wir $|f(P_0) - f_{n_k}(P_0)| < \delta_k$ für $k \geq N$; hieraus folgt $f_{n_k}(P_0) \rightarrow f(P_0)$ wegen $\delta_k \rightarrow 0$ für $k \rightarrow \infty$.

Bemerkung. Man könnte offenbar wie auch in Satz 6 annehmen, daß $f_n(P)$ und $f(P)$ nur fast überall auf $\mathcal{E}$ endlich sind und $f_n(P)$, wenn man die Mengen A und A_n aus $\mathcal{E}$ herausnimmt, auf der Restmenge dem Maß nach gegen $f(P)$ konvergiert.

Es existiert ein Satz, der den Zusammenhang zwischen der „Konvergenz fast überall“ und der „gleichmäßigen Konvergenz“ erklärt. Dieser Satz wurde 1911 von D. F. EGOROW bewiesen. Wir werden diesen Satz im weiteren nicht benutzen und ihn deshalb nur formulieren.

Satz 8. *Es sei $\mathcal{E}$ eine meßbare Menge endlichen Maßes und $f_n(P)$ eine Folge auf $\mathcal{E}$ meßbarer Funktionen, die fast überall auf $\mathcal{E}$ endliche Werte annehmen und fast*

überall auf $\mathcal{S}$ gegen eine Funktion $f(P)$ konvergieren, die gleichfalls fast überall auf $\mathcal{S}$ endliche Werte annimmt. Für jedes gegebene positive ε existiert dann eine solche ganz in $\mathcal{S}$ liegende abgeschlossene Menge F , daß $G(\mathcal{S} - F) < \varepsilon$ ist und daß $f_n(P)$ auf F gleichmäßig gegen $f(P)$ konvergiert.

45. Die Eigenschaft C. Man kann zeigen, daß die Meßbarkeit einer Funktion gleichbedeutend ist mit einer anderen Eigenschaft, der sogenannten Eigenschaft C, die sich mit Hilfe des Begriffs der Stetigkeit definieren läßt.

Zuerst werden wir einige neue Begriffe einführen.

Eine Funktion $f(P)$, die auf einer abgeschlossenen Menge F definiert ist, heißt *im Punkt* $P_0 \in F$ *stetig*, wenn für ein beliebig vorgegebenes positives ε eine solche positive Zahl η existiert, daß $|f(P_0) - f(P)| \leq \varepsilon$ gilt, wenn $P \in F$ in der η -Umgebung des Punktes P_0 liegt. Die Funktion $f(P)$ heißt *auf der abgeschlossenen Menge* F *stetig*, wenn sie in jedem Punkt dieser Menge stetig ist. Wir bemerken, daß im Sinne dieser Definition der Stetigkeit in einem Punkt jede Funktion in einem isolierten Punkt P_0 der Menge stetig ist, d. h. in einem solchen Punkt, für den eine beliebige η -Umgebung keine Punkte von F außer P_0 selbst enthält. Entsprechend kann man die Stetigkeit einer Funktion auch auf einer beliebigen, nicht unbedingt abgeschlossenen Menge definieren. Wir führen nun noch einen neuen Begriff ein.

Definition. Man sagt, eine Funktion $f(P)$, die auf einer meßbaren Menge $\mathcal{S}$ definiert ist, besitze auf dieser Menge die *Eigenschaft C*, wenn für ein beliebig vorgegebenes positives ε eine abgeschlossene Menge $F \subset \mathcal{S}$ existiert, für die erstens $G(\mathcal{S} - F) \leq \varepsilon$ und zweitens $f(P)$ auf F stetig ist.

Die Äquivalenz der Eigenschaft C mit der Meßbarkeit wird durch folgenden Satz gesichert, der schon 1913 von N. N. LUSIN bewiesen wurde.

Satz. Wenn eine Funktion $f(P)$ auf einer meßbaren Menge $\mathcal{S}$ endlichen Maßes definiert ist und fast überall auf $\mathcal{S}$ endliche Werte annimmt, ist für die Meßbarkeit dieser Funktion notwendig und hinreichend, daß sie auf $\mathcal{S}$ die Eigenschaft C hat.

Da wir diesen Satz nicht benutzen werden, beweisen wir ihn hier auch nicht.

46. Stückweise konstante Funktionen. Wir werden jetzt eine Klasse von Funktionen definieren, die in theoretischen Untersuchungen oft benutzt wird.

Definition. Eine Funktion $f(P)$, die auf einer meßbaren Menge $\mathcal{S}$ definiert ist, heißt auf dieser Menge *stückweise konstant*, wenn sie auf dieser Menge nur endlich oder abzählbar unendlich viele Werte annimmt.

Es seien c_k ($k=1, 2, \dots$) verschiedene Werte, die die Funktion $f(P)$ auf $\mathcal{S}$ annehmen kann, wobei unter diesen Werten auch die Werte ∞ und $-\infty$ auftreten können. Für die Meßbarkeit von $f(P)$ ist offenbar notwendig und hinreichend, daß die Mengen $\mathcal{E}_k$ der Punkte, in denen $f(P)$ gleich c_k ist, für alle k meßbar sind [42].

Wir führen jetzt einen neuen Begriff ein. Wenn $\mathcal{S}'$ eine Punktmenge ist, so heißt die Funktion $\omega_{\mathcal{S}'}(P)$, die auf der ganzen Ebene so definiert ist, daß $\omega_{\mathcal{S}'}(P) = 1$ ist, wenn P zu $\mathcal{S}'$ gehört, und $\omega_{\mathcal{S}'}(P) = 0$, wenn P nicht zu $\mathcal{S}'$ gehört, die *charakteristische Funktion* dieser Menge. Eine stückweise konstante Funktion $f(P)$ ist eine Linearkombination charakteristischer Funktionen:

$$f(P) = \sum_k c_k \omega_{\mathcal{E}_k}(P), \quad (9)$$

wobei P zu $\mathcal{S}$ gehört. Da die $\mathcal{E}_k$ keine gemeinsamen Punkte haben (c_k seien verschiedene Zahlen), ist in der Summe (9) nur ein Summand ungleich 0, außer in dem Fall, daß das dem Punkt P entsprechende c_k gleich 0 ist. In diesem Fall sind dann alle Summanden gleich 0.

Die charakteristische Funktion $\omega_{\mathcal{E}}(P)$ ist offenbar dann und nur dann meßbar, wenn $\mathcal{E}'$ eine meßbare Menge ist.

Wir zeigen jetzt, daß man meßbare Funktionen als Grenzwert stückweise konstanter Funktionen erhalten kann. Wir beschränken uns dabei auf den Fall nichtnegativer Funktionen.

Satz 1. Für jede nichtnegative, beschränkte und auf einer meßbaren Menge $\mathcal{E}$ meßbare Funktion existiert eine wachsende Folge aus $\mathcal{E}$ nichtnegativer, meßbarer, stückweise konstanter Funktionen $f_n(P)$ mit endlich vielen Werten, die auf $\mathcal{E}$ gleichmäßig in jedem Punkt von $\mathcal{E}$ gegen $f(P)$ strebt.

Beweis. Da $f(P)$ beschränkt ist, existiert eine solche positive Zahl L , für die $0 \leq f(P) < L$ ist. Wir teilen das Intervall $[0, L]$ in 2^n gleiche Teile durch die Teilpunkte

$$x_k = k \frac{L}{2^n} \quad (k = 1, 2, \dots, 2^n - 1).$$

Wir betrachten die meßbaren Mengen

$$\mathcal{E}_k^{(n)} = \mathcal{E} \left[k \frac{L}{2^n} \leq f(P) < (k+1) \frac{L}{2^n} \right]$$

und konstruieren die Folge der Funktionen $f_n(P)$ auf folgende Weise:

$$f_n(P) = k \frac{L}{2^n} \quad \text{für} \quad P \in \mathcal{E}_k^{(n)}. \quad (10)$$

Wie man leicht sieht, erfüllt die Folge $f_n(P)$ alle Voraussetzungen des Satzes. Jede dieser Funktionen $f_n(P)$ nimmt auf $\mathcal{E}$ endlich viele Werte an. Weiter zerfällt beim Übergang von n zu $n+1$ jedes Intervall

$$\left[k \frac{L}{2^n}, (k+1) \frac{L}{2^n} \right]$$

in die beiden Intervalle

$$\left[2k \frac{L}{2^{n+1}}, (2k+1) \frac{L}{2^{n+1}} \right] \quad \text{und} \quad \left[(2k+1) \frac{L}{2^{n+1}}, (2k+2) \frac{L}{2^{n+1}} \right],$$

und somit zerfällt auch jede der Mengen $\mathcal{E}_k^{(n)}$ in zwei Mengen:

$$\mathcal{E}_k^{(n)} = \mathcal{E}_{2k}^{(n+1)} + \mathcal{E}_{2k+1}^{(n+1)}.$$

Auf der Menge $\mathcal{E}_{2k}^{(n+1)}$ hat die Funktion $f_{n+1}(P)$ denselben Wert $k \frac{L}{2^n}$ wie die Funktion $f_n(P)$ auf der Gesamtmenge $\mathcal{E}_k^{(n)}$, und auf der Menge $\mathcal{E}_{2k+1}^{(n+1)}$ ist die Funktion $f_{n+1}(P)$ gleich

$$k \frac{L}{2^n} + \frac{L}{2^{n+1}};$$

folglich ist die Funktionenfolge wachsend. Weiter haben wir auf einer beliebigen Menge $\mathcal{E}_k^{(n)}$

$$f_n(P) = k \frac{L}{2^n} \quad \text{und} \quad k \frac{L}{2^n} \leq f(P) < (k+1) \frac{L}{2^n}.$$

Somit ist

$$0 \leq f(P) - f_n(P) < \frac{L}{2^n}$$

in allen Punkten P , die zu $\mathcal{E}$ gehören, und hieraus folgt, daß die Folge der $f_n(P)$ auf $\mathcal{E}$ gleichmäßig gegen $f(P)$ strebt.

In dem folgenden Satz betrachten wir den Fall, daß $f(P)$ auch unbeschränkt sein kann.

Satz 2. *Für jede nichtnegative, endliche Werte annehmende und auf der Menge $\mathcal{E}$ meßbare Funktion $f(P)$ existiert eine wachsende Folge auf $\mathcal{E}$ nichtnegativer, stückweise konstanter Funktionen $f_n(P)$, die auf $\mathcal{E}$ gleichmäßig gegen $f(P)$ strebt.*

Beweis. In diesem Fall unterteilen wir das unendliche Intervall $[0, \infty)$ durch die Punkte

$$x_k = \frac{k}{2^n} \quad (k = 1, 2, 3, \dots).$$

Wir definieren wieder die Mengen $\mathcal{E}_k^{(n)} = \mathcal{E} \left[\frac{k}{2^n} \leq f(P) < \frac{k+1}{2^n} \right]$ und die Funktionen

$$f_n(P) = \frac{k}{2^n} \quad \text{für } P \in \mathcal{E}_k^{(n)}. \quad (11)$$

Genauso wie im vorigen Satz kann man zeigen, daß die Folge der $f_n(P)$ allen Voraussetzungen des Satzes genügt.

Im Fall, daß die Funktion $f(P)$ nicht beschränkt ist, nehmen die Funktionen $f_n(P)$ abzählbar unendlich viele Werte an. Wenn man auf die gleichmäßige Annäherung verzichtet, kann man sich auf stückweise konstante Funktionen mit endlich vielen Werten beschränken. Im folgenden Satz werden wir jedoch auch ∞ als möglichen Wert der Funktion $f(P)$ zulassen.

Satz 3. *Für jede nichtnegative, auf der meßbaren Menge $\mathcal{E}$ meßbare Funktion existiert eine wachsende Folge auf $\mathcal{E}$ nichtnegativer, stückweise konstanter Funktionen $\varphi_n(P)$ mit endlich vielen endlichen Werten, die in jedem Punkt von $\mathcal{E}$ gegen $f(P)$ strebt.*

Außer den Mengen $\mathcal{E}_k^{(n)}$ des vorhergehenden Satzes bilden wir noch die Menge $\mathcal{E}_0 = \mathcal{E} [f(P) = \infty]$ und führen die Funktionenfolge $\varphi_n(P)$ auf folgende Weise ein:

$$\varphi_n(P) = f_n(P) \quad \text{für } f_n(P) \leq n$$

und

$$\varphi_n(P) = n \quad \text{für } f_n(P) > n \quad \text{oder } P \in \mathcal{E}_0.$$

Man sieht leicht, daß die Funktionen $\varphi_n(P)$ allen Voraussetzungen des Satzes genügen. Im folgenden werden wir diese Sätze vielfach benutzen.

47. Die Baireschen Klassen. In [40] definierten wir einen abgeschlossenen Körper von Punktmenge mit der Eigenschaft, daß jede Menge, die in diesem Körper liegt, auch in jedem Mengenkörper L_G liegt. Völlig analog kann man eine Familie von Funktionen so definieren, daß jede Funktion dieser Familie bei beliebiger Wahl der meßbaren Funktionen $G(\mathcal{A})$ eine meßbare Funktion ist.

Definition. Eine Funktion $f(P)$, die auf der Menge $\mathcal{E}$ definiert und B -meßbar ist, heißt B -Funktion¹⁾, wenn die Mengen

$$\mathcal{E} [f \geq a], \quad \mathcal{E} [f < a], \quad \mathcal{E} [f > a], \quad \mathcal{E} [f \leq a]$$

für beliebige reelle Zahlen a sämtlich B -meßbar sind.

¹⁾ nach dem französischen Mathematiker R. BAIRE

Aus dieser Definition folgt unmittelbar, daß jede B -Funktion bei beliebiger Wahl von $G(\mathcal{A})$ meßbar ist. Man kann auch eine andere Definition der B -Funktion geben, die völlig analog zu der in [40] betrachteten Definition der Mengen ist, die B -meßbar sind. Wir untersuchen alle Mengenfamilien, die die beiden folgenden Eigenschaften haben: 1. Die Familie enthält alle Funktionen, die auf $\mathcal{E}$ stetig sind; 2. enthält die Familie eine Folge von Funktionen $f_n(P)$, die in jedem Punkt von $\mathcal{E}$ konvergiert, so enthält diese Familie auch die Grenzfunktion. Die Familie der Funktionen, die allen Funktionenfamilien mit diesen beiden Eigenschaften angehören, heißt *Familie der B -Funktionen*. Wir werden die Äquivalenz dieser Definition mit der oben gegebenen Definition jedoch nicht beweisen.

Wir gehen jetzt auf einige Einzelheiten ein, die mit der letzten Definition zusammenhängen. Jede stetige Funktion ist eine B -Funktion, und man sagt gewöhnlich, daß eine solche Funktion der *nullten* Klasse angehört. Ist eine Funktion $f(P)$ Grenzfunktion einer Folge stetiger Funktionen, die in jedem Punkt von $\mathcal{E}$ konvergiert, wobei die Funktion $f(P)$ selbst nicht stetig zu sein braucht, so sagt man, daß eine solche Funktion $f(P)$ der *ersten* Klasse angehört. Alle Funktionen der ersten Klasse sind auch B -Funktionen. Ist die Funktion $f(P)$ Grenzfunktion einer Folge von Funktionen der ersten Klasse, die in jedem Punkt von $\mathcal{E}$ konvergiert, wobei die Funktion $f(P)$ keine Funktion der ersten Klasse zu sein braucht, so sagt man, daß $f(P)$ in der *zweiten* Klasse liegt. Alle Funktionen der zweiten Klasse sind auch B -Funktionen. Analog definiert man die höheren Funktionenklassen, und alle Funktionen dieser Klassen sind B -Funktionen. Diese Konstruktion kann man noch weiter fortsetzen. Es sei $f(P)$ Grenzfunktion einer Folge von Funktionen $f_n(P)$, wobei jede Funktion $f_n(P)$ zu irgendeiner endlichen Klasse und $f(P)$ selbst zu keiner Klasse mit endlicher Nummer gehört. Hierfür sagt man, daß die Funktion $f(P)$ zur transfiniten Funktionenklasse ω gehört. Alle Funktionen dieser Klasse sind B -Funktionen. Weiter kann man so die transfinite Funktionenklasse $\omega + 1$ usw. definieren. Nach dieser Methode erhält man alle B -Funktionen. Diese Behauptung erfordert zusätzliche Erläuterungen über transfinite Zahlen, auf die wir hier jedoch nicht eingehen können.

Man kann zeigen, daß jede Funktion $f(P)$, die auf einer B -Menge $\mathcal{E}$ meßbar ist, auf dieser Menge einer B -Funktion $\varphi(P)$ äquivalent ist.

§ 3. Das Lebesguesche Integral

48. Das Integral beschränkter Funktionen. Wir werden jetzt die übliche Definition des Integrals beschränkter Funktionen angeben, die sich auf die Zerlegung in alle möglichen meßbaren Mengen stützt, und beweisen, daß jede meßbare beschränkte Funktion integrierbar ist. Es sei auf einer meßbaren Menge $\mathcal{E}$ mit endlichem Maß eine beschränkte Punktfunktion $f(P)$ gegeben, d. h., auf $\mathcal{E}$ sei $|f(P)| \leq L$, wobei L eine gewisse positive Zahl ist. Wir zerlegen $\mathcal{E}$ in endlich viele meßbare Untermengen $\mathcal{E}_k$, die paarweise punktfremd sind:

$$\mathcal{E} = \sum_{k=1}^n \mathcal{E}_k. \quad (1)$$

Es seien m_k und M_k die untere bzw. die obere Grenze von $f(P)$ auf $\mathcal{E}_k$. Wir bilden die Summen

$$s_\delta = \sum_{k=1}^n m_k G(\mathcal{E}_k), \quad S_\delta = \sum_{k=1}^n M_k G(\mathcal{E}_k), \quad (2)$$

wobei δ die Unterteilung (1) der Menge $\mathfrak{E}$ bedeutet. Die Summen s_δ und S_δ sind für alle Zerlegungen beschränkt, und zwar ist $|s_\delta|$ und $|S_\delta|$ höchstens gleich $L \cdot G(\mathfrak{E})$. Es sei weiter i die obere Grenze der Summen s_δ und I die untere Grenze der Summen S_δ für alle möglichen Zerlegungen von $\mathfrak{E}$ in endlich viele meßbare Mengen.

Definition. Ist $i=I$, so sagt man, daß $f(P)$ bezüglich $G(\mathfrak{E})$ auf der Menge $\mathfrak{E}$ integrierbar ist, und nimmt als Wert des Integrals die Größe i :

$$i = \int_{\mathfrak{E}} f(P) G(d\mathfrak{E}) .$$

Dieses Integral heißt *Lebesgue-Stieltjessches Integral*. Ist δ eine Zerlegung für (1) und δ' irgendeine andere Zerlegung

$$\mathfrak{E} = \sum_{k=1}^{n'} \mathfrak{E}'_k , \quad (3)$$

so ist das Produkt $\delta\delta'$ der Zerlegungen wieder eine Zerlegung, die aus allen möglichen Untermengen $\mathfrak{E}_k\mathfrak{E}'_i$ besteht. Diese Mengen sind paarweise punktfremd. Einige von ihnen können auch leer sein. Die Zerlegung (3) heißt *Fortsetzung der Zerlegung* (1), wenn jede Menge $\mathfrak{E}'_i$ eine Teilmenge einer der Mengen $\mathfrak{E}_k$ ist. Ist δ_2 eine Fortsetzung von δ_1 , so schreiben wir $\delta_2 \geq \delta_1$.

Außer den Summen (2) stellen wir, wie auch für das Stieltjessche Integral, die Summe

$$\sigma_\delta = \sum_{k=1}^n f(P_k) G(\mathfrak{E}_k) \quad (4)$$

auf, wo die P_k gewisse Punkte aus $\mathfrak{E}_k$ sind. Für die Größen s_δ , S_δ , σ_δ , i und I gilt alles in [3] Gesagte.

Wir geben jetzt eine solche Folge von Zerlegungen für beliebige, beschränkte und auf $\mathfrak{E}$ meßbare Funktionen $f(P)$ an, daß $S_\delta - s_\delta \rightarrow 0$ ist, σ_δ also einen endlichen Grenzwert hat. Daraus folgt, daß das Integral i von $f(P)$ bezüglich $G(\mathfrak{E})$ existiert und daß s_δ , S_δ und σ_δ für die erwähnte Folge von Zerlegungen gegen i streben [3].

Es sei jetzt eine beschränkte Funktion $f(P)$, die auf $\mathfrak{E}$ definiert und meßbar ist, gegeben, und es sei m die untere und M die obere Grenze der Werte von $f(P)$ auf $\mathfrak{E}$. Wir zerlegen das Variabilitätsintervall $[m, M]$ von $f(P)$ durch die Punkte y_k in Teilintervalle,

$$m = y_0 < y_1 < y_2 < \dots < y_{n-1} < y_n = M , \quad (5)$$

und es sei η der größte Wert aller Differenzen $y_k - y_{k-1}$. Wir definieren folgende Zerlegung δ der Menge $\mathfrak{E}$ in meßbare Teilmengen $\mathfrak{E}_k$:

$$\begin{aligned} \mathfrak{E}_1 &= \mathfrak{E} [y_0 \leq f(x) \leq y_1] , \\ \mathfrak{E}_k &= \mathfrak{E} [y_{k-1} < f(x) \leq y_k] \quad (k=2, 3, \dots, n) . \end{aligned} \quad (6)$$

Aus dieser Definition der Mengen $\mathfrak{E}_k$ erhält man unmittelbar $y_{k-1} \leq m_k$ und $y_k \leq M_k$ [3] und somit

$$\sum_{k=1}^n y_{k-1} G(\mathfrak{E}_k) \leq s_\delta \leq S_\delta \leq \sum_{k=1}^n y_k G(\mathfrak{E}_k) , \quad (7)$$

also um so mehr

$$\sum_{k=1}^n y_{k-1} G(\mathfrak{E}_k) \leq i \leq I \leq \sum_{k=1}^n y_k G(\mathfrak{E}_k) . \quad (8)$$

Wir betrachten jetzt die Differenz

$$\sum_{k=1}^n y_k G(\mathcal{E}_k) - \sum_{k=1}^n y_{k-1} G(\mathcal{E}_k) = \sum_{k=1}^n (y_k - y_{k-1}) G(\mathcal{E}_k). \quad (9)$$

Beachten wir, daß $y_k - y_{k-1} \geq 0$ und $G(\mathcal{E})$ additiv ist, so erhalten wir

$$0 \leq \sum_{k=1}^n y_k G(\mathcal{E}_k) - \sum_{k=1}^n y_{k-1} G(\mathcal{E}_k) \leq \eta G(\mathcal{E}), \quad (10)$$

und daraus folgt, daß bei $\eta \rightarrow 0$ auch diese Differenz gegen 0 strebt. Hieraus folgt nach (7) und (8) unmittelbar $i = I$ und $S_\delta - s_\delta \rightarrow 0$. Die Zerlegung (6) der Ausgangsmenge $\mathcal{E}$ in Teilmengen $\mathcal{E}_k$ heißt *Lebesguesche Zerlegung*. Sie wird bestimmt durch die Zerlegung (5) des Variabilitätsintervalls $[m, M]$ von $f(P)$. Die der Zerlegung (6) entsprechenden Summen in den Ungleichungen (7) und (8) nennen wir *Lebesguesche Summen*. Aus dem oben Gesagten erhält man folgenden

Hauptsatz. *Eine beschränkte meßbare Funktion $f(P)$, die auf einer meßbaren Menge $\mathcal{E}$ endlichen Maßes gegeben ist, ist auf $\mathcal{E}$ integrierbar, und der Wert des Integrals ist gleich dem Grenzwert der Lebesgueschen Summen oder der Summen σ_δ bei beliebiger Wahl der Punkte P_k für die Lebesguesche Zerlegung bei unbeschränkter Verfeinerung der Zerlegungen des Variabilitätsbereichs $[m, M]$ von $f(P)$ in Teilintervalle.*

Wir bemerken, daß die Summen σ_δ auch für eine beliebige Fortsetzung jener Zerlegung, von der im Hauptsatz die Rede ist, denselben Grenzwert haben. Da das Integral wie üblich als Grenzwert der Summen σ_δ definiert ist, besitzt es auch die üblichen Eigenschaften des Riemannsches Integrals und des klassischen Stieltjes-schen Integrals. Diese Eigenschaften werden wir in [49] beweisen.

Das oben definierte Integral nannten wir Lebesgue-Stieltjessches Integral. Wir sprechen einfach von *Lebesgueschen Integral* in dem Spezialfall, daß $G(\Delta)$ der Flächeninhalt des Intervalls Δ ist.

Wir sahen, daß jede beschränkte Funktion mit endlich vielen Unstetigkeitsstellen meßbar ist. Es sei eine solche Funktion $f(P)$ in einem endlichen abgeschlossenen Intervall Δ gegeben. Wir wissen, daß diese Funktion in dem Intervall Δ im Riemannschen Sinne integrierbar ist. Sie ist als beschränkte meßbare Funktion auch im Lebesgueschen Sinne integrierbar. Wir zeigen jetzt, daß das Lebesguesche Integral mit dem Riemannschen übereinstimmt. Um nämlich das Lebesguesche Integral zu erhalten, genügt es, eine beliebige Folge von Zerlegungen des Intervalls Δ in meßbare Mengen zu nehmen, für welche die Summen (4) einen wohldefinierten Grenzwert haben, der gerade das Lebesguesche Integral liefert. Da die Funktion im Riemannschen Sinne integrierbar ist, führt schon die Zerlegung von Δ in Intervalle bei unbeschränkter Verfeinerung der Teilintervalle auf einen bestimmten Grenzwert der Summen (4), und dieser Grenzwert ist dann gerade das Riemannsche Integral. Aus diesen Überlegungen folgt sofort, daß das Lebesguesche und das Riemannsche Integral übereinstimmen.

Wie schon LEBESGUE zeigte, ist für die Existenz des Riemannsches Integrals in dem Intervall Δ notwendig und hinreichend, daß $f(P)$ beschränkt ist und die Menge ihrer Unstetigkeitsstellen ein Lebesguesches Maß hat, das gleich 0 ist (vgl. [10]). Eine solche Funktion ist, wie wir oben gesehen haben, auch im Lebesgueschen Sinne meßbar. Daß das Lebesguesche und das Riemannsche Integral übereinstimmen, kann genauso wie oben bewiesen werden. Somit erhalten wir: Jede Funktion, die in einem endlichen abgeschlossenen Intervall im Riemannschen Sinne (eigentlich) integrierbar ist, ist dort auch im Lebesgueschen Sinne integrierbar, und die beiden Integrale stimmen überein.

49. Eigenschaften des Integrals. Wir geben jetzt einige grundlegende Eigenschaften des Lebesgue-Stieltjesschen Integrals an. In jedem der folgenden Sätze wird vorausgesetzt, daß $\mathbb{E}$ eine meßbare Menge endlichen Maßes ist.

1. Ist c eine Konstante, so gilt

$$\int_{\mathbb{E}} cG(d\mathbb{E}) = cG(\mathbb{E}). \quad (11)$$

Für jede Zerlegung δ haben die Summen s_δ und S_δ den Wert $cG(\mathbb{E})$, und hieraus folgt (11) [3].

2. Sind $f_1(P)$ und $f_2(P)$ beschränkt und meßbar auf $\mathbb{E}$, so ist

$$\int_{\mathbb{E}} [f_1(P) + f_2(P)] G(d\mathbb{E}) = \int_{\mathbb{E}} f_1(P) G(d\mathbb{E}) + \int_{\mathbb{E}} f_2(P) G(d\mathbb{E}). \quad (12)$$

Es seien δ_n und δ'_n Zerlegungen, für welche σ_{δ_n} für die Funktion $f_1(P)$ und $\sigma_{\delta'_n}$ für $f_2(P)$ als Grenzwerte die entsprechenden Integrale haben. Für die Zerlegung $\delta''_n = \delta_n \delta'_n$ haben die Summen $\sigma_{\delta''_n}$ sowohl für $f_1(P)$ als auch für $f_2(P)$ als Grenzwerte die entsprechenden Integrale, und wir erhalten (12) unmittelbar aus dem Satz über den Grenzwert einer Summe.

Im folgenden werden wir nicht mehr besonders betonen, daß die Funktionen meßbar und beschränkt sind.

$$3. \quad \int_{\mathbb{E}} \sum_{k=1}^p c_k f_k(P) G(d\mathbb{E}) = \sum_{k=1}^p c_k \int_{\mathbb{E}} f_k(P) G(d\mathbb{E}). \quad (13)$$

Einen konstanten Faktor unter dem Integralzeichen kann man vor das Integralzeichen ziehen. Das folgt unmittelbar daraus, daß man einen solchen Faktor in den Summen σ_{δ_n} ausklammern kann. Ferner hat man die Eigenschaft 2 mehrmals anzuwenden.

4. Ist $f(P) \geq 0$ auf $\mathbb{E}$, so ist

$$\int_{\mathbb{E}} f(P) G(d\mathbb{E}) \geq 0. \quad (14)$$

Alle Summen σ_{δ_n} sind nämlich nicht negativ.

5. Ist $f_1(P) \geq f_2(P)$, so ist

$$\int_{\mathbb{E}} f_1(P) G(d\mathbb{E}) \geq \int_{\mathbb{E}} f_2(P) G(d\mathbb{E}). \quad (15)$$

Es genügt, die Eigenschaft 4 auf die Differenz $f_1(P) - f_2(P)$ anzuwenden und dann die Eigenschaft 3 zu benutzen.

$$6. \quad \left| \int_{\mathbb{E}} f(P) G(d\mathbb{E}) \right| \leq \int_{\mathbb{E}} |f(P)| G(d\mathbb{E}). \quad (16)$$

Um das zu beweisen, genügt es, das Produkt der Zerlegungen (6) für f und $|f|$ zu nehmen und die analoge Ungleichung für die Summen σ_δ aufzuschreiben.

7. Ist $a \leq f(P) \leq b$ auf $\mathbb{E}$, so ist

$$aG(\mathbb{E}) \leq \int_{\mathbb{E}} f(P) G(d\mathbb{E}) \leq bG(\mathbb{E}). \quad (17)$$

Das ist unmittelbar aus den Eigenschaften 5 und 1 abzulesen.

8. Ist $|f(P)| \leq L$, so ist

$$\left| \int_{\mathbb{E}} f(P) G(d\mathbb{E}) \right| \leq LG(\mathbb{E}). \quad (18)$$

Aus der Voraussetzung folgt $-L \leq f(P) \leq +L$, und die Ungleichung (18) ist eine Folgerung aus der Eigenschaft 7.

9. Ist $\mathfrak{E} = \mathfrak{E}' + \mathfrak{E}''$, wobei $\mathfrak{E}'$ und $\mathfrak{E}''$ meßbar und punktfremd sind, so ist

$$\int_{\mathfrak{E}} f(P) G(d\mathfrak{E}) = \int_{\mathfrak{E}'} f(P) G(d\mathfrak{E}) + \int_{\mathfrak{E}''} f(P) G(d\mathfrak{E}). \quad (19)$$

Um das zu beweisen, genügt es, die Zerlegungen (6) der Mengen $\mathfrak{E}'$ und $\mathfrak{E}''$ zu bilden, für diese Mengen die σ_s zu bilden und ihre Summe zu nehmen. Diese Summe hat einen wohldefinierten Grenzwert; damit ist die Formel (19) bewiesen.

10. Zerlegt man $\mathfrak{E}$ in endlich oder abzählbar unendlich viele meßbare Mengen $\mathfrak{E}_k$, so ist

$$\int_{\mathfrak{E}} f(P) G(d\mathfrak{E}) = \sum_k \int_{\mathfrak{E}_k} f(P) G(d\mathfrak{E}). \quad (20)$$

Für endlich viele Summanden folgt die Formel durch mehrfache Anwendung der Eigenschaft 9. Wir betrachten also nur den Fall unendlich vieler Mengen $\mathfrak{E}_k$. Es sei $|f(P)| \leq L$. Wir können

$$\mathfrak{E} = \mathfrak{E}_1 + \mathfrak{E}_2 + \dots + \mathfrak{E}_n + R_n$$

schreiben, wobei $R_n = \mathfrak{E} - (\mathfrak{E}_1 + \mathfrak{E}_2 + \dots + \mathfrak{E}_n)$ offenbar eine verschwindende Folge von Mengen ist; folglich gilt $G(R_n) \rightarrow 0$ [36]. Wenn wir die Additivität für endlich viele Summanden benutzen, erhalten wir

$$\int_{\mathfrak{E}} f(P) G(d\mathfrak{E}) = \sum_{k=1}^n \int_{\mathfrak{E}_k} f(P) G(d\mathfrak{E}) + \int_{R_n} f(P) G(d\mathfrak{E}). \quad (21)$$

Das letzte Integral kann man abschätzen,

$$\left| \int_{R_n} f(P) G(d\mathfrak{E}) \right| \leq LG(R_n),$$

und die Formel (21) liefert wegen $G(R_n) \rightarrow 0$ in der Grenze die Formel (20). Diese Eigenschaft heißt die *Volladditivität des Integrals*.

11. Für beliebig gegebenes $\varepsilon > 0$ existiert ein $\eta > 0$ derart, daß

$$\int_e f(P) G(d\mathfrak{E}) \leq \varepsilon \quad (22)$$

ist, wenn $e \subset \mathfrak{E}$ und $G(e) \leq \eta$ gilt.

Diese Eigenschaft folgt aus der Abschätzung

$$\left| \int_e f(P) G(d\mathfrak{E}) \right| \leq LG(e).$$

Man spricht gewöhnlich von der *absoluten Stetigkeit des Integrals*.

12. Ist $\mathfrak{E}$ eine Menge vom Maß 0, d. h. $G(\mathfrak{E}) = 0$, so gilt für jede auf $\mathfrak{E}$ beschränkte Funktion f

$$\int_{\mathfrak{E}} f(P) G(d\mathfrak{E}) = 0.$$

Die Funktion $f(P)$ ist auf $\mathfrak{E}$ meßbar [42], und für jede Zerlegung sind die Summen s_s und S_s gleich 0.

13. Sind $f(P)$ und $g(P)$ auf $\mathfrak{E}$ äquivalent, so ist

$$\int_{\mathfrak{E}} f(P) G(d\mathfrak{E}) = \int_{\mathfrak{E}} g(P) G(d\mathfrak{E}). \quad (23)$$

Es sei A diejenige Teilmenge von $\mathfrak{E}$, auf der $f \neq g$ ist. A ist nach Voraussetzung vom Maß 0. Auf der Menge $\mathfrak{E}' = \mathfrak{E} - A$ stimmen die Funktionen $f(P)$ und $g(P)$ überein. Somit erhalten wir die beiden Gleichungen

$$\int_A f(P) G(d\mathfrak{E}) = \int_A g(P) G(d\mathfrak{E}) = 0,$$

$$\int_{\mathfrak{E}'} f(P) G(d\mathfrak{E}) = \int_{\mathfrak{E}'} g(P) G(d\mathfrak{E});$$

wenn wir diese addieren, erhalten wir (23).

14. Ist $f(P) \equiv 0$ auf $\mathfrak{E}$ und

$$\int_{\mathfrak{E}} f(P) G(d\mathfrak{E}) = 0, \quad (24)$$

so ist $f(P)$ äquivalent 0.

Wir müssen beweisen, daß das Maß der Menge $\mathfrak{E} [f > 0]$ gleich 0 ist. Diese Menge kann als Summe

$$\mathfrak{E} [f > 0] = \sum_{n=1}^{\infty} \mathfrak{E} \left[f > \frac{1}{n} \right]$$

dargestellt werden, und wenn ihr Maß positiv wäre, wäre auch das Maß wenigstens eines der Summanden der Menge positiv. Es sei z. B. das Maß der Menge $B =$

$= \mathfrak{E} \left[f > \frac{1}{n_0} \right]$ positiv. Wir zerlegen das Integral in zwei Summanden:

$$\int_{\mathfrak{E}} f(P) G(d\mathfrak{E}) = \int_B f(P) G(d\mathfrak{E}) + \int_{\mathfrak{E}-B} f(P) G(d\mathfrak{E}). \quad (25)$$

Wegen $f \geq 0$ ist der zweite Summand nichtnegativ. Auf der Menge B ist $f > \frac{1}{n_0}$, und folglich ist der erste Summand mindestens gleich $\frac{1}{n_0} G(B)$. Wegen $G(B) > 0$ ist somit die linke Seite der Formel positiv, das widerspricht aber (24).

15. Strebt eine Folge von Funktionen $f_n(P)$, die auf $\mathfrak{E}$ meßbar und gleichmäßig beschränkt sind (d. h. für eine bestimmte positive, nicht von n abhängige Zahl L gilt die Abschätzung $|f_n(P)| \leq L$), auf $\mathfrak{E}$ fast überall gegen eine Grenzfunktion $f(P)$, so ist

$$\lim_{n \rightarrow \infty} \int_{\mathfrak{E}} f_n(P) G(d\mathfrak{E}) = \int_{\mathfrak{E}} f(P) G(d\mathfrak{E}). \quad (26)$$

Die Grenzfunktion genügt fast überall auf $\mathfrak{E}$ der Ungleichung $|f(P)| \leq L$. Durch Übergang zu einer äquivalenten Funktion kann man erreichen, daß diese Ungleichung überall auf $\mathfrak{E}$ gilt. Das Integral über die meßbare und beschränkte Funktion $f(P)$ auf $\mathfrak{E}$ ist sinnvoll.

Wir bilden das Integral der Differenz $f(P) - f_n(P)$ und wenden darauf die Eigenschaft 6 an:

$$\left| \int_{\mathfrak{E}} [f(P) - f_n(P)] G(d\mathfrak{E}) \right| \leq \int_{\mathfrak{E}} |f(P) - f_n(P)| G(d\mathfrak{E}). \quad (27)$$

Es sei ε eine gegebene positive Zahl und $\mathfrak{E}_n$ die Menge der Punkte von $\mathfrak{E}$, in denen $|f(P) - f_n(P)| \geq \varepsilon$ ist. Nach Satz 6 aus [44] strebt $G(\mathfrak{E}_n)$ gegen 0. In den Punkten der Menge $\mathfrak{E} - \mathfrak{E}_n$ gilt die Ungleichung $|f(P) - f_n(P)| < \varepsilon$. Außerdem gilt in jedem Punkt P von $\mathfrak{E}$

$$|f(P) - f_n(P)| \leq |f(P)| + |f_n(P)| \leq 2L.$$

Das rechts stehende Integral in Formel (27) kann man in zwei Integrale aufspalten:

$$\int_{\mathfrak{E}} |f(P) - f_n(P)| G(d\mathfrak{E}) = \int_{\mathfrak{E}_n} |f(P) - f_n(P)| G(d\mathfrak{E}) + \int_{\mathfrak{E} - \mathfrak{E}_n} |f(P) - f_n(P)| G(d\mathfrak{E}).$$

Hieraus folgt nach dem oben Gesagten

$$\int_{\mathfrak{E}} |f(P) - f_n(P)| G(d\mathfrak{E}) \leq 2LG(\mathfrak{E}_n) + \varepsilon G(\mathfrak{E} - \mathfrak{E}_n)$$

oder

$$\int_{\mathfrak{E}} |f(P) - f_n(P)| G(d\mathfrak{E}) \leq 2LG(\mathfrak{E}_n) + \varepsilon G(\mathfrak{E}).$$

Aus $G(\mathfrak{E}_n) \rightarrow 0$ folgt weiter, daß ein N existiert derart, daß $G(\mathfrak{E}_n) \leq \varepsilon$ für $n > N$ ist, und somit erhält man

$$\int_{\mathfrak{E}} |f(P) - f_n(P)| G(d\mathfrak{E}) \leq [2L + G(\mathfrak{E})] \varepsilon$$

für $n > N$. Vergleichen wir das mit (27), so erhalten wir

$$\left| \int_{\mathfrak{E}} f(P) G(d\mathfrak{E}) - \int_{\mathfrak{E}} f_n(P) G(d\mathfrak{E}) \right| \leq [2L + G(\mathfrak{E})] \varepsilon$$

für $n > N$. Da ε beliebig war, folgt hieraus (26).

Wir bemerken, daß es zum Beweis von (26) hinreicht, $|f_n(P)| \leq L$ fast überall auf $\mathfrak{E}$ vorauszusetzen. Durch Übergang zu äquivalenten Funktionen (falls notwendig) kann man erreichen, daß diese Ungleichung überall auf $\mathfrak{E}$ erfüllt ist. Unser Satz ermöglicht den Grenzübergang unter dem Integralzeichen, wenn nur vorausgesetzt wird, daß der absolute Betrag von $f_n(P)$ unabhängig vom Index n beschränkt ist. Die analoge Eigenschaft wurde in [11] für das Stieltjessche Integral gezeigt. Sie ist eine unmittelbare Folgerung des bewiesenen Satzes, da bei Stetigkeit von $f_n(P)$ und $f(P)$ das Lebesgue-Stieltjessche Integral sich auf das Stieltjessche Integral reduziert. Wir weisen noch darauf hin, daß man in der Formulierung von Eigenschaft 15 die „Konvergenz von $f_n(P)$ gegen $f(P)$ fast überall“ durch die „Konvergenz dem Maß nach“ ersetzen kann. Der Beweis bleibt im wesentlichen derselbe.

16. Gilt auf der Menge der Werte von $f(P)$ die Ungleichung $m < f(P) \leq M$, so ist

$$g(y) = G(\mathfrak{E} [m < f(P) \leq y])$$

eine in y wachsende Funktion, und das Lebesgue-Stieltjessche Integral läßt sich auf Grund der Formel

$$\int_{\mathfrak{E}} f(P) G(d\mathfrak{E}) = \int_m^M y dg(y) \quad (28)$$

auf ein Stieltjessches Integral zurückführen.

Um das zu beweisen, brauchen wir nur zu bemerken, daß die Lebesgueschen Summen (8) gewöhnliche Summen s_δ und S_δ für das Stieltjessche Integral sind, das auf der rechten Seite in Formel (28) steht.

Für den Fall des Lebesgueschen Integrals, d. h. in dem Fall, daß $G(\Delta)$ der Flächeninhalt des Intervalls Δ ist, bezeichnet man das Integral oft mit

$$\iint_{\mathfrak{E}} f(x, y) dx dy.$$

Im Fall einer Geraden und eines dreidimensionalen Raumes benutzt man für das Lebesguesche Integral die analogen Bezeichnungen

$$\int_{\mathfrak{E}} f(x) dx \quad \text{und} \quad \int_{\mathfrak{E}} f(x, y, z) dx dy dz.$$

50. Das Integral von unbeschränkten nichtnegativen Funktionen. Wir definieren jetzt das Integral für den Fall, daß $f(P)$ eine unbeschränkte und nichtnegative meß-

bare Funktion auf einer meßbaren Menge $\mathcal{E}$ endlichen Maßes ist. In diesem Fall werden wir $\mathcal{E}$ nicht nur in endlich viele, sondern auch in abzählbar unendlich viele meßbare Teilmengen $\mathcal{E}_k$ zerlegen. Im übrigen wird das Integral so gebildet wie auch im Fall einer beschränkten Funktion. Gegeben sei jetzt eine beliebige Zerlegung

$$\mathcal{E} = \sum_k \mathcal{E}_k. \quad (29)$$

Wir bilden die ihr entsprechenden Summen s_δ und S_δ :

$$s_\delta = \sum_k m_k G(\mathcal{E}_k), \quad S_\delta = \sum_k M_k G(\mathcal{E}_k). \quad (30)$$

In diesem Fall haben wir eine unendliche Reihe nichtnegativer Summanden, und einige der Zahlen m_k und M_k können gleich ∞ werden. Ist für einen Summanden aus (29) $G(\mathcal{E}_k) = 0$, so wollen wir dem entsprechenden Summanden in (30) auch dann den Wert 0 beilegen, wenn der erste Faktor m_k bzw. M_k gleich ∞ ist. Die Summen der Reihen (30) hängen nicht von der Reihenfolge der Summanden ab [I, 137]. Wenn eine endliche Zerlegung δ gegeben ist, wird in der Summe S_δ wenigstens eine der Zahlen gleich ∞ , da $f(P)$ unbeschränkt ist. Die Summen s_δ und S_δ können auch unendliche Werte annehmen. Sie haben dieselben Eigenschaften wie die endlichen Summen s_δ und S_δ , die wir früher benutzt haben. Es sei, wie auch schon früher, i die obere Grenze der s_δ und I die untere Grenze der S_δ . Diese Zahlen können auch gleich ∞ werden. Wie auch im Fall einer beschränkten Funktion zeigen wir, daß $i = I$ ist. Wir zerlegen $\mathcal{E}$ auf folgende Weise in Teilmengen: Wir wählen aus ihr diejenige Menge $\mathcal{E}_0$ aus, auf der $f(P) = \infty$ ist, wenn eine solche Menge existiert, und zerlegen die Restmenge auf folgende Weise in Mengen $\mathcal{E}_k$: Wir zerlegen das Intervall $[0, \infty)$ in Teilintervalle durch $0 = y_0 < y_1 < y_2 < \dots$ und bilden die Mengen

$$\mathcal{E}_1 = \mathcal{E} [y_0 \leq f \leq y_1], \quad \mathcal{E}_k = \mathcal{E} [y_{k-1} < f \leq y_k]. \quad (31)$$

Offenbar ist $m_k \leq y_{k-1}$, $M_k \leq y_k$ und

$$(+\infty) G(\mathcal{E}_0) + \sum_{k=1}^{\infty} y_{k-1} G(\mathcal{E}_k) \leq s_\delta \leq S_\delta \leq (+\infty) G(\mathcal{E}_0) + \sum_{k=1}^{\infty} y_k G(\mathcal{E}_k) \quad (32)$$

und um so mehr

$$(+\infty) G(\mathcal{E}_0) + \sum_{k=1}^{\infty} y_{k-1} G(\mathcal{E}_k) \leq i \leq I \leq (+\infty) G(\mathcal{E}_0) + \sum_{k=1}^{\infty} y_k G(\mathcal{E}_k). \quad (33)$$

Im Fall $G(\mathcal{E}_0) > 0$ ist $i = I = \infty$. Wir nehmen jetzt an, daß $G(\mathcal{E}_0) = 0$ ist. Dann schreiben wir die Ungleichungen (32) und (33) in der Gestalt

$$\sum_{k=1}^{\infty} y_{k-1} G(\mathcal{E}_k) \leq s_\delta \leq S_\delta \leq \sum_{k=1}^{\infty} y_k G(\mathcal{E}_k), \quad (34)$$

$$\sum_{k=1}^{\infty} y_{k-1} G(\mathcal{E}_k) \leq i \leq I \leq \sum_{k=1}^{\infty} y_k G(\mathcal{E}_k). \quad (34_1)$$

Wir werden annehmen, daß wir das Intervall $[0, \infty)$ so zerlegen, daß die Differenzen $y_k - y_{k-1}$ ($k = 1, 2, \dots$) beschränkt sind. Es sei η ihre obere Grenze. Wenn wir beachten, daß $y_k \leq y_{k-1} + \eta$ ist, können wir

$$\sum_{k=1}^{\infty} y_k G(\mathcal{E}_k) \leq \sum_{k=1}^{\infty} y_{k-1} G(\mathcal{E}_k) + \eta G(\mathcal{E}) \quad (35)$$

schreiben. Hieraus folgt unmittelbar: Ist bei einer gewissen Zerlegung mit endlichem η die in (34) rechts stehende Summe gleich ∞ , dann kann man dasselbe auch bezüglich der in (34) links stehenden Summe behaupten. Dann ist nach (34₁) $i = I = \infty$. Ist umgekehrt $I = \infty$, so ist nach (34₁) die in (35) links stehende Summe gleich ∞ für jede Zerlegung mit endlichem η ; daher ist auch die in (35) rechts stehende Summe gleich ∞ , und nach (34₁) ist $i = \infty$. Hieraus folgt nach (34): Ist für eine Zerlegung mit endlichem η die Summe S_δ gleich ∞ , so ist auch s_δ gleich ∞ ; also sind s_δ und S_δ gleich ∞ für jede Zerlegung mit endlichem η . In diesen Fällen ist $i = I = \infty$. Im Fall endlicher Summen haben wir wie in [49]

$$0 \leq \sum_{k=1}^{\infty} y_k G(\mathcal{E}_k) - \sum_{k=1}^{\infty} y_{k-1} G(\mathcal{E}_k) \leq \eta G(\mathcal{E}).$$

Für $\eta \rightarrow 0$ strebt diese Differenz gegen 0, und hieraus folgt $i = I$. Dabei streben beide Lebesgueschen Summen (34) und ebenso die Summen s_δ und S_δ für $\eta \rightarrow 0$ gegen den Wert des Integrals. Dasselbe kann man auch bezüglich der Summe

$$\sigma_\delta = \sum_{k=1}^{\infty} f(P_k) G(\mathcal{E}_k) \quad (36)$$

bei beliebiger Wahl der P_k aus $\mathcal{E}_k$ behaupten. Im Fall $i = I = \infty$ ist offenbar auch die Summe (36) gleich ∞ . Wenn s_{δ_n} , S_{δ_n} und σ_{δ_n} gegen den Wert des Integrals streben (im Fall eines endlichen Integrals) und $\delta'_n \geq \delta_n$ gilt, dann kann man dasselbe auch bezüglich der Summen $s_{\delta'_n}$, $S_{\delta'_n}$ und $\sigma_{\delta'_n}$ behaupten.

Den Hauptsatz aus [48] kann man genauso auf den Fall einer unbeschränkten nichtnegativen und meßbaren Funktion übertragen. Besonders wichtig ist der Fall, daß der Wert des Integrals endlich ist. In diesem Fall heißt die Funktion $f(P)$ auf der Menge $\mathcal{E}$ *summierbar*. Aus dem oben Gesagten folgt unmittelbar: *Für die Summierbarkeit einer Funktion ist notwendig, daß die Menge $\mathcal{E}_0 = \mathcal{E} [f = \infty]$ eine Menge vom Maß 0 ist.*

Wir geben nun eine andere Definition des Integrals einer unbeschränkten nichtnegativen und meßbaren Funktion, die mit der oben gegebenen Definition gleichbedeutend ist. Wir bezeichnen mit $[f(P)]_N$ die beschränkte nichtnegative Funktion, die wie folgt definiert ist („die bei N abgeschnittene Funktion“):

$$[f(P)]_N = \begin{cases} f(P) & \text{für } f(P) \leq N, \\ N & \text{für } f(P) > N. \end{cases} \quad (37)$$

Die Meßbarkeit dieser Funktion folgt unmittelbar aus $\mathcal{E} [[f]_N > a] = \mathcal{E} [f > a]$ für $a < N$ und $\mathcal{E} [[f]_N > a] = \emptyset$ für $a \geq N$, wobei $\emptyset$ die leere Menge ist. Wir betrachten die Integrale

$$i_N = \int_{\mathcal{E}} [f]_N G(d\mathcal{E}). \quad (38)$$

Sie wachsen mit wachsendem N , und den (endlichen oder unendlichen) Grenzwert dieser monotonen Veränderlichen für $N \rightarrow \infty$ nennen wir das *Integral* von $f(P)$. Wir beweisen, daß diese neue Definition des Integrals mit der früheren übereinstimmt. Zuerst nehmen wir an, daß die Menge $\mathcal{E}_0$, auf der $f(P) = \infty$ ist, das Maß 0 hat. Der Wert des Integrals (38) ist gleich der oberen Grenze i_N der Summen $s_\delta^{(N)}$ für die Funktion $[f]_N$. Diese Summen sind höchstens gleich den entsprechenden Summen s_δ für f , also ist $i_N \leq i$. Wir müssen zeigen, daß die monotone Veränderliche i_N für $N \rightarrow +\infty$ den Grenzwert i hat. Wir beweisen das indirekt. Es sei $i_N \rightarrow i' < i$ (i' ist endlich). Wir können dann eine solche Summe s_δ für $f(P)$ wählen, daß $s_\delta > i'$ ist.

Aus dieser Summe nehmen wir eine endliche Anzahl von Summanden heraus, so daß auch für diese endliche Summe s'_δ noch $s'_\delta > i'$ gilt:

$$s'_\delta = \sum'_k m_k G(\mathcal{E}_k) > i', \quad (*)$$

dabei ist die Summe endlich, und die Summation erstreckt sich über die restlichen Summanden, was durch den Strich am Summenzeichen angedeutet ist. Ist $m_k = \infty$, so ist $f(P) = \infty$ in allen Punkten von $\mathcal{E}_k$, also ist $\mathcal{E}_k \subset \mathcal{E}_0$, und daraus folgt $G(\mathcal{E}_k) = 0$, weil nach Voraussetzung $G(\mathcal{E}_0) = 0$ ist. Der entsprechende Summand der Summe ist, wie wir schon oben bemerkten, gleich 0, und wir brauchen ihn nicht aufzuschreiben. Somit kann man sagen, daß für alle Summanden der Summe (*) alle Zahlen m_k endlich sind. Wir stellen die analoge Summe für $[f]_N$ auf:

$$s'^{(N)}_\delta = \sum'_k m_k^{(N)} G(\mathcal{E}_k) \quad (m_k^{(N)} = \inf [f]_N \text{ auf } \mathcal{E}_k).$$

Ist N größer als alle m_k , die in der Summe (*) auftreten (das sind endlich viele), so ist $m_k^{(N)} = m_k$ und $s'^{(N)}_\delta = s'_\delta > i'$. Die ganze Summe $s'^{(N)}_\delta$ für $[f]_N$ wird größer als i' , also ist auch $i_N = \sup s'^{(N)}_\delta > i'$; doch das steht im Widerspruch dazu, daß i_N wachsend gegen i' strebt. Somit ist $i_N \rightarrow i$, und der Wert des Integrals bei der zweiten Integraldefinition stimmt mit dem bei der ersten Definition überein.

Ist $G(\mathcal{E}_0) > 0$, so ist nach der ersten Definition der Wert des Integrals gleich ∞ . Wir wollen zeigen, daß dies auch bei der zweiten Definition gilt. Beachten wir, daß die Funktion $[f]_N$ nichtnegativ ist, so gilt

$$i_N = \int_{\mathcal{E}_0} [f]_N G(d\mathcal{E}) \cong \int_{\mathcal{E}_0} [f]_N G(d\mathcal{E}) = NG(\mathcal{E}_0),$$

weil nach Definition $[f]_N = N$ in allen Punkten von $\mathcal{E}_0$ ist. Aus der Ungleichung $i_N \cong NG(\mathcal{E}_0)$ folgt $i_N \rightarrow \infty$ für $N \rightarrow \infty$, was zu beweisen war.

51. Eigenschaften des Integrals. Wenn wir die Summen σ_δ benutzen, können wir einige Eigenschaften des Integrals einer unbeschränkten nichtnegativen Funktion genauso beweisen wie in [49]. Außerdem können wir für den Beweis der Eigenschaften auch die zweite Definition des Integrals benutzen. Wir bemerken noch, daß $[f(P)]_N$ mit $f(P)$ für genügend großes N übereinstimmt, wenn $f(P)$ eine beschränkte nichtnegative Funktion ist, und die neue Definition ist der in [48] gegebenen äquivalent.

Wir werden jetzt einige Integraleigenschaften beweisen. Wie auch in [49] setzen wir voraus, daß $\mathcal{E}$ eine meßbare Menge endlichen Maßes ist.

1. Sind die $f_k(P)$ ($k=1, 2, \dots, m$) summierbare Funktionen, so ist auch ihre Linearkombination mit konstanten positiven Koeffizienten eine summierbare Funktion, und es gilt die Formel (13).

Dies kann man genauso wie in [49] beweisen.

2. Ist $f(P)$ auf $\mathcal{E}$ summierbar, so ist $f(P)$ auch auf jeder meßbaren Teilmenge $\mathcal{E}'$ von $\mathcal{E}$ summierbar.

Für $[f(P)]_N$ gilt, da $[f(P)]_N$ nichtnegativ ist, nach Eigenschaft 9 aus [49]

$$\int_{\mathcal{E}'} [f(P)]_N G(d\mathcal{E}) \leq \int_{\mathcal{E}} [f(P)]_N G(d\mathcal{E}).$$

In der Grenze erhalten wir

$$\int_{\mathcal{E}'} f(P) G(d\mathcal{E}) \leq \int_{\mathcal{E}} f(P) G(d\mathcal{E}), \quad (39)$$

und wenn die rechte Seite endlich ist, ist die linke Seite erst recht endlich.

3. Ist $f(P)$ auf $\mathcal{E}$ summierbar und wird die Menge $\mathcal{E}$ in endlich oder abzählbar unendlich viele meßbare Mengen zerlegt, so gilt die Formel (20).

Wir betrachten den Fall unendlich vieler Mengen $\mathcal{E}_k$. Für die beschränkte Funktion $[f]_N$ gilt

$$\int_{\mathcal{E}} [f]_N G(d\mathcal{E}) = \sum_{k=1}^{\infty} \int_{\mathcal{E}_k} [f]_N G(d\mathcal{E}) ; \quad (40)$$

hieraus folgt

$$\int_{\mathcal{E}} [f]_N G(d\mathcal{E}) \leq \sum_{k=1}^{\infty} \int_{\mathcal{E}_k} f(P) G(d\mathcal{E}) . \quad (41)$$

Bei unbeschränkt wachsendem N erhalten wir

$$\int_{\mathcal{E}} f(P) G(d\mathcal{E}) \leq \sum_{k=1}^{\infty} \int_{\mathcal{E}_k} f(P) G(d\mathcal{E}) . \quad (42)$$

Wir beweisen nun die entgegengesetzte Ungleichung. Da $f(P)$ nichtnegativ ist, können wir mit Hilfe von (40) für beliebiges endliches m

$$\int_{\mathcal{E}} [f]_N G(d\mathcal{E}) \geq \sum_{k=1}^m \int_{\mathcal{E}_k} [f]_N G(d\mathcal{E})$$

schreiben. Für unbeschränkt wachsendes N erhalten wir in der Grenze

$$\int_{\mathcal{E}} f(P) G(d\mathcal{E}) \geq \sum_{k=1}^m \int_{\mathcal{E}_k} f(P) G(d\mathcal{E}) .$$

Wenn jetzt m unbeschränkt wächst, kommen wir zu der der Ungleichung (42) entgegengesetzten Ungleichung

$$\int_{\mathcal{E}} f(P) G(d\mathcal{E}) \geq \sum_{k=1}^{\infty} \int_{\mathcal{E}_k} f(P) G(d\mathcal{E}) .$$

4. Wenn $\mathcal{E}$ in endlich oder abzählbar unendlich viele meßbare Mengen $\mathcal{E}_k$ zerlegt wird, die Funktion $f(P)$ auf jedem $\mathcal{E}_k$ meßbar ist und die Reihe

$$\sum_{k=1}^{\infty} \int_{\mathcal{E}_k} f(P) G(d\mathcal{E}) \quad (43)$$

mit nichtnegativen Summanden einen endlichen Wert hat (konvergiert), ist $f(P)$ auf $\mathcal{E}$ summierbar, und es gilt die Formel (20).

Für die Funktion $[f]_N$ gilt wie auch oben die Formel (40) und die Ungleichung (41); die rechte Seite von (41) ist endlich. Aus dieser Ungleichung folgt unmittelbar, daß die Integrale (38) einen endlichen Grenzwert haben, d. h., $f(P)$ ist auf $\mathcal{E}$ summierbar. Danach folgt (20) unmittelbar aus der vorigen Eigenschaft.

5. Ist $f(P)$ auf $\mathcal{E}$ summierbar, so existiert zu jedem $\varepsilon > 0$ ein $\eta > 0$ derart, daß

$$\int_{\mathcal{E}} f(P) G(d\mathcal{E}) \leq \varepsilon \quad (44)$$

für $e \subset \mathcal{E}$ und $G(e) \leq \eta$ ist.

Wir können ein solches N fest wählen, daß

$$\int_{\mathcal{E}} [f - [f]_N] G(d\mathcal{E}) \leq \frac{\varepsilon}{2} \quad (f \geq [f]_N)$$

ist. Dabei wird nach (39) für jedes $e \in \mathcal{E}$

$$\int_e [f - [f]_N] G(d\mathcal{E}) \leq \frac{\varepsilon}{2}, \quad \text{d. h.} \quad \int_e f G(d\mathcal{E}) \leq \int_e [f]_N F(d\mathcal{E}) + \frac{\varepsilon}{2},$$

und für $G(e) \leq \frac{\varepsilon}{2N}$ erhalten wir die geforderte Ungleichung.

Die letzten beiden Eigenschaften zeigen, daß das Integral einer unbeschränkten nichtnegativen Funktion, genau wie das Integral einer beschränkten Funktion, volladditiv und absolut stetig ist.

6. Ist $\mathcal{E}$ eine Menge vom Maß 0, so ist das Integral über $f(P)$ gleich 0.

Der Beweis wird wie in [49] geführt.

7. Die Integrale von auf $\mathcal{E}$ äquivalenten Funktionen sind gleich.

8. Ist das Integral über $f(P)$ gleich 0, so ist diese Funktion äquivalent 0.

9. Ist $f_2(P) \leq f_1(P)$ auf $\mathcal{E}$ und $f_1(P)$ summierbar, so ist auch $f_2(P)$ summierbar, und es gilt die Ungleichung

$$\int_{\mathcal{E}} f_2(P) G(d\mathcal{E}) \leq \int_{\mathcal{E}} f_1(P) G(d\mathcal{E}). \quad (45)$$

10. Ist $\omega_k(P)$ eine Folge nichtnegativer, auf $\mathcal{E}$ summierbarer Funktionen und gilt

$$\int_{\mathcal{E}} \omega_k(P) G(d\mathcal{E}) \rightarrow 0 \quad \text{für} \quad n \rightarrow \infty,$$

so konvergiert $\omega_n(P)$ auf $\mathcal{E}$ dem Maß nach gegen 0.

Die Eigenschaften 7, 8 und 9 können wir genauso wie in [49] beweisen. Die Ungleichung (45) ist richtig für $[f_2]_N$ und $[f_1]_N$, und wenn wir für $N \rightarrow \infty$ zur Grenze übergehen, erhalten wir (45). Die Eigenschaft 10 beweisen wir indirekt. Wenn $\omega_n(P)$ nicht dem Maße nach gegen 0 konvergiert, existiert ein solches $\delta > 0$, daß $G(\mathcal{E}_n)$ nicht gegen 0 strebt, wobei $\mathcal{E}_n = \mathcal{E} [\omega_n(P) \geq \delta]$ ist. Hieraus folgt, daß eine Teilfolge $\mathcal{E}_{n_k}$ existiert, für die $G(\mathcal{E}_{n_k}) \geq d$ ist, wobei d eine positive Zahl ist. Es ist

$$\int_{\mathcal{E}} \omega_{n_k}(P) G(d\mathcal{E}) \geq \int_{\mathcal{E}_{n_k}} \omega_{n_k}(P) G(d\mathcal{E}) \geq \delta G(\mathcal{E}_{n_k}) \geq \delta d.$$

Hieraus folgt, daß das links stehende Integral nicht gegen 0 geht, was ein Widerspruch zur Voraussetzung ist.

Wir definieren jetzt das Lebesgue-Stieltjesche Integral für eine unbeschränkte Funktion, die das Vorzeichen ändern kann. Für negative (nichtpositive) Funktionen können wir das Integral wie oben definieren.

52. Funktionen beliebigen Vorzeichens. Es sei auf einer meßbaren Menge $\mathcal{E}$ endlichen Maßes eine meßbare reelle Funktion $f(P)$ gegeben, die Werte beider Vorzeichen annehmen kann. Wir führen für den *positiven* bzw. *negativen Teil* der Funktion $f(P)$ folgende Bezeichnung ein:

$$\begin{aligned} f^+(P) &= \begin{cases} f(P) & \text{für } f(P) \geq 0, \\ 0 & \text{für } f(P) < 0, \end{cases} \\ f^-(P) &= \begin{cases} -f(P) & \text{für } f(P) \leq 0, \\ 0 & \text{für } f(P) > 0. \end{cases} \end{aligned} \quad (46)$$

Man kann diese Definition auch wie folgt schreiben:

$$f^+(P) = \frac{1}{2} [|f(P)| + f(P)], \quad f^-(P) = \frac{1}{2} [|f(P)| - f(P)]. \quad (46_1)$$

Die Funktion $f(P)$ kann man nun als eine Differenz zweier nichtnegativer Funktionen darstellen:

$$f(P) = f^+(P) - f^-(P).$$

Definition. Eine Funktion $f(P)$ heißt *summierbar auf $\mathfrak{E}$* , wenn $f^+(P)$ und $f^-(P)$ summierbar auf $\mathfrak{E}$ sind. Dabei ist der Wert des Integrals der Funktion $f(P)$ durch

$$\int_{\mathfrak{E}} f(P) G(d\mathfrak{E}) = \int_{\mathfrak{E}} f^+(P) G(d\mathfrak{E}) - \int_{\mathfrak{E}} f^-(P) G(d\mathfrak{E}) \quad (47)$$

definiert.

Bemerkung. Wenn nur eine der Funktionen $f^+(P)$ oder $f^-(P)$ summierbar ist, liefert (47) für das Integral von $f(P)$ einen bestimmten, aber unendlichen Wert. Wenn etwa $f^+(P)$ summierbar ist und $f^-(P)$ nicht, ist das Integral von $f(P)$ gleich $-\infty$.

Satz. Die Funktion $f(P)$ ist genau dann auf $\mathfrak{E}$ summierbar, wenn die nichtnegative Funktion $|f(P)|$ auf $\mathfrak{E}$ summierbar ist.

Ist $f(P)$ summierbar, so sind auch $f^+(P)$ und $f^-(P)$ und folglich auch ihre Summe $|f(P)| = f^+(P) + f^-(P)$ summierbar. Ist umgekehrt die Summe $f^+(P) + f^-(P)$ summierbar, so ist nach der Eigenschaft 9 aus [51] auch jeder Summand summierbar, und darum ist auch die Funktion $f(P)$ summierbar.

Bemerkung. Wir können auch beschränkte Funktionen in einen positiven und negativen Teil zerlegen; für das Integral gilt dann ebenfalls die Formel (47).

Im folgenden werden wir die Benennung „summierbare Funktion“ auch für eine beschränkte meßbare Funktion benutzen. Wir geben jetzt die grundlegenden Eigenschaften des Integrals summierbarer Funktionen beliebigen Vorzeichens an. Diese Eigenschaften folgen fast unmittelbar aus den analogen Eigenschaften des Integrals der nichtnegativen Funktionen $f^+(P)$ und $f^-(P)$.

1. Sind $f_k(P)$ ($k=1, 2, \dots, p$) summierbare Funktionen, so ist auch ihre Linearkombination mit konstanten Koeffizienten eine summierbare Funktion, und es gilt (13).

Die Summierbarkeit der Linearkombination folgt aus der Ungleichung

$$\left| \sum_{k=1}^p c_k f_k(P) \right| \leq \sum_{k=1}^p |c_k| |f_k(P)|,$$

dem oben bewiesenen Satz und der Eigenschaft 1 aus [51]. Beim Beweis der Formel (13) betrachten wir die beiden Fälle, Multiplikation einer Funktion mit einer Konstanten und Addition zweier Funktionen, einzeln. Es sei $f(P)$ summierbar und c konstant. Dann müssen wir die Formel

$$\int_{\mathfrak{E}} c f(P) G(d\mathfrak{E}) = c \int_{\mathfrak{E}} f(P) G(d\mathfrak{E})$$

beweisen. Für den Beweis nehmen wir c als negativ an. Dann ist $(cf)^+ = -cf^-$ und $(cf)^- = -cf^+$. Nach (47) erhalten wir

$$\int_{\mathfrak{E}} c f G(d\mathfrak{E}) = -c \int_{\mathfrak{E}} f^- G(d\mathfrak{E}) + c \int_{\mathfrak{E}} f^+ G(d\mathfrak{E}) = c \int_{\mathfrak{E}} f G(d\mathfrak{E}),$$

und die Formel (13) ist bewiesen. Wir nehmen jetzt an, daß $f_1(P)$ und $f_2(P)$ summierbare Funktionen sind. Zum Beweis der Formel

$$\int_{\mathfrak{E}} (f_1 + f_2) G(d\mathfrak{E}) = \int_{\mathfrak{E}} f_1 G(d\mathfrak{E}) + \int_{\mathfrak{E}} f_2 G(d\mathfrak{E}) \quad (48)$$

zerlegen wir die Funktionen f_1 , f_2 und $f = f_1 + f_2$ in den positiven und den negativen Teil:

$$f_1 = f_1^+ - f_1^-, \quad f_2 = f_2^+ - f_2^-, \quad f = f^+ - f^-.$$

Es ist

$$f_1^+ + f_2^+ + f^- = f_1^- + f_2^- + f^+.$$

Diese Funktionen sind alle nichtnegativ und summierbar. Wenn wir die Eigenschaft 1 aus [51] verwenden, erhalten wir

$$\int_{\mathfrak{E}} f_1^+ G(d\mathfrak{E}) + \int_{\mathfrak{E}} f_2^+ G(d\mathfrak{E}) + \int_{\mathfrak{E}} f^- G(d\mathfrak{E}) = \int_{\mathfrak{E}} f_1^- G(d\mathfrak{E}) + \int_{\mathfrak{E}} f_2^- G(d\mathfrak{E}) + \int_{\mathfrak{E}} f^+ G(d\mathfrak{E}),$$

woraus

$$\int_{\mathfrak{E}} f^+ G(d\mathfrak{E}) - \int_{\mathfrak{E}} f^- G(d\mathfrak{E}) = \int_{\mathfrak{E}} f_1^+ G(d\mathfrak{E}) - \int_{\mathfrak{E}} f_1^- G(d\mathfrak{E}) + \int_{\mathfrak{E}} f_2^+ G(d\mathfrak{E}) - \int_{\mathfrak{E}} f_2^- G(d\mathfrak{E})$$

folgt; damit ist (48) bewiesen.

2. Sind $f_1(P)$ und $f_2(P)$ auf $\mathfrak{E}$ summierbar und ist $f_1(P) \equiv f_2(P)$, so gilt die Formel (15).

Nach Eigenschaft 1 ist die nichtnegative Funktion $f_1(P) - f_2(P)$ summierbar und das (ebenfalls nichtnegative) Integral von $f_1(P) - f_2(P)$ gleich der Differenz der Integrale über $f_1(P)$ und $f_2(P)$; daraus folgt dann (15).

3. Wenn $f(P)$ summierbar ist, so gilt die Formel (16).

Die Ungleichung (16) ist der folgenden offenbar gültigen Ungleichung gleichwertig:

$$\left| \int_{\mathfrak{E}} [f^+(P) - f^-(P)] G(d\mathfrak{E}) \right| \leq \int_{\mathfrak{E}} [f^+(P) + f^-(P)] G(d\mathfrak{E}).$$

4. Ist die Funktion $f(P)$ auf $\mathfrak{E}$ summierbar, so ist sie auch auf jeder meßbaren Teilmenge $\mathfrak{E}'$ von $\mathfrak{E}$ summierbar.

5. Wenn $f(P)$ auf $\mathfrak{E}$ summierbar und die Menge $\mathfrak{E}$ in endlich oder abzählbar unendlich viele meßbare Mengen $\mathfrak{E}_k$ zerlegt ist, dann gilt die Formel (20).

Die letzten beiden Eigenschaften folgen daraus, daß diese Eigenschaften für $f^+(P)$ und $f^-(P)$ gelten.

6. Wenn $\mathfrak{E}$ in endlich oder abzählbar unendlich viele meßbare Mengen $\mathfrak{E}_k$ zerlegt ist, die Funktion $f(P)$ auf jeder Menge $\mathfrak{E}_k$ summierbar ist und die Reihe

$$\sum_{k=1}^{\infty} \int_{\mathfrak{E}_k} |f(P)| G(d\mathfrak{E}) \quad (49)$$

konvergiert, dann ist $f(P)$ auf $\mathfrak{E}$ summierbar, und es gilt die Formel (20).

Die nichtnegative Funktion $|f(P)|$ ist auf allen $\mathfrak{E}_k$ summierbar, und aus der Konvergenz der Reihe (49) folgt nach Eigenschaft 4 aus [51], daß $|f(P)|$ auf $\mathfrak{E}$ summierbar ist; folglich ist auch $f(P)$ auf $\mathfrak{E}$ summierbar. Dann folgt (20) aus der vorigen Eigenschaft. Wir bemerken, daß die Konvergenz der Reihe (43) allein nicht für die Summierbarkeit von $f(P)$ genügt.

7. Ist $f(P)$ auf $\mathfrak{E}$ summierbar, so existiert zu jedem $\varepsilon > 0$ ein solches $\eta > 0$, daß

$$\left| \int_e f(P) G(d\mathfrak{E}) \right| \leq \varepsilon$$

ist für $e \subset \mathfrak{E}$ und $G(e) \leq \eta$.

Diese Eigenschaft folgt unmittelbar aus ihrer Gültigkeit für $f^+(P)$ und $f^-(P)$.

Somit haben wir die Volladditivität und die absolute Stetigkeit für das Integral jeder summierbaren Funktion bewiesen.

8. Ist $\mathcal{E}$ eine Menge vom Maß 0, so ist das Integral jeder Funktion $f(P)$ auf $\mathcal{E}$ gleich 0.

9. Die Integrale von auf $\mathcal{E}$ äquivalenten Funktionen sind gleich.

Beide Eigenschaften folgen aus den analogen Eigenschaften für $f^+(P)$ und $f^-(P)$. Dazu muß man bemerken: Wenn zwei Funktionen äquivalent sind, sind auch ihre positiven bzw. negativen Teile äquivalent.

10. Ist $f(P)$ auf $\mathcal{E}$ meßbar, $F(P)$ meßbar, nichtnegativ und summierbar auf $\mathcal{E}$ und gilt $|f(P)| \leq F(P)$, so ist $f(P)$ summierbar, und es gilt

$$\left| \int_{\mathcal{E}} f(P) G(d\mathcal{E}) \right| \leq \int_{\mathcal{E}} F(P) G(d\mathcal{E}). \quad (50)$$

Nach der Eigenschaft 9 aus [51] können wir schließen, daß $|f(P)|$ summierbar ist; folglich ist auch $f(P)$ summierbar. Die Ungleichung (50) folgt unmittelbar aus den Eigenschaften 3 und 9 aus [51]. Aus dieser Ungleichung folgt, daß das Produkt einer summierbaren Funktion mit einer beschränkten meßbaren Funktion ebenfalls eine summierbare Funktion ist.

Wir beweisen jetzt noch zwei Integraleigenschaften, die wir im weiteren benötigen werden.

11. Wenn die Funktion $f(P)$ in einem endlichen Intervall Δ_0 summierbar ist und ihr Integral in jedem in Δ_0 enthaltenen Intervall gleich 0 ist, ist $f(P)$ in Δ_0 äquivalent 0.

Wir beweisen das indirekt. Ist $f(P)$ nicht äquivalent 0, so existiert eine solche positive Zahl a , daß eine der Mengen $\Delta_0 [f(P) \geq a]$ oder $\Delta_0 [f(P) \leq -a]$ ein positives Maß hat. Es sei dies die erste Menge, und wir bezeichnen sie mit $\mathcal{E}$. Dann ist

$$\int_{\mathcal{E}} f(P) G(d\mathcal{E}) \geq a G(\mathcal{E}) > 0.$$

Nun gibt es aber Mengen e_1 und e_2 mit beliebig kleinen Mäßen derart, daß $\mathcal{E} + e_1 = R + e_2$ ist, wobei R ein Elementarbereich [37], d. h. eine endliche Summe paarweise elementfremder Intervalle ist. Nach Voraussetzung muß das Integral von $f(P)$ über R gleich 0 sein, und wir können

$$\int_{\mathcal{E}} f(P) G(d\mathcal{E}) = \int_{e_2} f(P) G(d\mathcal{E}) - \int_{e_1} f(P) G(d\mathcal{E})$$

schreiben. Wegen der absoluten Stetigkeit des Integrals kann der absolute Betrag der rechten Seite beliebig klein gemacht werden; auf der linken Seite steht aber eine wohlbestimmte positive Zahl. Das ist aber sinnlos, und damit ist die obige Behauptung bewiesen.

12. Wenn $f(P)$ auf $\mathcal{E}$ summierbar ist und der Bedingung

$$\int_{\mathcal{E}} \varphi(P) f(P) G(d\mathcal{E}) = 0 \quad (51)$$

für beliebig gewähltes $\varphi(P)$ genügt, das auf $\mathcal{E}$ meßbar und beschränkt ist, dann ist $f(P)$ äquivalent 0. Ist die Bedingung (51) für beliebig gewähltes $\varphi(P)$ erfüllt, das meßbar und beschränkt auf $\mathcal{E}$ ist und für das

$$\int_{\mathcal{E}} \varphi(P) G(d\mathcal{E}) = 0 \quad (52)$$

gilt, so ist $f(P)$ äquivalent einer Konstanten.

Es sei $\mathcal{E}'$ der Teil von $\mathcal{E}$, auf dem $f(P) \geq 0$ ist. Wir wählen für $\varphi(P)$ eine Funktion, die gleich 1 auf $\mathcal{E}'$ und gleich 0 auf $\mathcal{E} - \mathcal{E}'$ ist. Nach der Bedingung (51) folgt, daß das Integral von $f^+(P)$ über $\mathcal{E}$ gleich 0 ist, und hieraus folgt nach Eigenschaft 8 aus [51], daß $f^+(P)$ äquivalent 0 ist. Analog kann man beweisen, daß $f^-(P)$ äquivalent 0 ist, und daher ist auch $f(P)$ äquivalent 0.

Wir beweisen jetzt den zweiten Teil der Behauptung. Wir bezeichnen mit $kG(\mathcal{E})$ den Wert des Integrals von $f(P)$ über $\mathcal{E}$. Nach (52) genügt die Funktion $f(P) - k$ der Bedingung (51), d. h.

$$\int_{\mathcal{E}} \varphi(P) [f(P) - k] G(d\mathcal{E}) = 0.$$

Außerdem gilt nach der Definition von k für beliebig gewähltes konstantes c

$$\int_{\mathcal{E}} c [f(P) - k] G(d\mathcal{E}) = 0. \quad (53)$$

Es sei $\psi(P)$ eine beliebige, auf $\mathcal{E}$ meßbare beschränkte Funktion und $cG(\mathcal{E})$ der Wert ihres Integrals. Die Funktion $\varphi(P) = \psi(P) - c$ genügt der Bedingung (52), und es ist

$$\int_{\mathcal{E}} [\psi(P) - c] [f(P) - k] G(d\mathcal{E}) = 0$$

oder nach (53)

$$\int_{\mathcal{E}} \psi(P) [f(P) - k] G(d\mathcal{E}) = 0.$$

Hieraus folgt nach dem oben Bewiesenen, daß $f(P) - k$ äquivalent 0 ist, d. h., $f(P)$ ist äquivalent k .

Abschließend betrachten wir noch Lebesgue-Stieltjessche Integrale einer Veränderlichen. Es sei $g(x)$ die dem Maß zugrunde liegende nichtfallende Funktion, $f(x)$ sei bezüglich $g(x)$ meßbar und auf einer meßbaren Menge $\mathcal{E}$ der x -Achse oder auf einem Intervall $[a, b]$ oder auf dem Intervall $(a, b]$ usw. summierbar. Die entsprechenden Integrale kann man dann wie folgt schreiben:

$$\int_{\mathcal{E}} f(x) dg(x) \quad \text{oder} \quad \int_{[a,b]} f(x) dg(x) \quad \text{oder} \quad \int_{(a,b]} f(x) dg(x) \quad \text{usw.}$$

Für $g(x) = x$ erhalten wir ein Lebesguesches Integral. In diesem Fall ist das Maß jedes Punktes gleich 0, und es ist unwichtig, ob wir zu einem Intervall seine Endpunkte hinzuzählen oder nicht; das Integral über das Intervall $[a, b]$ bezeichnet man gewöhnlich mit

$$\int_a^b f(x) dx.$$

Dabei wird $a < b$ angenommen. Außerdem ist vereinbarungsgemäß

$$\int_b^a f(x) dx = - \int_a^b f(x) dx.$$

53. Komplexe summierbare Funktionen. Es ist nicht schwer, den Begriff der summierbaren Funktion und den des Integrals auch für komplexwertige Funktionen $f(P)$ zu definieren. Wir zerlegen solche Funktionen in ihren Real- und Imaginärteil:

$$f(P) = f_1(P) + i f_2(P).$$

Die Funktion $f(P)$ heißt *summierbar*, wenn $f_1(P)$ und $f_2(P)$ summierbar sind, und das Integral von $f(P)$ ist in diesem Fall durch

$$\int_{\mathfrak{E}} f(P) G(d\mathfrak{E}) = \int_{\mathfrak{E}} f_1(P) G(d\mathfrak{E}) + i \int_{\mathfrak{E}} f_2(P) G(d\mathfrak{E}). \quad (54)$$

definiert.

Auch für diesen Fall gilt der oben bewiesene Satz: *Die Funktion $f(P)$ ist genau dann summierbar, wenn der Absolutbetrag $|f(P)|$ eine summierbare Funktion ist.*

Wir bemerken zunächst, daß infolge der Meßbarkeit von $f_1(P)$ und $f_2(P)$ auch die Quadratsumme dieser Funktionen meßbar ist; daher ist auch die positive Quadratwurzel aus dieser Summe, d. h. $|f| = \sqrt{f_1^2 + f_2^2}$ meßbar. Das folgt unmittelbar aus der Formel

$$\mathfrak{E}[\sqrt{f_1^2 + f_2^2} > a] = \mathfrak{E}[f_1^2 + f_2^2 > a^2].$$

Weiter folgt aus den Ungleichungen

$$|f_1| \leq \sqrt{f_1^2 + f_2^2}, \quad |f_2| \leq \sqrt{f_1^2 + f_2^2}, \quad \sqrt{f_1^2 + f_2^2} \leq |f_1| + |f_2|$$

und den Eigenschaften 9 und 1 aus [51], daß die Summierbarkeit von $|f_1|$ und $|f_2|$ mit der Summierbarkeit von $|f|$ gleichbedeutend ist; hieraus folgt dann unmittelbar die obige Behauptung.

Ferner gelten die oben bewiesenen Eigenschaften 1 und 3 bis 10, und wir können die Linearkombination der Funktionen $f_k(P)$ mit komplexen konstanten Koeffizienten bilden. Wir müssen noch näher auf den Beweis der Eigenschaft 3 eingehen:

$$\left| \int_{\mathfrak{E}} f G(d\mathfrak{E}) \right| \leq \int_{\mathfrak{E}} |f| G(d\mathfrak{E}). \quad (55)$$

Die Funktionen f_1, f_2 und $\sqrt{f_1^2 + f_2^2}$ sind summierbar; also streben für diese drei Funktionen die Summen $\sigma_{\delta_n}^{(1)}, \sigma_{\delta_n}^{(2)}, \sigma_{\delta_n}^{(3)}$, die ihren Folgen Lebesguescher Zerlegungen entsprechen, gegen die Integrale dieser Funktionen. Wenn wir die Folge der Zerlegungen $\delta_n = \delta_n^{(1)} \delta_n^{(2)} \delta_n^{(3)}$ nehmen, streben die Summen σ_{δ_n} für die Funktionen f_1, f_2 und $\sqrt{f_1^2 + f_2^2}$ erst recht gegen die entsprechenden Integrale. Ist δ_n eine Zerlegung von $\mathfrak{E}$ in die Teile $\mathfrak{E}_k^{(n)}$ und sind $P_k^{(n)}$ beliebige Punkte von $\mathfrak{E}_k^{(n)}$, so gilt

$$\left| \sum_k [f_1(P_k^{(n)}) + i f_2(P_k^{(n)})] G(\mathfrak{E}_k^{(n)}) \right| \leq \sum_k |f_1(P_k^{(n)}) + i f_2(P_k^{(n)})| G(\mathfrak{E}_k^{(n)}),$$

und beim Grenzübergang erhalten wir (55).

54. Der Grenzübergang unter dem Integralzeichen. Wir beweisen jetzt einige Sätze, die sich auf den Grenzübergang unter dem Integralzeichen für summierbare Funktionen beziehen.

Satz 1. *Bilden die $f_n(P)$ eine Folge von Funktionen, die auf einer Menge $\mathfrak{E}$ endlichen Maßes summierbar sind, wobei für alle Funktionen die Abschätzung*

$$|f_n(P)| \leq F(P) \quad (56)$$

mit einer auf $\mathfrak{E}$ summierbaren Funktion $F(P)$ gilt, und strebt $f_n(P)$ fast überall auf $\mathfrak{E}$ gegen $f(P)$, so ist $f(P)$ auf $\mathfrak{E}$ summierbar, und es gilt

$$\lim_{n \rightarrow \infty} \int_{\mathfrak{E}} f_n(P) G(d\mathfrak{E}) = \int_{\mathfrak{E}} f(P) G(d\mathfrak{E}). \quad (57)$$

Beweis. Aus der Voraussetzung des Satzes folgt, daß die Grenzfunktion fast überall auf $\mathfrak{E}$ der Ungleichung

$$|f(P)| \leq F(P) \quad (56_1)$$

genügt. Durch Übergang zu einer äquivalenten Funktion kann man erreichen, daß diese Ungleichung überall auf $\mathfrak{E}$ gilt. Nach der Eigenschaft 10 aus [52] sind die $f_n(P)$ und $f(P)$ summierbar auf $\mathfrak{E}$; daher haben sie auch fast überall auf $\mathfrak{E}$ endliche Werte. Wir betrachten das Integral der Differenz $f(P) - f_n(P)$ und wenden darauf die Eigenschaft 10 aus [52] an:

$$\left| \int_{\mathfrak{E}} [f(P) - f_n(P)] G(d\mathfrak{E}) \right| \leq \int_{\mathfrak{E}} |f(P) - f_n(P)| G(d\mathfrak{E}). \quad (58)$$

Es sei ε eine gegebene positive Zahl, und die Teilmengen $\mathfrak{E}_n$ von $\mathfrak{E}$ mögen dieselbe Bedeutung wie in Satz 6 aus [44] haben. Nach diesem Satz ist

$$G(\mathfrak{E}_r) \rightarrow 0 \quad (r \rightarrow \infty) \quad \text{und} \quad |f(P) - f_n(P)| \leq \varepsilon \quad \text{für} \quad P \in \mathfrak{E} - \mathfrak{E}_n. \quad (59)$$

Außerdem gilt in jedem Punkt P aus $\mathfrak{E}$

$$|f(P) - f_n(P)| \leq |f(P)| + |f_n(P)| \leq 2F(P). \quad (60)$$

Das in (58) rechts stehende Integral zerlegen wir in zwei Integrale:

$$\int_{\mathfrak{E}} |f(P) - f_n(P)| G(d\mathfrak{E}) = \int_{\mathfrak{E}_n} |f(P) - f_n(P)| G(d\mathfrak{E}) + \int_{\mathfrak{E} - \mathfrak{E}_n} |f(P) - f_n(P)| G(d\mathfrak{E}).$$

Hieraus ergibt sich nach (59) und (60)

$$\int_{\mathfrak{E}} |f(P) - f_n(P)| G(d\mathfrak{E}) \leq 2 \int_{\mathfrak{E}_n} F(P) G(d\mathfrak{E}) + \varepsilon G(\mathfrak{E} - \mathfrak{E}_n)$$

oder erst recht

$$\int_{\mathfrak{E}} |f(P) - f_n(P)| G(d\mathfrak{E}) \leq 2 \int_{\mathfrak{E}_n} F(P) G(d\mathfrak{E}) + \varepsilon G(\mathfrak{E}). \quad (61)$$

Aus $G(\mathfrak{E}_n) \rightarrow 0$ und der absoluten Stetigkeit des Integrals von $F(P)$ folgt, daß ein solches N existiert, für das

$$\int_{\mathfrak{E}_n} F(P) G(d\mathfrak{E}) \leq \varepsilon \quad \text{für} \quad n > N$$

gilt, und nach (61) ist

$$\int_{\mathfrak{E}} |f(P) - f_n(P)| G(d\mathfrak{E}) \leq [2 + G(\mathfrak{E})] \varepsilon \quad \text{für} \quad n > N.$$

Vergleichen wir dieses Ergebnis mit (58), so erhalten wir

$$\left| \int_{\mathfrak{E}} f(P) G(d\mathfrak{E}) - \int_{\mathfrak{E}} f_n(P) G(d\mathfrak{E}) \right| \leq [2 + G(\mathfrak{E})] \varepsilon.$$

Hieraus folgt, da ε beliebig war, (57). Wie beim Beweis der Eigenschaft 15 aus [49] genügt es vorauszusetzen, daß die Ungleichung (56) nur fast überall auf $\mathfrak{E}$ gilt.

Bemerkung. Man sieht leicht, daß beim Beweis des Satzes nur die Konvergenz von $f_n(P) \rightarrow f(P)$ dem Maß nach benutzt wurde; daher kann man bei der Formulierung des Satzes statt der Konvergenz von $f_n(P) \rightarrow f(P)$ fast überall die Konvergenz dem Maß nach setzen.

Satz 2. *Bilden die $f_n(P)$ eine nichtfallende Folge von Funktionen, die auf einer Menge $\mathfrak{E}$ endlichen Maßes summierbar sind, so ist das Integral der Grenzfunktion $f(P)$ über $\mathfrak{E}$ gleich einem endlichen Wert oder gleich ∞ , und es gilt die Formel (57).*

Beweis. Die summierbaren Funktionen $f_n(P)$ sind fast überall auf $\mathfrak{E}$ endlich, und die nichtfallende Folge der $f_n(P)$ hat in jedem Punkt einen Grenzwert, der auch ∞ werden kann. Wir betrachten die nichtfallende Folge der nichtnegativen

Funktionen $f_n(P) - f_1(P)$. Offenbar ist

$$0 \leq f_n(P) - f_1(P) \leq f(P) - f_1(P).$$

Ist die nichtnegative Funktion $f(P) - f_1(P)$ auf $\mathfrak{E}$ summierbar, so ist auch $f(P)$ summierbar. Mit der Differenz $f(P) - f_1(P)$ statt $F(P)$ wenden wir Satz 1 an und erhalten

$$\lim_{n \rightarrow \infty} \int_{\mathfrak{E}} [f_n(P) - f_1(P)] G(d\mathfrak{E}) = \int_{\mathfrak{E}} [f(P) - f_1(P)] G(d\mathfrak{E}).$$

Fügen wir auf beiden Seiten das Integral über $f_1(P)$ hinzu, so erhalten wir (57). Wir nehmen jetzt an, daß das Integral von $f(P) - f_1(P)$ gleich ∞ ist. Wegen der Summierbarkeit von $f_1(P)$ ist dann das Integral von $f(P)$ ebenfalls gleich ∞ . Wir bemerken weiter: Wenn eine Folge von Funktionen $\varphi_n(P)$ fast überall gegen $\varphi(P)$ strebt, gilt für beliebiges N fast überall

$$[\varphi_n(P)]_N \rightarrow [\varphi(P)]_N.$$

Um das zu beweisen, genügt es zu bemerken: Gilt in irgendeinem Punkt $\varphi_n(P) \rightarrow \varphi(P)$, so gilt in diesem Punkt $[\varphi_n(P)]_N \rightarrow [\varphi(P)]_N$. Hiervon kann man sich leicht überzeugen, wenn man die beiden Fälle $\varphi(P) \leq N$ und $\varphi(P) > N$ einzeln betrachtet. Also strebt die nichtfallende Folge nichtnegativer Funktionen $[f_n(P) - f_1(P)]_N$ fast überall gegen $[f(P) - f_1(P)]_N$. Die Grenzfunktion ist beschränkt und somit summierbar. Nach dem oben Bewiesenen ist

$$\lim_{n \rightarrow \infty} \int_{\mathfrak{E}} [f_n(P) - f_1(P)]_N G(d\mathfrak{E}) = \int_{\mathfrak{E}} [f(P) - f_1(P)]_N G(d\mathfrak{E}). \quad (62)$$

Es sei K eine beliebig gegebene positive Zahl. Weil das Integral von $[f(P) - f_1(P)]$ gleich ∞ ist, kann man ein solches N fest vorgeben, daß das in (62) rechts stehende Integral größer als K ist. Nach (62) gilt also für alle genügend großen n

$$\int_{\mathfrak{E}} [f_n(P) - f_1(P)]_N G(d\mathfrak{E}) > K$$

und um so mehr

$$\int_{\mathfrak{E}} [f_n(P) - f_1(P)] G(d\mathfrak{E}) > K.$$

Weil K beliebig war, folgt hieraus

$$\lim_{n \rightarrow \infty} \left[\int_{\mathfrak{E}} f_n(P) G(d\mathfrak{E}) - \int_{\mathfrak{E}} f_1(P) G(d\mathfrak{E}) \right] = \infty,$$

d. h.

$$\lim_{n \rightarrow \infty} \int_{\mathfrak{E}} f_n(P) G(d\mathfrak{E}) = \infty,$$

und die Formel (57) ist auch in dem Fall bewiesen, daß das Integral von $f(P)$ gleich ∞ ist.

Bemerkung. Ein analoger Satz gilt auch für fallende Folgen summierbarer Funktionen. Die Grenzfunktion kann hier ein Integral haben, das statt ∞ gleich $-\infty$ ist. Ist $f_n(P)$ eine fallende Folge, so erhalten wir, wenn wir $\varphi_n = -f_n$ setzen, eine wachsende Folge, und das Minuszeichen kann man vor das Integralzeichen ziehen.

Wir erwähnen eine im weiteren wichtige Folgerung aus diesem Satz.

Satz 3. Sind die Funktionen $u_k(P)$ ($k=1, 2, 3, \dots$) nichtnegativ und summierbar auf $\mathfrak{E}$ und konvergiert die Reihe

$$\sum_{k=1}^{\infty} \int_{\mathfrak{E}} u_k(P) G(d\mathfrak{E}) \quad (63)$$

(mit nichtnegativen Gliedern), so konvergiert die Reihe

$$\sum_{k=1}^{\infty} u_k(P) \quad (64)$$

fast überall auf $\mathfrak{S}$, und es gilt $u_k(P) \rightarrow 0$ fast überall auf $\mathfrak{S}$.

Beweis. Wir betrachten die nichtfallende Folge nichtnegativer und auf $\mathfrak{S}$ summierbarer Funktionen $f_n(P) = \sum_{k=1}^n u_k(P)$ und wenden auf diese Folge den oben bewiesenen Satz an. Da die Reihe (63) konvergent ist, haben die Integrale von $f_n(P)$ bei unbeschränkt wachsendem n einen endlichen Grenzwert. Folglich ist die Grenzfunktion, die in diesem Fall durch die Reihe (64) dargestellt wird, nämlich $f(P) = \sum_{k=1}^{\infty} u_k(P)$, auf $\mathfrak{S}$ summierbar und nimmt daher fast überall auf $\mathfrak{S}$ endliche Werte an, d. h., die Reihe (64) konvergiert tatsächlich fast überall auf $\mathfrak{S}$. Aber die Glieder einer konvergenten Reihe streben mit wachsendem Index gegen 0, d. h., es gilt $u_k(P) \rightarrow 0$ fast überall auf $\mathfrak{S}$. Damit ist der Satz bewiesen.

Satz 4. *Bilden die $f_n(P)$ eine Folge nichtnegativer und auf $\mathfrak{S}$ summierbarer Funktionen, die fast überall auf $\mathfrak{S}$ gegen eine Grenzfunktion $f(P)$ strebt, und übersteigen die Integrale von $f_n(P)$ bei beliebigen n eine Zahl A nicht, d. h., gilt*

$$\int_{\mathfrak{S}} f_n(P) G(d\mathfrak{S}) \leq A,$$

so ist $f(P)$ auf $\mathfrak{S}$ summierbar, und es gilt die Ungleichung

$$\int_{\mathfrak{S}} f(P) G(d\mathfrak{S}) \leq A. \quad (65)$$

Beweis. Es gilt die Ungleichung

$$\int_{\mathfrak{S}} [f_n]_N G(d\mathfrak{S}) \leq \int_{\mathfrak{S}} f_n G(d\mathfrak{S}) \leq A, \quad (66)$$

und nach Eigenschaft 15 aus [49] (mit N statt L) können wir

$$\lim_{n \rightarrow \infty} \int_{\mathfrak{S}} [f_n]_N G(d\mathfrak{S}) = \int_{\mathfrak{S}} [f]_N G(d\mathfrak{S})$$

schreiben. Gehen wir in (66) für $n \rightarrow \infty$ zur Grenze über, so erhalten wir

$$\int_{\mathfrak{S}} [f]_N G(d\mathfrak{S}) \leq A.$$

Hieraus folgt, daß $f(P)$ summierbar ist, und für $N \rightarrow \infty$ erhalten wir (65).

55. Die Klasse L_2 . Wir betrachten in diesem Abschnitt eine Klasse meßbarer Funktionen, die bei der Anwendung dieser Theorie auf verschiedene Fragen der Mathematik und der mathematischen Physik eine wichtige Rolle spielt.

Definition. Eine reelle Funktion $f(P)$, die auf einer meßbaren Menge $\mathfrak{S}$ endlichen Maßes meßbar ist, heißt *Funktion mit summierbarem Quadrat* oder *quadratisch summierbare Funktion* auf $\mathfrak{S}$, wenn ihr Quadrat $f^2(P)$ auf $\mathfrak{S}$ summierbar ist, d. h., wenn

$$\int_{\mathfrak{S}} f^2(P) G(d\mathfrak{S}) < \infty$$

gilt.

Die Klasse der Funktionen mit auf $\mathfrak{E}$ summierbarem Quadrat bezeichnet man mit L_2^q . Für den Fall des Lebesgueschen Integrals, d. h. für den Fall, daß $G(\Delta)$ der Flächeninhalt des Intervalls Δ ist, benutzt man das Symbol L_2 . Im weiteren werden wir jedoch statt L_2^q einfach L_2 schreiben. Es sei aber ausdrücklich darauf hingewiesen, daß alles im weiteren Gesagte auch bei beliebiger Wahl von $G(\Delta)$ richtig bleibt. Wir beweisen jetzt eine Reihe von Eigenschaften für die Klasse L_2 .

Satz 1. *Gehören $f(P)$ und $g(P)$ zu L_2 , so sind $f(P)$ und das Produkt $f(P)g(P)$ auf $\mathfrak{E}$ summierbar.*

Die Behauptung des Satzes folgt aus den Ungleichungen

$$|f| \leq \frac{1}{2} (1 + f^2), \quad |fg| \leq \frac{1}{2} (f^2 + g^2)$$

und den Eigenschaften 1 und 10 aus [52].

Wir bemerken noch, daß wir hier und im folgenden stets reelle Funktionen betrachten.

Satz 2. *Gehören $f(P)$ und $g(P)$ zu L_2 , so gehören auch $cf(P)$ und $f(P) + g(P)$ zu L_2 .*

Die Behauptung bezüglich $cf(P)$ ist klar, und für $f(P) + g(P)$ folgt sie aus der Formel

$$(f + g)^2 = f^2 + 2fg + g^2,$$

aus Satz 1 und der Eigenschaft 1 aus [52].

Satz 3. *Gehören f und g zu L_2 , so gilt die Ungleichung*

$$\left[\int_{\mathfrak{E}} fg G(d\mathfrak{E}) \right]^2 \leq \int_{\mathfrak{E}} f^2 G(d\mathfrak{E}) \int_{\mathfrak{E}} g^2 G(d\mathfrak{E}). \quad (67)$$

Den Beweis kann man genauso führen wie für das Riemannsches Integral. Wir wiederholen ihn noch einmal. Zuerst bemerken wir noch: Sind im quadratischen Trinom $au^2 + 2bu + c$ die Koeffizienten reell und ist $a > 0$, so folgt aus der Identität

$$au^2 + 2bu + c = \frac{1}{a} [(au + b)^2 + (ac - b^2)]$$

unmittelbar $b^2 \leq ac$, wenn der Ausdruck auf der linken Seite für alle reellen u nicht-negativ ist. Wir nehmen an, daß die Funktionen f und g nicht äquivalent 0 sind, denn in diesem Fall wäre die Ungleichung (67) trivial, weil ihre linke Seite dann gleich 0 wäre. Offenbar gilt

$$\int_{\mathfrak{E}} (fu + g)^2 G(d\mathfrak{E}) = u^2 \int_{\mathfrak{E}} f^2 G(d\mathfrak{E}) + 2u \int_{\mathfrak{E}} fg G(d\mathfrak{E}) + \int_{\mathfrak{E}} g^2 G(d\mathfrak{E}),$$

wobei u ein Parameter ist. Das links stehende Integral hat für alle reellen u nicht-negative Werte. Folglich hat auch die rechts stehende Summe diese Eigenschaft. Für dieses Trinom gilt also $b^2 \leq ac$, und damit erhalten wir die Ungleichung (67). Wir bemerken noch, daß der Koeffizient a in diesem Ausdruck sicher positiv ist, weil die Funktion f nicht äquivalent 0 ist.

Folgerung. Mit $f \in L_2$ ist offenbar auch $|f| \in L_2$, und wenn wir $|f|$ als Produkt $|f| = |f| \cdot 1$ darstellen, erhalten wir die Ungleichung

$$\left| \int_{\mathfrak{E}} f G(d\mathfrak{E}) \right| \leq \int_{\mathfrak{E}} |f| G(d\mathfrak{E}) \leq \sqrt{\int_{\mathfrak{E}} f^2 G(d\mathfrak{E}) \cdot G(\mathfrak{E})}. \quad (68)$$

Satz 4. *Gehören f und g zu L_2 , so gilt die Ungleichung*

$$\sqrt{\int_{\mathfrak{E}} (f+g)^2 G(d\mathfrak{E})} \leq \sqrt{\int_{\mathfrak{E}} f^2 G(d\mathfrak{E})} + \sqrt{\int_{\mathfrak{E}} g^2 G(d\mathfrak{E})}. \quad (69)$$

Aus der Ungleichung (67) erhalten wir

$$\int_{\mathfrak{E}} fg G(d\mathfrak{E}) \leq \sqrt{\int_{\mathfrak{E}} f^2 G(d\mathfrak{E})} \sqrt{\int_{\mathfrak{E}} g^2 G(d\mathfrak{E})}.$$

Wir multiplizieren beide Seiten der Ungleichung mit 2 und addieren auf beiden Seiten der erhaltenen Ungleichung das Integral von f^2 und das von g^2 . Die entstehende Ungleichung

$$\begin{aligned} & \int_{\mathfrak{E}} f^2 G(d\mathfrak{E}) + 2 \int_{\mathfrak{E}} fg G(d\mathfrak{E}) + \int_{\mathfrak{E}} g^2 G(d\mathfrak{E}) \\ & \leq \int_{\mathfrak{E}} f^2 G(d\mathfrak{E}) + 2 \sqrt{\int_{\mathfrak{E}} f^2 G(d\mathfrak{E})} \sqrt{\int_{\mathfrak{E}} g^2 G(d\mathfrak{E})} + \int_{\mathfrak{E}} g^2 G(d\mathfrak{E}) \end{aligned}$$

kann man in der Gestalt

$$\int_{\mathfrak{E}} (f+g)^2 G(d\mathfrak{E}) \leq \left[\sqrt{\int_{\mathfrak{E}} f^2 G(d\mathfrak{E})} + \sqrt{\int_{\mathfrak{E}} g^2 G(d\mathfrak{E})} \right]^2$$

schreiben, was unmittelbar auf (69) führt.

Bemerkung. Ist $f(P) \in L_2$, so nimmt, da $f^2(P)$ summierbar ist, die Funktion $f(P)$ fast überall auf $\mathfrak{E}$ endliche Werte an.

56. Konvergenz im Mittel. Wir führen einen neuen Konvergenzbegriff für die Klasse L_2 ein.

Definition. Man sagt, daß die Folge der Funktionen $f_n(P)$ aus L_2 im Mittel gegen die Funktion $f(P)$ aus L_2 konvergiert oder einfach, daß sie in L_2 gegen die Funktion $f(P)$ konvergiert, wenn

$$\lim_{n \rightarrow \infty} \int_{\mathfrak{E}} [f(P) - f_n(P)]^2 G(d\mathfrak{E}) = 0 \quad (70)$$

ist.

Zuerst bemerken wir, daß sich das in (70) auftretende Integral nicht ändert, wenn wir $f(P)$ durch eine äquivalente Funktion $g(P)$ ersetzen, und daß auch die Funktion $g(P)$ Grenzwert im Mittel für die Funktionen $f_n(P)$ ist. Im weiteren werden wir äquivalente Funktionen als *identisch* ansehen, d. h., wir werden äquivalente Funktionen, die zu L_2 gehören, als ein und dieselbe Funktion betrachten. Wir beweisen jetzt die Eindeutigkeit des Grenzwertes, d. h., wir beweisen den folgenden Satz.

Satz 1. *Konvergiert die Folge der $f_n(P)$ aus L_2 in L_2 gegen zwei Funktionen $f(P)$ und $g(P)$, so sind diese Funktionen äquivalent.*

Offenbar ist

$$f - g = (f - f_n) + (f_n - g).$$

Für die rechte Seite liefert Ungleichung (69)

$$\sqrt{\int_{\mathfrak{E}} (f-g)^2 G(d\mathfrak{E})} \leq \sqrt{\int_{\mathfrak{E}} (f-f_n)^2 G(d\mathfrak{E})} + \sqrt{\int_{\mathfrak{E}} (g-f_n)^2 G(d\mathfrak{E})}.$$

Für $n \rightarrow \infty$ strebt die rechte Seite gegen 0. Da die linke Seite nicht von n abhängt, ist

$$\int_{\mathfrak{E}} (f-g)^2 G(d\mathfrak{E}) = 0.$$

Nach der Eigenschaft 8 aus [51] folgt hieraus, daß die Differenz $f - g$ äquivalent 0 ist; folglich sind die Funktionen f und g äquivalent. Dieser Satz zeigt die Eindeutigkeit des Grenzwerts von $f_n(P)$ in L_2 , aber natürlich hat nicht jede Funktionenfolge einen Grenzwert im Mittel. Wir bemerken, daß aus der Konvergenz fast überall nicht die Konvergenz im Mittel und aus der Konvergenz im Mittel nicht die Konvergenz fast überall folgt. In Zusammenhang damit beweisen wir den

Satz 2. *Strebt eine Folge von Funktionen $f_n(P)$ aus L_2 auf $\mathfrak{E}$ im Mittel gegen $f(P)$, so kann man aus ihr eine Teilfolge $f_{n_k}(P)$ auswählen, die fast überall auf $\mathfrak{E}$ gegen $f(P)$ konvergiert.*

Aus (70) folgt nach der Eigenschaft 10 aus [51], daß die Funktionen $f_n(P)$ auf $\mathfrak{E}$ dem Maß nach gegen $f(P)$ streben. Satz 2 ist also eine Folgerung des Satzes 7 aus [44].

Folgerung. *Wenn eine Folge von Funktionen $f_n(P)$ im Mittel gegen $f(P)$ strebt und fast überall auf $\mathfrak{E}$ gegen $\varphi(P)$ konvergiert, sind $\varphi(P)$ und $f(P)$ auf $\mathfrak{E}$ äquivalent. Ist nämlich fast überall $f_n(P) \rightarrow \varphi(P)$, so gilt erst recht fast überall $f_{n_k}(P) \rightarrow \varphi(P)$. Da aber, wie wir gesehen haben, fast überall $f_{n_k}(P) \rightarrow f(P)$ gilt, folgt hieraus, daß $\varphi(P)$ und $f(P)$ äquivalent sind.*

Für die Konvergenz im Mittel kann man eine notwendige und hinreichende Bedingung angeben, die der Cauchyschen Bedingung für die Existenz des Grenzwerts von Zahlenfolgen [I, 31] analog ist. Zur Vorbereitung führen wir einen neuen Begriff ein.

Definition. Man sagt, daß eine Folge von Funktionen $f_n(P)$ aus L_2 im Mittel in sich konvergiert, wenn für beliebig gegebenes positives ε ein solches N existiert, daß

$$\int_{\mathfrak{E}} (f_n - f_m)^2 G(d\mathfrak{E}) \leq \varepsilon \quad \text{für } n > N \quad \text{und} \quad m > N \quad (71)$$

gilt.

Satz 3. *Dafür, daß die Folge der $f_n(P)$ im Mittel gegen eine Funktion aus L_2 konvergiert, ist notwendig, daß sie im Mittel in sich konvergent ist.*

Beweis. Die Folge konvergiere im Mittel gegen eine Funktion $f(P)$. Wir schreiben die Differenz von $f_n(P) - f_m(P)$ in der Form

$$f_n - f_m = (f_n - f) + (f - f_m)$$

und wenden die Ungleichung (69) an:

$$\sqrt{\int_{\mathfrak{E}} (f_n - f_m)^2 G(d\mathfrak{E})} \leq \sqrt{\int_{\mathfrak{E}} (f_n - f)^2 G(d\mathfrak{E})} + \sqrt{\int_{\mathfrak{E}} (f - f_m)^2 G(d\mathfrak{E})}.$$

Es sei ε eine gegebene positive Zahl. Infolge der Konvergenz im Mittel gegen $f(P)$ existiert ein solches N , daß für $n > N$ und $m > N$ die Integrale, die in der Ungleichung rechts unter dem Wurzelzeichen stehen, höchstens gleich $\varepsilon^2/4$ sind. Dadurch erhält man aus der letzten Ungleichung unmittelbar (71), und der Satz ist damit bewiesen. Wir beweisen jetzt die Umkehrung.

Satz 4. *Dafür, daß die Folge der $f_n(P)$ im Mittel gegen eine Funktion konvergiert, ist hinreichend, daß sie im Mittel in sich konvergent ist.*

Beweis. Die Folge $f_n(P)$ konvergiere in sich, und es muß nun gezeigt werden, daß sie im Mittel gegen eine bestimmte Funktion konvergiert. Infolge der Konvergenz in sich existiert eine wachsende Folge von Indizes $n_1 < n_2 < n_3 < \dots$ derart, daß

$$\int_{\mathfrak{E}} [f_{n_{k+1}}(P) - f_{n_k}(P)]^2 G(d\mathfrak{E}) \leq \frac{1}{2^{2k}}$$

gilt. Wenden wir die Ungleichung (67) für den Fall $f = |f_{n_{k+1}} - f_{n_k}|$ und $g \equiv 1$ an, so erhalten wir

$$\int_{\mathfrak{E}} |f_{n_{k+1}}(P) - f_{n_k}(P)| G(d\mathfrak{E}) \leq \sqrt{\int_{\mathfrak{E}} [f_{n_{k+1}}(P) - f_{n_k}(P)]^2 G(d\mathfrak{E})} \sqrt{\int_{\mathfrak{E}} G(d\mathfrak{E})}$$

oder nach der vorigen Ungleichung

$$\int_{\mathfrak{E}} |f_{n_{k+1}}(P) - f_{n_k}(P)| G(d\mathfrak{E}) \leq \frac{1}{2k} \sqrt{G(\mathfrak{E})}.$$

Hieraus folgt die Konvergenz der Reihe $\sum_{k=1}^{\infty} \int_{\mathfrak{E}} |f_{n_{k+1}}(P) - f_{n_k}(P)| G(d\mathfrak{E})$, und nach Satz 3 aus [54] konvergiert die Reihe $\sum_{k=1}^{\infty} |f_{n_{k+1}}(P) - f_{n_k}(P)|$ fast überall auf $\mathfrak{E}$. Um so mehr konvergiert die Reihe $f_{n_1}(P) + \sum_{k=1}^{\infty} [f_{n_{k+1}}(P) - f_{n_k}(P)]$ fast überall, und die Summe der ersten p Glieder dieser Reihe ist gleich $f_{n_p}(P)$, d. h., fast überall auf $\mathfrak{E}$ strebt die Folge $f_{n_1}(P), f_{n_2}(P), f_{n_3}(P), \dots$ gegen eine Funktion $f(P)$ mit endlichen Werten. Wir zeigen, daß $f(P) \in L_2$ ist und daß $f_n(P)$ im Mittel gegen $f(P)$ konvergiert. Da die Folge $f_n(P)$ in sich konvergiert, existiert für ein beliebig gegebenes positives ε ein solches N , daß

$$\int_{\mathfrak{E}} [f_{n_k}(P) - f_n(P)]^2 G(d\mathfrak{E}) \leq \varepsilon \quad \text{für } n_k > N \quad \text{und } n > N$$

ist. Wenn in dieser Ungleichung n_k gegen ∞ strebt und wir Satz 4 aus [54] benutzen, erhalten wir

$$\int_{\mathfrak{E}} [f(P) - f_n(P)]^2 G(d\mathfrak{E}) \leq \varepsilon \quad \text{für } n > N. \quad (72)$$

Hieraus folgt, daß die Differenz $f(P) - f_n(P)$ zu L_2 gehört; $f_n(P)$ gehört ebenfalls zu L_2 . Wenn wir $f(P) - f_n(P)$ und $f_n(P)$ addieren, erhalten wir nach Satz 2 aus [55], daß auch $f(P) \in L_2$ ist. Die Ungleichung (72) besagt schließlich, daß $f_n(P)$ im Mittel gegen $f(P)$ strebt.

Die letzten beiden Sätze liefern folgende Behauptung: *Die Konvergenz einer Folge von Funktionen $f_n(P)$ im Mittel in sich ist notwendig und hinreichend dafür, daß diese Folge im Mittel gegen eine Funktion aus L_2 konvergiert.*

Satz 5. *Gehören $f_n(P)$ und $g_n(P)$ zu L_2 und gilt im Mittel $f_n(P) \rightarrow f(P)$ und $g_n(P) \rightarrow g(P)$, so ist*

$$\int_{\mathfrak{E}} f_n(P) g_n(P) G(d\mathfrak{E}) \rightarrow \int_{\mathfrak{E}} f(P) g(P) G(d\mathfrak{E}).$$

Führen wir für beliebige Funktionen $\varphi(P)$ und $\psi(P)$ aus L_2 die Bezeichnung

$$(\varphi, \psi) = \int_{\mathfrak{E}} \varphi(P) \psi(P) G(d\mathfrak{E})$$

ein [IV/1, 22], so kann man die Schwarzsche Ungleichung in der Gestalt

$$(\varphi, \psi)^2 \leq (\varphi, \varphi) (\psi, \psi)$$

schreiben. Wir setzen nun

$$f_n(P) - f(P) = \varphi_n(P), \quad g_n(P) - g(P) = \psi_n(P);$$

nach Voraussetzung gilt $(\varphi_n, \varphi_n) \rightarrow 0$ und $(\psi_n, \psi_n) \rightarrow 0$. Wir bilden die Differenz

$$(f, g) - (f_n, g_n) = (f, g) - (f + \varphi_n, g + \psi_n) = -(f, \psi_n) - (\varphi_n, g) - (\varphi_n, \psi_n);$$

hieraus folgt

$$\begin{aligned} |(f, g) - (f_n, g_n)| &\leq |(f, \psi_n)| + |(\varphi_n, g)| + |(\varphi_n, \psi_n)| \\ &\leq \sqrt{(f, f)} \sqrt{(\psi_n, \psi_n)} + \sqrt{(\varphi_n, \varphi_n)} \sqrt{(g, g)} + \sqrt{(\varphi_n, \varphi_n)} \sqrt{(\psi_n, \psi_n)}. \end{aligned}$$

Die rechte Seite strebt für $n \rightarrow \infty$ gegen 0; somit ist $|(f, g) - (f_n, g_n)| \rightarrow 0$, also $(f_n, g_n) \rightarrow (f, g)$, was zu beweisen war.

57. Der Hilbertsche Funktionenraum. Die Familie der Funktionen aus L_2 stellt ebenso wie die Familie C [14] einen Funktionenraum dar. Die Elemente dieses Raumes sind reelle Funktionen mit auf $\mathfrak{E}$ summierbarem Quadrat. Äquivalente Funktionen werden dabei identifiziert, d. h., sie entsprechen ein und demselben Element aus L_2 .

Man definiert die Summe von Elementen und das Produkt mit einer reellen Zahl, wobei diese Operationen den gewöhnlichen Gesetzen der Algebra gehorchen. Unter der *Norm* des Elements x (Länge des Vektors) verstehen wir die nichtnegative Zahl, die durch

$$||f(P)|| = \sqrt{\int_{\mathfrak{E}} [f(P)]^2 G(d\mathfrak{E})} \quad (73)$$

definiert ist.

Man sagt, daß die Folge der Elemente $f_n(P)$ aus L_2 gegen das Element $f(P)$ aus L_2 konvergiert, wenn $||f(P) - f_n(P)|| \rightarrow 0$ für $n \rightarrow \infty$ gilt. Diese Konvergenz der Norm nach ist auf Grund von (73) gleichbedeutend mit der Konvergenz im Mittel.

Wir führen noch den Begriff des *skalaren* (oder *inneren*) *Produkts* zweier Elemente $f(P)$ und $g(P)$ ein. Es ist durch

$$(f, g) = \int_{\mathfrak{E}} fg G(d\mathfrak{E}) \quad (74)$$

definiert; offenbar ist

$$||f|| = \sqrt{(f, f)}. \quad (75)$$

Der *Abstand* zwischen zwei Elementen f und g ist durch

$$\varrho(f, g) = ||f - g|| = \sqrt{\int_{\mathfrak{E}} (f - g)^2 G(d\mathfrak{E})} = \sqrt{(f - g, f - g)} \quad (76)$$

definiert.

Es seien nun die Elemente f, g und h gegeben. Wir wenden auf die Formel

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

die Ungleichung (69) an und erhalten auf diese Weise nach der Definition (76) die sogenannte *Dreiecksungleichung*

$$\varrho(f, h) \leq \varrho(f, g) + \varrho(g, h). \quad (77)$$

Die *Null* des Raumes oder das *Nullelement* ist die Funktion, die in $\mathfrak{E}$ identisch 0 ist, sowie jede Funktion, die äquivalent 0 ist. Die Norm des Nullelements ist gleich 0, die Norm jedes anderen Elements ist nach der Eigenschaft 4 aus [49] positiv. Der Abstand $\varrho(f, g)$ ist nichtnegativ und gleich 0 genau dann, wenn die beiden Elemente übereinstimmen, d. h., wenn die Funktionen f und g äquivalent sind. Der Abstand und das skalare Produkt sind symmetrisch, d. h., es ist

$$\varrho(g, f) = \varrho(f, g), \quad (g, f) = (f, g).$$

Dafür, daß eine Folge von Elementen in diesem Funktionenraum einen Grenzwert hat, ist notwendig und hinreichend, daß sie in sich konvergent¹⁾ ist, d. h., daß zu jedem positiven ε ein N derart existiert, daß $\|f_m - f_n\| < \varepsilon$ für $n, m > N$ ist. Die letzte Eigenschaft heißt *Vollständigkeit des Raumes L_2* .

Die Ungleichung (77) gilt offenbar auch für eine beliebige endliche Anzahl von Summanden:

$$\varrho(f_1, f_m) \leq \varrho(f_1, f_2) + \varrho(f_2, f_3) + \dots + \varrho(f_{m-1}, f_m)$$

oder

$$\|f_1 - f_m\| \leq \|f_1 - f_2\| + \|f_2 - f_3\| + \dots + \|f_{m-1} - f_m\|. \quad (77_1)$$

Analog hierzu kann man auch einen Funktionenraum L_2 für komplexe Funktionen (54) definieren. Die Funktion $f(P) = f_1(P) + i f_2(P)$ heißt *Funktion aus L_2* , wenn $f_1(P)$ und $f_2(P)$ zu L_2 gehören. Dabei ist $|f(P)|^2$ eine summierbare Funktion. Die Sätze 1 und 2 aus [55] behalten ihre Gültigkeit. Die Ungleichungen (67) und (69) kann man wie folgt schreiben:

$$\left. \begin{aligned} \int_{\mathfrak{E}} |f g G(d\mathfrak{E})|^2 &\leq \int_{\mathfrak{E}} |f|^2 G(d\mathfrak{E}) \int_{\mathfrak{E}} |g|^2 G(d\mathfrak{E}), \\ \sqrt{\int_{\mathfrak{E}} |f + g|^2 G(d\mathfrak{E})} &\leq \sqrt{\int_{\mathfrak{E}} |f|^2 G(d\mathfrak{E})} + \sqrt{\int_{\mathfrak{E}} |g|^2 G(d\mathfrak{E})}. \end{aligned} \right\} \quad (78)$$

In der Definition der Konvergenz im Mittel und der Konvergenz in sich müssen die Quadrate der Abstände, $(f - f_n)^2$ und $(f_n - f_m)^2$, durch die Quadrate der Absolutbeträge der Abstände, $|f - f_n|^2$ und $|f_n - f_m|^2$, ersetzt werden. Die Sätze 3 und 4 aus [56] behalten ihre Gültigkeit. Bei der Definition des Funktionenraumes darf man nicht nur die Multiplikation mit reellen, sondern muß man auch die mit komplexen Zahlen zulassen. Die Norm eines Elements ist durch

$$\|f\| = \sqrt{\int_{\mathfrak{E}} |f|^2 G(d\mathfrak{E})} \quad (79)$$

und das Skalarprodukt durch

$$(f, g) = \int_{\mathfrak{E}} f \bar{g} G(d\mathfrak{E}) \quad (80)$$

definiert, wobei wir mit $\bar{\alpha}$, wie üblich, die zu α konjugiert komplexe Zahl bezeichnen. Wie vorher gilt auch hier die Formel (77). Der Abstand zwischen Elementen ist durch die Formel (76) definiert, in der $(f - g)^2$ durch $|f - g|^2$ zu ersetzen ist, und er hat dieselben Eigenschaften wie im reellen Raum. Für das skalare Produkt gilt

$$(g, f) = \overline{(f, g)}.$$

Alles über den Raum der komplexen Funktionen Gesagte folgt unmittelbar daraus, daß die Funktionen $f_1(P)$ und $f_2(P)$ zum reellen Raum L_2 gehören. Der Funktionenraum L_2 wird oft auch *Hilbertscher Funktionenraum* genannt.

Wir geben jetzt noch einen Spezialfall an. Wir nehmen an, die Funktion $G(\mathfrak{E})$ entspreche punktförmigen Massen, die sich in den Punkten $P_1, P_2, \dots, P_m$ befinden und gleich 1 sind. Dabei geht das Lebesgue-Stieltjessche Integral für jede Funktion $f(P)$, die in diesen Punkten endliche Werte annimmt, über einer beliebigen Menge $\mathfrak{E}$, die diese Punkte enthält, in die endliche Summe

$$\int_{\mathfrak{E}} f(P) G(d\mathfrak{E}) = \sum_{k=1}^m f(P_k)$$

über.

¹⁾ Hierfür sagt man auch, die Elementfolge ist eine *Cauchyfolge* (Anm. d. Red.).

Wenn wir die Werte jeder Funktion $f(P)$ in den Punkten P_k ($k=1, \dots, m$) als Komponente eines m -dimensionalen komplexen Vektors auffassen, erhalten wir den m -dimensionalen komplexen Raum R_m , dessen Eigenschaften wir in [III/1, 25] behandelt haben. Die oben angegebenen Definitionen der Addition, der Multiplikation mit einer Zahl, der Norm, des Skalarprodukts usw. entsprechen den früher definierten Begriffen.

58. Orthogonale Funktionensysteme. Unmittelbar mit dem Funktionenraum L_2 hängt die Theorie der orthogonalen Funktionensysteme zusammen. Wir haben diese Systeme bereits früher behandelt [IV/1, 36, 89]. Wir ergänzen diese Darlegungen, indem wir den Begriff des Lebesgue-Stieltjesschen oder des Lebesgueschen Integrals heranziehen. Wir werden zuerst nur reelle Funktionen betrachten.

Definition. Man sagt, daß die Funktionen

$$\varphi_1(P), \varphi_2(P), \dots \quad (81)$$

aus L_2 , die auf einer meßbaren Menge $\mathfrak{E}$ endlichen Maßes gegeben sind, ein *orthogonales und normiertes System* (oder *normiertes Orthogonalsystem* oder auch *Orthornormalsystem*) bilden, wenn die folgenden Bedingungen gelten:

$$\int_{\mathfrak{E}} \varphi_k(P) \varphi_l(P) G(d\mathfrak{E}) = \begin{cases} 0 & \text{für } k \neq l, \\ 1 & \text{für } k = l. \end{cases} \quad (82)$$

Für jede Funktion $f(P)$ aus L_2 können wir ihre *Fourierkoeffizienten* bezüglich des Systems (81) bilden:

$$a_n = \int_{\mathfrak{E}} f(P) \varphi_n(P) G(d\mathfrak{E}); \quad (83)$$

ihre *Fourierreihe* lautet

$$\sum_{n=1}^{\infty} a_n \varphi_n(P). \quad (84)$$

Bezüglich der Konvergenz dieser Reihe können wir nichts aussagen, aber wir können die Partialsummen dieser Reihe bilden:

$$S_n(f) = \sum_{k=1}^n a_k \varphi_k(P). \quad (85)$$

Der Ausdruck

$$\int_{\mathfrak{E}} \left[f(P) - \sum_{k=1}^n b_k \varphi_k(P) \right]^2 G(d\mathfrak{E}) \quad (86)$$

nimmt seinen kleinsten Wert an, wenn die Koeffizienten b_k gleich den Fourierkoeffizienten a_k sind. Dabei ergibt sich für (86)

$$\int_{\mathfrak{E}} [f(P) - S_n(f)]^2 G(d\mathfrak{E}) = \int_{\mathfrak{E}} f^2(P) G(d\mathfrak{E}) - \sum_{k=1}^n a_k^2. \quad (87)$$

Hieraus erhalten wir die *Besselsche Ungleichung*

$$\sum_{k=1}^{\infty} a_k^2 \leq \int_{\mathfrak{E}} f^2(P) G(d\mathfrak{E}) \quad (88)$$

und die Konvergenz der in dieser Ungleichung links stehenden Reihe. Wenn in (88) das Gleichheitszeichen steht, nennt man die so erhaltene Formel

$$\int_{\mathfrak{E}} f^2(P) G(d\mathfrak{E}) = \sum_{k=1}^{\infty} a_k^2 \quad (89)$$

die Vollständigkeitsrelation.¹⁾ Infolge (87) ist die Vollständigkeitsrelation damit gleichbedeutend, daß die Fourierreihe $S_n(f)$ im Mittel gegen die Funktion $f(P)$ strebt.

Wir beweisen jetzt den folgenden grundlegenden Satz.

Satz 1 (RIESZ-FISCHER). Ist c_n eine beliebige Folge reeller Zahlen und bilden die Quadrate dieser Zahlen eine konvergente Reihe,

$$\sum_{n=1}^{\infty} c_n^2 < \infty, \quad (90)$$

so existiert eine eindeutig bestimmte Funktion aus L_2 , für die die Zahlen c_n die Fourierkoeffizienten bezüglich des Systems (81) sind und für die die Vollständigkeitsrelation (89) gilt.

Beweis. Wir bilden die Funktionen

$$S_n(P) = \sum_{k=1}^n c_k \varphi_k(P). \quad (91)$$

Da das System (81) orthogonal und normiert ist, gilt

$$\int_{\mathfrak{E}} [S_q(P) - S_p(P)]^2 G(d\mathfrak{E}) = c_{p+1}^2 + c_{p+2}^2 + \dots + c_q^2 \quad (q > p),$$

und aus der Konvergenz der Reihe (90) folgt, daß die rechte Seite dieser Formel für unbeschränkt wachsendes p gegen 0 strebt, d. h., die Funktionenfolge (91) aus L_2 konvergiert in sich. Folglich existiert eine solche Funktion $f(P)$ aus L_2 , gegen die die Folge der $S_n(P)$ im Mittel konvergiert:

$$\lim_{n \rightarrow \infty} \int_{\mathfrak{E}} [f(P) - S_n(P)]^2 G(d\mathfrak{E}) = 0. \quad (92)$$

Wir zeigen jetzt, daß die c_k gleich den Fourierkoeffizienten a_k dieser Funktion sind. Wenn wir (83) berücksichtigen und ferner beachten, daß das System (81) orthogonal und normiert ist, so können wir

$$\int_{\mathfrak{E}} [f(P) - S_n(P)]^2 G(d\mathfrak{E}) = \left[\int_{\mathfrak{E}} f^2(P) G(d\mathfrak{E}) - \sum_{k=1}^n a_k^2 \right] + \sum_{k=1}^n (c_k - a_k)^2 \quad (93)$$

schreiben. Die auf der rechten Seite in eckigen Klammern stehende Differenz ist nach der Besselschen Ungleichung nichtnegativ. Die übrigen Summanden der rechten Seite sind ebenfalls nichtnegativ. Für $n \rightarrow \infty$ strebt die linke Seite gegen 0; also gilt dasselbe auch für die rechte Seite. Hieraus folgt unmittelbar, daß jeder der nichtnegativen Summanden $(c_k - a_k)^2$ gleich 0, d. h. $c_k = a_k$ ist, was wir beweisen wollten. Somit sind die Funktionen (91) Partialsummen der Fourierreihe von $f(P)$, und aus (93) folgt unmittelbar, daß für $f(P)$ die Vollständigkeitsrelation gilt. Wir müssen noch zeigen, daß die Funktion $f(P)$ mit den oben angegebenen Eigenschaften eindeutig bestimmt ist. Wenn außer $f(P)$ noch eine Funktion $g(P)$ mit den obigen Eigenschaften existiert, so ist (85) sowohl Partialsumme der Fourierreihe für $f(P)$ als auch für $g(P)$. Nach Voraussetzung gilt die Vollständigkeitsrelation sowohl für $f(P)$ als auch für $g(P)$, d. h., die Folge der $S_n(P)$ strebt im Mittel gegen $f(P)$ und gegen $g(P)$. Wegen der Eindeutigkeit des Grenzwerts in L_2 folgt hieraus, daß $g(P)$ und $f(P)$ äquivalent sind, d. h. ein und dasselbe Element aus L_2 darstellen. Damit ist der Satz vollständig bewiesen.

¹⁾ Die Formel (89) wird auch *Parsevalsche Gleichung* genannt. Bezüglich des Gebrauchs der Begriffe Vollständigkeit und Abgeschlossenheit von Funktionensystemen vgl. [II, 163] (Anm. d. Red.).

Definition 1. Das orthogonale und normierte System (81) heißt *vollständig*¹⁾, wenn für jede Funktion $f(P)$ aus L_2 die Vollständigkeitsrelation (89) gilt.

Beim Beweis von Satz 1 haben wir nicht vorausgesetzt, daß das System (81) vollständig ist. Wenn das der Fall ist, brauchen wir nicht zu betonen, daß für die Funktion die Vollständigkeitsrelation gilt, da dies nach Definition eines vollständigen Systems für jede Funktion aus L_2 gilt. Daher kann man für ein vollständiges System den Satz 1 folgendermaßen formulieren.

Satz 1'. Wenn das System (81) vollständig und c_n eine beliebig gegebene Folge reeller Zahlen ist, für die die Reihe (90) konvergiert, existiert eine eindeutig bestimmte Funktion aus L_2 , deren Fourierkoeffizienten die Zahlen c_n sind.

Außer dem Begriff des vollständigen Systems führen wir noch den Begriff des *abgeschlossenen*¹⁾ Systems ein.

Definition 2. Das System (81) heißt *abgeschlossen*, wenn in L_2 keine Funktion existiert, die verschieden von 0 (d. h. nicht äquivalent 0) und zu allen $\varphi_k(P)$ orthogonal ist.

Wir zeigen nun, daß die Begriffe Vollständigkeit und Abgeschlossenheit äquivalent sind.

Satz 2. Das System (81) ist genau dann abgeschlossen, wenn es vollständig ist.

Beweis. Wir beweisen die Notwendigkeit indirekt. Wir setzen also voraus, daß das System (81) abgeschlossen ist, und nehmen an, es sei nicht vollständig, d. h., es existiere eine Funktion $h(P)$ aus L_2 mit Fourierkoeffizienten a_k derart, daß

$$\int_{\mathfrak{E}} h^2(P) G(d\mathfrak{E}) > \sum_{k=1}^{\infty} a_k^2 \quad (94)$$

ist. Andererseits existiert nach Satz 1 eine Funktion $f(P)$ aus L_2 mit denselben Fourierkoeffizienten a_k , für die die Vollständigkeitsrelation (89) gilt. Wenn wir diese Formel mit (94) vergleichen, erhalten wir

$$\int_{\mathfrak{E}} h^2(P) G(d\mathfrak{E}) > \int_{\mathfrak{E}} f^2(P) G(d\mathfrak{E}). \quad (95)$$

Alle Fourierkoeffizienten der Differenz $f(P) - h(P)$ sind aber gleich 0, d. h., diese Differenz ist zu allen $\varphi_k(P)$ orthogonal, also infolge der Abgeschlossenheit äquivalent 0, d. h., $h(P)$ ist äquivalent zu $f(P)$, und das ist ein Widerspruch zu (95). Damit ist die Notwendigkeit bewiesen. Wir beweisen jetzt, daß die Bedingung auch hinreichend ist. Wir setzen voraus, daß das System vollständig ist, und müssen die Abgeschlossenheit beweisen, d. h., wir müssen zeigen: Sind alle Fourierkoeffizienten der Funktion $f(P)$ gleich 0, so ist diese Funktion äquivalent 0. Weil die Vollständigkeit des Systems vorausgesetzt wird, können wir für die Funktion $f(P)$ die Vollständigkeitsrelation (89) hinschreiben, die, da alle Fourierkoeffizienten der Funktion $f(P)$ gleich 0 sind,

$$\int_{\mathfrak{E}} f^2(P) G(d\mathfrak{E}) = 0$$

lautet. Hieraus folgt nach der Eigenschaft 8 aus [51], daß $f(P)$ äquivalent 0 ist.

Wir bemerken, daß wir jedes System linear unabhängiger Funktionen $\varphi_n(P)$ aus L_2 orthogonalisieren können, was wir bereits früher beschrieben haben [IV/1, 36].

¹⁾ Siehe die Fußnote auf S. 147 (Anm. d. Red.).

Alles oben Gesagte gilt unmittelbar auch für den Fall komplexer Funktionen aus L_2 . Für ein orthogonales und normiertes System (81) gelten dann die Gleichungen

$$\int_{\mathfrak{E}} \varphi_k(P) \overline{\varphi_l(P)} G(d\mathfrak{E}) = \begin{cases} 0 & \text{für } k \neq l, \\ 1 & \text{für } k = l, \end{cases} \quad (96)$$

und die Fourierkoeffizienten sind durch

$$a_n = \int_{\mathfrak{E}} f(P) \overline{\varphi_n(P)} G(d\mathfrak{E}) \quad (97)$$

definiert.

In den weiteren Formeln müssen wir überall für die Quadrate der Funktionen und Zahlen die Quadrate der Absolutbeträge schreiben. So hat z. B. die Vollständigkeitsrelation die Gestalt

$$\int_{\mathfrak{E}} |f(P)|^2 G(d\mathfrak{E}) = \sum_{k=1}^{\infty} |a_k|^2. \quad (98)$$

Die oben angegebenen Sätze gelten auch jetzt, nur müssen wir statt der Reihe (90) die Reihe der Zahlen $|c_n|^2$ betrachten.

Wir betrachten noch die sogenannte verallgemeinerte Vollständigkeitsrelation. Es seien a_n und b_n die Fourierkoeffizienten der Funktionen $f(P)$ und $g(P)$, und das System (81) sei vollständig. Die Funktion $f(P) + g(P)$ hat die Fourierkoeffizienten $a_n + b_n$, die Funktion $f(P) + ig(P)$ die Fourierkoeffizienten $a_n + ib_n$. Die Vollständigkeitsrelationen für sie lauten

$$\int_{\mathfrak{E}} |f+g|^2 G(d\mathfrak{E}) = \sum_{n=1}^{\infty} |a_n + b_n|^2$$

bzw.

$$\int_{\mathfrak{E}} |f+ig|^2 G(d\mathfrak{E}) = \sum_{n=1}^{\infty} |a_n + ib_n|^2$$

oder

$$\int_{\mathfrak{E}} [|f|^2 + |g|^2 + (fg + f\bar{g})] G(d\mathfrak{E}) = \sum_{n=1}^{\infty} [|a_n|^2 + |b_n|^2 + (\bar{a}_n b_n + a_n \bar{b}_n)],$$

$$\int_{\mathfrak{E}} [|f|^2 + |g|^2 + i(\bar{f}g - f\bar{g})] G(d\mathfrak{E}) = \sum_{n=1}^{\infty} [|a_n|^2 + |b_n|^2 + i(\bar{a}_n b_n - a_n \bar{b}_n)].$$

Wenn wir die Vollständigkeitsrelationen für f und g beachten und die zweite Gleichung mit i multiplizieren und zur ersten addieren, erhalten wir die *verallgemeinerte Vollständigkeitsrelation*

$$\int_{\mathfrak{E}} f\bar{g} G(d\mathfrak{E}) = \sum_{n=1}^{\infty} a_n \bar{b}_n. \quad (99)$$

Im Fall reeller Funktionen lautet die verallgemeinerte Vollständigkeitsrelation

$$\int_{\mathfrak{E}} fg G(d\mathfrak{E}) = \sum_{n=1}^{\infty} a_n b_n. \quad (100)$$

Aus der verallgemeinerten Vollständigkeitsrelation folgt unmittelbar, daß man die Fourierreihe einer beliebigen Funktion $f(P)$ aus L_2 über einer Menge $\mathfrak{E}$ oder über einer beliebigen meßbaren Teilmenge $\mathfrak{E}'$ von $\mathfrak{E}$ gliedweise integrieren kann

[II, 160]: Sind a_k ($k=1, 2, \dots$) die Fourierkoeffizienten von $f(P)$, so gilt

$$\int_{\mathfrak{E}} f(P) G(d\mathfrak{E}) = \sum_{k=1}^{\infty} a_k \int_{\mathfrak{E}} \varphi_k(P) G(d\mathfrak{E}). \quad (101)$$

Wir erwähnen noch eine Eigenschaft des Raumes L_2 , aus der die Existenz eines vollständigen orthogonalen normierten Systems in L_2 folgt. Diese Eigenschaft wird gewöhnlich *Separabilität* genannt; man kann sie folgendermaßen formulieren: *Es existiert eine abzählbare Menge von Elementen $\psi_k(P)$ ($k=1, 2, \dots$) aus L_2 , die in L_2 überall dicht ist, d. h. eine Menge mit der Eigenschaft, daß zu jedem $f(P)$ aus L_2 und beliebig gegebenem positivem ε ein solches Element $\psi_m(P)$ aus der abzählbaren Menge mit $\|f(P) - \psi_m(P)\| \leq \varepsilon$ existiert.* In einem der folgenden Abschnitte werden wir die Separabilität von L_2 beweisen. Wir werden jetzt zeigen, daß aus der Separabilität die Existenz eines vollständigen orthonormierten Systems folgt. Wenn wir das System der $\psi_k(P)$ orthonormalisieren [IV/1, 3], so erhalten wir ein orthonormiertes System $\varphi_k(P)$ ($k=1, 2, \dots$). Wir müssen beweisen, daß es vollständig ist. Nach dem oben Gesagten existiert zu jedem $f(P)$ aus L_2 und zu jedem positiven ε ein solches $\psi_m(P)$, daß $\|f - \psi_m\| \leq \varepsilon$ ist; $\psi_m(P)$ ist aber gemäß dem Orthogonalisierungsverfahren eine endliche Linearkombination von Funktionen $\varphi_k(P)$, d. h.

$$\psi_m(P) = c_1 \varphi_1(P) + c_2 \varphi_2(P) + \dots + c_l \varphi_l(P),$$

und somit gilt

$$\|f - \psi_m\|^2 = \int_{\mathfrak{E}} \left[f(P) - \sum_{k=1}^l c_k \varphi_k(P) \right]^2 G(d\mathfrak{E}) \leq \varepsilon^2.$$

Wenn wir die c_k durch die Fourierkoeffizienten von $f(P)$ bezüglich des Systems $\varphi_k(P)$ ersetzen, gilt erst recht die Ungleichung

$$\int_{\mathfrak{E}} [f(P) - S_l(f)]^2 G(d\mathfrak{E}) \leq \varepsilon^2$$

[II, 160]. $S_l(f)$ ist dabei eine Partialsumme der Fourierreihe der Funktion $f(P)$. Aus dieser Ungleichung folgt, da ε beliebig war, die Vollständigkeit des Systems $\varphi_k(P)$.

Bemerkung. Entspricht die Funktion $G(\mathfrak{E})$ punktförmigen Massen, die in den Punkten $P_1, P_2, \dots, P_m$ liegen, so enthält das vollständige System nur m Elemente. Dieser Fall ist jedoch uninteressant. Er führt, wie wir schon bemerkten, auf den endlichdimensionalen Raum R_m .

59. Der Raum l_2 . Neben dem Raum L_2 betrachten wir noch den Raum l_2 der unendlichen Folgen, der eng mit L_2 verknüpft ist, wobei wir sofort den komplexen Fall betrachten werden.

Eine unendliche Folge x komplexer Zahlen $(x_1, x_2, x_3, \dots)$ heißt, wenn die Reihe der $|x_n|^2$ konvergent ist, ein *Element* von l_2 . Jetzt kann man die Multiplikation eines Elements mit einer komplexen Zahl und die Addition von Elementen definieren. Das Element cx hat per definitionem die Koordinaten $(cx_1, cx_2, \dots)$, und die Summe der Elemente x und y mit den Koordinaten x_n bzw. y_n hat die Koordinaten $x_n + y_n$; dabei konvergiert die Reihe der Zahlen $|x_n + y_n|^2$. Dies folgt aus der Konvergenz der Reihen der Zahlen $|x_n|^2$ und $|y_n|^2$ und aus der Ungleichung $|x_n + y_n|^2 \leq 2(|x_n|^2 + |y_n|^2)$. Die Norm des Elements x ist durch

$$\|x\| = \sqrt{\sum_{n=1}^{\infty} |x_n|^2} \quad (102)$$

definiert und das Skalarprodukt der Elemente x und y durch

$$(x, y) = \sum_{n=1}^{\infty} x_n \bar{y}_n, \quad (103)$$

wobei die absolute Konvergenz der rechts stehenden Reihe unmittelbar aus der Ungleichung

$$|x_n \bar{y}_n| \leq \frac{1}{2} (|x_n|^2 + |y_n|^2)$$

folgt. Ferner gilt

$$||x||^2 = (x, x). \quad (104)$$

Der Abstand zwischen den Elementen x und y ist durch

$$\varrho(x, y) = \sqrt{(x - y, x - y)} = ||x - y|| = \sqrt{\sum_{n=1}^{\infty} |x_n - y_n|^2} \quad (105)$$

definiert.

Analog zu (67) und (69) erhält man die Ungleichungen

$$\left| \sum_{n=1}^{\infty} x_n \bar{y}_n \right|^2 \leq \sum_{n=1}^{\infty} |x_n|^2 \sum_{n=1}^{\infty} |y_n|^2, \quad (106)$$

$$\sqrt{\sum_{n=1}^{\infty} |x_n + y_n|^2} \leq \sqrt{\sum_{n=1}^{\infty} |x_n|^2} + \sqrt{\sum_{n=1}^{\infty} |y_n|^2}. \quad (107)$$

Sie werden genau so wie (67) und (69) bewiesen. Wir bemerken noch, daß man die Ungleichung (106) auch wie folgt schreiben kann:

$$|(x, y)|^2 \leq ||x||^2 ||y||^2. \quad (108)$$

Für den Abstand gilt nach (107) die Dreiecksungleichung. Das Element, dessen sämtliche Koordinaten gleich 0 sind, heißt *Nullelement* des Raumes. Wir sagen, daß die Folge der Elemente $x^{(n)}$ gegen das Element x konvergiert, wenn $||x - x^{(n)}|| \rightarrow 0$ gilt. Es seien $x_k^{(n)}$ die Koordinaten von $x^{(n)}$ und x_k die Koordinaten von x . Die Konvergenz von $x^{(n)}$ gegen x ist gleichbedeutend mit:

$$||x - x^{(n)}||^2 = \sum_{k=1}^{\infty} |x_k - x_k^{(n)}|^2 \rightarrow 0 \quad \text{für } n \rightarrow \infty. \quad (109)$$

Wir zeigen jetzt den Zusammenhang zwischen den Räumen L_2 und l_2 . Wir nehmen ein beliebiges vollständiges orthonormiertes System (81). Jeder Funktion $f(P)$ aus L_2 entspricht eine Folge komplexer Zahlen a_k (ihre Fourierkoeffizienten), wobei die Reihe der $|a_k|^2$ konvergiert. Umgekehrt entspricht jeder Folge komplexer Zahlen nach Satz 1 aus [58] eine bestimmte Funktion aus L_2 . Wenn wir irgendein vollständiges System (81) nehmen, erhalten wir auf diese Weise eine eindeutige Zuordnung zwischen den Elementen von L_2 und von l_2 . Jedem Element aus L_2 entspricht ein bestimmtes Element aus l_2 und umgekehrt. Da die Fourierkoeffizienten einer endlichen Linearkombination von Funktionen $\sum_{k=1}^m c_k f_k(P)$ gleich der entsprechenden Linearkombination der Fourierkoeffizienten der Summanden $f_k(P)$ sind, können wir behaupten, daß *die oben angegebene eindeutige Zuordnung distributiv ist*, d. h., wenn den Elementen $f_k(P)$ ($k = 1, 2, \dots, m$) aus L_2 die Elemente $x^{(k)}$ aus l_2 entsprechen, entspricht dem Element $\sum_{k=1}^m c_k f_k(P)$ das Element $\sum_{k=1}^m c_k x^{(k)}$. Nach der

allgemeinen Vollständigkeitsrelation (99) stimmen die Skalarprodukte entsprechender Elemente in L_2 und l_2 bei obiger Zuordnung überein. Die Normen der Elemente stimmen nach der Vollständigkeitsrelation (98) ebenfalls überein. Die Räume L_2 und l_2 sind verschiedene Realisierungen ein und desselben abstrakten Raumes. Im folgenden werden wir die Eigenschaften dieses abstrakten Raumes und Operatoren in ihm betrachten, wobei wir diesen Raum mit Hilfe eines Axiomensystems beschreiben.

Wir erwähnen noch den Begriff der Konvergenz in sich im Raum l_2 .¹⁾ Man sagt, daß eine Folge von Elementen $x^{(n)}$ in l_2 *in sich konvergiert*, wenn zu jedem gegebenen positiven ε ein solches N existiert, daß $\|x^{(n)} - x^{(m)}\| \leq \varepsilon$ für $m, n > N$ ist. Wenn wir die oben angegebene Zuordnung zwischen L_2 und l_2 und Satz 3 und 4 aus [56] beachten, können wir behaupten, daß der Grenzwert der Folge der $x^{(n)}$ in l_2 genau dann existiert, wenn die Folge in sich konvergiert. Dieser Grenzwert ist eindeutig bestimmt.

Wir betrachten die Menge K der Elemente von l_2 , die nur endlich viele von 0 verschiedene Koordinaten haben, wobei alle diese Koordinaten rationale komplexe Zahlen sind, d. h. Zahlen der Form $a + bi$, wobei a und b rationale reelle Zahlen sind. Beachten wir die Abzählbarkeit der Menge der rationalen Zahlen, so sehen wir, daß *die Menge K abzählbar* ist. Wir zeigen, daß *sie in l_2 dicht* ist. Es sei $x(x_1, x_2, x_3, \dots)$ ein Element aus l_2 und $z(c_1, c_2, \dots, c_n, 0, 0, \dots)$ ein Element aus der oben angegebenen Menge K . Dann gilt

$$\|x - z\|^2 = \sum_{k=1}^n |x_k - c_k|^2 + \sum_{k=n+1}^{\infty} |x_k|^2. \quad (110)$$

Es sei ε eine gegebene positive Zahl. Da die Reihe der $|x_k|^2$ konvergiert, können wir einen solchen Wert $n = n_0$ angeben, daß $\sum_{k=n_0+1}^{\infty} |x_k|^2 \leq \frac{1}{2} \varepsilon^2$ ist. In der endlichen Summe $\sum_{k=1}^{n_0} |x_k - c_k|^2$ können wir die rationalen Zahlen c_k so nahe bei den x_k wählen, daß diese Summe ebenfalls höchstens gleich $\varepsilon^2/2$ wird. Dadurch haben wir nach (110) $\|x - z\|^2 \leq \varepsilon^2$, d. h. $\|x - z\| \leq \varepsilon$. Damit ist gezeigt, daß die oben erwähnte abzählbare Menge K von Elementen aus l_2 in l_2 dicht ist. *Somit ist der Raum l_2 separabel*. Die Elemente

$$e_1(1, 0, 0, \dots), \quad e_2(0, 1, 0, 0, \dots), \quad e_3(0, 0, 1, 0, 0, \dots), \dots$$

dieses Raumes bilden ein vollständiges orthonormiertes System.

60. Lineare Mannigfaltigkeiten in L_2 . Wir führen jetzt einige neue Begriffe ein, die mit L_2 zusammenhängen.

Definition. Eine Menge U von Elementen aus L_2 (Funktionen aus L_2) heißt *lineare Mannigfaltigkeit*, wenn folgendes gilt: Gehören $\varphi(P)$ und $\psi(P)$ zu U , so gehören $c\varphi(P)$ bei beliebig gewähltem c und $\varphi(P) + \psi(P)$ ebenfalls zu U .

Aus dieser Definition folgt unmittelbar $c_1 f_1(P) + c_2 f_2(P) + \dots + c_m f_m(P) \in U$ für beliebige Zahlen c_k , wenn $f_k(P) \in U$ ($k = 1, 2, \dots, m$) ist.

Wir geben einige Beispiele von linearen Mannigfaltigkeiten an. Es sei M eine Menge auf $\mathfrak{E}$ beschränkter Funktionen, d. h. solcher Funktionen, für die für beliebiges $f(P) \in M$ eine Zahl d_f existiert derart, daß $|f(P)| \leq d_f$ auf $\mathfrak{E}$ ist. Die Menge M ist eine lineare Mannigfaltigkeit in L_2 .

¹⁾ Siehe die Fußnote auf S. 145 (Anm. d. Red.).

Man nennt eine Funktion $f(P)$ auf der Menge $\mathfrak{E}$ stetig, wenn folgendes gilt: Wenn die Punkte P und P_n ($n = 1, 2, \dots$) zu $\mathfrak{E}$ gehören und $P_n \rightarrow P$ gilt, strebt $f(P_n)$ gegen $f(P)$. Die Menge der Funktionen aus L_2 , die auf $\mathfrak{E}$ stetig sind, ist offenbar eine lineare Mannigfaltigkeit.

Satz 1. Die lineare Mannigfaltigkeit der auf einer beschränkten Menge $\mathfrak{E}$ stetigen Funktionen aus L_2 ist in L_2 überall dicht.

Beweis. Wir müssen zeigen, daß zu jedem Element $f(P)$ aus L_2 und jedem $\delta > 0$ eine auf $\mathfrak{E}$ stetige Funktion $\varphi(P)$ aus L_2 existiert derart, daß

$$\|f - \varphi\|^2 = \int_{\mathfrak{E}} [f(P) - \varphi(P)]^2 G(d\mathfrak{E}) \leq \delta \quad (111)$$

ist, wobei wir der Einfachheit halber, wie schon oben, den zweidimensionalen Fall betrachten. Wir haben $f(P) = f^+(P) - f^-(P)$, wobei $f^+(P)$ und $f^-(P)$ der positive bzw. der negative Teil von $f(P)$ sind. Man kann annehmen, daß diese Funktionen, die zu L_2 gehören und auf $\mathfrak{E}$ summierbar sind, nur endliche Werte annehmen und Grenzfunktionen einer wachsenden Folge stückweise konstanter Funktionen $\omega_n^+(P)$ bzw. $\omega_n^-(P)$ mit endlich vielen Werten sind, wobei $\omega_n^+(P) \leq f^+(P)$ und $\omega_n^-(P) \leq f^-(P)$ ist [46]. Es gilt (vgl. [54])

$$\lim_{n \rightarrow \infty} \int_{\mathfrak{E}} [f^+(P) - \omega_n^+(P)]^2 G(d\mathfrak{E}) = 0, \quad \lim_{n \rightarrow \infty} \int_{\mathfrak{E}} [f^-(P) - \omega_n^-(P)]^2 G(d\mathfrak{E}) = 0. \quad (112)$$

Weiter folgt aus $(x_1 + x_2)^2 \leq 2(x_1^2 + x_2^2)$ die Ungleichung

$$[f - (\omega_n^+ - \omega_n^-)]^2 \leq 2(f^+ - \omega_n^+)^2 + 2(f^- - \omega_n^-)^2,$$

und wir können, wenn wir die stückweise konstante Funktion $\omega_n(P) = \omega_n^+(P) - \omega_n^-(P)$ mit endlich vielen Werten einführen, auf Grund von (112) ein solches n wählen, daß $\|f - \omega_n\| \leq \varepsilon_0$ ist, wobei ε_0 eine beliebig gegebene positive Zahl ist. Wenn wir beachten, daß $\|f - \varphi\| \leq \|f - \omega_n\| + \|\omega_n - \varphi\|$ ist, genügt es zu zeigen, daß eine stetige Funktion $\varphi(P)$ derart existiert, daß $\|\omega_n - \varphi\| \leq \varepsilon$ ist, wobei ε eine beliebig gegebene positive Zahl und $\omega_n(P)$ eine gegebene Funktion mit endlich vielen Werten ist. Eine solche Funktion können wir in der Form

$$\omega(P) = \sum_{k=1}^m c_k \omega_{\mathfrak{E}_k}(P)$$

darstellen, wobei $\omega_{\mathfrak{E}_k}$ die charakteristische Funktion einer festen Menge $\mathfrak{E}_k$ ist, die zu $\mathfrak{E}$ gehört. Wenn $\varphi_k(P)$ ($k = 1, 2, \dots, m$) gewisse, auf $\mathfrak{E}$ stetige Funktionen sind und $\varphi(P) = c_1 \varphi_1(P) + c_2 \varphi_2(P) + \dots + c_m \varphi_m(P)$ ist, so gilt

$$\|\omega - \varphi\| \leq \sum_{k=1}^m |c_k| \|\omega_{\mathfrak{E}_k} - \varphi_k\|,$$

und der Beweis des Satzes reduziert sich auf den Beweis folgender Behauptung: Zu der charakteristischen Funktion $\omega_{\mathfrak{E}_0}(P)$ einer beliebigen meßbaren Menge $\mathfrak{E}_0$, die in $\mathfrak{E}$ liegt, und zu jedem $\varepsilon > 0$ existiert eine auf $\mathfrak{E}$ stetige Funktion $\varphi(P)$, so daß $\|\omega - \varphi\| \leq \varepsilon$ gilt. Bekanntlich existiert zu beliebig gegebenem $\varepsilon_0 > 0$ eine abgeschlossene Menge F , die in $\mathfrak{E}_0$ liegt, derart, daß $G(\mathfrak{E}_0 - F) \leq \varepsilon_0^2$ ist [35]. Dabei ist

$$\|\omega_{\mathfrak{E}_0} - \omega_F\|^2 = \int_{\mathfrak{E}} [\omega_{\mathfrak{E}_0}(P) - \omega_F(P)]^2 G(d\mathfrak{E}) = \int_{\mathfrak{E}_0 - F} G(d\mathfrak{E}) = G(\mathfrak{E}_0 - F) \leq \varepsilon_0^2,$$

und auf Grund der Ungleichung $\|\omega_{\mathfrak{E}_0} - \varphi\| \leq \|\omega_{\mathfrak{E}_0} - \omega_F\| + \|\omega_F - \varphi\|$ genügt es, die oben angegebene Behauptung für die charakteristische Funktion einer beschränkten abgeschlossenen Menge F , die zu $\mathfrak{E}$ gehört, zu zeigen. Wir bezeichnen

mit $r(P)$ den Abstand des Punktes P von der Menge F . Es gilt $r(Q_1) \leq r(Q) + |\overline{QQ_1}|$ und $r(Q) \leq r(Q_1) + |\overline{QQ_1}|$, wenn $|\overline{QQ_1}|$ der Abstand zwischen Q und Q_1 ist. Hieraus folgt die Stetigkeit der Funktion $r(P)$. Weiter ist $r(Q)$ dann und nur dann gleich 0, wenn $Q \in F$ ist [II, 92]. Man sieht leicht, daß $\omega_F(P)$ Grenzwert einer nicht-wachsenden Folge von auf $\mathcal{E}$ stetigen Funktionen

$$\varphi_n(P) = \frac{1}{1 + nr(P)} \quad (113)$$

ist, so daß $\|\omega_F - \varphi_n\| \rightarrow 0$ [54] für $n \rightarrow \infty$ ist; folglich existiert für ein beliebig gegebenes $\varepsilon > 0$ ein n , für das $\|\omega_F - \varphi_n\| \leq \varepsilon$ gilt, wobei $\varphi_n(P)$ auf $\mathcal{E}$ stetig ist. Damit ist der Satz bewiesen.

Folgerung 1. Wir beschränken uns der Einfachheit halber auf den Fall der Ebene und nehmen an, daß $\mathcal{E}$ ein abgeschlossenes Intervall Δ ($a_1 \leq x \leq b_1$, $a_2 \leq y \leq b_2$) ist. Zu jeder auf diesem Intervall stetigen Funktion $\varphi(x, y)$ kann man ein Polynom $p(x, y)$ so konstruieren, daß $|\varphi(x, y) - p(x, y)| \leq \varepsilon_0$ auf Δ ist, wobei ε_0 eine beliebig gegebene positive Zahl ist. Dabei ist

$$\|\varphi - p\|^2 = \int_{\Delta} [\varphi(x, y) - p(x, y)]^2 G(d\Delta) \leq \varepsilon_0^2 G(\Delta).$$

Wenn wir noch beachten, daß für $f(x, y) \in L_2(\Delta)$ die Ungleichung $\|f - p\| \leq \|f - \varphi\| + \|\varphi - p\|$ und daß der eben bewiesene Satz 1 gilt, können wir behaupten, daß die *lineare Mannigfaltigkeit der Polynome in $L_2(\Delta)$ überall dicht ist*.

An Stelle des Intervalls Δ können wir auch eine beliebige beschränkte abgeschlossene Menge F nehmen, da man eine auf F stetige Funktion unter Beibehaltung ihrer Stetigkeit auf ein abgeschlossenes Intervall Δ fortsetzen kann, das F enthält. Dabei muß man

$$\|\varphi - p\|_F^2 = \int_F [\varphi(x, y) - p(x, y)]^2 G(dF) \leq \int_{\Delta} [\varphi(x, y) - p(x, y)]^2 G(d\Delta)$$

berücksichtigen.

Folgerung 2. Wenn $p(x, y)$ ein beliebiges Polynom ist, existiert, weil die rationalen Zahlen auf der reellen Achse überall dicht liegen, ein solches Polynom $q(x, y)$ mit rationalen Koeffizienten, daß für beliebig gegebenes $\varepsilon_0 > 0$ auf Δ oder F die Ungleichung $|p(x, y) - q(x, y)| \leq \varepsilon_0$ gilt.

Hieraus folgt unmittelbar, daß die *Polynome mit rationalen Koeffizienten eine überall in $L_2(\Delta)$ oder $L_2(F)$ dichte Menge bilden*.

Wir zeigen, daß die Menge dieser Polynome abzählbar ist. Jedem Polynom $q(x, y)$ entspreche die positive Zahl $\sigma = n + r + s$, wo n der Grad von $q(x, y)$, r der Hauptnenner ihrer Koeffizienten (r wird immer positiv genommen) und s die Summe der absoluten Beträge der Zähler in ihren Koeffizienten ist, die zu dem Nenner r gehören (ausgenommen, wenn $q(x, y) \equiv 0$ ist; dann ordnen wir dem Polynom die Zahl $\sigma = 0$ zu). Wie man leicht sieht, ist die Anzahl der Polynome, denen die feste Zahl σ entspricht, endlich. Wir können alle Polynome mit rationalen Koeffizienten nach wachsenden Zahlen σ numerieren, wobei die Anordnung der Polynome mit gleichem σ gleichgültig ist. Also existiert eine abzählbare Menge von Elementen aus L_2 , die überall in L_2 dicht liegt, d. h., der Raum $L_2(\Delta)$ bzw. $L_2(F)$ ist separabel.

Wir nehmen jetzt an, daß $\mathcal{E}$ eine beliebige beschränkte meßbare Menge ist. Ihre abgeschlossene Hülle sei $\bar{\mathcal{E}}$. Das ist eine beschränkte abgeschlossene Menge, und wie wir bewiesen haben, existiert in $L_2(\bar{\mathcal{E}})$ eine überall dichte abzählbare Menge von

Funktionen $\varphi_k(x, y)$ ($k=1, 2, \dots$). Das sind Funktionen aus $L_2(\mathfrak{E})$ und dadurch auch aus $L_2(\mathfrak{G})$. Wir zeigen, daß $\varphi_k(x, y)$ auch auf $\mathfrak{E}$ überall dicht ist. Wir nehmen eine Funktion $f(x, y)$ aus $L_2(\mathfrak{G})$ und setzen sie durch 0 auf $\mathfrak{E}$ fort. Sie ist auch aus $L_2(\mathfrak{E})$, und daher existiert für jedes $\varepsilon > 0$ eine solche Funktion $\varphi_k(x, y)$ aus der obengenannten abzählbaren Menge, daß

$$\begin{aligned} \int_{\mathfrak{E}} [f(x, y) - \varphi_k(x, y)]^2 G(d\mathfrak{E}) \\ = \int_{\mathfrak{E}} [f(x, y) - \varphi_k(x, y)]^2 G(d\mathfrak{E}) + \int_{\mathfrak{E}-\mathfrak{E}} [f(x, y) - \varphi_k(x, y)]^2 G(d\mathfrak{E}) \leq \varepsilon \end{aligned}$$

und erst recht

$$\int_{\mathfrak{E}} [f(x, y) - \varphi_k(x, y)]^2 G(d\mathfrak{E}) \leq \varepsilon$$

ist, usw. Die Separabilität von L_2 auf einer beliebigen meßbaren Menge wird später bewiesen.

Wir beweisen noch einen Satz, den wir im folgenden benötigen.

Satz 2. *Wenn die Vollständigkeitsrelation für alle Funktionen gilt, die zu einer in L_2 dichten Menge K gehören, gilt sie auch für beliebige Funktionen von L_2 .*

Beweis. Es sei $f(P)$ ein Element aus L_2 und ε eine gegebene positive Zahl. Weil K in L_2 dicht ist, existiert in K ein solches Element $\varphi(P)$, daß $\|f - \varphi\| \leq \varepsilon/3$ ist. Nach Voraussetzung gilt die Vollständigkeitsrelation für $\varphi(P)$, und folglich kann man eine Partialsumme $s_n(\varphi)$ der Fourierreihe der Funktion $\varphi(P)$ bezüglich des orthonormierten Funktionensystems $\varphi_k(P)$ nehmen, so daß $\|\varphi - s_n(\varphi)\| \leq \varepsilon/3$ ist. Wenn wir die Beziehung

$$f - s_n(f) = (f - \varphi) + (\varphi - s_n(\varphi)) + (s_n(\varphi) - s_n(f))$$

und die Dreiecksungleichung benutzen, erhalten wir

$$\|f - s_n(f)\| \leq \|f - \varphi\| + \|\varphi - s_n(\varphi)\| + \|s_n(\varphi) - s_n(f)\|,$$

und hieraus folgt

$$\|f - s_n(f)\| \leq \frac{2}{3} \varepsilon + \|s_n(\varphi) - s_n(f)\|.$$

Die Differenz $s_n(\varphi) - s_n(f)$ der Partialsummen der Fourierreihe für φ und f ist gleich dem Abschnitt der Fourierreihe für die Differenz $\varphi - f$, d. h. $s_n(\varphi) - s_n(f) = s_n(\varphi - f)$, und nach der Besselschen Ungleichung ist $\|s_n(\varphi) - s_n(f)\| \leq \|\varphi - f\| \leq \frac{\varepsilon}{3}$. Also ist $\|f - s_n(f)\| \leq \frac{2}{3} \varepsilon + \frac{\varepsilon}{3} = \varepsilon$; hieraus folgt, da ε beliebig war, daß die Vollständigkeitsrelation für die Funktion $f(P)$ gilt. Damit ist der Satz bewiesen.

Alles oben Gesagte gilt auch unmittelbar für den Fall komplexer Funktionen aus L_2 [58].

61. Beispiele für vollständige Systeme. Wir führen einige einfache Beispiele für orthonormierte Systeme an, die auf dem endlichen Intervall $[a, b]$ vollständig sind. Wenn wir den Orthogonalisierungsprozeß auf die ganzen nichtnegativen Potenzen von x anwenden, 1, x , x^2 , ... [IV, 38], erhalten wir ein System orthogonaler Polynome $p_k(x)$ ($k=0, 1, 2, \dots$) im Intervall $[a, b]$, wobei $p_k(x)$ den Grad k hat. Jedes

Polynom $p(x)$ n -ten Grades kann dann als Linearkombination

$$p(x) = \sum_{k=0}^n c_k p_k(x) \quad (114)$$

dargestellt werden.

Um sich davon zu überzeugen, genügt es, die c_n so zu definieren, daß auf der rechten Seite der Koeffizient von x^n derselbe ist wie auch bei $p(x)$. Dann müssen wir c_{n-1} so definieren, daß der Koeffizient von x^{n-1} in dem Glied $c_{n-1}p_{n-1}(x)$ derselbe ist wie auch in $p(x) - c_n p_n(x)$ usw. Die Koeffizienten c_k in (114) sind offenbar gleich den Fourierkoeffizienten von $p(x)$ bezüglich $p_k(x)$. Aus (114) folgt, daß im Fall eines orthogonalen Systems $p_k(x)$ die Vollständigkeitsrelation für jedes Polynom $p(x)$ gilt, und hieraus folgt nach Satz 2 aus [60], daß *das System der orthogonalen Polynome vollständig ist*. Wir wissen schon, daß im Intervall $[-l, l]$ für das orthogonale System

$$\sin \frac{n\pi x}{l}, \quad \cos \frac{n\pi x}{l} \quad (n=0, 1, 2, \dots) \quad (115)$$

die Vollständigkeitsrelation für jede stetige Funktion erfüllt ist [II, 160]; hieraus folgt, daß das System (115) in L_2 vollständig ist. Genauso sind im Intervall $[0, l]$ die orthogonalen Funktionensysteme

$$\sin \frac{n\pi x}{l} \quad (n=1, 2, \dots) \quad \text{und} \quad \cos \frac{n\pi x}{l} \quad (n=0, 1, 2, \dots)$$

vollständig.

Wir wissen aus [IV/2, 77], daß wir im Fall der Eigenfunktionen $\varphi_k(x)$ ($k=1, 2, \dots$) eines linearen Randwertproblems jede Funktion, die stetige Ableitungen bis zur zweiten Ordnung hat und die Randbedingungen erfüllt, in eine gleichmäßig konvergente Fourierreihe nach den Funktionen $\varphi_k(x)$ entwickeln können. Für solche Funktionen gilt dann um so mehr die Vollständigkeitsrelation. Wenn wir die Werte der Funktion in der Nähe der Endpunkte des Intervalls ändern, können wir leicht feststellen, daß die Vollständigkeitsrelationen für alle Funktionen mit stetigen zweiten Ableitungen gelten, ohne zu fordern, daß die Randbedingungen in den Endpunkten gelten. Die Vollständigkeitsrelation gilt erst recht für alle Polynome; daher ist das System der Eigenfunktionen $\varphi_k(x)$ vollständig.

62. Die Höldersche und die Minkowskische Ungleichung. Neben der Klasse L_2 betrachtet man noch oft die Klasse L_p derjenigen meßbaren Funktionen $f(P)$, bei denen die p -te Potenz des absoluten Betrages, d. h. $|f(P)|^p$, auf $\mathcal{S}$ summierbar ist (vgl. [55]). Zunächst führen wir für die Summen und Integrale zu (67) und (69) analoge Ungleichungen für einen beliebigen Exponenten $p > 1$ ein.

Es sei α eine positive Zahl. Auf der x, y -Ebene betrachten wir die Kurve $y = x^\alpha$ und die achsenparallelen Geraden $x = a$ und $y = b$ (Abb. 3). Diese Geraden, die Koordinatenachsen und die erwähnte Kurve begrenzen zwei ebene Gebiete mit den Flächeninhalten

$$S_1 = \int_0^a x^\alpha dx = \frac{a^{1+\alpha}}{1+\alpha}, \quad S_2 = \int_0^b y^{\frac{1}{\alpha}} dy = \frac{b^{1+\frac{1}{\alpha}}}{1+\frac{1}{\alpha}}.$$

Die Summe dieser Flächeninhalte ist, wie unmittelbar in Abb. 3 zu sehen ist, nicht kleiner als der Flächeninhalt ab des Rechtecks mit den Seiten a und b , d. h.

$$ab \leq \frac{a^{1+\alpha}}{1+\alpha} + \frac{b^{\frac{1+\alpha}{\alpha}}}{\frac{1+\alpha}{\alpha}}.$$

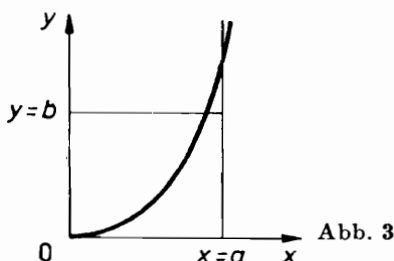


Abb. 3

Wenn wir $1+\alpha=p$ und $1+\frac{1}{\alpha}=p'$ setzen, können wir dafür

$$ab \leq \frac{a^p}{p} + \frac{b^{p'}}{p'} \quad (116)$$

schreiben, wobei die Zahlen p und p' durch die Beziehung

$$\frac{1}{p} + \frac{1}{p'} = 1 \quad (117)$$

verknüpft sind. Da die positive Zahl α beliebig war, gilt die Ungleichung (116) für beliebige positive p und p' , für die (117) erfüllt ist. Diese beiden Zahlen müssen größer als 1 sein. Wenn $p=2$ ist, ist $p'=2$, und (116) führt zu der bekannten Ungleichung $2ab \leq a^2 + b^2$. Aus Abb. 3 folgt, daß das Gleichheitszeichen genau dann gilt, wenn der Schnittpunkt der Geraden $x=a$ und $y=b$ auf der Kurve $y=x^\alpha$ liegt, d. h. für $b=a^\alpha$. Weiter mögen die positiven Zahlen a'_k und b'_k ($k=1, 2, \dots, n$) den Beziehungen

$$\sum_{k=1}^n a'_k{}^p = 1 \quad \text{und} \quad \sum_{k=1}^n b'_k{}^{p'} = 1 \quad (118)$$

genügen. Wir setzen in (116) $a=a'_k$ und $b=b'_k$. Wenn wir über k summieren und (117) und (118) beachten, erhalten wir

$$\sum_{k=1}^n a'_k b'_k \leq 1. \quad (119)$$

Wir betrachten jetzt beliebige positive Zahlen a_k und b_k und setzen

$$A = \left(\sum_{k=1}^n a_k^p \right)^{1/p}, \quad B = \left(\sum_{k=1}^n b_k^{p'} \right)^{1/p'}. \quad (120)$$

Die Zahlen $a'_k = \frac{a_k}{A}$ und $b'_k = \frac{b_k}{B}$ genügen den Beziehungen (118); folglich erhalten wir für sie die Ungleichung (119), die in diesem Fall in der Gestalt $\sum_{k=1}^n a_k b_k \leq AB$ ge-

geschrieben werden kann, d. h.

$$\sum_{k=1}^n a_k b_k \leq \left(\sum_{k=1}^n a_k^p \right)^{1/p} \left(\sum_{k=1}^n b_k^{p'} \right)^{1/p'}. \quad (121)$$

Für $n \rightarrow \infty$ erhalten wir für die unendlichen Summen die analoge Ungleichung

$$\sum_{k=1}^{\infty} a_k b_k \leq \left(\sum_{k=1}^{\infty} a_k^p \right)^{1/p} \left(\sum_{k=1}^{\infty} b_k^{p'} \right)^{1/p'}, \quad (122)$$

wenn die rechts stehenden Reihen konvergieren. Dabei ist dann auch nach obiger Ungleichung die links stehende Reihe konvergent. Einige der Zahlen a_k und b_k können auch gleich 0 sein. Für komplexe Zahlen können wir, wenn wir die Ungleichung

$$\left| \sum_k a_k b_k \right| \leq \sum_k |a_k| |b_k|$$

benutzen, die obigen Ungleichungen in der Form

$$\left| \sum_k a_k b_k \right| \leq \left(\sum_k |a_k|^p \right)^{1/p} \left(\sum_k |b_k|^{p'} \right)^{1/p'} \quad (123)$$

schreiben.

Diese Ungleichungen heißen *Höldersche Ungleichungen für Summen*. Für $p = p' = 2$ gehen sie in die gewöhnliche Ungleichung (106) aus [59] über. Analoge Ungleichungen gelten auch für Integrale. Wir nehmen nun $f(P) \in L_p$ und $g(P) \in L_{p'}$ an. Nach (116) gilt dann

$$|f(P) g(P)| \leq \frac{|f(P)|^p}{p} + \frac{|g(P)|^{p'}}{p'}.$$

Die rechte Seite ist nach Voraussetzung summierbar; also ist auch das Produkt $f(P) g(P)$ summierbar, d. h., für $f(P) \in L_p$ und $g(P) \in L_{p'}$ ist das Produkt $f(P) g(P)$ eine summierbare Funktion (siehe Satz 1 aus [55]). Für das Integral dieses Produktes gilt (analog zu (67) aus [55]) die Höldersche Ungleichung

$$\left| \int_{\mathfrak{E}} f g dx dy \right| \leq \left(\int_{\mathfrak{E}} |f|^p dx dy \right)^{1/p} \left(\int_{\mathfrak{E}} |g|^{p'} dx dy \right)^{1/p'}, \quad (124)$$

die wir nur für das Lebesguesche Integral aufschreiben (sie gilt auch für Lebesgue-Stieltjesche Integrale). Sie ergibt sich aus der Ungleichung (122) durch Grenzübergang. Es seien δ'_n und δ''_n zwei unbeschränkt abnehmende Folgen von Lebesgueschen Zerlegungen für $|f|$ und $|g|$, und $\delta_n = \delta'_n \delta''_n$ sei das Produkt dieser Zerlegungen. Ferner seien $\mathfrak{E}_k^{(n)}$ die Komponenten der Menge $\mathfrak{E}$ in der Zerlegung δ_n und $m'_{k,n}$ bzw. $m''_{k,n}$ die obere Grenze der Werte von $|f|$ bzw. $|g|$ auf $\mathfrak{E}_k^{(n)}$. Wenn wir (117) beachten, können wir

$$\sum_k m'_{k,n} m''_{k,n} m(\mathfrak{E}_k^{(n)}) = \sum_k m'_{k,n} m^{1/p}(\mathfrak{E}_k^{(n)}) m''_{k,n} m^{1/p'}(\mathfrak{E}_k^{(n)})$$

schreiben. Wir wenden jetzt auf die Zahlen

$$a_k = m'_{k,n} m^{1/p}(\mathfrak{E}_k^{(n)}) \quad \text{und} \quad b_k = m''_{k,n} m^{1/p'}(\mathfrak{E}_k^{(n)})$$

die Höldersche Ungleichung an:

$$\sum_k m'_{k,n} m''_{k,n} m(\mathfrak{E}_k^{(n)}) \leq \left(\sum_k m_{k,n}^p m(\mathfrak{E}_k^{(n)}) \right)^{1/p} \left(\sum_k m_{k,n}^{p'} m(\mathfrak{E}_k^{(n)}) \right)^{1/p'}. \quad (125)$$

Mit $m_{k,n}$ bezeichnen wir die obere Grenze der Werte von $|f| \cdot |g|$ auf $\mathfrak{E}_k^{(n)}$. Offenbar gilt $m_{k,n} \leq m'_{k,n} m''_{k,n}$, und aus (125) folgt

$$\sum_k m_{k,n} m(\mathfrak{E}_k^{(n)}) \leq \left(\sum_k m_{k,n}^p m(\mathfrak{E}_k^{(n)}) \right)^{1/p} \left(\sum_k m_{k,n}^{p'} m(\mathfrak{E}_k^{(n)}) \right)^{1/p'}.$$

Wenn wir für die Folge von Lebesgueschen Zerlegungen zur Grenze übergehen, erhalten wir

$$\int_{\mathfrak{E}} |f| |g| \, dxdy \leq \left(\int_{\mathfrak{E}} |f|^p \, dxdy \right)^{1/p} \left(\int_{\mathfrak{E}} |g|^{p'} \, dxdy \right)^{1/p'}. \quad (126)$$

Hieraus ergibt sich (124).

Wir beweisen noch eine Ungleichung, die analog der Ungleichung (69) aus [55] ist. Zuerst betrachten wir den Fall einer Summe. Es seien wieder a_k und b_k zwei Folgen positiver Zahlen. Indem wir die Gleichungen

$$(a_k + b_k)^p = (a_k + b_k)^{p-1} a_k + (a_k + b_k)^{p-1} b_k$$

summieren, erhalten wir

$$\sum_k (a_k + b_k)^p = \sum_k (a_k + b_k)^{p-1} a_k + \sum_k (a_k + b_k)^{p-1} b_k.$$

Wenn wir auf die rechts stehenden Summen die Höldersche Ungleichung anwenden, kommen wir zu der Ungleichung

$$\sum_k (a_k + b_k)^p \leq \left(\sum_k a_k^p \right)^{1/p} \left(\sum_k (a_k + b_k)^{(p-1)p'} \right)^{1/p'} + \left(\sum_k b_k^p \right)^{1/p} \left(\sum_k (a_k + b_k)^{(p-1)p'} \right)^{1/p'}.$$

Nach (117) ist aber $p' = \frac{p}{p-1}$, und damit können wir die letzte Ungleichung in der Gestalt

$$\sum_k (a_k + b_k)^p \leq \left(\sum_k (a_k + b_k)^p \right)^{1-1/p} \left[\left(\sum_k a_k^p \right)^{1/p} + \left(\sum_k b_k^p \right)^{1/p} \right]$$

schreiben.

Wenn wir beide Seiten durch den Faktor, der vor der eckigen Klammer steht, dividieren, erhalten wir die *Minkowskische Ungleichung für Summen*:

$$\left(\sum_k (a_k + b_k)^p \right)^{1/p} \leq \left(\sum_k a_k^p \right)^{1/p} + \left(\sum_k b_k^p \right)^{1/p}. \quad (127)$$

Aus dieser Ungleichung kann man genau so wie oben die Minkowskische Ungleichung für Integrale herleiten, wenn die Funktionen $f(P)$ und $g(P)$ zu L_p gehören:

$$\left(\int_{\mathfrak{E}} |f+g|^p \, dxdy \right)^{1/p} \leq \left(\int_{\mathfrak{E}} |f|^p \, dxdy \right)^{1/p} + \left(\int_{\mathfrak{E}} |g|^p \, dxdy \right)^{1/p}; \quad (128)$$

dabei war zu beachten, daß $|f+g| \leq |f| + |g|$ ist. Die Ungleichungen (127) und (128) wurden unter der Voraussetzung $p > 1$ hergeleitet. Sie gelten offenbar auch für $p=1$, aber für $p < 1$ sind sie im allgemeinen nicht erfüllt.

Wenn wir diese Ungleichungen benutzen, können wir für die Funktionenklasse L_p ($p > 1$) leicht Eigenschaften beweisen, die wir früher schon für L_2 bewiesen haben; dabei betrachten wir komplexe Funktionen. Wir folgen [55] und zählen diese Eigenschaften auf. Wenn $f(P) \in L_p$ und $g(P) \in L_{p'}$ ($p > 1$) ist, sind die Funktion $f(P)$ und das Produkt $f(P)g(P)$ auf $\mathfrak{E}$ summierbar. Das folgt aus (124). Gehören $f(P)$ und $g(P)$ zu $L_{p'}$, so ist für konstantes c auch $cf(P)$ und $f(P) + g(P)$ in L_p ($p \geq 1$) enthalten. Das folgt aus (128). Man sagt, daß die Folge der Funktionen $f_n(P)$ aus L_p in L_p ($p \geq 1$) im Mittel konvergiert oder mit dem Exponent p im Mittel gegen die Funktion $f(P)$ aus L_p konvergiert, wenn

$$\lim_{n \rightarrow \infty} \int_{\mathfrak{E}} |f(P) - f_n(P)|^p \, G(d\mathfrak{E}) = 0$$

ist. Der Grenzwert im Mittel in L_p ist bis auf äquivalente Funktionen eindeutig bestimmt. Wenn $f_n(P)$ im Mittel gegen $f(P)$ strebt, kann man aus der Folge der $f_n(P)$ eine Teilfolge $f_{n_k}(P)$ auswählen, die fast überall auf $\mathfrak{E}$ gegen $f(P)$ strebt. Die Kon-

vergenz in sich wird analog (72) definiert:

$$\int_{\mathfrak{E}} |f_n(P) - f_m(P)|^p G(d\mathfrak{E}) \leq \varepsilon$$

für $n > N$, $m \geq N$. Die Konvergenz der Folge der $f_n(P)$ in sich in L_p ist notwendig und hinreichend dafür, daß diese Folge im Mittel gegen eine Funktion aus L_p ($p \geq 1$) konvergiert.

Wenn $f_n(P) \rightarrow f(P)$ in L_p und $g_n(P) \rightarrow g(P)$ in $L_{p'}$ ($p > 1$) gilt, ist

$$\lim_{n \rightarrow \infty} \int_{\mathfrak{E}} f_n(P) g_n(P) G(d\mathfrak{E}) = \int_{\mathfrak{E}} f(P) g(P) G(d\mathfrak{E}).$$

Weiter kann man in L_p für $p \geq 1$ eine Norm

$$\|f\| = \left(\int_{\mathfrak{E}} |f(P)|^p G(d\mathfrak{E}) \right)^{1/p}$$

und den Abstand δ zwischen zwei Elementen, $\delta(f, g) = \|f - g\|$, einführen, wobei die Gleichung $\|cf(P)\| = |c| \cdot \|f(P)\|$ und die Dreiecksungleichung gelten.

Wir beweisen jetzt, daß $f(P)$ zu L_p gehört, wenn $q > p$ und $f(P) \in L_q$ ist.

Nach Voraussetzung gilt

$$\int_{\mathfrak{E}} |f(P)|^q G(d\mathfrak{E}) = A < \infty.$$

Wir betrachten das Integral

$$\begin{aligned} \int_{\mathfrak{E}} |f(P)|^p G(d\mathfrak{E}) &= \int_{\mathfrak{E}(|f(P)| \leq 1)} |f(P)|^p G(d\mathfrak{E}) + \int_{\mathfrak{E}(|f(P)| > 1)} |f(P)|^p G(d\mathfrak{E}) \\ &\leq \int_{\mathfrak{E}} G(d\mathfrak{E}) + \int_{\mathfrak{E}} |f(P)|^q G(d\mathfrak{E}) = G(\mathfrak{E}) + A; \end{aligned}$$

hieraus folgt $f(P) \in L_p$. Bei dem Beweis benutzten wir, daß das Maß $G(\mathfrak{E})$ der Menge $\mathfrak{E}$ endlich ist.

In L_p existiert für $p \neq 2$ kein skalares Produkt wie in L_2 .

Analog zu l_2 können wir den Raum l_p einführen. Die Elemente dieses Raumes sind unendliche Folgen komplexer Zahlen $(x_1, x_2, \dots)$, so daß die Reihe der Zahlen $|x_k|^p$ konvergiert. Für $p \geq 1$ besitzt l_p Eigenschaften, die den Eigenschaften von l_2 analog sind, genauso, wie oben die Eigenschaften von L_p und L_2 . In l_p ($p \neq 2$) existiert kein skalares Produkt, und es besteht kein Zusammenhang mit L_p wie zwischen l_2 und L_2 . Die Ungleichungen (106) und (107) sind durch die Ungleichungen (122) und (127) zu ersetzen, in denen $a_k = |x_k|$ und $b_k = |y_k|$ ist.

63. Das Integral über eine Menge unendlichen Maßes. Bisher betrachteten wir das Integral auf einer meßbaren Menge $\mathfrak{E}$ endlichen Maßes. Die Erweiterung auf den Fall von Mengen unendlichen Maßes geschieht im wesentlichen ebenso wie bei der Definition des Riemannschen Integrals in einem unendlichen Intervall. Es sei auf der meßbaren Menge $\mathfrak{E}$ unendlichen Maßes eine meßbare und nichtnegative Funktion $f(P)$ gegeben. Wir betrachten eine beliebige wachsende unendliche Folge von Mengen endlichen Maßes,

$$\mathfrak{E}_1 \subset \mathfrak{E}_2 \subset \mathfrak{E}_3 \subset \dots, \quad (129)$$

für die $\mathfrak{E}$ die Grenzmenge ist. Wir können die Mengen $\mathfrak{E}_n$ etwa als Produkt der Menge $\mathfrak{E}$ mit dem Intervall Δ_n ($-n \leq x \leq +n$, $-n \leq y \leq +n$) bilden. Für die beschränkten Mengen $\mathfrak{E}_k$ existieren die Integrale

$$\int_{\mathfrak{E}_k} f(P) G(d\mathfrak{E}), \quad (130)$$

die, da $f(P)$ nicht negativ ist, bei wachsendem n nicht abnehmen. Den Grenzwert der monotonen Folge (130) nennt man das *Integral von $f(P)$ über $\mathcal{E}$* :

$$\int_{\mathcal{E}} f(P) G(d\mathcal{E}) = \lim_{n \rightarrow \infty} \int_{\mathcal{E}_n} f(P) G(d\mathcal{E}). \quad (131)$$

Wir bemerken noch, daß die Integrale (130) auch gleich ∞ werden können. Dann wird offenbar auch das Integral von $f(P)$ über $\mathcal{E}$ gleich ∞ . Es kann aber auch vorkommen, daß alle Integrale (130) endlich sind, aber das Integral über $\mathcal{E}$ gleich ∞ ist. Um die oben gegebene Definition des Integrals zu rechtfertigen, müssen wir zeigen, daß der Grenzwert der Zahlenfolge (130) nicht von der Wahl der monoton wachsenden Mengenfolge $\mathcal{E}_n$ abhängt.

Satz. *Bei beliebiger Wahl der wachsenden Folge meßbarer Mengen $\mathcal{E}_n$ endlichen Maßes, die gegen $\mathcal{E}$ strebt, haben die Integrale (130) ein und denselben Grenzwert.*

Beweis. Wir beweisen diesen Satz indirekt. Es sei außer der Mengenfolge (129) noch eine andere wachsende Folge von Mengen endlichen Maßes $\mathcal{E}'_1 \subset \mathcal{E}'_2 \subset \mathcal{E}'_3 \subset \dots$ gegeben, die $\mathcal{E}$ als Grenzmenge hat, und zwar so, daß wir verschiedene Grenzwerte der Integralfolgen (130) für die Mengen $\mathcal{E}_n$ und $\mathcal{E}'_n$ haben:

$$\lim_{n \rightarrow \infty} \int_{\mathcal{E}_n} f(P) G(d\mathcal{E}) = a \quad \text{und} \quad \lim_{n \rightarrow \infty} \int_{\mathcal{E}'_n} f(P) G(d\mathcal{E}) = b > a. \quad (132)$$

Die Zahl a ist in jedem Fall endlich, und es gilt

$$\int_{\mathcal{E}_n} f(P) G(d\mathcal{E}) \leq a \quad (n = 1, 2, \dots). \quad (133)$$

Wir nehmen zuerst an, daß auch b endlich ist. Wenn wir eine positive Zahl $c < b - a$ wählen, können wir eine ganze positive Zahl m so wählen, daß

$$\int_{\mathcal{E}'_m} f(P) G(d\mathcal{E}) > a + c \quad (134)$$

ist. Da $f(P)$ nichtnegativ ist, gilt

$$\int_{\mathcal{E}'_m \mathcal{E}_n} f(P) G(d\mathcal{E}) \leq a. \quad (135)$$

Wir betrachten die Mengen $\mathcal{E}'_m \mathcal{E}_n$. Bei wachsendem n nehmen sie zu, und da $\mathcal{E}$ die Grenzmenge für $\mathcal{E}_n$ ist, ist $\mathcal{E}'_m$ die Grenzmenge für die Mengen $\mathcal{E}'_m \mathcal{E}_n$; hieraus folgt

$$\lim_{n \rightarrow \infty} G(\mathcal{E}'_m - \mathcal{E}'_m \mathcal{E}_n) = 0. \quad (136)$$

Da b endlich ist, ist $f(P)$ auf $\mathcal{E}'_m$ summierbar, und nach Formel (136) und der absoluten Stetigkeit des Integrals von $f(P)$ haben wir

$$\lim_{n \rightarrow \infty} \int_{\mathcal{E}'_m \mathcal{E}_n} f(P) G(d\mathcal{E}) = \int_{\mathcal{E}'_m} f(P) G(d\mathcal{E}),$$

im Widerspruch zu den Ungleichungen (134) und (135). Ist $b = \infty$, so nehmen wir $[f(P)]_N$ statt $f(P)$, wobei N und m so groß gewählt sind, daß die Ungleichung

$$\int_{\mathcal{E}'_m} [f(P)]_N G(d\mathcal{E}) > a + 1$$

gilt. Nach (133) ist auch

$$\int_{\mathcal{E}'_m \mathcal{E}_n} [f(P)]_N G(d\mathcal{E}) \leq a.$$

Dieselbe Überlegung wie vorhin führt zu einem Widerspruch. Damit ist der Satz bewiesen.

Wenn der Wert des Integrals einer nichtnegativen Funktion $f(P)$ über $\mathcal{E}$ endlich ist, sagt man, daß $f(P)$ auf $\mathcal{E}$ *summierbar* ist. Hieraus und aus der oben gegebenen Definition folgt unmittelbar: *Wenn $f(P)$ summierbar ist und die nichtnegative Funktion $\varphi(P)$ auf $\mathcal{E}$ der Ungleichung $\varphi(P) \leq f(P)$ genügt, ist auch $\varphi(P)$ summierbar.* Wir betrachten jetzt eine auf $\mathcal{E}$ meßbare Funktion, die das Vorzeichen wechseln kann, und zerlegen sie in den positiven und den negativen Teil:

$$f(P) = f^+(P) - f^-(P). \quad (137)$$

Die Funktion $f(P)$ heißt auf $\mathcal{E}$ *summierbar*, wenn $f^+(P)$ und $f^-(P)$ auf $\mathcal{E}$ summierbar sind. Dabei ist der Wert des Integrals definiert durch

$$\int_{\mathcal{E}} f(P) G(d\mathcal{E}) = \int_{\mathcal{E}} f^+(P) G(d\mathcal{E}) - \int_{\mathcal{E}} f^-(P) G(d\mathcal{E}). \quad (138)$$

Ist nur eine der Funktionen $f^+(P)$ und $f^-(P)$ summierbar, so ist, wie auch in [52], das Integral von $f(P)$ sinnvoll, aber sein Wert ist ∞ oder $-\infty$. Meist ist die Menge $\mathcal{E}$, auf der wir die Integration durchführen, die ganze Ebene oder die ganze Gerade oder allgemein der ganze n -dimensionale Raum.

Für die Integrale auf einer meßbaren Menge unendlichen Maßes gelten der Satz und die Eigenschaften 1 bis 7, 9 und 10 aus [52]. Wir beweisen nur die Volladditivität und die absolute Stetigkeit. Der Beweis des Satzes und der übrigen Eigenschaften ist äußerst einfach. Zuerst beweisen wir das folgende einfache

Lemma. *Wenn die nichtnegativen Zahlen $a_k^{(s)}$ für wachsendes s nicht fallen und $\lim_{s \rightarrow \infty} a_k^{(s)} = a_k$ ist, erhalten wir, wenn wir $a^{(s)} = \sum_{k=1}^{\infty} a_k^{(s)}$ setzen,*

$$\lim_{s \rightarrow \infty} a^{(s)} = \sum_{k=1}^{\infty} a_k. \quad (139)$$

Beweis. Wir führen den Beweis indirekt. Wir weisen darauf hin, daß die Summe (139) auch ∞ werden kann. Mit a bezeichnen wir den Grenzwert von $a^{(s)}$ und nehmen zuerst $a > \sum_{k=1}^{\infty} a_k$ an. Für hinreichend große Werte s gilt $a^{(s)} > c$, wobei c die Summe der Reihe (139) ist. Wenn wir s festhalten, können wir ein so großes m finden, daß $\sum_{k=1}^m a_k^{(s)} > \sum_{k=1}^{\infty} a_k$ ist. Daher gilt erst recht $\sum_{k=1}^m a_k > \sum_{k=1}^m a_k$, was $a_k \geq 0$ widerspricht. Wir nehmen jetzt $a < \sum_{k=1}^{\infty} a_k$ an. Für ein festes m ist also $\sum_{k=1}^m a_k > a$. Wir können jetzt ein so großes s auswählen, für welches $\sum_{k=1}^m a_k^{(s)} > a$ gilt. Diese endliche Summe ist offenbar höchstens gleich $a^{(s)}$; also gilt $a^{(s)} > a$. Das kann aber nicht sein, da die Folge der $a^{(s)}$ zwar gegen a strebt, aber nichtfallend ist. Damit ist das Lemma bewiesen.

Wir beweisen jetzt die Volladditivität des Integrals. Es sei $f(P)$ auf $\mathcal{E}$ summierbar, und diese Menge sei in endlich oder abzählbar unendlich viele meßbare Mengen $\mathcal{E}_k$ endlichen oder unendlichen Maßes zerlegt. Dabei ist dann $f(P)$ auf jedem $\mathcal{E}_k$ summierbar. Weiter nehmen wir an, $\mathcal{E}^{(1)} \subset \mathcal{E}^{(2)} \subset \dots$ sei eine wachsende Folge von Mengen endlichen Maßes, die gegen $\mathcal{E}$ strebt. Wir betrachten die Mengen $\mathcal{E}_k^{(s)} =$

$= \mathfrak{E}_k \mathfrak{E}^{(s)}$ endlichen Maßes. Sie wachsen für wachsendes s , und es gilt für sie

$$\lim_{s \rightarrow \infty} \mathfrak{E}_k^{(s)} = \mathfrak{E}_k \quad \text{und} \quad \mathfrak{E}^{(s)} = \mathfrak{E}_1^{(s)} + \mathfrak{E}_2^{(s)} + \mathfrak{E}_3^{(s)} + \dots,$$

wobei die rechts stehenden Mengen paarweise punktfremd sind. Für die Mengen $\mathfrak{E}^{(s)}$ endlichen Maßes gilt

$$\int_{\mathfrak{E}^{(s)}} f(P) G(d\mathfrak{E}) = \sum_{k=1}^{\infty} \int_{\mathfrak{E}_k^{(s)}} f(P) G(d\mathfrak{E}).$$

Wenn wir die Funktion $f(P)$ positiv annehmen und wenn wir in dieser Formel für $s \rightarrow \infty$ zur Grenze übergehen und das Lemma benutzen, erhalten wir die Formel (20) aus [49]. Im allgemeinen Fall gilt die Behauptung auf Grund von (137) und der Tatsache, daß sie für $f^+(P)$ und $f^-(P)$ einzeln gilt.

Entsprechend läßt sich die Eigenschaft 6 aus [49] nachweisen. Wir beweisen nun die absolute Stetigkeit des Integrals und nehmen dazu an, daß $f(P)$ nichtnegativ und auf $\mathfrak{E}$ summierbar ist. Es sei ein positives ε gegeben. Wir wählen m so groß, daß

$$\int_{\mathfrak{E} - \mathfrak{E}^{(m)}} fG(d\mathfrak{E}) \leq \frac{\varepsilon}{2} \quad (140)$$

gilt. Für eine beliebige Menge e , die in $\mathfrak{E}$ enthalten ist, können wir

$$\int_e fG(d\mathfrak{E}) = \int_{\mathfrak{E}^{(m)}e} fG(d\mathfrak{E}) + \int_{(\mathfrak{E} - \mathfrak{E}^{(m)})e} fG(d\mathfrak{E})$$

schreiben. Auf Grund der absoluten Stetigkeit des Integrals auf der Menge $\mathfrak{E}^{(m)}$ endlichen Maßes existiert ein solches $\eta > 0$, daß der absolute Betrag des Integrals über $\mathfrak{E}^{(m)}e$ nicht größer als $\varepsilon/2$ für $e \in \mathfrak{E}$ und $G(e) \leq \eta$ ist. Wenn wir (140) beachten, können wir dasselbe auch für das Integral über $(\mathfrak{E} - \mathfrak{E}^{(m)})e$ behaupten. Hieraus folgt, daß der absolute Betrag des Integrals über e nicht größer als ε ist, wenn $e \subset \mathfrak{E}$ und $G(e) \leq \eta$ ist. Damit ist auch die absolute Stetigkeit des Integrals bewiesen.

Die Sätze 1 bis 4 aus [54] kann man ganz leicht auf den Fall einer Menge $\mathfrak{E}$ unendlichen Maßes übertragen. Als Beispiel beweisen wir den Satz 1. Es sei ε eine gegebene positive Zahl. Wir wählen m so groß, daß die Ungleichung

$$\int_{\mathfrak{E} - \mathfrak{E}^{(m)}} F(P) G(d\mathfrak{E}) \leq \varepsilon \quad (141)$$

gilt, und schätzen das Integral der Differenz $f(P) - f_n(P)$ ab:

$$\left| \int_{\mathfrak{E}} (f - f_n) G(d\mathfrak{E}) \right| \leq \left| \int_{\mathfrak{E}^{(m)}} (f - f_n) G(d\mathfrak{E}) \right| + \left| \int_{\mathfrak{E} - \mathfrak{E}^{(m)}} (f - f_n) G(d\mathfrak{E}) \right|. \quad (142)$$

Auf der Menge $\mathfrak{E} - \mathfrak{E}^{(m)}$ gilt die Abschätzung $|f - f_n| \leq 2F$, und nach (141) erhalten wir

$$\left| \int_{\mathfrak{E} - \mathfrak{E}^{(m)}} (f - f_n) G(d\mathfrak{E}) \right| \leq \int_{\mathfrak{E} - \mathfrak{E}^{(m)}} |f - f_n| G(d\mathfrak{E}) \leq \int_{\mathfrak{E} - \mathfrak{E}^{(m)}} 2FG(d\mathfrak{E}) \leq 2\varepsilon.$$

Für die Mengen $\mathfrak{E}^{(m)}$ endlichen Maßes ist aber Satz 1 schon bewiesen; folglich existiert ein solches N , daß für $n > N$ der erste Summand auf der rechten Seite von (142) höchstens gleich ε ist. Somit erhalten wir

$$\left| \int_{\mathfrak{E}} (f - f_n) G(d\mathfrak{E}) \right| \leq 3\varepsilon \quad \text{für } n > N.$$

Damit ist, weil ε beliebig war, der Satz bewiesen. Ganz analog kann man auch die restlichen Sätze aus [54] beweisen.

64. Die Klasse L_2 auf einer Menge unendlichen Maßes. Die Konstruktion der Klasse L_2 und die Theorie der orthogonalen Funktionen lassen sich unschwer auf den Fall einer Menge $\mathfrak{E}$ unendlichen Maßes übertragen. Wir sagen, daß die Funktion $f(P)$ auf einer Menge $\mathfrak{E}$ unendlichen Maßes zu L_2 gehört, wenn sie auf $\mathfrak{E}$ meßbar ist und ihr Quadrat $f^2(P)$ bzw. das Quadrat ihres Absolutbetrags, $|f(P)|^2$, eine auf $\mathfrak{E}$ summierbare Funktion ist. Alle Sätze aus [55], außer Satz 1, gelten weiterhin. In Satz 1 benutzten wir wesentlich die Endlichkeit des Maßes von $\mathfrak{E}$. Man kann leicht ein Beispiel einer Funktion angeben, die zu L_2 gehört und nicht summierbar ist. Zum Beispiel gehört die Funktion $1/x$ in dem Intervall $[1, \infty]$ zu L_2 , weil $1/x^2$ summierbar ist, aber die Funktion $1/x$ selbst ist nicht summierbar. Außerdem benutzten wir die Endlichkeit des Maßes von $\mathfrak{E}$ noch beim Beweis des Satzes 4 aus [56]. Wir zeigen, daß dieser Satz auch in dem Fall gilt, daß das Maß von $\mathfrak{E}$ unendlich ist. Dabei nehmen wir an, daß $\mathfrak{E}$ die ganze Ebene $\mathfrak{E}_\infty$ ist. Gegeben sei eine Folge von Funktionen $f_n(P) \in L_2$ ($n = 1, 2, \dots$), die auf $\mathfrak{E}_\infty$ in sich konvergiert. Ferner sei Δ_m das Intervall, das durch die Ungleichungen $-m \leq x \leq m$, $-m \leq y \leq m$ definiert ist. Die Funktionen $f_n(P)$ gehören zu L_2 und konvergieren in jedem Δ_m in sich, weil das Integral einer nichtnegativen Funktion über Δ_m nicht größer als das Integral derselben Funktion über die ganze Ebene ist. Aus Satz 4 von [56] folgt, daß man aus der Folge der $f_n(P)$ eine Teilfolge $f_{n_1}^{(1)}(P)$, $f_{n_2}^{(1)}(P)$, ... auswählen kann, die fast überall auf Δ_1 konvergiert. Aus dieser Folge können wir wieder eine neue Teilfolge $f_{n_1}^{(2)}(P)$, $f_{n_2}^{(2)}(P)$, ... wählen, die fast überall auf Δ_2 konvergiert, usw. Man sieht leicht, daß die Folge $f_{n_1}^{(1)}(P)$, $f_{n_2}^{(2)}(P)$, ... fast überall auf $\mathfrak{E}$ konvergiert [IV/1, 15]. Es sei $f(P)$ die Grenzfunktion dieser Folge. Weil $f_n(P)$ auf $\mathfrak{E}_\infty$ in sich konvergiert, existiert für ein gegebenes positives ε ein solches N , daß

$$\int_{\mathfrak{E}_\infty} [f_{n_k}^{(k)}(P) - f_n(P)]^2 G(d\mathfrak{E}) \leq \varepsilon \quad \text{für } n_k > N, \quad n > N$$

ist. Für $k \rightarrow \infty$ erhalten wir wie in [56]

$$\int_{\mathfrak{E}_\infty} [f(P) - f_n(P)]^2 G(d\mathfrak{E}) \leq \varepsilon \quad \text{für } n > N,$$

womit die Verallgemeinerung von Satz 4 aus [56] bewiesen ist.

Wenn $f(P) \in L_2(\mathfrak{E}_\infty)$ und $\varepsilon_0 > 0$ gegeben ist, existiert ein N derart, daß

$$0 \leq \int_{\mathfrak{E}_\infty - \Delta_N} f^2(P) G(d\mathfrak{E}) \leq \varepsilon_0^2$$

ist.

Wir definieren eine Funktion $\psi(P)$ durch

$$\psi(P) = \begin{cases} f(P) & \text{in } \Delta_N, \\ 0 & \text{außerhalb } \Delta_N. \end{cases}$$

Offenbar ist $\psi(P) \in L_2(\mathfrak{E}_\infty)$ und

$$\|f - \psi\|_{\mathfrak{E}_\infty}^2 = \int_{\mathfrak{E}_\infty} [f(P) - \psi(P)]^2 G(d\mathfrak{E}) = \int_{\mathfrak{E}_\infty - \Delta_N} f^2(P) G(d\mathfrak{E}) \leq \varepsilon_0^2,$$

und hieraus sieht man, daß die lineare Mannigfaltigkeit der Funktionen $\psi(P)$, die nur in einem endlichen Intervall verschieden von 0 sind, in $L_2(\mathfrak{E}_\infty)$ überall dicht ist.

Wir beweisen nun die Separabilität von $L_2(\mathfrak{E}_\infty)$. Auf Δ_m ($m = 1, 2, \dots$) existiert, wie oben bewiesen wurde, eine in L_2 überall dichte Menge von Funktionen $\varphi_{k,m}(P)$ ($k, m = 1, 2, \dots$). Wir setzen diese Funktionen durch 0 auf die ganze Ebene $\mathfrak{E}_\infty$ fort

und erhalten somit eine abzählbare Menge von Funktionen $\varphi_{k,m}(P)$ aus $L_2(\mathcal{E}_\infty)$. Man sieht leicht, daß sie in $L_2(\mathcal{E}_\infty)$ überall dicht ist. Ist nämlich $f(P) \in L_2(\mathcal{E}_\infty)$ und $\varepsilon > 0$ gegeben, so existiert ein m_0 , für welches

$$\int_{\mathcal{E}_\infty - \Delta_{m_0}} f^2(P) G(d\mathcal{E}) \leq \frac{\varepsilon^2}{2}$$

gilt, und man kann nach dem oben Gesagten eine solche Funktion $\varphi_{k,m_0}(P)$ aus der oben angegebenen abzählbaren Menge auswählen, daß

$$\int_{\Delta_{m_0}} [f(P) - \varphi_{k,m_0}(P)]^2 G(d\mathcal{E}) \leq \frac{\varepsilon^2}{2}$$

ist; hieraus erhält man, da $\varphi_{k,m_0}(P) = 0$ außerhalb Δ_{m_0} ist,

$$\int_{\mathcal{E}_\infty} [f(P) - \varphi_{k,m_0}(P)]^2 G(d\mathcal{E}) \leq \varepsilon^2.$$

Damit ist die Separabilität von $L_2(\mathcal{E}_\infty)$ bewiesen.

Wir zeigen jetzt, daß die lineare Mannigfaltigkeit der stetigen Funktionen $\varphi(P)$, die außerhalb eines endlichen Intervalls (das für verschiedene Funktionen $\varphi(P)$ verschieden sein kann) gleich 0 sind, in $L_2(\mathcal{E}_\infty)$ überall dicht ist.

Diese lineare Mannigfaltigkeit nennt man die *lineare Mannigfaltigkeit der finiten stetigen Funktionen*.

Es seien $f(P) \in L_2(\mathcal{E}_\infty)$ und $\varepsilon > 0$ gegeben. Wir zeigen, daß eine finite stetige Funktion $\varphi(P)$ existiert, für die $\|f - \varphi\|_{\mathcal{E}_\infty} \leq \varepsilon$ gilt. Wie wir schon sahen, existiert ein solches m , daß $\|f\|_{\mathcal{E}_\infty - \Delta_m} \leq \varepsilon/2$ ist. Wenn wir Δ_m festhalten, können wir behaupten, daß in Δ_m eine stetige Funktion $\varphi(P)$ existiert derart, daß $\|f - \varphi\|_{\Delta_m} \leq \varepsilon/2$ gilt. Wir setzen $M = \max_{P \in \Delta_m} |\varphi(P)|$, wählen ein $h > 0$, nehmen $\varphi(P) = 0$ auf dem Rand von Δ_{m+h} an und setzen $\varphi(P)$ auf das ganze Intervall Δ_{m+h} unter Beibehaltung der Stetigkeit und ohne Vergrößerung von $\max |\varphi(P)|$ fort. Außerhalb von Δ_{m+h} setzen wir $\varphi(P) = 0$, so daß $\varphi(P)$ eine finite stetige Funktion ist. Unter Berücksichtigung des oben Gesagten erhalten wir

$$\begin{aligned} \|f - \varphi\|_{\mathcal{E}_\infty}^2 &= \|f - \varphi\|_{\Delta_m}^2 + \|f - \varphi\|_{\mathcal{E}_\infty - \Delta_m}^2 \\ &\leq \frac{\varepsilon^2}{4} + 2\|f\|_{\mathcal{E}_\infty - \Delta_m}^2 + 2\|\varphi\|_{\mathcal{E}_\infty - \Delta_m}^2 \leq \frac{\varepsilon^2}{4} + \frac{\varepsilon^2}{2} + 2\|\varphi\|_{\mathcal{E}_\infty - \Delta_m}^2. \end{aligned}$$

Da $\|\varphi\|_{\mathcal{E}_\infty - \Delta_m}^2$ gleich dem Integral von $|\varphi(P)|^2$ über $\Delta_{m+h} - \Delta_m$ ist, gilt für das Lebesguesche Integral

$$\|\varphi\|_{\mathcal{E}_\infty - \Delta_m}^2 = \int_{\Delta_{m+h} - \Delta_m} |\varphi(P)|^2 dx dy \leq M^2 (h^2 + 2mh),$$

und wenn wir h so wählen, daß $M^2 (h^2 + 2mh) \leq \frac{\varepsilon^2}{8}$ ist, erhalten wir danach $\|f - \varphi\|_{\mathcal{E}_\infty}^2 \leq \varepsilon^2$, was zu beweisen war. Eine analoge Behauptung gilt auch für das Lebesgue-Stieltjesche Integral.

Entsprechend läßt sich Satz 2 aus [60] beweisen. Wir bemerken, daß für das Lebesguesche Integral die Polynome nicht zu $L_2(\mathcal{E}_\infty)$ gehören. Wenn $\mathcal{E}$ eine beliebige meßbare Menge ist, so erhalten wir, wenn wir die Funktionen aus $L_2(\mathcal{E})$ auf $\mathcal{E}_\infty$ durch 0 fortsetzen, Funktionen aus $L_2(\mathcal{E}_\infty)$. Wenn wir hiervon ausgehen, können

wir alles oben Gesagte auch auf den Fall einer beliebigen, nicht beschränkten meßbaren Menge ausdehnen.

Als Beispiel für ein vollständiges orthogonales System in dem Intervall $(-\infty, \infty)$ führen wir die Hermiteschen Funktionen [III/2, 157]

$$\eta_k(x) = (-1)^k e^{x^2/2} \frac{d^k}{dx^k} (e^{-x^2})$$

und in dem Intervall $(0, \infty)$ die Laguerreschen Funktionen [III/2, 161]

$$\psi_k(x) = e^{x/2} \frac{d^k}{dx^k} (x^k e^{-x})$$

an.¹⁾

Das oben über $L_2(\mathbb{R}_\infty)$ Gesagte kann unmittelbar auch auf $L_p(\mathbb{R}_\infty)$, $p \geq 1$, übertragen werden, so wie das für den Fall einer beschränkten Menge geschah, und zwar auch für den Fall komplexer Funktionen.

65. Integrierbare Funktionen von beschränkter Variation. Bisher haben wir bei der Untersuchung des Lebesgue-Stieltjeschen Integrals vorausgesetzt, daß die Funktion $G(\mathcal{E})$ nichtnegativ ist. Wir behandeln jetzt den Fall, daß man die integrierbare Funktion $G(\mathcal{E})$ aus einer Intervallfunktion $G(\Delta)$ erhält, die eine Funktion von beschränkter Variation ist. Für diese Funktionen gilt die kanonische Darstellung als Differenz zweier nichtnegativer Funktionen,

$$G(\Delta) = G_1(\Delta) - G_2(\Delta)$$

mit

$$G_1(\Delta) = \frac{1}{2} [V(\Delta) + G(\Delta)], \quad G_2(\Delta) = \frac{1}{2} [V(\Delta) - G(\Delta)]$$

und $V(\Delta)$ die Totalvariation von $G(\Delta)$ im Intervall Δ ist. Jede der Funktionen $G_1(\Delta)$ und $G_2(\Delta)$ liefert eine nichtnegative, additive und normale Funktion $G_1(\mathcal{E})$ bzw. $G_2(\mathcal{E})$ auf dem abgeschlossenen Mengenkörper L_{G_1} bzw. L_{G_2} . Wir bezeichnen mit L_G den abgeschlossenen Mengenkörper, welcher der Durchschnitt von L_{G_1} und L_{G_2} ist. Auf diesem Körper definieren wir die volladditive und normale Funktion

$$G(\mathcal{E}) = G_1(\mathcal{E}) - G_2(\mathcal{E}).$$

Wir betrachten die nichtnegative, additive und normale Intervallfunktion

$$V(\Delta) = G_1(\Delta) + G_2(\Delta).$$

Ihre Fortsetzung führt zu einer Funktion $V(\mathcal{E})$, die auf dem abgeschlossenen Körper L_V definiert ist. Wenn wir die letzte Formel und die Nichtnegativität der Funktionen $G_i(\Delta)$ benutzen, können wir leicht zeigen, daß L_V der Durchschnitt von L_{G_1} und L_{G_2} ist, d. h., L_V stimmt mit L_G überein. Zuerst müssen wir zeigen, daß für jede Menge $\mathcal{E}$ das äußere Maß bezüglich der Funktion $V(\Delta)$, also $|\mathcal{E}|_V$, gleich der Summe der äußeren Maße bezüglich $G_1(\Delta)$ und $G_2(\Delta)$ ist, also $|\mathcal{E}|_V = |\mathcal{E}|_{G_1} + |\mathcal{E}|_{G_2}$. Wenn wir dann die Definition der Meßbarkeit benutzen, folgt leicht, daß die Menge $\mathcal{E}$ bezüglich $G_1(\Delta)$ und $G_2(\Delta)$ meßbar ist, wenn sie bezüglich $V(\Delta)$ meßbar ist; ebenso gilt umgekehrt, daß sie auch bezüglich $V(\Delta)$ meßbar ist, wenn sie bezüglich $G_1(\Delta)$ und $G_2(\Delta)$ meßbar ist. Bei der Integration muß man die Klasse der Funktionen $f(P)$ be-

¹⁾ Einen einfachen Beweis der Vollständigkeit dieser Systeme findet man bei R. COURANT und D. HILBERT, Methoden der mathematischen Physik, Band 1, 3. Aufl., Springer-Verlag, Berlin/Heidelberg/New York 1968, S. 81.

trachten, die bezüglich $V(\Delta)$ meßbar sind, d. h. die Klasse der Funktionen, die bezüglich $G_1(\Delta)$ und $G_2(\Delta)$ gleichzeitig meßbar sind. Das Integral wird durch

$$\int_{\mathfrak{E}} F(P) G(d\mathfrak{E}) = \int_{\mathfrak{E}} f(P) G_1(d\mathfrak{E}) - \int_{\mathfrak{E}} f(P) G_2(d\mathfrak{E})$$

definiert, und seine Existenz folgt aus der Existenz der rechts stehenden Integrale, wobei wir annehmen, daß diese beiden Integrale endliche Werte haben; anderenfalls kann die rechte Seite dieser Formel zu einem unbestimmten Ausdruck führen. Zwei Funktionen heißen äquivalent, wenn sie bezüglich $V(\mathfrak{E})$ äquivalent sind. Die Eigenschaften 1, 4, 5 und 7 bis 9 aus [52] für das Integral bleiben gültig. In Eigenschaft 3 gilt an Stelle der Ungleichung (16) die Ungleichung

$$\left| \int_{\mathfrak{E}} f(P) G(d\mathfrak{E}) \right| \leq \int_{\mathfrak{E}} |f(P)| V(d\mathfrak{E}).$$

In Eigenschaft 6 müssen wir an Stelle der Konvergenz der Reihe (49) die Konvergenz der Reihe

$$\sum_{k=1}^{\infty} \int_{\mathfrak{E}} |f(P)| V(d\mathfrak{E})$$

fordern, und schließlich ist in Eigenschaft 10 die Ungleichung (50) durch

$$\left| \int_{\mathfrak{E}} f(P) G(d\mathfrak{E}) \right| \leq \int_{\mathfrak{E}} F(P) V(d\mathfrak{E})$$

zu ersetzen.

Nach Definition des Integrals sind summierbare Funktionen diejenigen Funktionen, die bezüglich $V(\mathfrak{E})$ summierbar sind. Die Sätze 1 und 2 aus [54] für den Grenzübergang bleiben unverändert.

Man kann den Begriff des Integrals auch leicht auf den Fall ausdehnen, daß die Funktion $G(\Delta)$ eine komplexe Funktion ist:

$$G(\Delta) = G'(\Delta) + G''(\Delta)i,$$

wobei $G'(\Delta)$ und $G''(\Delta)$ Funktionen beschränkter Variation sind. Wenn wir die kanonische Zerlegung dieser Funktionen benutzen,

$$G'(\Delta) = G'_1(\Delta) - G'_2(\Delta), \quad G''(\Delta) = G''_1(\Delta) - G''_2(\Delta),$$

erhalten wir die Formel

$$G(\Delta) = (G'_1(\Delta) - G'_2(\Delta)) + (G''_1(\Delta) - G''_2(\Delta))i.$$

Die Funktion $G(\Delta)$ führt auf die Funktion $G(\mathfrak{E})$, die auf dem abgeschlossenen Körper L_G definiert ist, welcher der Durchschnitt der abgeschlossenen Körper L'_{G_i} und L''_{G_i} ($i=1, 2$) ist. Die Definitionen der bezüglich $G(\mathfrak{E})$ meßbaren Funktionen und des Integrals entsprechen den obigen, wobei die integrierbare Funktion auch eine komplexe Funktion sein kann.

Im Fall einer Veränderlichen haben wir für eine Funktion von beschränkter Variation die kanonische Darstellung

$$g(x) = g_1(x) - g_2(x)$$

als Differenz zweier nichtabnehmender Funktionen, und das Integral läßt sich in der Gestalt

$$\int_{\mathfrak{E}} f(x) dg(x) = \int_{\mathfrak{E}} f(x) dg_1(x) - \int_{\mathfrak{E}} f(x) dg_2(x)$$

schreiben.

Wenn wir die Totalvariation $v(x) = g_1(x) + g_2(x)$ einführen, erhalten wir

$$\left| \int_a^b f(x) dg(x) \right| \leq \int_a^b |f(x)| dv(x);$$

die Summierbarkeit von $f(x)$ bezüglich $g_1(x)$ und $g_2(x)$ ist also gleichbedeutend mit der Summierbarkeit von $|f(x)|$ bezüglich $v(x)$.

66. Die Reduktion mehrfacher Integrale. Wir behandeln jetzt ein grundlegendes Resultat aus der Theorie der mehrfachen Lebesgueschen Integrale, das die Berechnung eines mehrfachen Integrals durch aufeinanderfolgende einfache Quadraturen betrifft. Wir erinnern an die entsprechenden Resultate aus der Theorie der mehrfachen Integrale [II, 107]. Wenn z. B. die Funktion $f(x, y)$ in dem endlichen abgeschlossenen Intervall $\Delta [a \leq x \leq b, c \leq y \leq d]$ stetig ist, gilt für die Zurückführung eines Doppelintegrals auf zwei Quadraturen die Formel

$$\iint_{\Delta} f(x, y) dx dy = \int_a^b \left[\int_c^d f(x, y) dy \right] dx = \int_c^d \left[\int_a^b f(x, y) dx \right] dy.$$

Wir formulieren jetzt einen analogen Satz für das Lebesguesche Integral. Er wurde zuerst von dem italienischen Mathematiker FUBINI im Jahre 1907 bewiesen.

Satz von FUBINI. *Es sei $f(x, y)$ in dem endlichen Intervall $\Delta [a \leq x \leq b, c \leq y \leq d]$ summierbar. Dann ist die Funktion $f(x, y)$ bezüglich y in dem Intervall $[c, d]$ für fast alle x aus dem Intervall $[a, b]$ summierbar, die Funktion*

$$h(x) = \int_c^d f(x, y) dy \quad (143)$$

ist fast überall in dem Intervall $[a, b]$ definiert und bezüglich x in diesem Intervall summierbar, und es gilt die Beziehung

$$\iint_{\Delta} f(x, y) dx dy = \int_a^b \left[\int_c^d f(x, y) dy \right] dx. \quad (144)$$

Eine analoge Behauptung gilt auch bei Änderung der Reihenfolge der Integration. Dafür erhalten wir

$$\iint_{\Delta} f(x, y) dx dy = \int_c^d \left[\int_a^b f(x, y) dx \right] dy. \quad (145)$$

Wir weisen darauf hin, daß in diesem Satz die Integrale im Lebesgueschen Sinne aufzufassen sind und daß man in demselben Sinne natürlich auch die Summierbarkeit der Funktionen verstehen muß. Die Behauptung über die Summierbarkeit schließt offenbar auch die Behauptung über die Meßbarkeit der Funktionen ein. Wir betonen, daß die Funktion (143) nicht in allen Punkten des Intervalls $[a, b]$ definiert zu sein braucht, aber in jedem Fall ist sie dort fast überall definiert. Die analoge Bemerkung bezieht sich auch auf die Funktion

$$l(y) = \int_a^b f(x, y) dx. \quad (146)$$

Um den ziemlich komplizierten Beweis des Satzes von FUBINI zu vereinfachen, formulierten wir den Satz für einen Spezialfall [67]. Im weiteren werden wir verschiedene allgemeine Formulierungen dieses Satzes geben. Dem Beweis dieses Satzes schicken wir mehrere Hilfssätze voraus.

Hilfssatz 1. *Wenn für die Funktionen $f_1(x, y)$, $f_2(x, y)$, ..., $f_m(x, y)$, die in dem Intervall Δ summierbar sind, der Satz von Fubini gilt, so gilt er auch für jede Linearkombination*

$$f(x, y) = \sum_{k=1}^m c_k f_k(x, y) \quad (147)$$

dieser Funktionen.

Jede dieser Funktionen $f_k(x, y)$ ist nach Voraussetzung bezüglich y im Intervall $[c, d]$ summierbar, wenn man aus dem Intervall $[a, b]$, in dem x variiert, eine Menge A_k vom Maß 0 ausschließt. Wenn wir aus $[a, b]$ die Menge $A = A_1 + A_2 + \dots + A_m$, die ebenfalls das Maß 0 hat, ausschließen, ist die Funktion (147) für die restlichen Werte von x bezüglich y in dem Intervall $[c, d]$ summierbar. Alle Funktionen

$$h_k(x) = \int_c^d f_k(x, y) dy$$

sind in $[a, b]$ mit Ausnahme der Punkte der Menge A definiert. Weiter gilt nach Voraussetzung für die Funktionen $f_k(x, y)$ die Formel (144). Wenn wir die Integrationsregeln für eine Summe beachten und daran denken, daß wir einen konstanten Faktor vor das Integralzeichen ziehen können, ergibt sich, daß die Formel (144) auch für die Funktion (147) gilt. Damit ist Hilfssatz 1 bewiesen.

Bemerkung. Wenn von den Funktionen $f_k(x, y)$ nur bekannt ist, daß sie bezüglich y in dem Intervall $[c, d]$ für fast alle Werte $x \in [a, b]$ meßbar sind, kann man offenbar auch dasselbe bezüglich der Funktion (147) behaupten, weil die Summe meßbarer Funktionen wieder eine meßbare Funktion ist. Dabei wird angenommen, daß die Summe sinnvoll ist [43].

Hilfssatz 2. *Es sei in dem Intervall Δ eine monotone Folge summierbarer Funktionen $f_n(x, y)$ gegeben, die gegen die in Δ summierbare Funktion $f(x, y)$ konvergiert. Gilt dabei für jede Funktion $f_n(x, y)$ der Satz von Fubini, so gilt er auch für die Grenzfunktion $f(x, y)$.*

Zum Beweis nehmen wir an, daß die $f_n(x, y)$ eine nichtfallende Folge bilden. Der Fall einer nichtwachsenden Folge läßt sich hierauf zurückführen, indem $f_n(x, y)$ durch $-f_n(x, y)$ ersetzt wird. Nach Voraussetzung ist jede der Funktionen $f_n(x, y)$ meßbar und bezüglich y in dem Intervall $[c, d]$ summierbar, wenn wir aus dem Intervall $[a, b]$ der Veränderlichen x eine Menge A_n vom Maß 0 ausschließen.

Wenn wir aus $[a, b]$ die Menge $A = A_1 + A_2 + \dots$ vom Maß 0 ausschließen, ist für die Restmenge der Werte von x auch die Grenzfunktion $f(x, y)$ bezüglich y in $[c, d]$ meßbar. Nach Voraussetzung ist jede der Funktionen

$$h_n(x) = \int_c^d f_n(x, y) dy \quad (148)$$

in $[a, b]$ definiert, wenn wir aus $[a, b]$ eine Menge A_n vom Maß 0 ausschließen. Wenn wir aus $[a, b]$ die Menge A ausschließen, die das Maß 0 hat, sind alle Funktionen (148) für die restlichen x -Werte, d. h. auf der Menge $[a, b] - A$ definiert und nach Voraussetzung auf $[a, b]$ summierbar. Die Folge der $h_n(x)$ ist eine wachsende Folge, und wir können fast überall auf $[a, b]$ die meßbare Grenzfunktion $h(x) = \lim_{n \rightarrow \infty} h_n(x)$ definieren. Wenn wir beachten, daß für die Funktionen $f_n(x, y)$ nach Voraussetzung der Satz von FUBINI gilt und daß die Grenzfunktion $f(x, y)$ nach

Voraussetzung in $\mathcal{A}$ summierbar ist, können wir

$$\int_a^b h_n(x) dx = \iint_{\mathcal{A}} f_n(x, y) dx dy \leq \iint_{\mathcal{A}} f(x, y) dx dy$$

schreiben. Hieraus folgt nach Satz 2 aus [54], daß $h(x)$ in $[a, b]$ summierbar ist, und es gilt

$$\int_a^b h(x) dx = \lim_{n \rightarrow \infty} \int_a^b h_n(x) dx = \lim_{n \rightarrow \infty} \iint_{\mathcal{A}} f_n(x, y) dx dy.$$

Andererseits erhalten wir aus Satz 2 aus [54]

$$\lim_{n \rightarrow \infty} \iint_{\mathcal{A}} f_n(x, y) dx dy = \iint_{\mathcal{A}} f(x, y) dx dy,$$

und somit gilt

$$\int_a^b h(x) dx = \iint_{\mathcal{A}} f(x, y) dx dy. \quad (149)$$

Wir definieren die Funktion $h(x)$ folgendermaßen:

$$h(x) = \lim_{n \rightarrow \infty} h_n(x) = \lim_{n \rightarrow \infty} \int_c^d f_n(x, y) dy.$$

Damit der Hilfssatz vollständig bewiesen ist, müssen wir noch zeigen, daß die Funktion $f(x, y)$ bezüglich y in $[c, d]$ für fast alle $x \in [a, b]$ summierbar ist und daß die Funktion $h(x)$ fast überall auf $[a, b]$ durch die Formel

$$h(x) = \int_c^d f(x, y) dy \quad (150)$$

ausgedrückt werden kann. Wenn wir das bewiesen haben, erhalten wir nach (149) für die Funktion $f(x, y)$ den Satz von FUBINI. Es sei also B die Menge derjenigen Punkte aus $[a, b]$, in denen die Funktion $h(x)$ definiert und gleich ∞ ist. Da die Funktion $h(x)$ summierbar ist, hat B das Maß 0. Wenn aus $[a, b]$ die Menge $A + B$ vom Maß 0 ausgeschlossen wird, strebt auf der übrigbleibenden Menge, d. h. fast überall auf $[a, b]$, die wachsende Folge $h_n(x)$ gegen die Funktion $h(x)$, die endliche Werte annimmt, d. h., für jeden Punkt x , der in der Menge $[a, b] - (A + B)$ liegt, sind die Integrale der nichtfallenden Folge der Funktionen $f_n(x, y)$ über das Intervall $[c, d]$ nach y durch die Zahl $h(x)$ beschränkt. Nach Satz 2 aus [54] ist $f(x, y)$ für diese x -Werte bezüglich y in $[c, d]$ summierbar, und es gilt die Formel

$$\int_c^d f(x, y) dy = \lim_{n \rightarrow \infty} \int_c^d f_n(x, y) dy.$$

Nach (148) kann die Funktion $h(x)$ fast überall in $[a, b]$ durch (150) ausgedrückt werden. Damit ist Hilfssatz 2 bewiesen.

Bemerkung. Aus dem Anfang dieses Beweises folgt unmittelbar: Wird für die Funktionen $f_n(x, y)$ nur vorausgesetzt, daß sie bezüglich y in $[c, d]$ für fast alle x aus $[a, b]$ meßbar sind, so ist auch die Grenzfunktion $f(x, y)$ bezüglich y für fast alle x aus $[a, b]$ meßbar.

67. Der Fall einer charakteristischen Funktion. In diesem Abschnitt werden wir den Satz von FUBINI für den Fall beweisen, daß der Integrand die charakteristische Funktion einer meßbaren Menge $\mathcal{E}$ ist, die zum Intervall $\mathcal{A}$ gehört, von dem im Satz

die Rede ist. Das Integral von $\omega_{\mathfrak{E}}(P) = \omega_{\mathfrak{E}}(x, y)$ liefert offenbar das Maß $m(\mathfrak{E})$ der Menge $\mathfrak{E}$. Es sei $\mathfrak{E}_{x_0}$ die Menge derjenigen Punkte von $\mathfrak{E}$, die eine gegebene Abszisse x_0 haben, d. h., $\mathfrak{E}_{x_0}$ ist der Schnitt von $\mathfrak{E}$ mit der Geraden $x = x_0$. Die charakteristische Funktion dieser Menge ist gleich $\omega_{\mathfrak{E}}(x_0, y)$. Die Meßbarkeit von $\mathfrak{E}_{x_0}$ bezüglich y ist gleichbedeutend mit der Meßbarkeit der Funktion $\omega_{\mathfrak{E}}(x_0, y)$ in dem Intervall $[c, d]$, und wenn diese meßbar ist, ist das lineare Maß von $\mathfrak{E}_{x_0}$, das wir mit $m'(\mathfrak{E}_{x_0})$ bezeichnen, gleich dem Integral von $\omega_{\mathfrak{E}}(x_0, y)$ in dem genannten Intervall. Die Summierbarkeit der Funktion $\omega_{\mathfrak{E}}(x_0, y)$ garantiert ihre Beschränktheit. Der Satz von FUBINI für die charakteristische Funktion $\omega_{\mathfrak{E}}(x, y)$ besteht somit in folgender Behauptung: *Die Funktion $\omega_{\mathfrak{E}}(x, y)$ ist in dem Intervall $[c, d]$ bezüglich y für fast alle x aus $[a, b]$ meßbar; die beschränkte Funktion*

$$h(x) = m'(\mathfrak{E}_x) = \int_c^d \omega_{\mathfrak{E}}(x, y) dy$$

ist in $[a, b]$ meßbar, und es gilt

$$m(\mathfrak{E}) = \int_a^b \left[\int_c^d \omega_{\mathfrak{E}}(x, y) dy \right] dx. \quad (151)$$

Kurz gesagt, $\mathfrak{E}_x$ ist bezüglich y für fast alle x -Werte meßbar, und es gilt

$$m(\mathfrak{E}) = \int_a^b m'(\mathfrak{E}_x) dx. \quad (152)$$

Wir werden den Satz von FUBINI für charakteristische Funktionen in mehreren Schritten beweisen.

Hilfssatz 3. *Der Satz von Fubini gilt für die charakteristische Funktion jedes halboffenen Intervalls, jeder offenen Menge und jeder G_{δ} -Menge, die zu Δ gehören.*

Wenn es in Δ ein halboffenes Intervall Δ' ($\alpha < x \leq \beta$, $\gamma < y \leq \delta$) gibt, ist $\omega_{\Delta'}(x, y)$ bezüglich y für jedes x meßbar, es gilt

$$h(x) = \int_c^d \omega_{\Delta'}(x, y) dy = \delta - \gamma$$

für $\alpha < x \leq \beta$, ferner ist $h(x) = 0$, wenn x außerhalb Δ' liegt, und der Hilfssatz gilt offenbar, weil das Maß von Δ' gleich dem Produkt $(\beta - \alpha)(\delta - \gamma)$ ist. Eine offene Menge $\mathfrak{E}_0$ ist die Summe abzählbar vieler halboffener punktfremder Intervalle Δ_k , und es ist

$$\omega_{\mathfrak{E}_0}(x, y) = \sum_{k=0}^{\infty} \omega_{\Delta_k}(x, y).$$

Nach dem Hilfssatz 1 aus [66] gilt der Satz von FUBINI für endliche Summen $\sum_{k=1}^m \omega_{\Delta_k}(x, y)$. Für wachsendes m bilden diese Summen eine nichtfallende Folge, die gegen die beschränkte und deshalb summierbare Funktion $\omega_{\mathfrak{E}_0}(x, y)$ strebt; nach dem Hilfssatz 2 aus [66] gilt der Satz von FUBINI auch für $\omega_{\mathfrak{E}_0}(x, y)$. Wir nehmen schließlich an, $\mathfrak{E}'_0$ sei eine G_{δ} -Menge, die zu dem offenen Intervall Δ gehört. Wir können sie als Produkt

$$\mathfrak{E}'_0 = \prod_{k=1}^{\infty} O_k \quad (153)$$

darstellen, wobei O_k offene Mengen sind, die zu dem offenen Intervall Δ gehören. Lagen einige der O_k nicht in dem offenen Intervall Δ , so könnten wir O_k durch das Produkt dieser O_k mit dem offenen Intervall Δ ersetzen. Nach (153) ist $\omega_{\mathcal{E}'_0}(x, y)$ der Limes der nichtwachsenden Folge charakteristischer Funktionen $\omega_{\mathcal{E}_m}(x, y)$ der offenen Mengen

$$\mathcal{E}_m = \prod_{k=1}^m O_k,$$

und da für $\omega_{\mathcal{E}_m}(x, y)$ der Satz von FUBINI schon bewiesen ist, gilt er auf Grund von Hilfssatz 2 aus [66] auch für $\omega_{\mathcal{E}'_0}(x, y)$. Wenn einige der Punkte der Menge $\mathcal{E}'_0$ vom Typ G_δ auf dem Rand von Δ liegen, erweitern wir Δ etwas, so daß $\mathcal{E}'_0$ im Innern des erweiterten Intervalls Δ_0 liegt. Der Satz von FUBINI gilt für $\omega_{\mathcal{E}'_0}(x, y)$ in Δ_0 . Hieraus erhalten wir, wenn wir beachten, daß $\omega_{\mathcal{E}'_0}(x, y) = 0$ außerhalb von Δ ist, den Satz von FUBINI für $\omega_{\mathcal{E}'_0}(x, y)$ in dem Intervall Δ . Wir bemerken, daß in allen in diesem Hilfssatz betrachteten Fällen $\omega_{\mathcal{E}}(x, y)$ bezüglich y für alle x -Werte meßbar ist.

Hilfssatz 4. *Wenn $\mathcal{E}$ eine Menge ist, die zu Δ gehört und das ebene Maß 0 hat, ist für fast alle x aus $[a, b]$ das lineare Maß von $\mathcal{E}_x$ gleich 0, und für $\omega_{\mathcal{E}}(x, y)$ gilt der Satz von Fubini.*

Wir konstruieren eine Menge $\mathcal{E}'_0$ vom Typ G_δ , die in Δ liegt und die Menge $\mathcal{E}$ überdeckt derart, daß $m(\mathcal{E}_0 - \mathcal{E}) = 0$ ist [40]. Wir haben $\mathcal{E}'_0 = \mathcal{E} + (\mathcal{E}'_0 - \mathcal{E})$, und weil $m(\mathcal{E}) = 0$ und $m(\mathcal{E}_0 - \mathcal{E}) = 0$ ist, ist auch $m(\mathcal{E}'_0) = 0$. Für $\mathcal{E}'_0$ gilt der Satz von FUBINI, und wir können

$$\int_a^b \left[\int_c^d \omega_{\mathcal{E}'_0}(x, y) dy \right] dx = 0$$

schreiben. Die in den eckigen Klammern stehende Größe ist nichtnegativ, und auf Grund der Eigenschaft 14 aus [49] gilt fast überall in dem Intervall $a \leq x \leq b$

$$\int_c^d \omega_{\mathcal{E}'_0}(x, y) dy = 0.$$

Hieraus sieht man, daß das lineare Maß der Menge $\mathcal{E}'_0$ der Punkte, die auf fast allen zur y -Achse parallelen Geraden liegen, gleich 0 ist.

Wegen $\mathcal{E} \subset \mathcal{E}'_0$ gilt das auch für die Menge $\mathcal{E}$, d. h., für fast alle x aus $[a, b]$ gilt

$$\int_c^d \omega_{\mathcal{E}}(x, y) dy = 0;$$

daher gilt für $\omega_{\mathcal{E}}(x, y)$ der Satz von FUBINI:

$$m(\mathcal{E}) = 0 = \int_a^b \left[\int_c^d \omega_{\mathcal{E}}(x, y) dy \right] dx.$$

Hilfssatz 5. *Der Satz von Fubini gilt für die charakteristische Funktion $\omega_{\mathcal{E}}(x, y)$ jeder meßbaren Menge $\mathcal{E}$, die zu Δ gehört.*

Wir konstruieren eine Menge $\mathcal{E}'_0$ vom Typ G_δ , die zu Δ gehört und die $\mathcal{E}$ überdeckt, derart, daß $m(\mathcal{E}'_0 - \mathcal{E}) = 0$ ist. Nach Hilfssatz 3 und 4 gilt der Satz von FUBINI für die charakteristischen Funktionen der Mengen $\mathcal{E}'_0$ und $\mathcal{E}'_0 - \mathcal{E}$. Da aber $\omega_{\mathcal{E}}(x, y) = \omega_{\mathcal{E}'_0}(x, y) + \omega_{\mathcal{E}'_0 - \mathcal{E}}(x, y)$ ist, gilt nach Hilfssatz 1 der Satz von FUBINI auch für $\omega_{\mathcal{E}}(x, y)$.

Hat eine meßbare nichtbeschränkte Menge $\mathcal{E}$ endliches Maß, so ist $\mathcal{E}_x$ für fast alle x meßbar, und es gilt die Formel (152). Das erhält man unmittelbar durch Grenzübergang von beschränkten Mengen. Auf dieselbe Weise kann man leicht zeigen: Ist $\mathcal{E}$ meßbar, so ist $\mathcal{E}_x$ für fast alle x meßbar. Wenn $m(\mathcal{E}_x)$ außerdem summierbar ist, hat $\mathcal{E}$ endliches Maß, und es gilt die Formel (152). Hilfssatz 4 gilt offenbar auch für nichtbeschränkte Mengen.

68. Der Satz von Fubini. Um den Satz von FUBINI vollständig beweisen zu können, benötigen wir noch einen einfachen Hilfssatz.

Hilfssatz 6. Der Satz von Fubini gilt für jede meßbare Funktion $f(x, y)$, die in Δ endlich viele endliche Werte annimmt.

Es sei $f(x, y) = c_k$ ($k = 1, 2, \dots, m$), wenn der Punkt (x, y) zur Menge $\mathcal{E}_k$ gehört, wobei $\Delta = \mathcal{E}_1 + \mathcal{E}_2 + \dots + \mathcal{E}_m$ ist. Wir können $f(x, y)$ als Linearkombination der charakteristischen Funktionen der Mengen $\mathcal{E}_k$ darstellen:

$$f(x, y) = \sum_{k=1}^m c_k \omega_{\mathcal{E}_k}(x, y);$$

jetzt erhält man den Satz von FUBINI für $f(x, y)$ unmittelbar aus den Hilfssätzen 1 und 5.

Auf Grund von Hilfssatz 6 ist der Satz von FUBINI ganz einfach zu beweisen. Es sei $f(x, y)$ auf Δ summierbar. Wir zerlegen $f(x, y)$ in den positiven und den negativen Teil:

$$f(x, y) = f^+(x, y) - f^-(x, y).$$

Nach Hilfssatz 1 genügt es, den Satz für f^+ und f^- zu beweisen, d. h., bei dem Beweis können wir annehmen, daß die summierbare Funktion $f(x, y)$ nichtnegativ ist. Wie wir aus [46] wissen, kann eine solche Funktion als Grenzfunktion einer nichtfallenden Folge meßbarer nichtnegativer Funktionen $f_n(x, y)$ mit endlich vielen Werten dargestellt werden. Auf Grund von Hilfssatz 6 gilt der Satz von FUBINI für die Funktionen $f_n(x, y)$, und nach Hilfssatz 2 gilt er dann auch für die Funktion $f(x, y)$. Damit ist der Satz von FUBINI bewiesen.

Wir weisen darauf hin, daß in dem Satz von FUBINI die Funktion $f(x, y)$ in dem Intervall Δ als summierbar vorausgesetzt wurde. Unter dieser Voraussetzung haben nach dem Satz von FUBINI die in den Formeln (144) und (145) rechts stehenden Integrale einen Sinn und liefern das Doppelintegral von $f(x, y)$ über Δ . Die umgekehrte Schlußfolgerung, daß das Doppelintegral existiert, wenn die rechts stehenden Integrale Sinn haben, kann falsch sein. Es gibt Beispiele dafür, daß die iterierten Integrale, die auf den rechten Seiten der Formeln (144) und (145) stehen, sinnvoll und die Resultate einander gleich sind, aber die Funktion $f(x, y)$ auf Δ nicht meßbar oder meßbar, jedoch nicht summierbar ist. Wenn aber $f(x, y)$ auf dem Intervall Δ nichtnegativ ist, gilt auch die Umkehrung, und es gilt der

Satz. Ist $f(x, y)$ in dem Intervall Δ meßbar und nichtnegativ, so folgt aus der Existenz des iterierten Integrals, das in (144) rechts steht, daß die Funktion $f(x, y)$ auf Δ summierbar ist; damit gilt für sie der Satz von Fubini.

Wir nehmen an, daß das in (144) rechts stehende iterierte Integral sinnvoll ist, d. h., fast überall existiert bezüglich x in dem Intervall $[a, b]$ die Funktion (143), die in dem Intervall $[a, b]$ summierbar ist. Wir führen die Funktionen

$$[f(x, y)]_n = \begin{cases} f(x, y) & \text{für } f(x, y) \leq n, \\ n & \text{für } f(x, y) > n \end{cases}$$

ein. Sie sind beschränkt, meßbar und bilden eine nichtfallende Folge, die gegen $f(x, y)$ strebt. Sie sind offenbar auf Δ summierbar, und für sie gilt der Satz von FUBINI. Wir können

$$\iint_{\Delta} [f(x, y)]_n \, dx dy = \int_a^b \left[\int_c^d [f(x, y)]_n \, dy \right] dx$$

schreiben. Nun ist $[f(x, y)]_n \leq f(x, y)$; folglich gilt

$$\iint_{\Delta} [f(x, y)]_n \, dx dy \leq \int_a^b h(x) \, dx.$$

Hieraus folgt nach [50], daß $f(x, y)$ auf Δ summierbar ist.

Folgerung 1. Wenn $f(x, y)$ das Vorzeichen wechselt, aber für $|f(x, y)|$ die rechte Seite von (144) existiert, ist nach dem Satz $|f(x, y)|$ summierbar; daher ist auch $f(x, y)$ auf Δ summierbar, und wir können auf $f(x, y)$ den Satz von Fubini anwenden.

Folgerung 2. Ist $f(x, y)$ auf Δ meßbar und bezüglich y für fast alle x summierbar, so ist die durch (143) definierte Funktion $h(x)$ meßbar. Wie üblich können wir $f(x, y)$ als nichtnegativ annehmen. Für die beschränkten Funktionen $[f(x, y)]_n$ gilt der Satz von FUBINI, und

$$h_n(x) = \int_c^d [f(x, y)]_n \, dy$$

ist meßbar. Wenn n gegen ∞ strebt, sehen wir, daß die Grenzfunktion $h(x)$ meßbar ist.

Wir geben noch einige Verallgemeinerungen des Satzes von FUBINI an. Wenn $f(x, y)$ auf einer meßbaren beschränkten Menge $\mathfrak{E}$ summierbar ist, gilt

$$\iint_{\mathfrak{E}} f(x, y) \, dx dy = \int_{B_x} \left[\int_{\mathfrak{E}_x} f(x, y) \, dy \right] dx = \int_{B_y} \left[\int_{\mathfrak{E}_y} f(x, y) \, dx \right] dy, \quad (154)$$

wobei $\mathfrak{E}_x$ die Menge der Punkte von $\mathfrak{E}$ ist, die die Abszisse x haben, und $\mathfrak{E}_y$ die analoge Menge für y ; B_x und B_y sind die Projektionen von $\mathfrak{E}$ auf die x - bzw. y -Achse. Die Integrale über $\mathfrak{E}_x$ und $\mathfrak{E}_y$ brauchen für gewisse Werte von x und y , die je eine Menge vom Maß 0 bilden, nicht zu existieren. Um (154) zu beweisen, genügt es, $\mathfrak{E}$ durch ein endliches Intervall Δ zu überdecken, und die Funktion $f_0(x, y)$ zu konstruieren, die in den Punkten von $\mathfrak{E}$ gleich $f(x, y)$ und in allen anderen Punkten von Δ , die nicht zu $\mathfrak{E}$ gehören, gleich 0 ist.

Wir beweisen jetzt, daß man den Satz von FUBINI auch auf den Fall nichtbeschränkter Mengen ausdehnen kann. Als Beispiel betrachten wir die ganze Ebene. Es genügt, nichtnegative Funktionen zu betrachten. Wir nehmen also an, daß die Funktion $f(x, y)$ meßbar, nichtnegativ und summierbar in der Ebene ist, d. h., es existiere das Doppelintegral

$$A = \iint_{-\infty}^{\infty} f(x, y) \, dx dy. \quad (155)$$

Auf jedem endlichen Intervall $\Delta_{m,n} [-m \leq x \leq m, -n \leq y \leq n]$ gilt der Satz von FUBINI, d. h.

$$\iint_{\Delta_{m,n}} f(x, y) \, dx dy = \int_{-m}^m \left[\int_{-n}^n f(x, y) \, dy \right] dx \leq A.$$

Es sei B_{mn} ($m, n=1, 2, \dots$) die Menge der Punkte x aus $[-m, m]$, für welche $f(x, y)$ bezüglich y im Intervall $[-n, n]$ nicht meßbar ist. Wir haben $B_{mn} \subset B_{m'n'}$ für $m' \geq m$, $n' \geq n$ und $m(B_{mn})=0$. Die Grenzmenge B hat ebenfalls das Maß 0 [35], und die Funktion $f(x, y)$ ist bezüglich y in jedem Intervall $[-n, n]$ ($n=1, 2, \dots$) für fast alle x aus $(-\infty, \infty)$ meßbar. Wenn n unbeschränkt wächst und wir Satz 4 aus [54] und die Definition des Integrals über $(-\infty, \infty)$ anwenden, erhalten wir

$$\int_{-m}^m \left[\int_{-\infty}^{\infty} f(x, y) dy \right] dx \leq A.$$

Wenn schließlich m wächst, erhalten wir die Ungleichung

$$\int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} f(x, y) dy \right] dx \leq A. \quad (156)$$

Wir zeigen jetzt, daß in dieser Ungleichung das Zeichen $<$ nicht auftreten kann. Wenn es gelten würde, dann würde eine solche positive Zahl α existieren, daß

$$\int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} f(x, y) dy \right] dx < A - \alpha$$

und erst recht

$$\int_{\Delta_{n,n}} f(x, y) dx dy = \int_{-n}^n \left[\int_{-n}^n f(x, y) dy \right] dx < A - \alpha$$

gilt, was unsinnig ist, weil das Integral über das Intervall $\Delta_{n,n}$ bei unbeschränkt wachsendem n gegen das Integral (155) streben muß. Somit muß in der Formel (156) das Gleichheitszeichen stehen, und wenn wir dies mit (155) vergleichen, erhalten wir für die ganze Ebene die Formel, die im Satz von FUBINI angegeben ist. Aus diesem Beweis folgt offenbar auch die Existenz der aufeinanderfolgenden Quadraturen, die in dieser Formel auftreten.

Der Satz von FUBINI kann auch für Integrale beliebiger Vielfachheit definiert werden. Wir formulieren das entsprechende Ergebnis. Es sei Δ_{m+n} das Intervall im $(m+n)$ -dimensionalen Raum R_{m+n} , das durch die Ungleichungen

$$a_1 \leq x_1 \leq b_1, \quad a_2 \leq x_2 \leq b_2, \quad \dots, \quad a_{m+n} \leq x_{m+n} \leq b_{m+n}$$

definiert ist, und Δ_m und Δ_n seien Intervalle in den Räumen R_m und R_n , die durch die Ungleichungen

$$\Delta_m: \quad a_1 \leq x_1 \leq b_1; \quad a_2 \leq x_2 \leq b_2; \quad \dots; \quad a_m \leq x_m \leq b_m;$$

$$\Delta_n: \quad a_{m+1} \leq x_{m+1} \leq b_{m+1}; \quad \dots; \quad a_{m+n} \leq x_{m+n} \leq b_{m+n}$$

definiert ist. Weiter sei $f(P)$ eine im Intervall Δ_{m+n} summierbare Funktion. Wenn wir einen Punkt $P_0(x_1^0, x_2^0, \dots, x_m^0)$ aus Δ_m festhalten, ist die Funktion $f(P)$ eine in Δ_n summierbare Funktion für beliebige P_0 , außer eventuell einer Menge von Punkten P_0 , die in R_m das Maß 0 hat. Das Integral von $f(P)$ über Δ_n ,

$$h(x_1, x_2, \dots, x_m) = \int_{\Delta_n} f(P) dx_{m+1} \dots dx_{m+n},$$

liefert in Δ_m eine summierbare Funktion, und es gilt

$$\int_{\Delta_{m+n}} f(P) dx_1 dx_2 \dots dx_{m+n} = \int_{\Delta_m} \left[\int_{\Delta_n} f(P) dx_{m+1} dx_{m+2} \dots dx_{m+n} \right] dx_1 dx_2 \dots dx_m. \quad (157)$$

Der Satz von FUBINI gestattet noch eine einfache Verallgemeinerung für den Fall des Lebesgue-Stieltjesschen Integrals. Wir nehmen an, es seien zwei wachsende und beschränkte Funktionen $g(x)$ und $k(y)$ gegeben. Wenn wir diese Funktionen benutzen, können wir die Maße $G(\mathcal{A})$ und $K(\mathcal{A})$ halboffener Intervalle definieren, und dann diese Funktionen auf die abgeschlossenen Körper L_G und L_K erweitern. Somit erhalten wir additive, nichtnegative und normale Funktionen $G(\mathcal{E})$ und $K(\mathcal{E})$ auf L_G und L_K . Ausgehend von der Funktion $g(x)k(y)$, die in der Ebene definiert ist, können wir eine additive, nichtnegative und normale Funktion $M(\mathcal{E})$ auf einem abgeschlossenen Mengenkörper L_M in der Ebene konstruieren. Wenn $f(P) = f(x, y)$ bezüglich $M(\mathcal{E})$ meßbar und in einem Intervall $\mathcal{A}$ der Ebene summierbar ist, dann ist diese Funktion bezüglich y in dem Intervall $\mathcal{A}_y$ der y -Achse, das dem Intervall $\mathcal{A}$ der Ebene entspricht, in bezug auf die Funktion $K(\mathcal{E})$ summierbar:

$$h(x) = \int_{\mathcal{A}_y} f(x, y) K(d\mathcal{E}) = \int_{\mathcal{A}_y} f(x, y) dk(y),$$

wenn (nötigenfalls) aus dem Intervall $\mathcal{A}_x$ der x -Achse, das dem Intervall $\mathcal{A}$ in der Ebene entspricht, eine Menge von Werten, die in bezug auf $G(\mathcal{E})$ das Maß 0 haben, ausgeschlossen wird. Die Funktion $h(x)$ ist bezüglich $G(\mathcal{E})$ auf $\mathcal{A}_x$ summierbar, und es gilt

$$\iint_{\mathcal{A}} f(P) M(d\mathcal{E}) = \iint_{\mathcal{A}} f(x, y) dd[g(x)k(y)] = \int_{\mathcal{A}_x} \left[\int_{\mathcal{A}_y} f(x, y) dk(y) \right] dg(x).$$

Wenn wir die Reihenfolge der Integration vertauschen, erhalten wir eine analoge zweite Formel. Diese Verallgemeinerung des Satzes von FUBINI wird genauso bewiesen wie der Satz von FUBINI, nur muß man das Lebesguesche Integral überall durch das Lebesgue-Stieltjessche Integral ersetzen, und die Meßbarkeit im Lebesgueschen Sinne muß man durch die Meßbarkeit bezüglich der Funktionen $G(\mathcal{E})$, $K(\mathcal{E})$ und $M(\mathcal{E})$ ersetzen.

Bemerkung. Wenn bezüglich der Funktion $f(x, y)$ nur vorausgesetzt wird, daß sie in dem Intervall $\mathcal{A}$ der Ebene meßbar ist, folgt hieraus, daß sie für fast alle x -Werte aus $[a, b]$ bezüglich y in $[c, d]$ und für fast alle y -Werte aus $[c, d]$ bezüglich x in $[a, b]$ meßbar ist. Diese Bemerkung folgt unmittelbar aus der Bemerkung, die wir nach dem Beweis der Hilfssätze 1 und 2 [66] machten, und aus dem ganzen weiteren Beweis des Satzes von FUBINI.

69. Das Vertauschen der Reihenfolge der Integration. Wir erwähnen noch einen Satz, der sich mit dem Vertauschen der Reihenfolge der Integration beschäftigt.

Satz. Es sei die Funktion $g(x, t)$ bezüglich t auf dem Intervall $[c, d]$ für alle $x \in [a, b]$ summierbar und bezüglich x auf $[a, b]$ für alle Werte $t \in [c, d]$ mit Ausnahme eventuell einer Menge, deren Lebesguesches Maß gleich 0 ist, von beschränkter Variation. Es sei ferner die Totalvariation von $g(x, t)$ bezüglich x in dem Intervall $[a, b]$ für alle Werte t nicht größer als eine nichtnegative und in $[c, d]$ meßbare Funktion $F(t)$, für die das Integral

$$\int_c^d F(t) dt \quad (158)$$

existiert. Dann ist die Funktion

$$\int_c^d g(x, t) dt \quad (159)$$

eine Funktion von beschränkter Variation von x in $[a, b]$, und für jede in $[a, b]$ stetige Funktion $f(x)$ gilt

$$\int_c^d \left[\int_a^b f(x) dx g(x, t) \right] dt = \int_a^b f(x) dx \left[\int_c^d g(x, t) dt \right], \quad (160)$$

wobei die Integrale bezüglich t Lebesguesche Integrale sind.

Beweis. Um zu beweisen, daß (159) eine Funktion von beschränkter Variation ist, zerlegen wir das Intervall $[a, b]$ in Teilintervalle durch $a = x_0 < x_1 < x_2 < \dots < x_{n-1} < x_n = b$ und bilden für diese Zerlegung die Summe t_δ [8]. Wir erhalten

$$t_\delta = \sum_{k=1}^n \left| \int_c^d g(x_k, t) dt - \int_c^d g(x_{k-1}, t) dt \right| = \sum_{k=1}^n \left| \int_c^d [g(x_k, t) - g(x_{k-1}, t)] dt \right|;$$

hieraus folgt

$$t_\delta \leq \sum_{k=1}^n \int_c^d |g(x_k, t) - g(x_{k-1}, t)| dt.$$

Nach Voraussetzung gilt aber $\sum_{k=1}^n |g(x_k, t) - g(x_{k-1}, t)| \leq F(t)$, folglich ist

$$t_\delta \leq \int_c^d F(t) dt;$$

hieraus ergibt sich, daß die Funktion (159) eine Funktion von beschränkter Variation ist.

Offensichtlich gilt

$$\int_c^d \sum_{k=1}^n f(\xi_k) [g(x_k, t) - g(x_{k-1}, t)] dt = \sum_{k=1}^n f(\xi_k) \left[\int_c^d g(x_k, t) dt - \int_c^d g(x_{k-1}, t) dt \right], \quad (161)$$

wobei ξ_k ein beliebiger Punkt im Intervall $[x_{k-1}, x_k]$ ist. Wenn die Teilintervalle beliebig klein werden, strebt die rechte Seite dieser Formel gegen das auf der rechten Seite von (160) stehende Integral. Für die in dem Intervall $[a, b]$ stetige Funktion $f(x)$ gilt die Ungleichung $|f(x)| \leq L$, wobei L eine positive Zahl ist. Der in (161) links stehende Integrand läßt sich folgendermaßen abschätzen:

$$\left| \sum_{k=1}^n f(\xi_k) [g(x_k, t) - g(x_{k-1}, t)] \right| \leq L \sum_{k=1}^n |g(x_k, t) - g(x_{k-1}, t)| \leq LF(t).$$

Wenn wir Satz 1 aus [54] anwenden, sehen wir, daß wir in dem in (161) links stehenden Integral bei beliebig kleiner werdenden Intervallen unter dem Integralzeichen zur Grenze übergehen können, wobei der Integrand im Limes das nachstehende Stieltjessche Integral liefert:

$$\int_a^b f(x) dg(x, t).$$

Der Grenzübergang in (161) führt aber dann zu Formel (160). Dieser Satz läßt einige elementare Verallgemeinerungen zu. Es kann z. B. das Intervall $[a, b]$ unendlich groß und die Funktion $f(x)$ innerhalb dieses Intervalls als stetig und beschränkt angenommen werden. Das Lebesguesche Integral über t kann durch ein Lebesgue-Stieltjessches Integral ersetzt werden. Das ursprüngliche Stieltjessche Integral über $g(x, t)$ läßt sich durch das allgemeine Stieltjessche Integral ersetzen, und die Funktion $f(x)$ braucht in dem Intervall $[a, b]$ nur als beschränkt angenommen zu werden. Dabei folgt aus der Existenz des in (160) rechts stehenden Integrals die Existenz des links stehenden Integrals und die Gleichheit dieser Integrale.

70. Die Stetigkeit im Mittel. Wir wenden uns wieder L_p ($p \geq 1$) zu und beweisen, daß jede Funktion aus L_p stetig ist, wenn ihr Zuwachs in der Norm von L_p untersucht wird. Wir betrachten den Fall einer beschränkten meßbaren Menge $\mathfrak{E}$. Wir bilden die Integrale im Lebesgueschen Sinne, und der Einfachheit halber betrachten wir den zweidimensionalen Fall.

Satz. Wenn auf einer beschränkten meßbaren Menge $\mathcal{E}$ die Funktion $f(x, y)$ zu L_p gehört, existiert zu beliebig gegebenem $\varepsilon > 0$ ein solches $\eta > 0$, daß für $|h| \leq \eta$ und $|k| \leq \eta$

$$\|f(x+h, y+k) - f(x, y)\|_{\mathcal{E}}^p = \int_{\mathcal{E}} |f(x+h, y+k) - f(x, y)|^p dx dy \leq \varepsilon^p \quad (162)$$

gilt.

Beweis. Der Punkt $(x+h, y+k)$ braucht nicht zu $\mathcal{E}$ zu gehören, und wir setzen $f(x, y)$ außerhalb von $\mathcal{E}$ gleich 0. Der Index an der Norm zeigt an, bezüglich welcher Menge die Norm gebildet ist. Wir können $\mathcal{E}$ in ein endliches abgeschlossenes Intervall Δ_0 ($a_1 \leq x \leq b_1$, $a_2 \leq y \leq b_2$) einschließen. Nach Satz 1 aus [60] existiert eine in Δ_0 stetige Funktion $\varphi(x, y)$ derart, daß $\|f - \varphi\|_{\Delta_0} \leq \varepsilon/4$ ist. Wir können $\varphi(x, y)$, ohne daß die Funktion unstetig wird, auf ein größeres Intervall fortsetzen, z. B. auf das Intervall Δ_1 ($a_1 - 1 \leq x \leq b_1 + 1$, $a_2 - 1 \leq y \leq b_2 + 1$). Wir schreiben die Differenz $f(x+h, y+k) - f(x, y)$ in der Form

$$f(x+h, y+k) - \varphi(x+h, y+k) + \varphi(x+h, y+k) - \varphi(x, y) + \varphi(x, y) - f(x, y);$$

dann folgt

$$\begin{aligned} \|f(x+h, y+k) - f(x, y)\|_{\Delta_0} &\leq \|f(x+h, y+k) - \varphi(x+h, y+k)\|_{\Delta_0} \\ &\quad + \|\varphi(x+h, y+k) - \varphi(x, y)\|_{\Delta_0} + \|\varphi(x, y) - f(x, y)\|_{\Delta_0}. \end{aligned}$$

Nun war $\|\varphi(x, y) - f(x, y)\|_{\Delta_0} \leq \varepsilon/4$. Wegen der gleichmäßigen Stetigkeit von $\varphi(x, y)$ in Δ_1 existiert ein solches $\eta_1 > 0$ (wir wählen $\eta_1 < 1$), daß

$$\|\varphi(x+h, y+k) - \varphi(x, y)\|_{\Delta_0} \leq \frac{\varepsilon}{4} \quad \text{für } |h| \leq \eta_1 \quad \text{und} \quad |k| \leq \eta_1$$

ist, also

$$\|f(x+h, y+k) - f(x, y)\|_{\Delta_0} \leq \frac{\varepsilon}{2} + \|f(x+h, y+k) - \varphi(x+h, y+k)\|_{\Delta_0} \quad (163)$$

für

$$|h|, |k| \leq \eta_1.$$

Nun ist aber

$$\|f(x+h, y+k) - \varphi(x+h, y+k)\|_{\Delta_0}^p = \int_{\Delta_0} |f(x+h, y+k) - \varphi(x+h, y+k)|^p dx dy$$

oder

$$\|f(x+h, y+k) - \varphi(x+h, y+k)\|_{\Delta_0}^p = \int_{\Delta_0(h, k)} |f(x, y) - \varphi(x, y)|^p dx dy,$$

wobei $\Delta_0(h, k)$ das Intervall ist, das aus Δ_0 durch Parallelverschiebung um den Vektor (h, k) entsteht. Wenn wir etwa $h > 0$ und $k > 0$ annehmen, erhalten wir

$$\begin{aligned} \|f(x+h, y+k) - \varphi(x+h, y+k)\|_{\Delta_0}^p &= \|f(x, y) - \varphi(x, y)\|_{\Delta'}^p + \int_{\substack{a_1+h \leq x \leq b_1+h, \\ b_2 \leq y \leq b_2+k}} |\varphi(x, y)|^p dx dy + \int_{\substack{b_1 \leq x \leq b_1+h, \\ a_2+k \leq y \leq b_2+k}} |\varphi(x, y)|^p dx dy, \end{aligned}$$

wobei Δ' ein Teil von Δ_0 ist, so daß $\|f(x, y) - \varphi(x, y)\|_{\Delta'} \leq \varepsilon/4$ gilt. Die letzten beiden Integrale streben offenbar für $h, k \rightarrow 0$ gegen 0; folglich existiert ein solches $\eta_2 > 0$, daß

$$\|f(x+h, y+k) - \varphi(x+h, y+k)\|_{\Delta_0} \leq \frac{\varepsilon}{2} \quad \text{für } |h| \leq \eta \quad \text{und} \quad |k| \leq \eta$$

mit $\eta = \min(\eta_1, \eta_2)$ ist. Aus (163) erhalten wir

$$||f(x+h, y+k) - f(x, y)||_{\mathcal{A}_0} \leq \varepsilon$$

für $|h| \leq \eta$ und $|k| \leq \eta$ und somit um so mehr

$$||f(x+h, y+k) - f(x, y)||_{\mathcal{B}} \leq \varepsilon$$

für $|h| \leq \eta$ und $|k| \leq \eta$, was zu beweisen war.

Der Satz kann auch für nichtbeschränkte meßbare Mengen bewiesen werden.

71. Mittelfunktionen. Wir führen jetzt einen gewissen Mittelungsprozeß für eine beliebige summierbare Funktion $f(P)$ ein. Er führt zu Folgen von Funktionen, welche Ableitungen beliebiger Ordnung haben und in einem gewissen Sinne gegen $f(P)$ streben. Der Bestimmtheit wegen betrachten wir den Fall der Ebene und definieren die Punkte durch ihre Koordinaten. Es sei $\omega_1(P, Q) = \omega_1(x, y; \xi, \eta)$ eine Funktion, die nur vom Abstand

$$r = |PQ| = \sqrt{(x-\xi)^2 + (y-\eta)^2}$$

abhängt und die für $r \geq 1$ gleich 0 ist. Weiter sei diese Funktion stetig und besitze stetige Ableitungen beliebiger Ordnung nach allen vier Koordinaten. Wir bemerken dabei, daß die Differentiation von $\omega_1(P; Q)$ nach x und y durch die Differentiation nach ξ bzw. η mit entgegengesetztem Vorzeichen im Ergebnis ersetzt werden kann.

Wir nehmen außerdem an, daß

$$\int \omega_1(x, y; \xi, \eta) \, dx dy = 1 \quad (164)$$

ist. Wir schreiben das Integrationsgebiet nicht dazu, da wir annehmen, daß sich die Integration über die ganze Ebene erstreckt. Nach dem oben Gesagten integriert man faktisch über den Kreis $(x-\xi)^2 + (y-\eta)^2 \leq 1$. Wenn in (164) die Integration nach (x, y) durch die Integration nach (ξ, η) ersetzt wird, erhalten wir dasselbe Ergebnis.

Wir führen jetzt die Bezeichnung

$$\omega_\varrho(x, y; \xi, \eta) = \omega_1\left(\frac{x}{\varrho}, \frac{y}{\varrho}; \frac{\xi}{\varrho}, \frac{\eta}{\varrho}\right) \quad (\varrho > 0) \quad (165)$$

ein. Die Funktion ω_ϱ hat dieselben Differentiationseigenschaften wie ω_1 , jedoch ist $\omega_\varrho = 0$ für $r \geq \varrho$ sowie

$$\int \omega_\varrho(x, y; \xi, \eta) \, dx dy = \varrho^2. \quad (166)$$

Im folgenden geben wir eine der Möglichkeiten an, die Funktion $\omega_1(P; Q)$ auszuwählen. Mit $C_\varrho(\xi, \eta)$ bezeichnen wir im folgenden den Kreis mit dem Mittelpunkt (ξ, η) und dem Radius ϱ . Wir führen noch die Bezeichnung

$$\int |\omega_1(x, y; \xi, \eta)| \, dx dy = c \quad (167)$$

ein, woraus

$$\int |\omega_\varrho(x, y; \xi, \eta)| \, dx dy = c\varrho^2 \quad (168)$$

folgt. In einem beschränkten offenen oder abgeschlossenen Bereich D_0 sei eine summierbare Funktion $f(x, y)$ gegeben (statt D_0 könnte man auch eine beliebige beschränkte meßbare Menge betrachten). Wir setzen die Funktion $f(x, y)$ auf die ganze Ebene durch 0 fort und bilden von ihr die *Mittelfunktion*

$$f_\varrho(\xi, \eta) = \frac{1}{\varrho^2} \int f(x, y) \omega_\varrho(x, y; \xi, \eta) \, dx dy. \quad (169)$$

Die positive Zahl ϱ wird der *Radius der Mittlung* genannt. Der Integrand ist außerhalb des Kreises $C_\varrho(\xi, \eta)$ gleich 0, und wenn der Abstand d des Punktes (ξ, η) von D_0 größer als 0 ist, ist $f_\varrho(\xi, \eta) = 0$ für $\varrho \leq d$.

Satz 1. Die Funktion $f_\varrho(\xi, \eta)$ ist stetig und hat stetige partielle Ableitungen beliebiger Ordnung in der ganzen Ebene.

Beweis. Da ω_ϱ nur von den Differenzen $x - \xi$ und $y - \eta$ abhängt, gilt

$$|f_\varrho(\xi + h, \eta + k) - f_\varrho(\xi, \eta)| \leq \frac{1}{\varrho^2} \int |f(x, y)| \cdot |\omega_\varrho(x - h, y - k; \xi, \eta) - \omega_\varrho(x, y; \xi, \eta)| \, dx dy. \quad (170)$$

Aus der gleichmäßigen Stetigkeit von ω_ϱ als Funktion von x, y, ξ, η folgt, daß für ein beliebig gegebenes $\varepsilon > 0$ ein $\eta > 0$ existiert derart, daß

$$|\omega_\varrho(x - h, y - h; \xi, \eta) - \omega_\varrho(x, y; \xi, \eta)| \leq \varepsilon \quad \text{für } |h| \leq \eta \quad \text{und} \quad |k| \leq \eta$$

gilt und folglich

$$|f_\varrho(\xi + h, \eta + k) - f_\varrho(\xi, \eta)| \leq \frac{\varepsilon}{\varrho^2} \int |f(x, y)| \, dx dy \quad \text{für } |h| \leq \eta \quad \text{und} \quad |k| \leq \eta$$

ist. Hieraus folgt, da ε beliebig war, die Stetigkeit von $f_\varrho(\xi, \eta)$. Wir beweisen nun die Existenz und Stetigkeit der Ableitung nach ξ :

$$\frac{f_\varrho(\xi + h, \eta) - f_\varrho(\xi, \eta)}{h} = \frac{1}{\varrho^2} \int f(x, y) \frac{\omega_\varrho(x - h, y; \xi, \eta) - \omega_\varrho(x, y; \xi, \eta)}{h} \, dx dy. \quad (171)$$

Nach dem Mittelwertsatz ist

$$\frac{\omega_\varrho(x - h, y; \xi, \eta) - \omega_\varrho(x, y; \xi, \eta)}{h} = - \frac{\partial}{\partial x} \omega_\varrho(x - \theta h, y; \xi, \eta) \quad (0 < \theta < 1).$$

Für beliebiges h überschreitet der Absolutbetrag der rechten Seite eine Zahl K nicht, und in dem Integral (171) wird der Absolutbetrag des Integranden nicht größer als die summierbare Funktion $K|f(x, y)|$, so daß wir unter dem Integralzeichen zur Grenze übergehen können. Wir erhalten

$$\frac{\partial f_\varrho(\xi, \eta)}{\partial \xi} = - \frac{1}{\varrho^2} \int f(x, y) \frac{\partial \omega_\varrho(x, y; \xi, \eta)}{\partial x} \, dx dy = \frac{1}{\varrho^2} \int f(x, y) \frac{\partial \omega_\varrho(x, y; \xi, \eta)}{\partial \xi} \, dx dy.$$

Die Stetigkeit dieser partiellen Ableitung wird genauso bewiesen wie die Stetigkeit von $f_\varrho(\xi, \eta)$. Diesen Beweis kann man auch für höhere Ableitungen anwenden, und diese ergeben sich durch Differentiation unter dem Integralzeichen:

$$\frac{\partial^k f_\varrho(\xi, \eta)}{\partial \xi^p \partial \eta^q} = \frac{1}{\varrho^2} \int f(x, y) \frac{\partial^k \omega_\varrho(x, y; \xi, \eta)}{\partial \xi^p \partial \eta^q} \, dx dy. \quad (172)$$

Satz 2. Gehört $f(x, y)$ zu $L_2(D_0)$, so ist

$$\int |f_\varrho(x, y)|^2 \, dx dy \leq c_1 \int |f(x, y)|^2 \, dx dy, \quad (173)$$

wobei c_1 eine Konstante ist, und

$$\lim_{\varrho \rightarrow 0} \|f - f_\varrho\|^2 = \lim_{\varrho \rightarrow 0} \int |f(x, y) - f_\varrho(x, y)|^2 \, dx dy = 0. \quad (174)$$

Beweis. Es gilt

$$|f_\varrho(\xi, \eta)| \leq \frac{1}{\varrho^2} \int \sqrt{|\omega_\varrho(x, y; \xi, \eta)|} \sqrt{|\omega_\varrho(x, y; \xi, \eta)|} |f(x, y)| \, dx dy.$$

Wenden wir die Schwarzsche Ungleichung an, so erhalten wir

$$|f_\varrho(\xi, \eta)|^2 \leq \frac{1}{\varrho^2} \int |\omega_\varrho(x, y; \xi, \eta)| \, dx dy \frac{1}{\varrho^2} \int |f(x, y)|^2 |\omega_\varrho(x, y; \xi, \eta)| \, dx dy. \quad (175)$$

Wenn wir (168) benutzen, nach (ξ, η) integrieren und die Reihenfolge der Integration vertauschen [69], erhalten wir

$$\int |f_\varrho(\xi, \eta)|^2 \, d\xi d\eta \leq c \int |f(x, y)|^2 \left[\frac{1}{\varrho^2} \int |\omega_\varrho(x, y; \xi, \eta)| \, d\xi d\eta \right] dx dy;$$

hieraus ergibt sich wegen (168) die Beziehung (173) ($c_1 = c^2$), wobei sich die Integration auf der rechten Seite über D_0 erstreckt, weil $f(x, y)$ außerhalb von D_0 verschwindet.

Wir beweisen jetzt die Formel (174). Wenn wir (166) und (169) beachten, können wir

$$\begin{aligned} \|f - f_\varrho\|^2 &= \int |f(\xi, \eta) - f_\varrho(\xi, \eta)|^2 \, d\xi d\eta \\ &= \int \left| \frac{1}{\varrho^2} \int [f(\xi, \eta) - f(x, y)] \omega_\varrho(x, y; \xi, \eta) \, dx dy \right|^2 d\xi d\eta \end{aligned} \quad (176)$$

schreiben. Im inneren Integral führen wir an Stelle von x und y die neuen Integrationsvariablen u und v durch $x = \xi + u$, $y = \eta + v$ ein, und wenn wir $\omega_\varrho(\xi + u, \eta + v; \xi, \eta) = 0$ für $u^2 + v^2 \geq \varrho^2$ beachten, erhalten wir mit Hilfe der Schwarzschen Ungleichung

$$\begin{aligned} & \left| \int [f(\xi, \eta) - f(x, y)] \omega_\varrho(x, y; \xi, \eta) \, dx dy \right|^2 \\ & \leq \int_{u^2+v^2 \leq \varrho^2} |f(\xi, \eta) - f(\xi + u, \eta + v)|^2 \, dudv \int_{u^2+v^2 \leq \varrho^2} \omega_\varrho^2(\xi + u, \eta + v; \xi, \eta) \, dudv. \end{aligned}$$

Wegen $\omega_\varrho^2 \leq c_2$ (wobei c_2 eine nicht von ϱ abhängende Konstante ist) wird die rechte Seite nicht größer als

$$c_2 \pi \varrho^2 \int_{u^2+v^2 \leq \varrho^2} |f(\xi, \eta) - f(\xi + u, \eta + v)|^2 \, dudv.$$

Im Ergebnis liefert die Formel (176)

$$\|f - f_\varrho\|^2 \leq \frac{c_2 \pi}{\varrho^2} \int \left[\int_{u^2+v^2 \leq \varrho^2} |f(\xi, \eta) - f(\xi + u, \eta + v)|^2 \, dudv \right] d\xi d\eta. \quad (177)$$

Wenn wir die Reihenfolge der Integration vertauschen, erhalten wir

$$\|f - f_\varrho\|^2 \leq \frac{c_2 \pi}{\varrho^2} \int \left[\int |f(\xi, \eta) - f(\xi + u, \eta + v)|^2 \, d\xi d\eta \right] dudv. \quad (178)$$

Auf Grund der Stetigkeit im Mittel existiert für beliebig gegebenes $\varepsilon > 0$ ein $\eta > 0$ derart, daß

$$\int |f(\xi, \eta) - f(\xi + u, \eta + v)|^2 \, d\xi d\eta \leq \varepsilon$$

für $u^2 + v^2 \leq \eta^2$ ist. Daher liefert die Ungleichung (178)

$$\|f - f_\varrho\|^2 \leq \frac{c_2 \pi \varepsilon}{\varrho^2} \iint_{u^2+v^2 \leq \varrho^2} dudv = c_2 \pi^2 \varepsilon \quad \text{für } 0 < \varrho < \eta.$$

Da ε beliebig war, folgt hieraus (174).

Wir beweisen noch einen Satz, der uns im folgenden nützlich sein wird.

Satz 3. *Es sei U eine Menge von Funktionen $f(x, y)$ aus $L_2(D_0)$, die bezüglich der Norm durch ein und dieselbe Zahl beschränkt sind. Dann sind die entsprechenden Funktionen $f_\varrho(x, y)$ für festes ϱ dem Absolutbetrag nach durch ein und dieselbe Zahl beschränkt und gleichgradig stetig.*

Beweis. Nach Voraussetzung existiert ein solches positives m , daß für alle Funktionen $f(x, y)$ aus U

$$\int_{D_0} |f(x, y)|^2 dx dy \leq m \quad (179)$$

gilt. Wenden wir auf die rechte Seite von (169) die Schwarzsche Ungleichung an, so erhalten wir

$$|f_\varrho(\xi, \eta)|^2 \leq \frac{1}{\varrho^4} \int \omega_\varrho^2(x, y; \xi, \eta) dx dy \int |f(x, y)|^2 dx dy \leq \frac{c_2 \pi}{\varrho^2} m.$$

Wir müssen noch die gleichgradige Stetigkeit beweisen. Dazu wenden wir auf die rechte Seite von (170) die Schwarzsche Ungleichung an und benutzen (179):

$$|f_\varrho(\xi + h, \eta + k) - f_\varrho(\xi, \eta)| \leq \frac{m}{\varrho^4} \int |\omega_\varrho(x - h, y - k; \xi, \eta) - \omega_\varrho(x, y; \xi, \eta)|^2 dx dy,$$

und aus der gleichmäßigen Stetigkeit der Funktion ω_ϱ folgt die gleichgradige Stetigkeit der Funktionen $f_\varrho(\xi, \eta)$ bei beliebiger Wahl der Funktion $f(x, y)$ aus U .

Den oben angegebenen Beweis können wir auch für L_p ($p \geq 1$) anwenden. Ist $p > 1$, so müssen statt der Schwarzschen Ungleichung die Höldersche Ungleichung und die Formel

$$|\omega_\varrho(x, y; \xi, \eta)| = |\omega_\varrho(x, y; \xi, \eta)|^{1/p} |\omega_\varrho(x, y; \xi, \eta)|^{1/p'} \quad \left(\frac{1}{p} + \frac{1}{p'} = 1 \right)$$

benutzt werden.

Für $p = 1$ läßt sich alles genauso beweisen, nur wird hierbei weder die Schwarzsche Ungleichung noch die Höldersche Ungleichung benutzt. Somit gilt der folgende Satz.

Satz 2'. *Gehört $f(x, y)$ zu $L_p(D_0)$ ($p \geq 1$), so ist*

$$\int |f_\varrho(x, y)|^p dx dy \leq c_1 \int |f(x, y)|^p dx dy, \quad (180)$$

$$\lim_{\varrho \rightarrow 0} \int |f(x, y) - f_\varrho(x, y)|^p dx dy = 0. \quad (181)$$

Es gilt auch Satz 3, wenn man L_p statt L_2 setzt. Der Beweis bleibt auch dann gültig, wenn D_0 eine nichtbeschränkte meßbare Menge, z. B. die ganze Ebene $\mathbb{E}_\infty$ ist. Wenn wir Satz 3 benutzen, können wir leicht zeigen, daß die lineare Mannigfaltigkeit der finiten stetigen Funktionen mit stetigen Ableitungen beliebiger Ordnung in $L_p(\mathbb{E}_\infty)$ überall dicht ist.

Wir bemerken noch, daß alles oben Gesagte sowohl für reelle als auch für komplexe Räume L_p gilt. Wir betrachteten die Mittelfunktionen unter der Voraussetzung, daß $f(x, y)$ zu L_p ($p \geq 1$) gehört. Wir nehmen jetzt an, $f(x, y)$ sei in einem offenen Gebiet D_0 stetig, und es sei D' ein beliebiger, aber fester abgeschlossener Bereich, der im Innern von D_0 liegt. Da $f(x, y)$ in jedem abgeschlossenen Bereich, der im Innern von D_0 liegt, gleichmäßig stetig ist, existiert für gegebenes $\varepsilon > 0$ ein solches $\eta > 0$, daß $|f(x, y) - f(\xi, \eta)| \leq \varepsilon$ ist, wenn $(x, y) \in C_\varrho(\xi, \eta)$ für $\varrho \leq \eta$ und (ξ, η) ein beliebiger Punkt aus D' ist. Aus der Ungleichung

$$|f(\xi, \eta) - f_\varrho(\xi, \eta)| \leq \frac{1}{\varrho^2} \int |f(\xi, \eta) - f(x, y)| \cdot |\omega_\varrho(x, y; \xi, \eta)| dx dy$$

erhalten wir

$$|f(\xi, \eta) - f_\varrho(\xi, \eta)| \leq \varepsilon c \quad \text{für } \varrho \leq \eta \quad \text{und} \quad (\xi, \eta) \in D'. \quad (182)$$

Wenn die Funktion $f(x, y)$ in D_0 außerdem stetige Ableitungen bis zu einer bestimmten Ordnung hat und wenn wir in (172) die Differentiation nach ξ und η durch die Differentiation nach x und y ersetzen, partiell integrieren und die Eigenschaften der Funktion ω_ϱ benutzen, erhalten wir für $(\xi, \eta) \in D'$ und hinreichend kleine ϱ

$$\frac{\partial^k f_\varrho(\xi, \eta)}{\partial \xi^p \partial \eta^q} = \frac{1}{\varrho^2} \int \frac{\partial^k f(x, y)}{\partial x^p \partial y^q} \omega_\varrho(x, y; \xi, \eta) \, dx dy, \quad (183)$$

d. h., die Ableitung einer Mittelfunktion ist gleich der Mittelfunktion der Ableitung für $(\xi, \eta) \in D'$ und hinreichend kleines ϱ .

Wir bemerken noch, daß unter den erwähnten Voraussetzungen $\omega_\varrho(x, y; \xi, \eta)$ in der Nähe des Randes von D gleich 0 ist; daher erhalten wir nach der partiellen Integration für die Kurvenintegrale den Wert 0. Als Integrationsweg in diesen Integralen kann man hinreichend glatte Kurven nehmen, die in hinreichender Nähe des Randes von D liegen.

Aus dem oben Gesagten folgt

Satz 4. Ist $f(x, y)$ mit seinen Ableitungen bis zu einer bestimmten Ordnung l im Innern von D_0 stetig, so streben in jedem abgeschlossenen Bereich D' , der im Innern von D_0 liegt, für $\varrho \rightarrow 0$ die Mittelfunktion und ihre Ableitungen bis zur Ordnung l gleichmäßig gegen $f(x, y)$ und gegen ihre entsprechenden Ableitungen.

Wir bemerken noch: Wenn $f(x, y)$ beschränkt ist, $|f(x, y)| \leq m$, gilt

$$|f_\varrho(\xi, \eta)| \leq \frac{1}{\varrho^2} \int |f(x, y)| |\omega_\varrho(x, y; \xi, \eta)| \, dx dy \leq cm. \quad (184)$$

Satz 5. Ist $F(x, y)$ in jedem abgeschlossenen Bereich D' , der im Innern von D_0 liegt, summierbar und gilt

$$\int_{D_0} F(x, y) \varphi(x, y) \, dx dy = 0 \quad (185)$$

für jede Funktion $\varphi(x, y)$, die im Innern von D_0 stetig ist und stetige Ableitungen bis zur Ordnung l besitzt und außerhalb eines bestimmten abgeschlossenen Bereichs, der im Innern von D_0 liegt, gleich 0 ist, so ist F äquivalent 0.

Beweis. Es genügt zu zeigen, daß aus der Voraussetzung des Satzes die Gültigkeit von (185) für jede beschränkte meßbare finite Funktion $\varphi(x, y)$ folgt, d. h. für jede Funktion, die außerhalb eines abgeschlossenen Bereichs, der im Innern von D_0 liegt, gleich 0 ist (dieser Bereich kann wie in den Voraussetzungen des Satzes für verschiedene $\varphi(x, y)$ verschieden sein). Jetzt können wir den Beweis, daß $F(x, y)$ äquivalent 0 ist, genauso wie den Beweis von Satz 12 aus [52] führen. Es sei $\varphi(x, y)$ eine Funktion mit $|\varphi(x, y)| \leq m$, und $\varphi_\varrho(x, y)$ seien die Mittelfunktionen, so daß $|\varphi_\varrho(x, y)| \leq cm$ ist. Die Funktion $\varphi_\varrho(x, y)$ ist für hinreichend kleine ϱ finit und hat stetige Ableitungen beliebiger Ordnung, so daß nach Voraussetzung

$$\int_{D_0} F(x, y) \varphi_\varrho(x, y) \, dx dy = 0 \quad (186)$$

gilt. Die Funktionen $\varphi_\varrho(x, y)$ konvergieren für $\varrho \rightarrow 0$ in L_1 gegen $\varphi(x, y)$ auf D_0 , und es existiert eine solche Folge $\varphi_{\varrho_n}(x, y)$, die fast überall gegen $\varphi(x, y)$ strebt [62].

Außerdem ist

$$|F(x, y) \varphi_\varrho(x, y)| \leq c m |F(x, y)|,$$

wobei die $F(x, y)$ $\varphi_\varrho(x, y)$ finit sind und rechts eine summierbare Funktion steht. Wenn wir in (186) für $\varrho = \varrho_n$ zur Grenze übergehen, erhalten wir (185). Damit ist der Satz bewiesen.

Wir weisen noch auf eine Möglichkeit hin, wie sich die Funktion $\omega_1(x, y; \xi, \eta)$ wählen läßt, und zwar setzen wir

$$\omega_1(x, y; \xi, \eta) = \begin{cases} c_0 e^{r^2/(r^2-1)} & \text{für } r < 1, \\ 0 & \text{für } r \geq 1 \end{cases} \quad (r^2 = (x - \xi)^2 + (y - \eta)^2), \quad (187)$$

wobei die Konstante c_0 so gewählt ist, daß die Bedingung (164) gilt. Es ist klar, daß stetige Ableitungen beliebiger Ordnung für $r < 1$ und $r > 1$ vorhanden sind. Für $r \rightarrow 1$, $r < 1$, strebt $e^{r^2/(r^2-1)}$ gegen 0. Durch Induktion kann man sich leicht davon überzeugen, daß die Ableitungen beliebiger Ordnung für $r < 1$ die Form

$$\frac{\partial^n e^{r^2/(r^2-1)}}{\partial \xi^l \partial \eta^m} = \frac{p_{l,m}(x - \xi, y - \eta)}{(r^2 - 1)^{2n}} e^{r^2/(r^2-1)}$$

haben, wobei $p_{l,m}(u, v)$ ein Polynom ist. Nähert sich der Punkt (ξ, η) der Kreislinie $r = 1$, so strebt dieser Ausdruck gegen 0. Der Mittelwertsatz lehrt, daß die Ableitung in jedem Punkt dieser Kreislinie existiert und gleich 0 ist. In diesem Beispiel (187) ist $\omega_1 > 0$ für $r < 1$, also die Konstante c aus (167) gleich 1.

III. Mengenfunktionen. Absolute Stetigkeit. Verallgemeinerung des Integralbegriffs

72. Additive Mengenfunktionen. Die Punktfunktion $f(P)$ sei bezüglich der nicht-negativen, additiven und normalen Funktion $G(\mathfrak{E})$ meßbar. Wir bilden das *unbestimmte Integral*

$$\varphi(\mathfrak{E}) = \int_{\mathfrak{E}} f(P) G(d\mathfrak{E}). \quad (1)$$

Es ist für alle Mengen $\mathfrak{E}$ definiert, die dem abgeschlossenen Körper L_G angehören und auf denen $f(P)$ summierbar ist. Wenn die Funktion $f(P)$ auf $\mathfrak{E}$ summierbar ist, ist sie auch auf jeder meßbaren Teilmenge $\mathfrak{E}'$ der Menge $\mathfrak{E}$ summierbar, und wenn wir $\mathfrak{E}$ in endlich oder abzählbar unendlich viele paarweise elementfremde Mengen $\mathfrak{E}_k$ zerlegen, ist $\varphi(\mathfrak{E})$ gleich der Summe der $\varphi(\mathfrak{E}_k)$ (Volladditivität). Wir untersuchen jetzt die Eigenschaften der volladditiven Funktionen, die nicht als unbestimmtes Integral, sondern irgendwie gegeben sind. Die Funktion $\varphi(\mathfrak{E})$ nehme endliche reelle Werte für die Mengen an, die zu einer Familie C von Mengen aus einem abgeschlossenen Mengenkörper T gehören, der alle abgeschlossenen und alle offenen Mengen enthält. Dabei nehmen wir an: Wenn $\mathfrak{E}$ zu C gehört, gehört auch jede Teilmenge von $\mathfrak{E}$, die in T liegt, zu C . Außerdem setzen wir voraus, daß $\varphi(\mathfrak{E})$ volladditiv ist, d. h., wenn die Menge $\mathfrak{E} \in C$ in endlich oder abzählbar unendlich viele paarweise elementfremde Mengen $\mathfrak{E}_k$ zerlegt ist, wobei alle $\mathfrak{E}_k$ zu T und somit zu C gehören,

$$\mathfrak{E} = \sum_k \mathfrak{E}_k, \quad (2)$$

ist

$$\varphi(\mathfrak{E}) = \sum_k \varphi(\mathfrak{E}_k). \quad (3)$$

Bei unendlich vielen Summanden muß diese Reihe absolut konvergent sein. Ist $\varphi(\mathfrak{E})$ als unbestimmtes Integral im Sinne von (1) gegeben, so ist für den abgeschlossenen Mengenkörper T der Körper L_G zu nehmen, und die Familie C besteht aus denjenigen Mengen $\mathfrak{E}$ aus L_G , auf denen $f(P)$ summierbar ist. Der für das Weitere wichtigste Fall wird sein, daß C aus einer Menge $\mathfrak{E}_0$, die zu B [40] gehört, und aus denjenigen Mengen $\mathfrak{E}$ aus B , die Teilmengen von $\mathfrak{E}_0$ sind, besteht. Dabei ist C selbst ein abgeschlossener Körper.

Ist $\varphi(\mathfrak{E})$ für solche Mengen aus T , die z. B. zu einem abgeschlossenen Intervall A gehören, definiert, so kann man $\varphi(\mathfrak{E})$ für alle Mengen $\mathfrak{E}$ aus T durch

$$\varphi(\mathfrak{E}) = \varphi(\mathfrak{E}A) \quad (4)$$

definieren. Dabei nimmt $\varphi(\mathfrak{E})$ endliche Werte an und ist auf dem ganzen Körper T volladditiv. Wenn wir im folgenden von $\varphi(\mathfrak{E})$ sprechen, werden wir natürlich annehmen, daß $\mathfrak{E}$ zu C gehört. Auf Grund der Additivität gilt $\varphi(\mathfrak{E}) = 0$, wenn $\mathfrak{E}$ die leere Menge ist. Aus der Additivität folgt unmittelbar

$$\varphi(\mathfrak{E}'' - \mathfrak{E}') = \varphi(\mathfrak{E}'') - \varphi(\mathfrak{E}'), \quad (5)$$

wenn $\mathfrak{E}'$ und $\mathfrak{E}''$ zu C gehören und $\mathfrak{E}' \subset \mathfrak{E}''$ ist.

Weiter folgt aus der Volladditivität: Wenn $\mathfrak{E}_n$ ($n=1, 2, \dots$) eine monotone Mengenfolge aus C ist und die Grenzmenge $\mathfrak{E}$ ebenfalls zu C gehört, strebt die Folge der $\varphi(\mathfrak{E}_n)$ gegen $\varphi(\mathfrak{E})$. Bei einer nichtfallenden Mengenfolge ist nämlich

$$\mathfrak{E} = \mathfrak{E}_1 + (\mathfrak{E}_2 - \mathfrak{E}_1) + (\mathfrak{E}_3 - \mathfrak{E}_2) + \dots,$$

und auf Grund der Volladditivität folgt

$$\varphi(\mathfrak{E}) = \varphi(\mathfrak{E}_1) + [\varphi(\mathfrak{E}_2) - \varphi(\mathfrak{E}_1)] + [\varphi(\mathfrak{E}_3) - \varphi(\mathfrak{E}_2)] + \dots,$$

also $\varphi(\mathfrak{E}_n) \rightarrow \varphi(\mathfrak{E})$. Im Fall einer nichtwachsenden Folge $\mathfrak{E}_n$ verläuft der Beweis analog, wobei aber $\mathfrak{E}$ unbedingt zu C gehören muß. Wir bemerken noch, daß eine endliche Linearkombination

$$c_1\varphi_1(\mathfrak{E}) + c_2\varphi_2(\mathfrak{E}) + \dots + c_p\varphi_p(\mathfrak{E})$$

volladditiver Funktionen offenbar wieder eine volladditive Funktion ist.

Wir beweisen jetzt einige grundlegende Sätze dieser Theorie:

Satz 1. Die Werte von $\varphi(\mathfrak{E})$ für alle Mengen $\mathfrak{E}$, die zu einer beliebigen Menge $\mathfrak{E}_1$ aus C gehören, sind dem absoluten Betrag nach beschränkt.

Beweis. Wir beweisen diesen Satz indirekt. Wenn die Behauptung nicht stimmt, existiert ein $\mathfrak{E}_2 \subset \mathfrak{E}_1$ derart, daß $|\varphi(\mathfrak{E}_2)| \geq 2$ und $|\varphi(\mathfrak{E}_1 - \mathfrak{E}_2)| \geq 2$ ist. Dabei müssen wir beachten, daß

$$\varphi(\mathfrak{E}_1) = \varphi(\mathfrak{E}_2) + \varphi(\mathfrak{E}_1 - \mathfrak{E}_2)$$

ist und daß dann, wenn einer der Summanden nicht beschränkt ist, auch der zweite Summand nicht beschränkt ist, da $\varphi(\mathfrak{E}_1)$ eine gegebene Zahl ist. Für $\mathfrak{E}_2$ oder $\mathfrak{E}_1 - \mathfrak{E}_2$ gilt der Satz nicht. Wir können annehmen, daß er für $\mathfrak{E}_2$ nicht gilt, und es existiert ein $\mathfrak{E}_3 \subset \mathfrak{E}_2$ derart, daß $|\varphi(\mathfrak{E}_3)| \geq 3$ und $|\varphi(\mathfrak{E}_2 - \mathfrak{E}_3)| \geq 3$ ist, usw. Wir erhalten $\mathfrak{E}_1 \supset \mathfrak{E}_2 \supset \mathfrak{E}_3 \supset \dots$ und, wenn wir $\mathfrak{E} = \mathfrak{E}_1 \mathfrak{E}_2 \dots$ setzen, nach dem oben Gesagten $\varphi(\mathfrak{E}_n) \rightarrow \varphi(\mathfrak{E})$, was mit dem unbeschränkten Anwachsen des Absolutbetrages von $\varphi(\mathfrak{E}_n)$ unvereinbar ist. Damit ist der Satz bewiesen.

Es sei nun δ eine Zerlegung von $\mathfrak{E}$ in endlich viele $\mathfrak{E}_k$. Wir bilden die Summe

$$t_\delta = \sum_k |\varphi(\mathfrak{E}_k)| \quad (6)$$

und zeigen, daß die Menge der Werte t_δ für beliebige δ beschränkt ist. Es sei $\mathfrak{E}'_\delta$ die Summe derjenigen $\mathfrak{E}_k$, für die $\varphi(\mathfrak{E}_k) \geq 0$ ist, und $\mathfrak{E}''_\delta$ die Summe derjenigen $\mathfrak{E}_k$, die für die $\varphi(\mathfrak{E}_k) < 0$ ist. Wenn wir beachten, daß $\varphi(\mathfrak{E})$ additiv ist, können wir

$$t_\delta = \varphi(\mathfrak{E}'_\delta) - \varphi(\mathfrak{E}''_\delta) \quad (7)$$

schreiben. Wenn wir beachten, daß $\mathfrak{E}'_\delta + \mathfrak{E}''_\delta = \mathfrak{E}$ und folglich $\varphi(\mathfrak{E}) = \varphi(\mathfrak{E}'_\delta) + \varphi(\mathfrak{E}''_\delta)$ ist, so können wir (7) in der Form

$$t_\delta = 2\varphi(\mathfrak{E}'_\delta) - \varphi(\mathfrak{E}) = \varphi(\mathfrak{E}) - 2\varphi(\mathfrak{E}''_\delta) \quad (8)$$

schreiben.

Wir bezeichnen mit $\overline{\varphi}(\mathfrak{E})$ und $\underline{\varphi}(\mathfrak{E})$ die obere bzw. die untere Grenze der Werte $\varphi(e)$ für $e \subset \mathfrak{E}$, wobei die leere Menge auch zu $\mathfrak{E}$ gezählt wird:

$$\overline{\varphi}(\mathfrak{E}) = \sup \varphi(e), \quad \underline{\varphi}(\mathfrak{E}) = \inf \varphi(e) \quad (e \subset \mathfrak{E}). \quad (9)$$

Nach Satz 1 können wir behaupten, daß $\overline{\varphi}(\mathfrak{E})$ und $\underline{\varphi}(\mathfrak{E})$ endlich sind. Aus der ersten Hälfte der Formel (8) folgt, daß t_δ für beliebig gewähltes δ beschränkt ist: $t_\delta \leq 2\overline{\varphi}(\mathfrak{E}) - \varphi(\mathfrak{E})$. Die obere Grenze der Summen t_δ für alle Zerlegungen δ heißt die

Totalvariation von $\varphi(\mathfrak{E})$ auf der Menge $\mathfrak{E}$. Wir werden sie mit $\Phi(\mathfrak{E})$ bezeichnen. Ist δ_n eine Folge von Zerlegungen, so daß t_{δ_n} gegen $\Phi(\mathfrak{E})$ strebt, so folgt aus der ersten Hälfte der Formel (8), daß dann auch $\varphi(\mathfrak{E}'_{\delta_n})$ gegen $\bar{\varphi}(\mathfrak{E})$ strebt, und aus der zweiten Hälfte der Formel (8), die wir in der Gestalt

$$2\varphi(\mathfrak{E}'_{\delta_n}) = \varphi(\mathfrak{E}) - t_{\delta_n}$$

schreiben können, folgt $\varphi(\mathfrak{E}'_{\delta_n}) \rightarrow \underline{\varphi}(\mathfrak{E})$, so daß (8) beim Grenzübergang von δ_n gegen δ die Beziehung

$$\Phi(\mathfrak{E}) = 2\bar{\varphi}(\mathfrak{E}) - \varphi(\mathfrak{E}) = \varphi(\mathfrak{E}) - 2\underline{\varphi}(\mathfrak{E})$$

liefert. Hieraus erhält man

$$\bar{\varphi}(\mathfrak{E}) = \frac{1}{2} [\Phi(\mathfrak{E}) + \varphi(\mathfrak{E})], \quad \underline{\varphi}(\mathfrak{E}) = -\frac{1}{2} [\Phi(\mathfrak{E}) - \varphi(\mathfrak{E})],$$

$$\Phi(\mathfrak{E}) = \bar{\varphi}(\mathfrak{E}) - \underline{\varphi}(\mathfrak{E}), \quad (10)$$

$$\varphi(\mathfrak{E}) = \bar{\varphi}(\mathfrak{E}) + \underline{\varphi}(\mathfrak{E}) = \bar{\varphi}(\mathfrak{E}) - [-\underline{\varphi}(\mathfrak{E})]. \quad (11)$$

Aus der Definition von $\bar{\varphi}(\mathfrak{E})$ und $\underline{\varphi}(\mathfrak{E})$ folgt $\bar{\varphi}(\mathfrak{E}) \geq 0$ und $\underline{\varphi}(\mathfrak{E}) \leq 0$. Die Funktionen $\bar{\varphi}(\mathfrak{E})$ und $-\underline{\varphi}(\mathfrak{E})$ nennt man meistens die *positive* bzw. die *negative Variation* von $\varphi(\mathfrak{E})$ auf $\mathfrak{E}$.

Satz 2. *Die positive Variation, die negative Variation und die Totalvariation sind auf $\mathcal{C}$ volladditive Funktionen.*

Beweis. Es sei eine beliebige Zerlegung von $\mathfrak{E}$ in Teilmengen $\mathfrak{E}_k$ gegeben. Wir bilden die Reihe

$$S = \sum_k \bar{\varphi}(\mathfrak{E}_k)$$

mit nichtnegativen Gliedern und zeigen, daß $S = \bar{\varphi}(\mathfrak{E})$ ist. Zunächst beweisen wir, daß $S < \infty$ ist. Wäre nämlich $S = \infty$, so würde die Summe $\sum \varphi(e_k)$ für entsprechend gewähltes $e_k \subset \mathfrak{E}_k$ einen beliebig großen Wert annehmen. Aber diese Summe ist gleich $\varphi(e)$ mit $e = \sum e_k$ ($e \subset \mathfrak{E}$), und wir erhalten einen Widerspruch zu Satz 1. Es sei nun ein $\varepsilon > 0$ gegeben. Für ein passend gewähltes e_k erhalten wir $\varphi(e) \geq S - \varepsilon$ und folglich $\bar{\varphi}(\mathfrak{E}) \geq S - \varepsilon$; hieraus ergibt sich, da ε beliebig war, $\bar{\varphi}(\mathfrak{E}) \geq S$. Um jetzt die entgegengesetzte Ungleichung zu beweisen, wählen wir $e \subset \mathfrak{E}$ derart, daß $\varphi(e) \geq \bar{\varphi}(\mathfrak{E}) - \varepsilon$ ist. Es sei $e_k = e \cap \mathfrak{E}_k$. Wir haben

$$\varphi(e) = \sum_k \varphi(e_k), \quad \text{d. h.} \quad \bar{\varphi}(\mathfrak{E}) - \varepsilon \leq \sum_k \varphi(e_k);$$

hieraus erhält man erst recht $\bar{\varphi}(\mathfrak{E}) - \varepsilon \leq \sum_k \bar{\varphi}(\mathfrak{E}_k)$, und da ε beliebig ist, finden wir $\bar{\varphi}(\mathfrak{E}) \leq \sum_k \bar{\varphi}(\mathfrak{E}_k)$ und schließlich

$$\bar{\varphi}(\mathfrak{E}) = \sum_k \bar{\varphi}(\mathfrak{E}_k).$$

Analog ist auch die negative Variation volladditiv und folglich auf Grund von (10) auch die Totalvariation. Die Formel (11) besagt, daß *jede volladditive Funktion die Differenz volladditiver nichtnegativer Funktionen ist*. Wenn wir bei der Bildung der Summen (6) eine Zerlegung von $\mathfrak{E}$ in unendlich viele Mengen benutzt hätten, so hätten wir dieselbe obere Grenze bekommen.

Wie man leicht sieht, ist die Funktion $\Phi(\mathfrak{E})$ auf derselben Menge wie $\varphi(\mathfrak{E})$ definiert.

Aber wir können den Körper L_{Φ} bilden, und dann läßt sich $\varphi(\mathfrak{E})$ auf diesen Körper L_{Φ} fortsetzen.

Jeder Punkt (als abgeschlossene Menge) gehört zu T . Ist $\mathfrak{E} \in C$, so gehört jeder Punkt P aus $\mathfrak{E}$ zu C , und wir können von dem Wert $\varphi(P)$ der Funktion $\varphi(\mathfrak{E})$ im Punkt P sprechen. Ist $\varphi(P) \neq 0$, so heißt der Punkt P *Unstetigkeitsstelle* von $\varphi(\mathfrak{E})$. Andernfalls spricht man von einer *Stetigkeitsstelle* von $\varphi(\mathfrak{E})$. Wenn $\varphi(P) > 0$ ist, folgt aus der oben gegebenen Definition $\varphi(P) = \bar{\varphi}(P)$ und $\underline{\varphi}(P) = 0$, und ist $\varphi(P) < 0$, so ist $\bar{\varphi}(P) = 0$ und $\underline{\varphi}(P) = \varphi(P)$. Stetigkeitsstellen von $\varphi(\mathfrak{E})$ sind also auch Stetigkeitsstellen von $\bar{\varphi}(\mathfrak{E})$ und $\underline{\varphi}(\mathfrak{E})$. Weil $\bar{\varphi}(\mathfrak{E})$ und $\underline{\varphi}(\mathfrak{E})$ endlich sind, ist die Anzahl der Unstetigkeitsstellen, die zu $\mathfrak{E}$ gehören und so beschaffen sind, daß $\varphi(P) \geq \alpha$ oder $\varphi(P) \leq -\alpha$ ist (α eine gegebene positive Zahl), endlich, und die Anzahl aller Unstetigkeitsstellen ist endlich oder abzählbar. Es seien P_k diese Punkte. Ist die Menge der Punkte P_k abzählbar, so ist die aus $\varphi(P_k)$ bestehende Reihe absolut konvergent. Wir führen eine neue Mengenfunktion ein, die auf der Familie C durch

$$\varphi_d(\mathfrak{E}) = \sum_{P_k \in \mathfrak{E}} \varphi(P_k) \quad (12)$$

definiert ist, wobei über alle Punkte P_k aus $\mathfrak{E}$ summiert wird. Diese Funktion ist ebenfalls volladditiv und heißt *Sprungfunktion*. Die Differenz

$$\varphi_c(\mathfrak{E}) = \varphi(\mathfrak{E}) - \varphi_d(\mathfrak{E}) \quad (13)$$

ist eine volladditive Funktion ohne Unstetigkeitsstellen.

73. Die singuläre Funktion. Es zeigt sich, daß nicht jede auf der Familie C aus L_G volladditive Funktion $\varphi(\mathfrak{E})$ in Gestalt eines Integrals (1) dargestellt werden kann. Hier und in [78] bis [80] nehmen wir als Körper T den Körper B [40].

Wir beweisen erst später den folgenden grundlegenden Satz, den wir aber jetzt schon anwenden werden.

Satz. Jede auf C volladditive Funktion $\varphi(\mathfrak{E})$ kann für alle Mengen $\mathfrak{E}$, die zu einer beliebigen festen Menge $\mathfrak{E}_0$ aus C gehören, durch die Formel

$$\varphi(\mathfrak{E}) = \varphi(\mathfrak{E}H) + \int_{\mathfrak{E}} f(P) G(d\mathfrak{E}) \quad (14)$$

dargestellt werden, wobei H eine bestimmte Menge aus $\mathfrak{E}_0$ mit $G(H) = 0$ ist und $f(P)$ meßbar und auf $\mathfrak{E}_0$ summierbar ist. Der Summand $\varphi(\mathfrak{E}H)$ heißt *singulärer Teil* von $\varphi(\mathfrak{E})$. Der singuläre Teil ist durch die Werte von $\varphi(\mathfrak{E})$ auf Mengen vom Maß 0 definiert. Der zweite Summand, den wir den absolut stetigen Teil nennen, ist auf jeder Menge vom Maß 0 gleich 0. Wir beweisen jetzt die Eindeutigkeit der Zerlegung in den singulären und den absolut stetigen Teil. Neben (14) möge für die $\mathfrak{E}$ aus C , die zu $\mathfrak{E}_0$ gehören,

$$\varphi(\mathfrak{E}) = \varphi(\mathfrak{E}H_1) + \int_{\mathfrak{E}} f_1(P) G(d\mathfrak{E}),$$

gelten, wobei $G(H_1) = 0$ ist. Aus dieser Formel und aus (14) erhalten wir

$$\varphi(\mathfrak{E}H) - \varphi(\mathfrak{E}H_1) = \int_{\mathfrak{E}} f_1(P) G(d\mathfrak{E}) - \int_{\mathfrak{E}} f(P) G(d\mathfrak{E}).$$

Wir ersetzen $\mathfrak{E}$ durch die Menge $\mathfrak{E}H + \mathfrak{E}H_1$, die zu $\mathfrak{E}_0$ gehört. Wenn wir beachten, daß $G(\mathfrak{E}H + \mathfrak{E}H_1) = 0$ und folglich das Integral über $\mathfrak{E}H + \mathfrak{E}H_1$ gleich 0 ist und daß $(\mathfrak{E}H + \mathfrak{E}H_1)H = \mathfrak{E}H$ und $(\mathfrak{E}H + \mathfrak{E}H_1)H_1 = \mathfrak{E}H_1$ ist, dann erhalten wir $\varphi(\mathfrak{E}H) = \varphi(\mathfrak{E}H_1)$. Hieraus folgt, daß die absolut stetigen Teile identisch sein müssen, d. h.

$$\int_{\mathfrak{E}} f(P) G(d\mathfrak{E}) = \int_{\mathfrak{E}} f_1(P) G(d\mathfrak{E}). \quad (15)$$

Die Formel (14) bleibt sinnvoll, wenn wir die Familie C auf das Produkt von L_G und L_Φ fortsetzen.

Bei dem Beweis des Satzes werden wir von einer beliebigen, aber festen Menge $\mathfrak{E}_0$, die zu C gehört, ausgehen, und wir werden annehmen, daß alle $\mathfrak{E}$ in $\mathfrak{E}_0$ liegen, wie es im Satz formuliert ist. Bei der Zerlegung von $\varphi(\mathfrak{E})$ in den singulären und den absolut stetigen Teil gingen wir von einer Menge $\mathfrak{E}_0$ aus und nahmen an, daß alle $\mathfrak{E}$ zu $\mathfrak{E}_0$ gehören. Dabei erhielten wir die Eindeutigkeit dieser Zerlegung. Wären wir von einer Menge $\mathfrak{E}'_0$ ausgegangen, die von $\mathfrak{E}_0$ verschieden ist, und zu C gehört, so hätten wir für alle Mengen, die gleichzeitig zu $\mathfrak{E}_0$ und $\mathfrak{E}'_0$ gehören, dieselbe Zerlegung erhalten, die wir früher mit Hilfe der Menge $\mathfrak{E}_0$ erhielten. Andernfalls würden wir nämlich für die Mengen, die zu dem Produkt $\mathfrak{E}'_0 = \mathfrak{E}_0 \mathfrak{E}'_0$ gehören, das ebenfalls in der Familie C liegt, zwei verschiedene Zerlegungen von $\varphi(\mathfrak{E})$ erhalten, was aber nicht möglich ist, wie wir bereits oben sahen.

In diesem Sinne können wir sagen, daß die Zerlegung von $\varphi(\mathfrak{E})$ in den singulären und den absolut stetigen Teil auf der ganzen Familie C eindeutig ist.

Wir zeigen jetzt, daß die in Formel (14) unter dem Integralzeichen stehende Funktion $f(P)$ vollständig bestimmt ist, wobei wir wie bisher Funktionen, die bezüglich $G(\mathfrak{E})$ äquivalent sind, identifizieren. Wir müssen zeigen: Wenn für alle zu $\mathfrak{E}_0$ gehörenden $\mathfrak{E}$ die Beziehung (15) gilt, so ist die Differenz $\psi(P) = f_1(P) - f(P)$ auf $\mathfrak{E}_0$ äquivalent 0.

Es sei $\mathfrak{E}_0^+$ derjenige Teil von $\mathfrak{E}_0$, auf dem $\psi(P) \geq 0$ ist, und $\mathfrak{E}_0^- = \mathfrak{E}_0 - \mathfrak{E}_0^+$. Die Mengen $\mathfrak{E}_0^+$ und $\mathfrak{E}_0^-$ gehören zu C , und wir erhalten, wenn wir in (15) $\mathfrak{E}$ durch $\mathfrak{E}_0^+$ und durch $\mathfrak{E}_0^-$ ersetzen,

$$\int_{\mathfrak{E}_0^+} \psi(P) G(d\mathfrak{E}) = \int_{\mathfrak{E}_0^-} \psi(P) G(d\mathfrak{E}) = 0 ;$$

hieraus folgt, daß $\psi(P)$ auf $\mathfrak{E}_0^+$ und $\mathfrak{E}_0^-$ und damit auch auf $\mathfrak{E}_0$ äquivalent 0 ist. Bilden wir die Funktion $f(P)$ für die beiden Mengen $\mathfrak{E}_0$ bzw. $\mathfrak{E}'_0$ aus C , so sind, wie auch oben, diese beiden Funktionen auf $\mathfrak{E}''_0 = \mathfrak{E}_0 \mathfrak{E}'_0$ äquivalent. In diesem Sinne kann man von der *eindeutigen Bestimmtheit der Funktion $f(P)$* sprechen. Gehören z. B. alle endlichen Intervalle zu C , so ist $f(P)$, wenn wir die vorigen Überlegungen auf die wachsenden Intervalle $-n \leq x \leq +n$, $-n \leq y \leq +n$ ($n = 1, 2, \dots$) anwenden, in der ganzen Ebene eindeutig bestimmt. Die Funktion $f(P)$ nennt man die *Ableitung von $\varphi(\mathfrak{E})$ bezüglich $G(\mathfrak{E})$* . Ist $k_P^{(\varepsilon)}$ der Kreis (oder die Kugel) mit dem Mittelpunkt P und dem Radius ε , so kann man zeigen, daß für alle P (eventuell mit Ausnahme einer Menge vom Maß 0 bezüglich $G(\mathfrak{E})$) der Quotient $\varphi(k_P^{(\varepsilon)}) : G(k_P^{(\varepsilon)})$ für $\varepsilon \rightarrow 0$ gegen eine Funktion strebt, die bezüglich $G(\mathfrak{E})$ äquivalent $f(P)$ ist. Dabei wird vorausgesetzt, daß $\varphi(\mathfrak{E})$ auf dem Kreis $k_P^{(\varepsilon)}$ für hinreichend kleines ε definiert ist. Im weiteren werden wir diese Behauptung jedoch nicht benutzen, und aus diesem Grunde beweisen wir sie auch nicht.

Definition. Eine Funktion $\varphi(\mathfrak{E})$ heißt *absolut stetig bezüglich $G(\mathfrak{E})$* , wenn für ein beliebiges festes $\mathfrak{E}_0$ aus C und ein beliebig gegebenes positives ε ein solches $\eta > 0$ existiert, daß $|\varphi(e)| \leq \varepsilon$ für $e \in \mathfrak{E}_0$ ist, sobald $|G(e)| \leq \eta$ gilt.

Ist $\varphi(\mathfrak{E})$ bezüglich $G(\mathfrak{E})$ absolut stetig, so ist $\varphi(\mathfrak{E}) = 0$ für $\mathfrak{E} \in C$ und $G(\mathfrak{E}) = 0$. Der zweite Summand der Formel (14) ist, wie wir wissen, eine auf C absolut stetige Funktion. Ist umgekehrt bekannt, daß $\varphi(\mathfrak{E})$ absolut stetig ist, so ist $\varphi(\mathfrak{E}H) = 0$; denn es ist $G(\mathfrak{E}H) = 0$, und $\varphi(\mathfrak{E})$ läßt sich durch die Formel

$$\varphi(\mathfrak{E}) = \int_{\mathfrak{E}} f(P) G(d\mathfrak{E}) \quad (16)$$

darstellen, d. h., der singuläre Teil ist nicht vorhanden. Aus diesen Überlegungen erhalten wir als Folgerung des Satzes:

Folgerung. Ist $\varphi(\mathfrak{E}) = 0$ für $G(\mathfrak{E}) = 0$, so ist $\varphi(\mathfrak{E})$ durch die Formel (16) darstellbar und bezüglich $G(\mathfrak{E})$ auf jeder Menge $\mathfrak{E}_0$ aus C eine absolut stetige Funktion.

Bemerkung. Ist $G(\mathfrak{E})$ nicht stetig, so ist auch das durch (16) definierte $\varphi(\mathfrak{E})$ im allgemeinen nicht stetig. Ist etwa $G(P_0) = a \neq 0$, so ist $\varphi(P_0) = a f(P_0)$. Die Funktion $\varphi(\mathfrak{E})$ ist aber bezüglich $G(\mathfrak{E})$ in dem obigen Sinne absolut stetig.

Ist $G(\mathfrak{E})$ stetig, so ist auch das durch (16) definierte $\varphi(\mathfrak{E})$ stetig. Wenn $G(\Delta)$ der Flächeninhalt des Intervalls Δ und L_G der Körper L der im Lebesgueschen Sinne meßbaren Mengen ist, hat (14) die Gestalt

$$\varphi(\mathfrak{E}) = \varphi(\mathfrak{E}H) + \iint_{\mathfrak{E}} f(x, y) \, dx dy ;$$

dabei hat H das Lebesguesche Maß 0. Die Formel (16) erhält also die Gestalt

$$\varphi(\mathfrak{E}) = \iint_{\mathfrak{E}} f(x, y) \, dx dy ;$$

in diesem Fall ist $\varphi(\mathfrak{E})$ in jedem Punkt stetig.

Die obere Grenze der Werte von $\varphi(e)$ für Mengen e , die zu $\mathfrak{E}$ gehören, erhalten wir in dem Fall der Formel (16), wenn wir $f(P)$ über die Menge integrieren, auf der $f(P) \geq 0$ ist, und die untere Grenze erhalten wir, wenn wir über die Menge integrieren, auf der $f(P) \leq 0$ ist. Wir erhalten auf diese Weise folgende Formeln für die positive Variation, die negative Variation und die Totalvariation der durch (16) definierten Funktion $\varphi(\mathfrak{E})$:

$$\bar{\varphi}(\mathfrak{E}) = \int_{\mathfrak{E}} f^+(P) G(d\mathfrak{E}), \quad \underline{\varphi}(\mathfrak{E}) = \int_{\mathfrak{E}} f^-(P) G(d\mathfrak{E}), \quad \Phi(\mathfrak{E}) = \int_{\mathfrak{E}} |f(P)| G(d\mathfrak{E}). \quad (17)$$

Wenn wir aus der Funktion $\varphi(\mathfrak{E})$ die Sprungfunktion $\varphi_d(\mathfrak{E})$ aussondern und die übrigebleibende Funktion nach Formel (14) zerlegen, erhalten wir für $\varphi(\mathfrak{E})$ eine Zerlegung in drei Summanden:

$$\varphi(\mathfrak{E}) = \varphi_d(\mathfrak{E}) + \varphi_c(\mathfrak{E}H) + \int_{\mathfrak{E}} f_c(P) G(d\mathfrak{E}). \quad (18)$$

74. Der Fall einer Veränderlichen. Es sei $\mathfrak{E}$ jetzt z. B. das endliche Intervall $[a, b]$. Statt $G(\mathfrak{E})$ führen wir eine entsprechende nichtfallende Punktfunktion $g(x)$ ein. An Stelle der Mengenfunktion nehmen wir die Punktfunktion $\omega(x) = \varphi([a, x])$, und die Formel (14) lautet nun

$$\omega(x) = \varphi([a, x] H) + \int_{[a, x]} f(x) \, dg(x); \quad (19)$$

für Lebesguesche Integrale erhält man

$$\omega(x) = \varphi([a, x] H) + \int_a^x f(x) \, dx. \quad (20)$$

Ist $\varphi(\mathfrak{E})$ absolut stetig, so erhalten wir

$$\omega(x) = \int_{[a, x]} f(x) \, dg(x) \quad \text{bzw.} \quad \omega(x) = \int_a^x f(x) \, dx. \quad (21)$$

Wir betrachten nun den Fall des Lebesgueschen Integrals näher. Beim Übergang von der Intervallfunktion zur Punktfunktion können wir zu der Punktfunktion konstante Summanden addieren und

$$\omega(x) = \int_a^x f(x) \, dx + \omega(a) \quad (22)$$

schreiben, wobei $f(x)$ eine meßbare und in $[a, b]$ summierbare Funktion ist. Wenn wir beachten, daß das Integral von $f(x)$ absolut stetig ist, ergibt sich für die Funktion (22) folgende Eigenschaft: Für beliebig gegebenes positives ε existiert ein $\eta > 0$ derart, daß

$$\left| \sum_{k=1}^n [\omega(b_k) - \omega(a_k)] \right| \leq \varepsilon \quad (23)$$

gilt, wenn (a_k, b_k) ($k=1, 2, \dots, n$) sich nicht überschneidende Intervalle sind mit

$$\sum_{k=1}^n (b_k - a_k) \leq \eta. \quad (24)$$

Wir werden von einer Punktfunktion ausgehen und die Punktfunktion $\omega(x)$, die in dem Intervall $[a, b]$ definiert ist, in diesem Intervall *absolut stetig* nennen, wenn sie die eben erwähnte Eigenschaft hat. Eine absolut stetige Funktion ist schlechthin stetig, da man im besonderen $n=1$ setzen kann. Im folgenden werden wir sehen, daß monotone stetige Punktfunktionen existieren, die nicht absolut stetig sind. Aus der obigen Eigenschaft folgt: Zu beliebig gegebenem $\varepsilon > 0$ existiert ein solches positives η , daß

$$\sum_{k=1}^n |\omega(b_k) - \omega(a_k)| \leq \varepsilon \quad (25)$$

gilt, wenn (24) erfüllt ist.

Besitzt nämlich $\omega(x)$ die oben angegebene Eigenschaft (23), d. h., ist $\omega(x)$ absolut stetig, so gibt es zu gegebenem $\varepsilon > 0$ ein solches $\eta > 0$, daß wir bei Gültigkeit von (24)

$$\left| \sum_{k=1}^n [\omega(b_k) - \omega(a_k)] \right| \leq \frac{\varepsilon}{2} \quad (26)$$

erhalten.

Jedes System von Intervallen (a_k, b_k) , das (24) befriedigt, können wir in zwei Klassen zerlegen. Zur Klasse I rechnen wir diejenigen Intervalle, für die $\omega(b_k) - \omega(a_k) \geq 0$ ist, und zur Klasse II diejenigen, für die $\omega(b_k) - \omega(a_k) < 0$ ist. Nach (26) erhalten wir

$$\begin{aligned} \sum_I |\omega(b_k) - \omega(a_k)| &= \sum_I [\omega(b_k) - \omega(a_k)] \leq \frac{\varepsilon}{2}, \\ \sum_{II} |\omega(b_k) - \omega(a_k)| &= \left| \sum_{II} [\omega(b_k) - \omega(a_k)] \right| \leq \frac{\varepsilon}{2}; \end{aligned}$$

hieraus folgt (25). Wenn wir beachten, daß die Summanden von (25) nicht negativ sind und n beliebig ist, erhalten wir für absolut stetige Funktionen die folgende Eigenschaft: Jedem gegebenen positiven ε entspricht ein positives η derart, daß

$$\sum_k |\omega(b_k) - \omega(a_k)| \leq \varepsilon \quad (27)$$

gilt, wenn die paarweise elementfremden Intervalle (a_k, b_k) , die eine endliche oder abzählbare Menge bilden, der Bedingung

$$\sum_k (b_k - a_k) \leq \eta \quad (28)$$

genügen.

Ist umgekehrt diese Bedingung erfüllt, so ist auch die ursprüngliche Bedingung (23) erfüllt, und $\omega(x)$ ist absolut stetig.

Satz 1. *Die Summe, die Differenz und das Produkt zweier absolut stetiger Funktionen sind absolut stetige Funktionen. Der Quotient $\frac{\omega_1(x)}{\omega_2(x)}$ zweier absolut stetiger Funktionen ist ebenfalls eine absolut stetige Funktion, wenn $\omega_2(x)$ nicht 0 wird.*

Beweis. Wir beweisen nur die absolute Stetigkeit des Produktes $\omega_1(x) \omega_2(x)$. Die Funktionen $\omega_1(x)$ und $\omega_2(x)$ sind in $[a, b]$ beschränkt, etwa $|\omega_1(x)| \leq l_1$ und $|\omega_2(x)| \leq l_2$. Es gilt

$$|\omega_1(b_k) \omega_2(b_k) - \omega_1(a_k) \omega_2(a_k)| \leq |\omega_2(b_k)| |\omega_1(b_k) - \omega_1(a_k)| + |\omega_1(a_k)| |\omega_2(b_k) - \omega_2(a_k)| \\ \leq l_2 |\omega_1(b_k) - \omega_1(a_k)| + l_1 |\omega_2(b_k) - \omega_2(a_k)|.$$

Wenn wir über k summieren und die absolute Stetigkeit von $\omega_1(x)$ und $\omega_2(x)$ beachten, ist die Eigenschaft (25) für das Produkt $\omega_1(x) \omega_2(x)$ bewiesen.

Satz 2. *Eine absolut stetige Funktion $\omega(x)$ ist eine Funktion von beschränkter Variation, und ihre Totalvariation $v(x)$ ist ebenfalls eine absolut stetige Funktion.*

Beweis. Es sei η_0 eine positive Zahl, für welche die Ungleichung

$$\sum_k |\omega(b_k) - \omega(a_k)| \leq 1 \quad (29)$$

gilt, wenn die Bedingung (23) für $\eta = \eta_0$ erfüllt ist. Wir zerlegen $[a, b]$ durch die festen Punkte

$$a = c_0 < c_1 < c_2 < \dots < c_{N-1} < c_N = b$$

so in Teilintervalle, daß $c_k - c_{k-1} \leq \eta_0$ ($k = 1, 2, \dots, N$) ist. Bei beliebiger Zerlegung des Intervalls gilt für $[c_{k-1}, c_k]$ die Ungleichung (29); die Summen t_δ und infolgedessen auch die Totalvariation von $\omega(x)$ sind in jedem der Intervalle $[c_{k-1}, c_k]$ nicht größer als 1, also im ganzen Intervall $[a, b]$ nicht größer als N . Die Zahl η möge dem ε entsprechen, so daß (25) gilt, wenn (24) erfüllt ist. Wir wollen jedes der Intervalle $[a_k, b_k]$, die in (24) auftreten, in Teilintervalle zerlegen. Die Summe der Längen der erhaltenen Intervalle genügt wie früher der Bedingung (24), und die den Teilintervallen zugeordnete Summe (25) ist wie früher höchstens gleich ε . Die obere Grenze der Summe derjenigen Summanden, die den Teilintervallen aus $[a_k, b_k]$ entsprechen, liefert offenbar $v(b_k) - v(a_k)$. Somit gilt, wenn die Bedingung (24) erfüllt ist,

$$\sum_{k=1}^n [v(b_k) - v(a_k)] \leq \varepsilon;$$

hieraus folgt, daß $v(x)$ absolut stetig ist. Damit ist der Satz bewiesen.

Wenn wir die Funktionen

$$\omega_1(x) = \frac{1}{2} [v(x) + \omega(x)], \quad \omega_2(x) = \frac{1}{2} [v(x) - \omega(x)] \quad (30)$$

bilden, die nichtfallend [8] und absolut stetig sind, können wir $\omega(x)$ nach Satz 1 als Differenz zweier nichtabnehmender absolut stetiger Funktionen darstellen:

$$\omega(x) = \omega_1(x) - \omega_2(x). \quad (31)$$

Wie wir bereits sagten, liefert das unbestimmte Integral (22) einer summierbaren Funktion $f(x)$ eine absolut stetige Punktfunktion $\omega(x)$ im Sinne der obigen Definition (23), (24) oder (23), (25). Wir beweisen jetzt die Umkehrung des Satzes.

Satz 3. Jede absolut stetige Funktion $\omega(x)$ ist durch ein unbestimmtes Integral darstellbar:

$$\omega(x) = \int_a^x f(x) \, dx + \omega(a). \quad (32)$$

Beweis. Wenn wir von der Funktion $\omega_1(x)$ ausgehen und $\omega_1(x) = \omega_1(a)$ für $x < a$ und $\omega_1(x) = \omega_1(b)$ für $x > b$ setzen, können wir dem Intervall $\Delta[\alpha, \beta]$ die nicht-negative Zahl $\varphi_1(\Delta) = \omega_1(\beta) - \omega_1(\alpha)$ zuordnen, wobei wegen der Stetigkeit von $\omega_1(x)$ unwichtig ist, ob das Intervall Δ abgeschlossen oder offen ist. Ist eine lineare Menge Lebesgue-meßbar, so existiert eine solche offene Menge $O \supset \mathfrak{E}$, daß die Menge $O - \mathfrak{E}$ durch endlich oder abzählbar unendlich viele Intervalle $[a_k, b_k]$ überdeckt werden kann, wobei die Summe der Längen dieser Intervalle beliebig klein ist [35]. Da die Funktion $\omega(x)$ in $[a, b]$ absolut stetig ist und ihre Fortsetzung außerhalb $[a, b]$ konstant ist, können wir diese Überdeckung so durchführen, daß die Summe der nichtnegativen Summanden $\omega_1(b_k) - \omega_1(a_k)$ für die überdeckenden Intervalle beliebig klein ist, d. h., wenn $\mathfrak{E}$ Lebesgue-meßbar ist, so ist $\mathfrak{E}$ auch bezüglich $\omega_1(x)$ meßbar. Somit können wir $\varphi_1(\Delta)$ auf alle zu $[a, b]$ gehörenden Mengen aus L fortsetzen, ohne die Volladditivität zu verletzen. Aus diesen Überlegungen folgt sofort: Ist das Lebesguesche Maß $m(\mathfrak{E})$ gleich 0, so ist auch $\varphi_1(\mathfrak{E}) = 0$, und es gilt

$$\varphi_1(\mathfrak{E}) = \int_{\mathfrak{E}} f_1(x) \, dx. \quad (33)$$

Ganz analog erhalten wir

$$\varphi_2(\mathfrak{E}) = \int_{\mathfrak{E}} f_2(x) \, dx, \quad (34)$$

wenn wir $\varphi_2(\Delta) = \omega_2(\beta) - \omega_2(\alpha)$ setzen. Dabei sind $f_1(x)$ und $f_2(x)$ in $[a, b]$ summierbar, und es ist

$$\varphi(\mathfrak{E}) = \varphi_1(\mathfrak{E}) - \varphi_2(\mathfrak{E}) = \int_{\mathfrak{E}} [f_1(x) - f_2(x)] \, dx = \int_{\mathfrak{E}} f(x) \, dx. \quad (35)$$

Wenn wir für $\mathfrak{E}$ das Intervall $[a, x]$ nehmen, so kommen wir zu der Formel (22).

Man kann beweisen, daß die Funktion $f(x)$ in (22) bis auf einen Summanden, der fast überall gleich 0 ist, eindeutig bestimmt ist. Hätten wir nämlich neben (22) für $\omega(x)$ eine entsprechende Formel mit dem Integranden $g(x)$, so wäre das Integral der Differenz $f(x) - g(x)$ für jedes in $[a, b]$ liegende Intervall gleich 0, und nach der Eigenschaft 11 aus [52] können wir behaupten, daß die erwähnte Differenz äquivalent 0 ist. Die Funktion $f(x)$ in Formel (22) heißt *Ableitung* von $\omega(x)$, und man bezeichnet sie mit dem üblichen Symbol $f(x) = \omega'(x)$. Man kann zeigen, was hier jedoch nicht geschehen soll, daß für alle x aus $[a, b]$, eventuell mit Ausnahme einer Menge von Werten mit dem Lebesgueschen Maß 0, der Grenzwert

$$\lim_{h \rightarrow 0} \frac{\omega(x+h) - \omega(x)}{h} = F(x)$$

existiert, wobei $F(x)$ äquivalent $f(x)$ ist. Ist $f(x)$ eine in $[a, b]$ stetige Funktion, so existiert für alle x aus $[a, b]$ die gewöhnliche Ableitung $\omega'(x) = f(x)$ des Integrals bezüglich der oberen Grenze. Ist $f'(x)$ nicht nur stetig, sondern in $[a, b]$ auch absolut stetig, so gilt

$$\omega'(x) = f(x) = \int_a^x h(x) \, dx + c,$$

wobei $h(x)$ summierbar ist. Es gilt $f'(x) = h(x)$, und $h(x)$ heißt *Ableitung zweiter Ordnung* von $\omega(x)$; man bezeichnet sie mit dem üblichen Symbol $h(x) = \omega''(x)$.

Ganz analog kann $\omega(x)$ absolut stetige Ableitungen bis zur k -ten Ordnung haben und besitzt damit eine summierbare Ableitung $(k+1)$ -ter Ordnung; $\omega(x)$ läßt sich in der Form

$$\omega(x) = \int_a^x dx \int_a^x dx \dots \int_a^x \omega^{(k+1)}(x) dx + \omega(a) + \frac{\omega'(a)}{1!} (x-a) + \dots + \frac{\omega^{(k)}(a)}{k!} (x-a)^k$$

darstellen.

Die hier dargelegte Theorie kann leicht auf den Fall übertragen werden, daß $\omega(x)$ bezüglich einer nichtfallenden Funktion $g(x)$, die als stetig angenommen werden kann, absolut stetig ist, d. h., daß es zu jedem gegebenen $\varepsilon > 0$ ein $\eta > 0$ gibt derart, daß die Bedingung

$$\left| \sum_{k=1}^n [\omega(b_k) - \omega(a_k)] \right| \leq \varepsilon \quad (36)$$

erfüllt ist, wenn die (a_k, b_k) elementfremde Intervalle sind, für welche

$$\sum_{k=1}^n [g(b_k) - g(a_k)] \leq \eta \quad (37)$$

gilt.

Genauso wie oben können wir statt (36) auch (27) schreiben, und die Funktion $\omega(x)$ ist in $[a, b]$ stetig. Statt (32) erhalten wir

$$\omega(x) = \int_a^x f(x) dg(x) + \omega(a). \quad (38)$$

75. Absolut stetige Mengenfunktionen. Wir behandeln jetzt den allgemeinen Fall von Mengen in der Ebene und untersuchen eingehend die durch (16) gegebene Transformation, wobei wir annehmen, daß $f(P)$ auf der ganzen Ebene nichtnegativ und summierbar ist. Ist $f(P)$ auf einer meßbaren Menge $\mathfrak{E}_0$ definiert und summierbar, so erhalten wir, wenn wir $f(P)$ außerhalb von $\mathfrak{E}_0$ durch 0 fortsetzen, eine Funktion, die in der ganzen Ebene summierbar ist. Die Formel (16) definiert eine auf dem Körper L_G volladditive Funktion $\varphi(\mathfrak{E})$. Diese Funktion ist für alle halboffenen Intervalle definiert, und wir können diese Intervallfunktion $\varphi(\Delta)$ auf den Körper L_ϕ fortsetzen, wie wir das früher bezüglich $G(\Delta)$ getan haben.

Satz 1. *Jede Menge $\mathfrak{E}$ aus L_G gehört zu L_ϕ , und die Formel (16) liefert das Maß $\varphi(\mathfrak{E})$ dieser Menge, das wir bei der angegebenen Fortsetzung von $\varphi(\Delta)$ erhalten.*

Jede offene Menge O ist Summe abzählbar vieler paarweise punktfremder halboffener Intervalle Δ_k ($k=1, 2, \dots$). Wenn wir beide Seiten der Formel

$$\varphi(\Delta_k) = \int_{\Delta_k} f(P) G(d\mathfrak{E})$$

über k summieren, erhalten wir links das Maß $\varphi(O)$, weil das Maß additiv ist, und rechts das Integral über O , weil das Integral additiv ist, d. h., die Formel (16) liefert das Maß $\varphi(O)$ jeder offenen Menge. Wenn wir beachten, daß jede abgeschlossene Menge F die Differenz der ganzen Ebene $\mathfrak{E}'$ (einer offenen Menge) und einer offenen Menge O ist (nämlich der Differenz $O = \mathfrak{E}' - F$), und wenn wir die Formeln

$$\varphi(\mathfrak{E}') = \int_{\mathfrak{E}'} f(P) G(d\mathfrak{E}), \quad \varphi(O) = \int_O f(P) G(d\mathfrak{E})$$

voneinander subtrahieren, erhalten wir, wie bereits oben, daß die Formel (16) das Maß $\varphi(F)$ jeder abgeschlossenen Menge F liefert. Es sei nun $\mathfrak{E}$ eine Menge aus L_G

und ε_n eine Folge positiver Zahlen, die gegen 0 strebt. Wir wissen, daß Folgen abgeschlossener bzw. offener Mengen F_n bzw. O_n existieren derart, daß $F_n \subset \mathcal{E} \subset O_n$ und $G(O_n - F_n) = G(O_n) - G(F_n) \leq \varepsilon_n$ ist. Da das Integral (16) absolut stetig ist, gilt $\varphi(O_n - F_n) \rightarrow 0$; folglich gehört $\mathcal{E}$ zu L_φ [38]. Dabei ist das Maß von $\mathcal{E}$ in L_φ durch den Grenzwert von $\varphi(F_n)$ oder von $\varphi(O_n)$ gegeben, d. h. durch den Grenzwert der Integrale

$$\int_{F_n} f(P) G(d\mathcal{E}),$$

wobei $F_n \subset \mathcal{E}$ und $G(\mathcal{E} - F_n) \rightarrow 0$ gilt, und auf Grund der absoluten Stetigkeit des Integrals ist dieser Grenzwert gleich dem Integral über $\mathcal{E}$, d. h., das Maß von $\mathcal{E}$ in L_φ wird durch das Integral (16) ausgedrückt. Damit ist der Satz bewiesen.

Der Satz gilt auch unter der Voraussetzung, daß die nichtnegative Funktion $f(P)$ nur auf jeder beschränkten Menge summierbar ist; der Beweis bleibt im Prinzip derselbe. Im folgenden Satz präzisieren wir die Struktur von L_φ .

Satz 2. Eine Menge $\mathcal{E}$ gehört dann und nur dann zu L_φ , wenn $\mathcal{E}$ in Form einer Summe

$$\mathcal{E} = \mathcal{E}^{(1)} + \mathcal{E}^{(2)} \quad (39)$$

dargestellt werden kann, wobei $\mathcal{E}^{(1)}$ zu L_G gehört und in den Punkten $P \in \mathcal{E}^{(2)}$ die Beziehung $f(P) = 0$ gilt.

Beweis. Wir beweisen zunächst die Notwendigkeit. Es gehöre $\mathcal{E}$ zu L_φ . Wir führen die Mengen

$$\begin{aligned} \mathcal{E}_0 &= \mathcal{E} [f(P) = 0], \quad \mathcal{E}_1 = \mathcal{E} [f(P) > 1], \\ \mathcal{E}_n &= \mathcal{E} \left[\frac{1}{n} < f(P) \leq \frac{1}{n-1} \right] \quad (n = 2, 3, \dots) \end{aligned} \quad (40)$$

ein und bezeichnen mit $\mathcal{E}'$ die Menge der Punkte, für die $f(P)$ nicht definiert oder gleich ∞ ist. Die Menge $\mathcal{E}'$ ist bezüglich $G(\mathcal{E})$ meßbar, und es ist $G(\mathcal{E}') = 0$. Dasselbe gilt auch für jede Teilmenge von $\mathcal{E}'$. Die Funktion $f(P)$, die bezüglich $G(\mathcal{E})$ meßbar ist, ist dann auch bezüglich $\varphi(\mathcal{E})$ meßbar, und alle Mengen (40) gehören ebenso wie $\mathcal{E}'$ zu L_φ . Wir definieren weiter die Mengen

$$\mathcal{E}^{(1)} = \mathcal{E}\mathcal{E}' + \sum_{n=1}^{\infty} \mathcal{E}\mathcal{E}_n, \quad \mathcal{E}^{(2)} = \mathcal{E}\mathcal{E}_0. \quad (41)$$

Dafür gilt die Formel (39). In den Punkten von $\mathcal{E}^{(2)}$ wird $f(P)$ gleich 0, und wir müssen beweisen, daß $\mathcal{E}^{(1)}$ zu L_G gehört. Die Menge $\mathcal{E}\mathcal{E}'$ hat bezüglich $G(\mathcal{E})$ das Maß 0; also brauchen wir nur zu zeigen, daß die Mengen $\mathcal{E}\mathcal{E}_n$ bezüglich $G(\mathcal{E})$ meßbar sind. Weil die $\mathcal{E}\mathcal{E}_n$ bezüglich $\varphi(\mathcal{E})$ meßbar sind, existieren abgeschlossene Mengen F_n und offene Mengen O_n derart, daß

$$F_n \subset \mathcal{E}\mathcal{E}_n \subset O_n \quad \text{und} \quad \varphi(O_n - F_n) \leq \frac{\varepsilon}{n} \quad (42)$$

gilt, wo ε eine gegebene positive Zahl ist. Wir bilden für $n = 2, 3, \dots$ und für $n = 1$ die Mengen

$$D_n = \mathcal{E}_n(O_n - F_n) = \mathcal{E}_n O_n - \mathcal{E}_n F_n, \quad (43)$$

d. h., D_n ist die Menge derjenigen Punkte von $O_n - F_n$, in denen $\frac{1}{n} < f(P) \leq \frac{1}{n-1}$ bzw. $f(P) > 1$ (für $n = 1$) ist. Weil $O_n - F_n$ zu L_G gehört und $f(P)$ bezüglich $G(\mathcal{E})$ meß-

bar ist, liegen die Mengen D_n in L_G . Weiter folgt aus $F_n \subset \mathcal{E}\mathcal{E}_n$, daß $F_n \subset \mathcal{E}_n$ und $\mathcal{E}_n F_n = F_n$ ist. Aus $\mathcal{E}\mathcal{E}_n \subset O_n$ folgt $\mathcal{E}\mathcal{E}_n \subset O_n \mathcal{E}_n$ und nach (43) $\mathcal{E}\mathcal{E}_n - F_n \subset D_n$; daher ist $|\mathcal{E}\mathcal{E}_n - F_n|_G \leq |D_n|_G$. Da D_n zu L_G gehört, können wir auch $|D_n|_G = G(D_n)$ schreiben, d. h., es ist

$$|\mathcal{E}\mathcal{E}_n - F_n|_G \leq G(D_n). \quad (44)$$

Weiter gilt wegen (43) und $\mathcal{E}_n F_n = F_n$

$$D_n \subset O_n - F_n,$$

und wenn wir (42) benutzen, erhalten wir

$$\int_{D_n} f(P) G(d\mathcal{E}) = \varphi(D_n) \leq \varphi(O_n - F_n) \leq \frac{\varepsilon}{n}. \quad (45)$$

Wegen (43) gehört die Menge D_n zu $\mathcal{E}_n$, und auf $\mathcal{E}_n$ ist $f(P) > \frac{1}{n}$. Somit erhalten wir die Abschätzung

$$\int_{D_n} f(P) G(d\mathcal{E}) \geq \frac{1}{n} G(D_n),$$

und die Ungleichung (45) reduziert sich auf

$$\frac{G(D_n)}{n} < \frac{\varepsilon}{n},$$

d. h. $G(D_n) < \varepsilon$. Auf Grund von (44) erhalten wir $|\mathcal{E}\mathcal{E}_n - F_n|_G < \varepsilon$, und da ε beliebig war, sehen wir hieraus, daß $\mathcal{E}\mathcal{E}_n$ zu L_G gehört. Damit ist die Notwendigkeit der Bedingung (39) gezeigt.

Wir beweisen jetzt, daß sie auch hinreichend ist. Es sei die Formel (39) gegeben, wobei $\mathcal{E}^{(1)} \in L_G$ gilt und $f(P) = 0$ in den Punkten von $\mathcal{E}^{(2)}$ ist. Wir müssen zeigen, daß $\mathcal{E}$ zu L_φ gehört. Die Menge $\mathcal{E}^{(1)}$ ist nach Satz 1 in L_φ enthalten. Es ist jetzt noch zu zeigen, daß dasselbe auch für $\mathcal{E}^{(2)}$ gilt. Die Menge H aller Punkte, in denen $f(P) = 0$ ist, ist bezüglich $G(\mathcal{E})$ meßbar; daher gilt $H \in L_\varphi$ und wegen (16) $\varphi(H) = 0$. Auf Grund von $\mathcal{E}^{(2)} \subset H$ ist $\mathcal{E}^{(2)}$ bezüglich $\varphi(\mathcal{E})$ meßbar und hat das Maß 0. Damit ist der Satz bewiesen.

Wenn wir noch beachten, daß $\mathcal{E}^{(2)} \subset H$ und folglich $\varphi(\mathcal{E}^{(2)}) = 0$ ist, können wir behaupten, daß $\varphi(\mathcal{E}) = \varphi(\mathcal{E}^{(1)})$ ist, d. h., zur Berechnung von $\varphi(\mathcal{E})$ müssen wir die Formel (16) benutzen und darin $\mathcal{E}$ durch $\mathcal{E}^{(1)}$ ersetzen. Wir bemerken noch, daß man in der Formel (39) alle Punkte der Menge $\mathcal{E}^{(1)}$, in denen $f(P) = 0$ ist, zu $\mathcal{E}^{(2)}$ rechnen kann. Die Menge $\mathcal{E}^{(1)} \mathcal{E}_0$ dieser Punkte ist bezüglich $G(\mathcal{E})$ meßbar.

Wir beweisen jetzt einen Satz, der es ermöglicht, das Lebesgue-Stieltjessche Integral bezüglich $\varphi(\mathcal{E})$ auf das Integral bezüglich $G(\mathcal{E})$ zurückzuführen.

Satz 3. Wenn $F(P)$ auf einer Menge $\mathcal{E}$, die bezüglich $\varphi(\mathcal{E})$ meßbar und von endlichem Maß ist, definiert, bezüglich $\varphi(\mathcal{E})$ meßbar und summierbar ist, so ist das Produkt $F(P) f(P)$ bezüglich $G(\mathcal{E})$ in $\mathcal{E}^{(1)}$ meßbar, und es gilt die Formel

$$\int_{\mathcal{E}} F(P) \varphi(d\mathcal{E}) = \int_{\mathcal{E}^{(1)}} F(P) f(P) G(d\mathcal{E}), \quad (46)$$

wofür man auch

$$\int_{\mathcal{E}} F(P) \left[\int_{d\mathcal{E}} f(P) G(d\mathcal{E}) \right] = \int_{\mathcal{E}^{(1)}} F(P) f(P) G(d\mathcal{E}) \quad (46_1)$$

schreiben kann.

Beweis. Wenn wir $F(P)$ außerhalb $\mathfrak{E}$ durch 0 fortsetzen, können wir annehmen, daß $F(P)$ genauso wie $f(P)$ überall definiert ist. Außerdem können wir annehmen, daß $F(P)$ und $f(P)$ in allen Punkten endliche Werte haben. Die Funktion $F(P)$ ist bezüglich $\varphi(\mathfrak{E})$ und die Funktion $f(P)$ bezüglich $G(\mathfrak{E})$, folglich auch bezüglich $\varphi(\mathfrak{E})$ meßbar. Wir führen eine neue Funktion $F_0(P)$ ein, indem wir $F_0(P) = F(P)$ für $f(P) \neq 0$ und $F_0(P) = 0$ für $f(P) = 0$ setzen. Mit anderen Worten, $F_0(P)$ ist gleich $F(P) - F(P) \omega_H(P)$, wobei $\omega_H(P)$ die charakteristische Funktion der Menge H derjenigen Punkte ist, in denen $f(P) = 0$ ist. Wie wir schon erwähnten, ist $H \in L_G$ und folglich $H \in L_\varphi$, d. h., sowohl $F(P)$ als auch $\omega_H(P)$ ist bezüglich $\varphi(\mathfrak{E})$ meßbar; daher ist auch $F_0(P)$ bezüglich $\varphi(\mathfrak{E})$ meßbar. Wir zeigen, daß $F_0(P)$ auch bezüglich $G(\mathfrak{E})$ meßbar ist. Weil $F_0(P)$ bezüglich $\varphi(\mathfrak{E})$ meßbar ist, können wir für beliebiges a die Menge $\mathfrak{E}_a$ derjenigen Punkte, in denen $F_0(P) > 0$ ist, durch $\mathfrak{E}_a = \mathfrak{E}_a^{(1)} + \mathfrak{E}_a^{(2)}$ darstellen, wobei $\mathfrak{E}_a^{(1)}$ in L_G liegt und $f(P)$ in den Punkten von $\mathfrak{E}_a^{(2)}$ gleich 0 ist. Ist $a \geq 0$, so ist nach der Definition von $F_0(P)$ die Menge $\mathfrak{E}_a^{(2)}$ leer, und folglich gilt $\mathfrak{E}_a \in L_G$. Für $a < 0$ enthält die Menge $\mathfrak{E}_a$ die ganze Menge H , und, wie wir schon oben erwähnten, kann man für diesen Fall annehmen, daß $\mathfrak{E}_a^{(2)}$ mit H übereinstimmt, d. h., daß $f(P) > 0$ in allen Punkten von $\mathfrak{E}_a^{(1)}$ ist. Wegen $H \in L_G$ liegt auch $\mathfrak{E}_a = \mathfrak{E}_a^{(1)} + H$ in L_G . Somit ist $\mathfrak{E}_a \in L_G$ für alle a , und weil $F_0(P)$ bezüglich $G(\mathfrak{E})$ meßbar ist, ist auch das Produkt $F_0(P) f(P)$ bezüglich $G(\mathfrak{E})$ meßbar. Wir betrachten jetzt die in dem Satz erwähnte Menge $\mathfrak{E}^{(1)}$. In den Punkten dieser Menge stimmt das Produkt $F_0(P) f(P)$ mit $F(P) f(P)$ überein, d. h., $F(P) f(P)$ ist in $\mathfrak{E}^{(1)}$ bezüglich $G(\mathfrak{E})$ meßbar. Für den Beweis der Formel (46) beschränken wir uns auf den Fall, daß $F(P)$ beschränkt ist. Für nichtbeschränkte Funktionen kann man das analog beweisen. Es sei $|F(P)| < L$, und $\mathfrak{E}_{n,k}$ sei die Menge derjenigen Punkte P aus $\mathfrak{E}$, in denen die Ungleichung

$$\frac{k}{2^n} L < F(P) \leq \frac{k+1}{2^n} L \quad (k = -2^n, -2^n + 1, \dots, 2^n - 1)$$

gilt. Wir bilden die stückweise konstanten Funktionen

$$F_n(P) = \frac{k}{2^n} L \quad \text{für } P \in \mathfrak{E}_{n,k}.$$

Die Folge der $F_n(P)$ wächst für wachsendes n und ist dem absoluten Betrag nach durch die Zahl L beschränkt.

Es gilt

$$\int_{\mathfrak{E}} F_n(P) \varphi(d\mathfrak{E}) = \sum_{k=-2^n}^{2^n-1} \frac{k}{2^n} L \varphi(\mathfrak{E}_{n,k}) = \sum_{k=-2^n}^{2^n-1} \frac{k}{2^n} L \int_{\mathfrak{E}_{n,k}^{(1)}} f(P) G(d\mathfrak{E}),$$

wobei wir $\mathfrak{E}_{n,k}^{(1)}$ aus $\mathfrak{E}_{n,k}$ nach Formel (39) erhalten. Wenn wir über alle Werte k summieren, erhalten wir die Menge $\mathfrak{E}^{(1)}$, und ziehen wir in der letzten Formel den

Faktor $\frac{k}{2^n} L$ unter das Integralzeichen, so kommen wir zu der Formel

$$\int_{\mathfrak{E}} F_n(P) \varphi(d\mathfrak{E}) = \int_{\mathfrak{E}^{(1)}} F_n(P) f(P) G(d\mathfrak{E}).$$

Der Integrand auf der rechten Seite ist dem absoluten Betrag nach nicht größer als die summierbare Funktion $L f(P)$, und auf beiden Seiten können wir unter dem Integralzeichen zur Grenze übergehen, was die Formel (46) liefert. Wenn die Menge $\mathfrak{E}$ auch bezüglich $G(\mathfrak{E})$ meßbar ist, kann man in (46) $\mathfrak{E}^{(1)}$ durch $\mathfrak{E}$ er-

setzen, weil $\mathcal{E}^{(2)} = \mathcal{E} - \mathcal{E}^{(1)}$ und $f(P) = 0$ in $\mathcal{E}^{(2)}$ ist, so daß das Integral über $\mathcal{E}^{(2)}$ gleich 0 ist.¹⁾

Ist $G(\mathcal{A})$ eine Funktion von beschränkter Variation und benutzen wir die kanonische Darstellung als Differenz zweier nichtnegativer Funktionen, so erhalten wir

$$G(\mathcal{E}) = G_1(\mathcal{E}) - G_2(\mathcal{E})$$

und können den Satz auf $G_1(\mathcal{E})$ und $G_2(\mathcal{E})$ anwenden. An Stelle von L_G haben wir in diesem Fall L_V mit $V(\mathcal{E}) = G_1(\mathcal{E}) + G_2(\mathcal{E})$.

Die Funktion $\varphi(\mathcal{E})$ können wir in Form einer Differenz zweier nichtnegativer Funktionen, $\varphi(\mathcal{A}) = \varphi_1(\mathcal{A}) - \varphi_2(\mathcal{A})$, darstellen; statt L_φ haben wir L_{V_1} , wobei $V_1(\mathcal{A}) = \varphi_1(\mathcal{A}) + \varphi_2(\mathcal{A})$ ist. Analog kann man den Fall behandeln, daß $f(P)$ das Vorzeichen wechselt. Dabei muß man $f(P)$ als Differenz des positiven und des negativen Teiles darstellen:

$$f(P) = f^+(P) - f^-(P)$$

und dann den Satz und die Formel für jeden Summanden einzeln beweisen. Dabei wird $\varphi(\mathcal{A})$ als Differenz zweier nichtnegativer Funktionen dargestellt, und der Mengenkörper L_φ wird durch den Körper der Mengen ersetzt, die bezüglich der Funktion

$$\psi(\mathcal{E}) = \int_{\mathcal{E}} |f(P)| G(d\mathcal{E})$$

meßbar sind.

Analog kann man den Satz auch auf den Fall komplexer Funktionen $f(P)$ und $G(\mathcal{E})$ übertragen.

Wir behandeln jetzt noch diesen Satz im Fall einer Veränderlichen. Es sei $g(x)$ eine nichtfallende und beschränkte Funktion in dem endlichen Intervall $[a, b]$, und $f(x)$ sei eine nichtnegative und bezüglich $g(x)$ in $[a, b]$ summierbare Funktion. Wir betrachten die Funktion

$$\omega(x) = \int_{[a, x]} f(x) dg(x).$$

Jede bezüglich $g(x)$ meßbare Menge ist auch bezüglich $\omega(x)$ meßbar, und für die Meßbarkeit einer Menge $\mathcal{E}$ bezüglich $\omega(x)$ ist notwendig und hinreichend, daß $\mathcal{E}$ in der Form (39) darstellbar ist, wobei $\mathcal{E}^{(1)}$ bezüglich $g(x)$ meßbar und $f(x)$ in allen Punkten von $\mathcal{E}^{(2)}$ gleich 0 ist. Wenn $F(x)$ auf einer bezüglich $\omega(x)$ meßbaren Menge $\mathcal{E}$ bezüglich $\omega(x)$ summierbar ist, so gilt

$$\int_{\mathcal{E}} F(x) d \left[\int_{[a, x]} f(x) dg(x) \right] = \int_{\mathcal{E}^{(1)}} F(x) f(x) dg(x), \quad (47)$$

wobei $\mathcal{E}^{(1)}$ die in (39) vorkommende Teilmenge von $\mathcal{E}$ ist. Ist $F(x)$ auf $\mathcal{E}$ bezüglich $g(x)$ meßbar, so kann man $\mathcal{E}^{(1)}$ durch $\mathcal{E}$ ersetzen. Im Fall $g(x) = x$ erhalten wir

$$\int_{\mathcal{E}} F(x) d \left[\int_a^x f(x) dx \right] = \int_{\mathcal{E}^{(1)}} F(x) f(x) dx. \quad (48)$$

Wenn also $\omega(x)$ keinen singulären Summanden hat, kann man das Lebesgue-Stieltjessche Integral von $F(x)$ bezüglich $\omega(x)$ durch das Lebesguesche Integral ausdrücken. Ist $\mathcal{E}$ das Intervall $[a, b]$, so kann man für die Formeln (47) und (48)

¹⁾ Die Beweise dieser letzten drei Sätze sind dem Buch von M. H. STONE, *Linear transformations in Hilbert space and their applications to analysis*, Amer. Math. Soc. Coll. Publ., Vol. XV, New York 1932, entnommen.

auch

$$\int_{[a,b]} F(x) d\left[\int_{[a,x]} f(x) dg(x)\right] = \int_{[a,b]} F(x) f(x) dg(x), \quad (49)$$

$$\int_a^b F(x) d\left[\int_a^x f(x) dx\right] = \int_a^b F(x) f(x) dx \quad (49_1)$$

schreiben, wobei z. B. in der letzten Formel $F(x)$ bezüglich des unbestimmten Integrals von $f(x)$ als summierbar vorausgesetzt wird. Wir nehmen an, daß $\Phi(x)$ und $\Psi(x)$ absolut stetige Funktionen sind, d. h.

$$\Phi(x) = \int_a^x \Phi'(x) dx + c_1, \quad \Psi(x) = \int_a^x \Psi'(x) dx + c_2, \quad (50)$$

wobei $\Phi'(x)$ und $\Psi'(x)$ auf $[a, b]$ summierbar sind. Wenn wir die Formel (49₁) benutzen, können wir

$$\int_a^b \Phi(x) \Psi'(x) dx + \int_a^b \Psi(x) \Phi'(x) dx = \int_a^b \Phi(x) d\Psi(x) + \int_a^b \Psi(x) d\Phi(x)$$

schreiben, wobei die rechts stehenden Integrale gewöhnliche Stieltjessche Integrale sind, weil $\Phi(x)$ und $\Psi(x)$ stetig und von beschränkter Variation sind. Für die rechte Seite gilt [2]

$$\int_a^b \Phi(x) d\Psi(x) + \int_a^b \Psi(x) d\Phi(x) = [\Phi(x) \Psi(x)]_{x=a}^{x=b};$$

und setzen wir dies in die vorhergehende Formel ein, so erhalten wir die Forme für die partielle Integration:

$$\int_a^b \Phi(x) \Psi'(x) dx + \int_a^b \Psi(x) \Phi'(x) dx = [\Phi(x) \Psi(x)]_{x=a}^{x=b}. \quad (51)$$

Aus (50) folgt, daß für die Summe $\Phi(x) + \Psi(x)$ der Integrand gleich $\Phi'(x) + \Psi'(x)$ ist, d. h.

$$[\Phi(x) + \Psi(x)]' = \Phi'(x) + \Psi'(x).$$

Wenn wir in (51) $b = x$ setzen, erhalten wir

$$\Phi(x) \Psi(x) = \int_a^x [\Phi'(x) \Psi(x) + \Phi(x) \Psi'(x)] dx + \Phi(a) \Psi(a),$$

d. h.

$$[\Phi(x) \Psi(x)]' = \Phi'(x) \Psi(x) + \Phi(x) \Psi'(x).$$

Analoge Betrachtungen kann man auch für den Fall eines unendlichen Intervalls anstellen.

76. Beispiel. Wir geben jetzt ein Beispiel einer nichtabnehmenden stetigen Funktion an, die nicht absolut stetig ist und für die in der Formel (20) der zweite, d. h. der absolut stetige Summand fehlt. Vorbereitend konstruieren wir in dem Intervall $[0, 1]$ eine abgeschlossene Menge $F_{0.1}$ ¹⁾ Wir zerlegen das Intervall $[0, 1]$ durch die Punkte $\frac{1}{3}$ und $\frac{2}{3}$ in drei gleiche Teile und nehmen das mittlere offene Intervall

¹⁾ Diese abgeschlossene Menge heißt *Cantorsches Diskontinuum* oder *Cantorsche „Wischmenge“* (Anm. d. Red.).

$\left(\frac{1}{3}, \frac{2}{3}\right)$ heraus. Jedes der übrigbleibenden Intervalle $\left[0, \frac{1}{3}\right]$ und $\left[\frac{2}{3}, 1\right]$ zerlegen wir in drei gleiche Teile: das eine durch die Punkte $\frac{1}{9}$ und $\frac{2}{9}$, das zweite durch die Punkte $\frac{7}{9}$ und $\frac{8}{9}$. Wir nehmen wieder aus jedem dieser Intervalle ihre mittleren offenen Teile heraus, d. h. die Intervalle $\left(\frac{1}{9}, \frac{2}{9}\right)$ und $\left(\frac{7}{9}, \frac{8}{9}\right)$. Jedes der vier übrigbleibenden Intervalle

$$\left[0, \frac{1}{9}\right], \quad \left[\frac{2}{9}, \frac{3}{9}\right], \quad \left[\frac{6}{9}, \frac{7}{9}\right], \quad \left[\frac{8}{9}, 1\right]$$

zerlegen wir wieder in drei gleiche Teile und entfernen wieder aus jedem dieser vier Intervalle die offenen, mittleren Intervalle usw. Auf diese Weise entfernen wir schließlich aus dem Intervall $[0, 1]$ abzählbar viele offene Intervalle, die paarweise punktfremd sind und auch keine gemeinsamen Endpunkte haben:

$$\left(\frac{1}{3}, \frac{2}{3}\right), \quad \left(\frac{1}{9}, \frac{2}{9}\right), \quad \left(\frac{7}{9}, \frac{8}{9}\right), \quad \left(\frac{1}{27}, \frac{2}{27}\right), \quad \left(\frac{7}{27}, \frac{8}{27}\right), \quad \left(\frac{19}{27}, \frac{20}{27}\right), \quad \left(\frac{26}{27}, \frac{26}{27}\right), \dots, (52)$$

d. h., wir entfernen eine offene Menge H_0 , und es bleibt eine Menge übrig, die wir mit F_0 bezeichnen und die abgeschlossen ist. Beim ersten Schritt entfernten wir ein offenes Intervall der Länge $\frac{1}{3}$, beim zweiten Schritt zwei Intervalle der Länge $\frac{1}{3^2}$, beim dritten Schritt 2^2 Intervalle der Länge $\frac{1}{3^3}$ und allgemein beim n -ten Schritt 2^{n-1} Intervalle der Länge $\frac{1}{3^n}$. Somit ist das Lebesguesche Maß der offenen Menge H_0 gleich

$$\sum_{n=1}^{\infty} \frac{2^{n-1}}{3^n} = \frac{\frac{1}{3}}{1 - \frac{2}{3}} = 1,$$

und die von dem Intervall $[0, 1]$ übrigbleibende Menge F_0 hat folglich das Maß 0. Wir definieren jetzt in dem Intervall $[0, 1]$ eine Funktion $f(x)$:

$$f(x) = \frac{1}{2} \quad \text{für } x \in \left(\frac{1}{3}, \frac{2}{3}\right); \quad f(x) = \frac{1}{4} \quad \text{für } x \in \left(\frac{1}{9}, \frac{2}{9}\right); \quad f(x) = \frac{3}{4} \quad \text{für } x \in \left(\frac{7}{9}, \frac{8}{9}\right);$$

allgemein setzen wir $f(x)$ gleich $\frac{1}{1^n}, \frac{3}{2^n}, \frac{5}{2^n}, \dots, \frac{2^n-1}{2^n}$ auf den von links nach rechts aufeinanderfolgenden Intervallen, die wir beim n -ten Schritt entfernen. Die Funktion $f(x)$ ist auf diese Weise in den Punkten der Menge H_0 definiert und bleibt in jedem der offenen Intervalle (52), aus denen diese Menge besteht, konstant. Wir definieren $f(x)$ noch in den Randpunkten von $[0, 1]$, indem wir $f(0) = 0$ und $f(1) = 1$ setzen. Das Prinzip, nach dem wir die Funktion in jedem der Intervalle definierten, besteht in folgendem: Wir setzen $f(x)$ in einem Intervall der Menge H_0 , das wir beim n -ten Schritt erhalten, gleich dem arithmetischen Mittel der Werte von $f(x)$ in den früher erhaltenen Nachbarintervallen oder in den Endpunkten des Intervalls $[0, 1]$, wenn auf der einen Seite des betrachteten neuen Intervalls aus H_0 keine der früher erhaltenen Intervalle von H_0 liegen. Hieraus folgt, daß $f(x)$ eine nichtabnehmende Funktion auf der Menge H_0 ist. Wir setzen nun die Definition von $f(x)$ auf die Menge F_0 fort. Es sei $x_0 \in F_0$. Da F_0 das Maß 0 hat, liegen in jeder ε -Umgebung von x_0 Punkte von H_0 , und strebt der Punkt x bezüglich H_0 von links gegen x_0 , so

nimmt $f(x)$ nicht ab und hat einen Grenzwert, den wir als Wert von $f(x)$ für $x=x_0$ nehmen. Mit anderen Worten, die obige Fortsetzung reduziert sich darauf, daß wir $f(x_0)$ als obere Grenze der Werte von $f(x)$ für von links gegen x_0 strebende $x \in H_0$ annehmen. Im Punkt $x=1$ liefert diese Definition offenbar den Wert $f(1)=1$. Somit ist die in dem ganzen Intervall $[0, 1]$ definierte Funktion nicht abnehmend. Man kann leicht zeigen, daß sie stetig ist. Hätte sie nämlich eine Unstetigkeitsstelle $x=x'$, so würde sich wenigstens eines der Intervalle $[f(x'-0), f(x')]$ oder $[f(x'), f(x'+0)]$ nicht auf einen Punkt reduzieren und im Innern keine Werte von $f(x)$ enthalten, da $f(x)$ monoton ist. Aber die oben nur für die Menge H_0 definierten Werte von $f(x)$ liegen in dem Intervall $[0, 1]$ überall dicht. Aus der angenommenen Unstetigkeit von $f(x)$ ergibt sich also ein Widerspruch.

Wir erinnern daran, daß $f(x)$ in jedem der Intervalle (52) einen konstanten Wert hat. Mit Hilfe der nichtfallenden stetigen Funktion $f(x)$ können wir eine volladditive nichtnegative Mengenfunktion $\varphi(\mathfrak{E})$ konstruieren, die jedenfalls auf B -Mengen definiert ist. Nach dem oben Gesagten ist $\varphi(H_0)=0$, und daher ist $\varphi(\mathfrak{E})$ in jeder B -Menge, die eine Teilmenge von H_0 darstellt, gleich 0. Wenn wir das Intervall $[0, x]$ nehmen, so können wir

$$[0, x] = [0, x] H_0 + [0, x] F_0$$

und folglich

$$f(x) = \varphi([0, x]) = \varphi([0, x] H_0) + \varphi([0, x] F_0)$$

schreiben. Der erste Summand ist nach dem oben Gesagten gleich 0, und das Maß von F_0 ist gleich 0; folglich ist $f(x)$ gleich dem singulären Teil [74]:

$$f(x) = \varphi([0, x] F_0),$$

wobei F_0 die Rolle von H in Formel (20) und $f(x)$ die von $\omega(x)$ spielt. Wir untersuchen noch die Menge F_0 . Die stetige nichtabnehmende Funktion $f(x)$ nimmt alle reellen Werte zwischen 0 und 1 an. In jedem der „ausgewischten“ Intervalle inklusive ihrer Randpunkte hat $f(x)$ einen konstanten Wert, wobei die Menge dieser Intervalle abzählbar ist. Die Menge aller Werte von $f(x)$ ist nicht abzählbar. (Sie hat die Mächtigkeit des Kontinuums.) Somit enthält F_0 offenbar Punkte, die von den Randpunkten der „ausgewischten“ Intervalle verschieden sind. Man kann zeigen, daß F_0 die Mächtigkeit des Kontinuums hat.

77. Absolut stetige Funktionen mehrerer Veränderlicher. Analog wie wir die absolut stetigen Punktfunktionen einer Veränderlichen konstruiert haben [74], führen wir jetzt den Begriff der absolut stetigen Funktion mehrerer Veränderlicher ein. Wir beschränken uns dabei auf Funktionen zweier Veränderlicher. Es sei in dem zweidimensionalen Intervall $\Delta_0 [a \leq x \leq b, c \leq y \leq d]$ eine stetige Funktion $F(x, y)$ gegeben. Wir können mit ihrer Hilfe eine Funktion $\varphi(\delta)$ der Intervalle konstruieren, die in Δ_0 liegen, und zwar können wir, wenn wir das Intervall δ durch die Ungleichungen $x_1 \leq x \leq x_2, y_1 \leq y \leq y_2$ definieren, wie früher

$$\varphi(\delta) = F(x_2, y_2) - F(x_1, y_2) - F(x_2, y_1) + F(x_1, y_1) \quad (53)$$

setzen, wobei δ abgeschlossen oder nicht abgeschlossen sein kann, da $F(x, y)$ nach Voraussetzung stetig ist. Wenn wir zu $F(x, y)$ die Summe $f_1(x) + f_2(y)$ addieren, wobei der erste Summand nur von x und der zweite nur von y abhängt, wird $\varphi(\delta)$ davon nicht beeinflusst. Die Intervallfunktion $\varphi(\delta)$ nennen wir absolut stetig, wenn sie eine zu (23) aus [74] analoge Bedingung erfüllt, d. h., wenn es zu beliebig gegebenem positivem ε ein positives η gibt derart, daß

$$\left| \sum_{k=1}^n \varphi(\delta_k) \right| \leq \varepsilon$$

gilt, sobald die Summe der Flächeninhalte der sich paarweise nicht überlappenden Intervalle δ_k ($k = 1, 2, \dots, n$) höchstens gleich η ist.

Definition. Eine Funktion $F(x, y)$ nennt man *absolut stetig in den beiden Veränderlichen* x, y , wenn eine mit Hilfe von (53) definierte Funktion $\varphi(\delta)$ eine absolut stetige Intervallfunktion ist und wenn außerdem $F(a, y)$ und $F(x, c)$ absolut stetige Funktionen von y bzw. x sind.

Der letzte Vorbehalt über die absolute Stetigkeit von $F(x, y)$ auf der unteren und der linken Seite des Intervalls Δ_0 muß gemacht werden, damit man die Möglichkeit hat, zu $F(x, y)$ die obenerwähnte Summe $f_1(x) + f_2(y)$ hinzufügen. Offensichtlich gilt

$$F(x, y) = [F(x, y) - F(a, y) - F(x, c) + F(a, c)] \\ + [F(x, c) - F(a, c)] + [F(a, y) - F(a, c)] + F(a, c).$$

Der erste Summand rechts ist $\varphi(\delta_{x,y})$, wo $\delta_{x,y}$ das Intervall $a \leq x' \leq x$, $c \leq y' \leq y$ ist, und diese Funktion kann wie in [75] durch ein unbestimmtes Doppelintegral einer summierbaren Funktion dargestellt werden. Der zweite und der dritte Summand rechts sind absolut stetige Funktionen von x bzw. y ; folglich können sie durch einfache unbestimmte Integrale dargestellt werden. Somit läßt sich jede absolut stetige Funktion $F(x, y)$ nach der Formel

$$F(x, y) = \int_a^x \int_c^y f(x, y) \, dx dy + \int_a^x g(x) \, dx + \int_c^y h(y) \, dy + F(a, c) \quad (54)$$

darstellen.

Umgekehrt ist leicht zu sehen, daß jede durch eine solche Formel darstellbare Funktion eine absolut stetige Funktion ist. Wir können diese Formel auf Grund des Satzes von FUBINI in der Gestalt

$$F(x, y) = \int_a^x \left[\int_c^y f(x, y) \, dy + g(x) \right] dx + \int_c^y f(y) \, dy + F(a, c) \quad (55)$$

oder

$$F(x, y) = \int_c^y \left[\int_a^x f(x, y) \, dx + h(y) \right] dy + \int_a^x g(x) \, dx + F(a, c) \quad (56)$$

schreiben. Hieraus sieht man: Ist $F(x, y)$ eine absolut stetige Funktion zweier Veränderlicher, so ist sie auch eine absolut stetige Funktion von x für jedes feste y und eine absolut stetige Funktion von y für jedes feste x . Die umgekehrte Behauptung ist nicht richtig, d. h., eine Funktion kann für jede Veränderliche absolut stetig sein, braucht aber nicht in beiden Veränderlichen absolut stetig zu sein.

Der Integrand der ersten Summanden aus (55) und (56) liefert nach Definition [74] die partiellen Ableitungen der absolut stetigen Funktion $F(x, y)$:

$$\frac{\partial F(x, y)}{\partial x} = \int_c^y f(x, y) \, dy + g(x), \quad \frac{\partial F(x, y)}{\partial y} = \int_a^x f(x, y) \, dx + h(y). \quad (57)$$

Der Integrand in diesen Formeln definiert die gemischte Ableitung zweiter Ordnung:

$$\frac{\partial}{\partial y} \left[\frac{\partial F(x, y)}{\partial x} \right] = \frac{\partial}{\partial x} \left[\frac{\partial F(x, y)}{\partial y} \right] = f(x, y).$$

Sind die partiellen Ableitungen F_x und F_y selbst absolut stetige Funktionen zweier Veränderlicher, so können wir alle partiellen Ableitungen zweiter Ordnung definieren. Wenn diese ebenfalls absolut stetige Funktionen zweier Veränderlicher sind, so können wir alle Ableitungen dritter Ordnung definieren, usw.

Man kann zeigen, daß die partiellen Ableitungen fast überall in Δ_0 die Grenzwerte der entsprechenden Differenzenquotienten sind. So ist etwa F_x der Grenzwert des Quotienten

$$\frac{F(x+h, y) - F(x, y)}{h}.$$

Man kann eine absolut stetige Funktion $F(x, y)$ als Punktfunktion $f(M)$ in der Ebene deuten. Wenn wir in diese Ebene an Stelle des ursprünglichen kartesischen x, y -Koordinatensystems ein neues x', y' -System einführen, erhalten wir eine neue Funktion $F(x', y')$, die in den neuen Veränderlichen nicht absolut stetig zu sein braucht. Als Beispiel betrachten wir die absolut stetige Funktion

$$F(x, y) = \int_0^x f(t) dt,$$

wobei $f(t)$ die stetige, jedoch nicht absolut stetige Funktion ist, die wir in [76] für das Intervall $[0, 1]$ konstruiert haben. Wir setzen sie ins Äußere von $[0, 1]$ fort, indem wir $f(x) = 0$ für $x < 0$ und $f(x) = 1$ für $x > 1$ setzen. Die obige Formel definiert die absolut stetige Funktion $F(x, y)$ in der ganzen Ebene (sie hängt faktisch nur von x ab). Drehen wir das Koordinatensystem um den Ursprung um 45° , so erhalten wir in den neuen Koordinaten

$$F(x', y') = \int_0^{(x'+y')/\sqrt{2}} f(t) dt.$$

Die partielle Ableitung dieser Funktion nach x' , die durch

$$F_{x'} = \frac{1}{\sqrt{2}} f\left(\frac{1}{\sqrt{2}} x' + \frac{1}{\sqrt{2}} y'\right)$$

ausgedrückt wird, ist bei gegebenem x' keine absolut stetige Funktion von y' , was wegen (57) der Fall sein müßte, wenn $F(x', y')$ in beiden Veränderlichen absolut stetig wäre. Wir weisen darauf hin, daß diese Funktion für beliebig gewählte kartesische Koordinaten bezüglich jeder Veränderlichen für alle Werte der anderen Veränderlichen absolut stetig ist. Analog zum Vorhergehenden wird auch die Theorie der absolut stetigen Funktionen beliebig vieler Veränderlicher aufgebaut.

Im folgenden Kapitel werden wir eine allgemeinere Definition der partiellen Ableitungen geben, die nicht nur auf absolut stetige Funktionen mehrerer Veränderlicher angewendet werden kann.

78. Hilfssätze. In den nächsten beiden Abschnitten werden wir einen neuen Begriff einführen und einige Hilfssätze beweisen, die zum Beweis des grundlegenden Satzes aus [73] und zur weiteren Verallgemeinerung des Integralbegriffs notwendig sind.

Wir betrachten Funktionen $\varphi(\mathfrak{S})$, die auf der Familie C , die aus einer Menge $\mathfrak{S}_0 \in B$ und allen Teilmengen $\mathfrak{S}$ von $\mathfrak{S}_0 \in B$ besteht, volladditiv sind; dabei sei $G(\mathfrak{S}_0)$ endlich und von 0 verschieden. Wir bezeichnen mit V_1 die Menge dieser Funktionen. Liegen $\varphi_1(\mathfrak{S})$ und $\varphi_2(\mathfrak{S})$ in V_1 , so gehört auch $c_1\varphi_1(\mathfrak{S}) + c_2\varphi_2(\mathfrak{S})$ zu V_1 . Wir wissen, daß für jedes $\varphi(\mathfrak{S})$ aus V_1 die Summen

$$t_\delta(\varphi) = \sum_k |\varphi(\mathfrak{S}_k)| \quad (58)$$

bei beliebiger Zerlegung von $\mathfrak{S}_0$ in endlich viele Mengen $\mathfrak{S}_k$ beschränkt bleiben [72]. Die obere Grenze der Summen t_δ , d. h. die Totalvariation von $\varphi(\mathfrak{S})$ auf $\mathfrak{S}_0$, bezeichnen wir mit $\|\varphi\|_1$. Es gilt offenbar $\|c\varphi\|_1 = |c| \cdot \|\varphi\|_1$, wobei c eine Konstante ist. Ist die Zerlegung δ' eine Fortsetzung der Zerlegung δ , so schreiben wir $\delta' \supseteq \delta$. Wenn wir beachten, daß für jede Zerlegung $e = e' + e''$

$$|\varphi(e)| \leq |\varphi(e')| + |\varphi(e'')|$$

gilt, können wir behaupten, daß $t_{\delta'}(\varphi) \geq t_\delta(\varphi)$ für $\delta' \supseteq \delta$ ist. Ist $\{\delta_n\}$ eine solche Folge von Zerlegungen, daß $t_{\delta_n}(\varphi) \rightarrow \|\varphi\|_1$ gilt, und $\delta'_n \supseteq \delta_n$, so gilt auch $t_{\delta'_n}(\varphi) \rightarrow \|\varphi\|_1$. Für

$t_{\delta_n}(\varphi) \rightarrow \|\varphi\|_1$, $t_{\delta'_n}(\psi) \rightarrow \|\psi\|_1$ und $t_{\delta''_n}(\varphi + \psi) \rightarrow \|\varphi + \psi\|_1$ erhalten wir, wenn $\delta''_n = \delta_n \delta'_n$ gesetzt wird,

$$t_{\delta''_n}(\varphi) \rightarrow \|\varphi\|_1, \quad t_{\delta''_n}(\psi) \rightarrow \|\psi\|_1, \quad t_{\delta''_n}(\varphi + \psi) \rightarrow \|\varphi + \psi\|_1.$$

Ist $\mathcal{E}''_{k,n}$ eine Teilmenge in δ''_n , so erhalten wir aus der Ungleichung

$$|\varphi(\mathcal{E}''_{k,n}) + \psi(\mathcal{E}''_{k,n})| \leq |\varphi(\mathcal{E}''_{k,n})| + |\psi(\mathcal{E}''_{k,n})|$$

durch Summation und Grenzübergang

$$\|\varphi + \psi\|_1 \leq \|\varphi\|_1 + \|\psi\|_1. \quad (59)$$

Neben der Summe (58) betrachten wir für die Funktionen $\varphi(\mathcal{E})$, die der Bedingung

$$\varphi(\mathcal{E}) = 0 \quad \text{für} \quad G(\mathcal{E}) = 0 \quad (60)$$

genügen, noch die Summe

$$S_\delta(\varphi) = \sum_k \frac{\varphi^2(\mathcal{E}_k)}{G(\mathcal{E}_k)}. \quad (61)$$

Ist $G(\mathcal{E}_k) = 0$, so haben die entsprechenden Summanden die Form $\frac{0}{0}$, und wir setzen sie gleich 0. Wir brauchen dies nicht zu tun, wenn wir nur solche Zerlegungen δ betrachten würden, bei denen alle $G(\mathcal{E}_k)$ von 0 verschieden sind, was darauf hinauslaufen würde, daß diejenigen $\mathcal{E}_k$, für die $G(\mathcal{E}_k) = 0$ ist, zu anderen Teilmengen hinzugenommen werden. Die Menge der Werte von $S_\delta(\varphi)$ ist nicht notwendig beschränkt. Wir bezeichnen mit V_2 die Menge derjenigen Funktionen $\varphi(\mathcal{E})$, die der Bedingung (60) genügen und für die die Menge der $S_\delta(\varphi)$ beschränkt ist. Die Menge V_2 ist eine Teilmenge von V_1 . Man sieht leicht, daß mit $\varphi(\mathcal{E})$ und $\psi(\mathcal{E})$ auch $c\varphi(\mathcal{E})$ und $\varphi(\mathcal{E}) + \psi(\mathcal{E})$ zu V_2 gehören. Die obere Grenze der Summen $S_\delta(\varphi)$ bezeichnen wir mit $\|\varphi\|_2$. Wir zeigen, daß für $\delta' \supseteq \delta$ die Ungleichung $S_{\delta'}(\varphi) \geq S_\delta(\varphi)$ gilt. Hierfür genügt es zu zeigen: Ist $\mathcal{E} = \mathcal{E}' + \mathcal{E}''$ eine Zerlegung von $\mathcal{E}$, so gilt

$$\frac{\varphi^2(\mathcal{E}')}{G(\mathcal{E}')} + \frac{\varphi^2(\mathcal{E}'')}{G(\mathcal{E}'')} \geq \frac{\varphi^2(\mathcal{E})}{G(\mathcal{E})}. \quad (62)$$

Diese Ungleichung ist äquivalent mit

$$G(\mathcal{E}) G(\mathcal{E}'') \varphi^2(\mathcal{E}') + G(\mathcal{E}) G(\mathcal{E}') \varphi^2(\mathcal{E}'') - G(\mathcal{E}') G(\mathcal{E}'') \varphi^2(\mathcal{E}) \geq 0$$

oder wegen $G(\mathcal{E}) = G(\mathcal{E}') + G(\mathcal{E}'')$ und $\varphi(\mathcal{E}) = \varphi(\mathcal{E}') + \varphi(\mathcal{E}'')$ mit

$$[G(\mathcal{E}'') \varphi(\mathcal{E}') - G(\mathcal{E}') \varphi(\mathcal{E}'')]^2 \geq 0.$$

Für $S_{\delta_n}(\varphi) \rightarrow \|\varphi\|_2$ und $\delta'_n \supseteq \delta_n$ gilt $S_{\delta'_n}(\varphi) \rightarrow \|\varphi\|_2$.

Wir stellen jetzt für die Funktionen aus V_2 eine Ungleichung auf, die den Zusammenhang zwischen $\|\varphi\|_1$ und $\|\varphi\|_2$ herstellt. Wenn wir die Cauchysche Ungleichung anwenden, erhalten wir

$$\begin{aligned} t_\delta(\varphi) &= \sum_k |\varphi(\mathcal{E}_k)| = \sum_k \frac{|\varphi(\mathcal{E}_k)|}{\sqrt{G(\mathcal{E}_k)}} \sqrt{G(\mathcal{E}_k)} \\ &\leq \sqrt{\sum_k \frac{\varphi^2(\mathcal{E}_k)}{G(\mathcal{E}_k)}} \sqrt{\sum_k G(\mathcal{E}_k)} = \sqrt{\sum_k \frac{\varphi^2(\mathcal{E}_k)}{G(\mathcal{E}_k)}} \sqrt{G(\mathcal{E}_0)}, \end{aligned}$$

d. h.

$$t_\delta(\varphi) \leq \sqrt{S_\delta(\varphi)} \sqrt{G(\mathcal{E}_0)}.$$

Ganz entsprechend wie bei der Herleitung von (59) kann man eine solche Folge von Zerlegungen δ_n angeben, daß $t_{\delta_n}(\varphi) \rightarrow \|\varphi\|_1$ und $S_{\delta_n}(\varphi) \rightarrow \|\varphi\|_2$ ist und die Ungleichung $t_{\delta_n}(\varphi) \leq \sqrt{S_{\delta_n}(\varphi)} \sqrt{G(\mathbb{E}_0)}$ in der Grenze die Beziehung

$$\|\varphi\|_1 \leq \sqrt{\|\varphi\|_2} \sqrt{G(\mathbb{E}_0)} \quad (63)$$

liefert.

Wir betrachten noch eine Familie von Funktionen $\varphi(\mathbb{E})$. Eine Funktion $\varphi(\mathbb{E})$ heißt *wesentlich beschränkt*, wenn eine solche Konstante c existiert, daß für jedes in $\mathbb{E}_0$ enthaltene $\mathbb{E}$ aus B

$$|\varphi(\mathbb{E})| \leq cG(\mathbb{E}) \quad (64)$$

gilt. Die Menge aller wesentlich beschränkten Funktionen (für verschiedene Werte von c) bezeichnen wir mit V_0 . Aus (61) folgt unmittelbar: Ist $\varphi(\mathbb{E}) \in V_0$, so gilt $S_\delta(\varphi) \leq c^2 G(\mathbb{E}_0)$ für jedes δ , d. h., V_0 ist eine Teilmenge von V_2 .

Wir betrachteten früher stückweise konstante Punktfunktionen [46]. Wir führen jetzt stückweise konstante Mengenfunktionen ein. Eine Funktion $\psi(\mathbb{E})$, die der Bedingung (60) genügt, heißt auf $\mathbb{E}_0$ *stückweise konstant*, wenn eine solche Zerlegung von $\mathbb{E}_0$ in endlich viele Mengen $\mathbb{E}_k$ existiert, daß

$$\psi(\mathbb{E}) = \begin{cases} \frac{\psi(\mathbb{E}_k)}{G(\mathbb{E}_k)} G(\mathbb{E}) & \text{für } \mathbb{E} \subset \mathbb{E}_k \text{ und } G(\mathbb{E}_k) \neq 0, \\ 0 & \text{für } \mathbb{E} \subset \mathbb{E}_k \text{ und } G(\mathbb{E}_k) = 0 \end{cases} \quad (65)$$

gilt. Wenn wir den Quotienten $\psi(\mathbb{E}_k) : G(\mathbb{E}_k)$ mit a_k bezeichnen ($a_k = 0$ für $G(\mathbb{E}_k) = 0$), kann die stückweise konstante Funktion $\psi(\mathbb{E})$ durch das Integral

$$\psi(\mathbb{E}) = \int_{\mathbb{E}} \sum_k a_k \omega_{\mathbb{E}_k}(P) G(d\mathbb{E}) \quad (66)$$

dargestellt werden, wobei $\omega_{\mathbb{E}_k}(P)$ die charakteristische Funktion der Menge $\mathbb{E}_k$ ist. Unter dem Integralzeichen steht eine stückweise konstante Punktfunktion, die auf der Menge $\mathbb{E}_k$ gleich a_k ist. Umgekehrt liefert jedes Integral der obigen Form eine stückweise konstante Mengenfunktion $\psi(\mathbb{E})$.

Für jede beliebig gegebene Zerlegung δ der Menge $\mathbb{E}_0$ in endlich viele Mengen $\mathbb{E}_k$ ordnen wir jetzt jeder auf $\mathbb{E}_0$ volladditiven Mengenfunktion $\varphi(\mathbb{E})$, die der Bedingung (60) genügt, und jeder Punktfunktion $f(P)$, die auf $\mathbb{E}_0$ meßbar und summierbar ist, die stückweise konstanten Funktionen $\varphi_\delta(\mathbb{E})$ und $f_\delta(P)$ zu. Wir definierten $\varphi_\delta(\mathbb{E})$ für ein zu $\mathbb{E}_k$ gehöriges $\mathbb{E}$ durch

$$\frac{\varphi_\delta(\mathbb{E})}{G(\mathbb{E})} = \frac{\varphi(\mathbb{E}_k)}{G(\mathbb{E}_k)}, \quad (67)$$

d. h. für jedes $\mathbb{E}$ aus $\mathbb{E}_0$ durch die Formel

$$\varphi_\delta(\mathbb{E}) = \int_{\mathbb{E}} \sum_k a_k \omega_{\mathbb{E}_k}(P) G(d\mathbb{E}) \quad \text{mit} \quad a_k = \frac{\varphi(\mathbb{E}_k)}{G(\mathbb{E}_k)}. \quad (68)$$

Die Funktion $f_\delta(P)$ nehmen wir als konstant an,

$$f_\delta(P) = \frac{1}{G(\mathbb{E}_k)} \int_{\mathbb{E}_k} f(P) G(d\mathbb{E}) \quad (69)$$

für $P \in \mathbb{E}_k$. Ist $G(\mathbb{E}_k) = 0$, so sei auch der Ausdruck (69) gleich 0. Wenn $\varphi(\mathbb{E})$ durch ein Integral dargestellt wird,

$$\varphi(\mathbb{E}) = \int_{\mathbb{E}} f(P) G(d\mathbb{E}), \quad (70)$$

so ist offenbar

$$\varphi_\delta(\mathbb{E}) = \int_{\mathbb{E}} f_\delta(P) G(d\mathbb{E}). \quad (71)$$

Aus der Definition von $\varphi_\delta(\mathbb{E})$ und $f_\delta(P)$ folgt

$$\begin{aligned} (c_1\varphi(\mathbb{E}) + c_2\psi(\mathbb{E}))_\delta &= c_1\varphi_\delta(\mathbb{E}) + c_2\psi_\delta(\mathbb{E}), \\ (c_1f(P) + c_2F(P))_\delta &= c_1f_\delta(P) + c_2F_\delta(P). \end{aligned} \quad (72)$$

Wir zeigen, daß für $\delta' \cong \delta$ die Beziehung

$$(\varphi_\delta(\mathbb{E}))_{\delta'} = \varphi_{\delta'}(\mathbb{E}) \quad (73)$$

gilt. Für δ' gibt es eine Zerlegung eines jeden $\mathbb{E}_k$ in Mengen $\mathbb{E}'_{k,s}$, wobei nach (67)

$$a_k = \frac{\varphi_\delta(\mathbb{E}'_{k,s})}{G(\mathbb{E}'_{k,s})} = \frac{\varphi(\mathbb{E}_k)}{G(\mathbb{E}_k)}$$

ist; hieraus folgt

$$\begin{aligned} (\varphi_\delta(\mathbb{E}))_{\delta'} &= \int_{\mathbb{E}} \sum_{k,s} a_k \omega_{\mathbb{E}'_{k,s}}(P) G(d\mathbb{E}) \\ &= \int_{\mathbb{E}} \sum_k a_k \sum_s \omega_{\mathbb{E}'_{k,s}}(P) G(d\mathbb{E}) = \int_{\mathbb{E}} \sum_k a_k \omega_{\mathbb{E}_k}(P) G(d\mathbb{E}) = \varphi_{\delta'}(\mathbb{E}). \end{aligned}$$

Wenn c das größte der a_k ist, gilt nach (65) $|\psi(\mathbb{E})| \leq cG(\mathbb{E})$, d. h., jede stückweise konstante Mengenfunktion ist wesentlich beschränkt. Wir bemerken noch, daß wir stückweise konstante Funktionen nur für eine Zerlegung von $\mathbb{E}_0$ in endlich viele Mengen betrachten.

Es sei $\delta' \cong \delta$, und $\mathbb{E}'_{k,s}$ seien die oben angegebenen Teilmengen. Aus der Definition von $\varphi_{\delta'}(\mathbb{E})$ folgt

$$\frac{\varphi_{\delta'}(\mathbb{E}'_{k,s})}{G(\mathbb{E}'_{k,s})} = \frac{\varphi_{\delta'}(\mathbb{E}_k)}{G(\mathbb{E}_k)}, \quad \varphi_{\delta'}(\mathbb{E}_k) = \varphi(\mathbb{E}_k)$$

und daraus

$$S_{\delta'}(\varphi_\delta) = \sum_{k,s} \frac{\varphi_\delta^2(\mathbb{E}'_{k,s})}{G(\mathbb{E}'_{k,s})} = \sum_{k,s} \frac{\varphi_\delta(\mathbb{E}_k)}{G(\mathbb{E}_k)} \varphi_\delta(\mathbb{E}'_{k,s}) = \sum_k \frac{\varphi(\mathbb{E}_k)}{G(\mathbb{E}_k)} \varphi_\delta(\mathbb{E}_k) = \sum_k \frac{\varphi^2(\mathbb{E}_k)}{G(\mathbb{E}_k)},$$

d. h.

$$S_{\delta'}(\varphi_\delta) = S_\delta(\varphi). \quad (74)$$

Wenn $S_{\delta_n}(\varphi_\delta)$ gegen $\|\varphi_\delta\|_2$ strebt, gilt für $\delta'_n = \delta_n \delta$ erst recht $S_{\delta'_n}(\varphi_\delta) \rightarrow \|\varphi_\delta\|_2$. Nun ist $\delta'_n \cong \delta$, und nach (74) hängt $S_{\delta'_n}(\varphi_\delta) = S_\delta(\varphi)$ nicht von n ab; daher folgt

$$\|\varphi_\delta\|_2 = S_\delta(\varphi). \quad (75)$$

Wenn $\varphi(\mathbb{E})$ zu V_2 gehört, existiert eine solche Folge von Zerlegungen $\delta^{(n)}$, daß $S_{\delta^{(n)}}(\varphi) \rightarrow \|\varphi\|_2$ gilt; folglich existiert eine solche Folge von Zerlegungen, daß

$$\|\varphi_{\delta^{(n)}}\|_2 \rightarrow \|\varphi\|_2 \quad (76)$$

gilt.

Die Größe (75) ist leicht durch ein Integral auszudrücken. Es sei $g_\delta(P)$ der Integrand in dem Integral (66) für die stückweise konstante Funktion $\varphi_\delta(\mathbb{E})$,

$$g_\delta(P) = a_k = \frac{\varphi(\mathbb{E}_k)}{G(\mathbb{E}_k)} \quad \text{für } P \in \mathbb{E}_k.$$

Es gilt

$$\varphi_{\delta}^2(\mathfrak{E}_k) = \varphi^2(\mathfrak{E}_k) = a_k^2 G^2(\mathfrak{E}_k) = G(\mathfrak{E}_k) \int_{\mathfrak{E}_k} g_{\delta}^2(P) G(d\mathfrak{E}),$$

also

$$S_{\delta}(\varphi) = \sum_k \frac{\varphi^2(\mathfrak{E}_k)}{G(\mathfrak{E}_k)} = \sum_k \int_{\mathfrak{E}_k} g_{\delta}^2(P) G(d\mathfrak{E}),$$

d. h.

$$S_{\delta}(\varphi) = \|\varphi_{\delta}\|_2 = \int_{\mathfrak{E}_0} g_{\delta}^2(P) G(d\mathfrak{E}), \quad (77)$$

und, wenn wir die gewöhnliche Bezeichnung der Norm in L_2 benutzen,

$$\|\varphi_{\delta}\|_2 = \|g_{\delta}(P)\|_{\mathfrak{E}_0}^2. \quad (78)$$

Wir geben noch zwei Formeln an, die wir im weiteren benötigen werden. Nach (73) gilt

$$[\varphi(\mathfrak{E}) - \varphi_{\delta}(\mathfrak{E})]_{\delta'} = \varphi_{\delta'}(\mathfrak{E}) - \varphi_{\delta}(\mathfrak{E})$$

für $\delta' \geq \delta$. Die Differenz $g_{\delta'}(P) - g_{\delta}(P)$, die konstante Werte in den Punkten der Teilmengen $\mathfrak{E}'_{k,s}$ der Zerlegung δ' annimmt, ist der Integrand in der Formel (66) für $\varphi_{\delta'}(\mathfrak{E}) - \varphi_{\delta}(\mathfrak{E})$, d. h.

$$\varphi_{\delta'}(\mathfrak{E}) - \varphi_{\delta}(\mathfrak{E}) = \int_{\mathfrak{E}_0} [g_{\delta'}(P) - g_{\delta}(P)] G(d\mathfrak{E});$$

nach (78) können wir dafür

$$\|(\varphi - \varphi_{\delta})_{\delta'}\|_2 = \|\varphi_{\delta'} - \varphi_{\delta}\|_2 = \|g_{\delta'}(P) - g_{\delta}(P)\|_{\mathfrak{E}_0}^2 \quad (79)$$

schreiben. Die letzte Formel betrifft die Funktion $g_{\delta'}(P)$. Sie nimmt in den Punkten von $\mathfrak{E}'_{k,s}$ den konstanten Wert $a'_{k,s} = \frac{\varphi(\mathfrak{E}'_{k,s})}{G(\mathfrak{E}'_{k,s})}$ an, und nach Definition von $f_{\delta}(P)$ nimmt die Funktion $(g_{\delta'}(P))_{\delta}$ in den Punkten von $\mathfrak{E}_k$ den konstanten Wert

$$\frac{1}{G(\mathfrak{E}_k)} \sum_s \int_{\mathfrak{E}'_{k,s}} \frac{\varphi(\mathfrak{E}'_{k,s})}{G(\mathfrak{E}'_{k,s})} G(d\mathfrak{E}) = \frac{1}{G(\mathfrak{E}_k)} \sum_s \frac{\varphi(\mathfrak{E}'_{k,s})}{G(\mathfrak{E}'_{k,s})} G(\mathfrak{E}'_{k,s}) = \frac{\varphi(\mathfrak{E}_k)}{G(\mathfrak{E}_k)}$$

an, d. h.

$$(g_{\delta'}(P))_{\delta} = g_{\delta}(P). \quad (80)$$

79. Hilfssätze (Fortsetzung). Wir führen einen neuen Begriff ein, der für unsere weiteren Überlegungen wichtig ist. Es sei $\varphi(\mathfrak{E}) \in V_1$ und $\psi(\mathfrak{E}) \in V_1$; $\mathfrak{E} = \mathfrak{E}' + \mathfrak{E}''$ sei eine Zerlegung von $\mathfrak{E}$ in zwei elementfremde Teilmengen. Wir bezeichnen mit $\inf[\varphi, \psi]$ die untere Grenze der Summen $\varphi(\mathfrak{E}') + \psi(\mathfrak{E}'')$ bei allen möglichen Zerlegungen von $\mathfrak{E}$:

$$\inf[\varphi, \psi] = \inf_{\mathfrak{E} = \mathfrak{E}' + \mathfrak{E}''} [\varphi(\mathfrak{E}') + \psi(\mathfrak{E}'')] = \omega(\mathfrak{E}). \quad (81)$$

Demnach ist

$$\varphi(\mathfrak{E}') + \psi(\mathfrak{E}'') \geq \omega(\mathfrak{E}), \quad (82)$$

und zu beliebig gegebenem positivem ε existiert eine solche Zerlegung $\mathfrak{E} = \mathfrak{E}' + \mathfrak{E}''$, daß

$$\varphi(\mathfrak{E}') + \psi(\mathfrak{E}'') \leq \omega(\mathfrak{E}) + \varepsilon \quad (83)$$

gilt.

Die Funktion $\omega(\mathfrak{E})$ hat für jedes $\mathfrak{E}$ aus $\mathfrak{E}_0$ einen endlichen Wert, weil $\varphi(\mathfrak{E}')$ und $\psi(\mathfrak{E}'')$ beschränkt sind. Nehmen wir nämlich für $\mathfrak{E}'$ die ganze Menge $\mathfrak{E}$ und für $\mathfrak{E}''$ die leere Menge oder umgekehrt für $\mathfrak{E}'$ die leere Menge und für $\mathfrak{E}''$ ganz $\mathfrak{E}$, so erhalten wir

$$\omega(\mathfrak{E}) \leq \varphi(\mathfrak{E}), \quad \omega(\mathfrak{E}) \leq \psi(\mathfrak{E}). \quad (84)$$

Wir beweisen nun, daß $\omega(\mathfrak{E})$ volladditiv ist. Gegeben sei eine Zerlegung von $\mathfrak{E}$ in endlich oder abzählbar unendlich viele paarweise punktfremde Mengen $\mathfrak{E}_k$. Es ist

$$\varphi(\mathfrak{E}) = \sum_k \varphi(\mathfrak{E}_k), \quad \psi(\mathfrak{E}) = \sum_k \psi(\mathfrak{E}_k),$$

wobei diese Reihen absolut konvergent sind. Nach (84) bilden die positiven $\omega(\mathfrak{E}_k)$ eine konvergente Reihe, weil die Reihen mit den Gliedern $\varphi(\mathfrak{E}_k)$ und $\psi(\mathfrak{E}_k)$ absolut konvergent sind; daher hat auch die ganze Reihe, die aus $\omega(\mathfrak{E}_k)$ besteht, eine wohlbestimmte Summe, die nicht von der Reihenfolge der Summanden abhängt [I, 142]. Es sei $\varepsilon > 0$ gegeben. Dann existiert eine solche Zerlegung $\mathfrak{E} = \mathfrak{E}' + \mathfrak{E}''$ derart, daß

$$\varphi(\mathfrak{E}') + \psi(\mathfrak{E}'') \leq \omega(\mathfrak{E}); \quad \varphi(\mathfrak{E}') + \psi(\mathfrak{E}'') \leq \omega(\mathfrak{E}) + \varepsilon \quad (85)$$

gilt. Wir setzen $\mathfrak{E}'_k = \mathfrak{E}_k \mathfrak{E}'$ und $\mathfrak{E}''_k = \mathfrak{E}_k \mathfrak{E}''$, so daß

$$\mathfrak{E}'_k + \mathfrak{E}''_k = \mathfrak{E}_k, \quad \sum_k \mathfrak{E}'_k = \mathfrak{E}', \quad \sum_k \mathfrak{E}''_k = \mathfrak{E}''$$

ist. Unter Benutzung der Definition von $\omega(\mathfrak{E})$ können wir $\varphi(\mathfrak{E}'_k) + \psi(\mathfrak{E}''_k) \leq \omega(\mathfrak{E}_k)$ schreiben, und summieren wir über k , so erhalten wir, da $\varphi(\mathfrak{E})$ und $\psi(\mathfrak{E})$ volladditiv sind,

$$\sum_k \omega(\mathfrak{E}_k) \leq \varphi(\mathfrak{E}') + \psi(\mathfrak{E}'');$$

die zweite Ungleichung von (85) liefert $\sum_k \omega(\mathfrak{E}_k) \leq \omega(\mathfrak{E}) + \varepsilon$. Hieraus folgt, weil ε beliebig war,

$$\sum_k \omega(\mathfrak{E}_k) \leq \omega(\mathfrak{E}). \quad (86)$$

Wir beweisen jetzt die entgegengesetzte Ungleichung. Dazu wählen wir eine solche Zerlegung $\mathfrak{E}_k = \mathfrak{E}_{k,1} + \mathfrak{E}_{k,2}$, daß die Ungleichung

$$\varphi(\mathfrak{E}_{k,1}) + \psi(\mathfrak{E}_{k,2}) < \omega(\mathfrak{E}_k) + \frac{\varepsilon}{2^k} \quad (87)$$

gilt. Wenn wir über k summieren und die Summe der $\mathfrak{E}_{k,1}$ mit $\mathfrak{E}_1$ und die der $\mathfrak{E}_{k,2}$ mit $\mathfrak{E}_2$ bezeichnen, erhalten wir

$$\varphi(\mathfrak{E}_1) + \psi(\mathfrak{E}_2) \leq \sum_k \omega(\mathfrak{E}_k) + \varepsilon, \quad (88)$$

wobei $\mathfrak{E}_1 \mathfrak{E}_2 = 0$ und $\mathfrak{E}_1 + \mathfrak{E}_2 = \mathfrak{E}$ ist. Nun ist aber $\varphi(\mathfrak{E}_1) + \psi(\mathfrak{E}_2) \leq \omega(\mathfrak{E})$ nach der Definition von $\omega(\mathfrak{E})$; daher liefert (88) die Ungleichung $\omega(\mathfrak{E}) \leq \sum_k \omega(\mathfrak{E}_k) + \varepsilon$, und weil ε beliebig war, erhalten wir eine zu (86) entgegengesetzte Ungleichung. Hieraus folgt die Volladditivität von $\omega(\mathfrak{E})$:

$$\omega(\mathfrak{E}) = \sum_k \omega(\mathfrak{E}_k).$$

Aus dem oben Gesagten folgt, daß die durch (81) definierte Funktion $\omega(\mathfrak{E})$ zu V_1 gehört.

Wir führen noch die folgende Bezeichnung für eine beliebige Funktion $\varphi(\mathfrak{E})$ aus V_1 ein:

$$\varphi_n(\mathfrak{E}) = \inf [\varphi, nG]. \quad (89)$$

Beachten wir, daß $G(\mathfrak{E}) \geq 0$ ist, so können wir behaupten, daß $\varphi_{n+1}(\mathfrak{E}) \geq \varphi_n(\mathfrak{E})$ und außerdem nach (84) $\varphi_n(\mathfrak{E}) \leq \varphi(\mathfrak{E})$ ist. Somit hat die Folge der $\varphi_n(\mathfrak{E})$ für jedes $\mathfrak{E}$, das zu $\mathfrak{E}_0$ gehört, für $n \rightarrow \infty$ einen endlichen Grenzwert.

Wir erinnern noch einmal an die Definition der *absoluten Stetigkeit*: Eine Funktion $\varphi(\mathfrak{E})$ heißt in $\mathfrak{E}_0$ *absolut stetig*, wenn zu jedem gegebenen positiven ε ein solches positives η existiert, daß $|\varphi(\mathfrak{E})| \leq \varepsilon$ ist, wenn $\mathfrak{E} \subset \mathfrak{E}_0$ gilt, $\mathfrak{E}$ zu B gehört und $G(\mathfrak{E}) \leq \eta$ ist.

Hilfssatz 1. *Wenn $\varphi(\mathfrak{E})$ in $\mathfrak{E}_0$ absolut stetig ist, gilt für jedes $\mathfrak{E}$*

$$\varphi_n(\mathfrak{E}) \rightarrow \varphi(\mathfrak{E}).$$

Beweis. Es sei $\varepsilon > 0$ gegeben. Für eine Zerlegung $\mathfrak{E} = \mathfrak{E}'_n + \mathfrak{E}''_n$ gilt nach (83)

$$\varphi(\mathfrak{E}'_n) + nG(\mathfrak{E}''_n) < \varphi_n(\mathfrak{E}) + \varepsilon \leq \varphi(\mathfrak{E}) + \varepsilon. \quad (90)$$

Aber nach Satz 1 aus [72] ist $\varphi(\mathfrak{E}'_n) \leq l$, wobei l eine bestimmte Zahl ist, und aus (90) folgt

$$G(\mathfrak{E}''_n) < \frac{\varphi(\mathfrak{E}) + \varepsilon - l}{n}, \quad (91)$$

wobei die rechte Seite für alle genügend großen n höchstens gleich η ist; hieraus erhalten wir wegen der absoluten Stetigkeit von $\varphi(\mathfrak{E})$

$$|\varphi(\mathfrak{E}''_n)| \leq \varepsilon$$

für genügend großes n . Die erste der Ungleichungen (90) liefert $\varphi(\mathfrak{E}'_n) \leq \varphi_n(\mathfrak{E}) + \varepsilon$. Wegen $\varphi(\mathfrak{E}'_n) = \varphi(\mathfrak{E}) - \varphi(\mathfrak{E}''_n)$ ist folglich $\varphi(\mathfrak{E}) \leq \varphi_n(\mathfrak{E}) + \varepsilon + \varphi(\mathfrak{E}''_n)$, und auf Grund von $|\varphi(\mathfrak{E}''_n)| \leq \varepsilon$ erhalten wir $\varphi(\mathfrak{E}) - \varphi_n(\mathfrak{E}) \leq 2\varepsilon$. Hieraus folgt, da ε beliebig war, $\varphi_n(\mathfrak{E}) \rightarrow \varphi(\mathfrak{E})$.

Hilfssatz 2. *Für eine nichtnegative volladditive Funktion $\varphi(\mathfrak{E})$ ist der Grenzwert von $\varphi_n(\mathfrak{E})$ eine volladditive, auf $\mathfrak{E}_0$ absolut stetige Funktion.*

Beweis. Wir setzen $\lim_{n \rightarrow \infty} \varphi_n(\mathfrak{E}) = \varphi^{(ac)}(\mathfrak{E})$. Wenn wir beachten, daß $\varphi_n(\mathfrak{E})$ volladditiv und nichtnegativ ist und das Lemma aus [63] gilt, können wir behaupten, daß $\varphi^{(ac)}(\mathfrak{E})$ volladditiv ist. Aus (84) folgt $0 \leq \varphi_n(\mathfrak{E}) \leq nG(\mathfrak{E})$, also ist jedes der $\varphi_n(\mathfrak{E})$ absolut stetig. Weiter folgt aus $\varphi^{(ac)}(\mathfrak{E}_0 - \mathfrak{E}) \geq \varphi_n(\mathfrak{E}_0 - \mathfrak{E})$ für $\mathfrak{E} \subset \mathfrak{E}_0$ die Ungleichung $\varphi^{(ac)}(\mathfrak{E}_0) - \varphi^{(ac)}(\mathfrak{E}) \geq \varphi_n(\mathfrak{E}_0) - \varphi_n(\mathfrak{E})$, d. h.

$$0 \leq \varphi^{(ac)}(\mathfrak{E}) - \varphi_n(\mathfrak{E}) \leq \varphi^{(ac)}(\mathfrak{E}_0) - \varphi_n(\mathfrak{E}_0). \quad (92)$$

Hieraus erkennt man, daß die Folge der $\varphi_n(\mathfrak{E})$ für alle $\mathfrak{E}$ aus $\mathfrak{E}_0$ gleichmäßig gegen $\varphi^{(ac)}(\mathfrak{E})$ strebt. Wenn wir die absolute Stetigkeit von $\varphi_n(\mathfrak{E})$ beachten, können wir auch die absolute Stetigkeit von $\varphi^{(ac)}(\mathfrak{E})$ auf $\mathfrak{E}_0$ beweisen. In der Tat, es sei ε eine gegebene positive Zahl. Wir können ein solches $n = n_0$ fest wählen, daß $\varphi^{(ac)}(\mathfrak{E}_0) - \varphi_{n_0}(\mathfrak{E}_0) \leq \varepsilon/2$ ist. Dann gilt nach (92) $\varphi^{(ac)}(\mathfrak{E}) \leq \varphi_{n_0}(\mathfrak{E}) + \varepsilon/2$. Da $\varphi_{n_0}(\mathfrak{E})$ absolut stetig ist, existiert ein solches positives η , daß $\varphi_{n_0}(\mathfrak{E}) \leq \varepsilon/2$ für $\mathfrak{E} \subset \mathfrak{E}_0$ und $G(\mathfrak{E}) \leq \eta$ gilt, und aus der oben hingeschriebenen Ungleichung folgt $\varphi^{(ac)}(\mathfrak{E}) \leq \varepsilon$ für $\mathfrak{E} \subset \mathfrak{E}_0$ und $G(\mathfrak{E}) \leq \eta$. Damit ist Hilfssatz 2 bewiesen.

Hilfssatz 3. *Ist $\varphi(\mathfrak{E}) \in V_1$ nichtnegativ und absolut stetig auf $\mathfrak{E}_0$, so existiert für ein beliebig gegebenes $\varepsilon > 0$ eine wesentlich beschränkte Funktion $\psi(\mathfrak{E})$ mit*

$$||\varphi - \psi||_1 \leq \varepsilon. \quad (93)$$

Beweis. Aus (84) folgt, wie wir bereits sagten, $0 \leq \varphi_n(\mathfrak{E}) \leq nG(\mathfrak{E})$, d. h., jede der Funktionen $\varphi_n(\mathfrak{E})$ ist wesentlich beschränkt. Außerdem gilt $\varphi(\mathfrak{E}) - \varphi_n(\mathfrak{E}) \geq 0$; folg-

lich ist die Totalvariation dieser Differenz auf $\mathfrak{E}_0$ gleich ihrem Wert für $\mathfrak{E}_0$, d. h.

$$\|\varphi(\mathfrak{E}) - \varphi_n(\mathfrak{E})\|_1 = \varphi(\mathfrak{E}_0) - \varphi_n(\mathfrak{E}_0).$$

Da aber $\varphi(\mathfrak{E})$ absolut stetig ist, gilt $\varphi_n(\mathfrak{E}_0) \rightarrow \varphi(\mathfrak{E}_0)$ auf Grund von Hilfssatz 1, d. h.

$$\|\varphi(\mathfrak{E}) - \varphi_n(\mathfrak{E})\|_1 \rightarrow 0,$$

und um die Ungleichung (93) zu befriedigen, genügt es, für $\varphi(\mathfrak{E})$ eine Funktion $\varphi_n(\mathfrak{E})$ mit genügend großem n zu nehmen. Die Ungleichung (93) kann befriedigt werden, wenn wir für $\varphi(\mathfrak{E})$ eine auf $\mathfrak{E}_0$ stückweise konstante Funktion wählen.

Vorläufig beweisen wir das für die Funktion $\varphi(\mathfrak{E})$ aus V_2 .

Satz 1. Ist $\varphi(\mathfrak{E}) \in V_2$, so existiert für ein beliebig gegebenes $\varepsilon > 0$ eine solche stückweise konstante Funktion $\omega(\mathfrak{E})$, daß

$$\|\varphi - \omega\|_2 \leq \varepsilon \quad (94)$$

gilt.

Beweis. Es sei $f(P)$ eine meßbare Funktion aus $L_2(\mathfrak{E}_0)$, und es gelte

$$a = \frac{1}{G(\mathfrak{E})} \int_{\mathfrak{E}} f(P) G(d\mathfrak{E}),$$

wobei dieser Ausdruck gleich 0 gesetzt wird, wenn $G(\mathfrak{E}) = 0$ ist. Offensichtlich gilt

$$\int_{\mathfrak{E}} [f(P) - a]^2 G(d\mathfrak{E}) = \int_{\mathfrak{E}} f^2(P) G(d\mathfrak{E}) - a^2 G(\mathfrak{E}).$$

Wir setzen in dieser Beziehung $\mathfrak{E} = \mathfrak{E}_k$ und $a = \int_{\mathfrak{E}_k} f(P) \frac{G(d\mathfrak{E})}{G(\mathfrak{E}_k)}$, wobei $\mathfrak{E}_k$ die Teilmengen der Menge $\mathfrak{E}_0$ sind, die durch die Zerlegung δ entstehen; summieren wir über k , so erhalten wir

$$\|f(P) - f_\delta(P)\|_{\mathfrak{E}_0}^2 = \|f(P)\|_{\mathfrak{E}_0}^2 - \|f_\delta(P)\|_{\mathfrak{E}_0}^2. \quad (95)$$

Wenn $f(P) = g_{\delta'}(P)$ gesetzt wird, wobei $\delta' > \delta$ sei und $g_\delta(P)$ die Funktion aus (77) ist, erhalten wir auf Grund von (80) die Beziehung $f_\delta(P) = g_\delta(P)$, und (95) liefert

$$\|g_{\delta'}(P) - g_\delta(P)\|_{\mathfrak{E}_0}^2 = \|g_{\delta'}(P)\|_{\mathfrak{E}_0}^2 - \|g_\delta(P)\|_{\mathfrak{E}_0}^2. \quad (96)$$

Wenn wir (78) und (79) beachten, können wir

$$\|(\varphi - \varphi_\delta)_{\delta'}\|_2 = \|\varphi_{\delta'}\|_2 - \|\varphi_\delta\|_2 \quad (97)$$

schreiben.

Nach (76) existieren solche Folgen von Zerlegungen δ_n und δ'_n , daß

$$\|(\varphi - \varphi_\delta)_{\delta_n}\|_2 \rightarrow \|\varphi - \varphi_\delta\|_2 \quad \text{und} \quad \|\varphi_{\delta'_n}\|_2 \rightarrow \|\varphi\|_2$$

ist. Für die Folge $\delta''_n = \delta_n \delta'_n$ gilt erst recht

$$\|(\varphi - \varphi_\delta)_{\delta''_n}\|_2 \rightarrow \|\varphi - \varphi_\delta\|_2 \quad \text{und} \quad \|\varphi_{\delta''_n}\|_2 \rightarrow \|\varphi\|_2.$$

Wenn wir in (97) $\delta' = \delta''_n$ setzen und zur Grenze übergehen, erhalten wir

$$\|\varphi - \varphi_\delta\|_2 = \|\varphi\|_2 - \|\varphi_\delta\|_2. \quad (98)$$

Wenn wir wieder (76) beachten, können wir behaupten, daß eine solche Zerlegung δ existiert, daß die rechte Seite von (98) höchstens gleich ε ist, und setzen wir $\omega(\mathfrak{E}) = \varphi_\delta(\mathfrak{E})$, so erhalten wir (94).

Satz 2. Ist $\varphi(\mathfrak{E}) \in V_1$ und auf $\mathfrak{E}_0$ absolut stetig, so existiert für ein beliebig gegebenes $\varepsilon > 0$ eine solche auf $\mathfrak{E}_0$ stückweise konstante Funktion $\omega(\mathfrak{E})$, daß

$$\|\varphi - \omega\|_1 \leq \varepsilon \quad (99)$$

gilt.

Beweis. Wir können φ als Differenz zweier nichtnegativer Funktionen aus V_1 darstellen: $\varphi(\mathfrak{E}) = \varphi_1(\mathfrak{E}) - \varphi_2(\mathfrak{E})$, und wenn zwei stückweise konstante Funktionen $\omega_1(\mathfrak{E})$ und $\omega_2(\mathfrak{E})$ existieren derart, daß $\|\varphi_1 - \omega_1\|_1 \leq \varepsilon/2$ und $\|\varphi_2 - \omega_2\|_1 \leq \varepsilon/2$ ist, erhalten wir nach (59), wenn wir die stückweise konstante Funktion $\omega(\mathfrak{E}) = \omega_1(\mathfrak{E}) + \omega_2(\mathfrak{E})$ einführen,

$$\|\varphi - \omega\|_1 \leq \|\varphi_1 - \omega_1\|_1 + \|\varphi_2 - \omega_2\|_1 \leq \varepsilon,$$

und somit genügt es, den Satz für den Fall nichtnegativer Funktionen $\varphi(\mathfrak{E})$ zu beweisen.

Wegen Hilfssatz 3 existiert eine solche wesentlich beschränkte Funktion $\psi(\mathfrak{E})$, daß $\|\varphi - \psi\|_1 \leq \varepsilon/2$ gilt. Die Funktion $\psi(\mathfrak{E})$ ist in V_2 enthalten [78]; folglich existiert eine solche stückweise konstante Funktion $\omega(\mathfrak{E})$, daß $\|\psi - \omega\|_2 \leq \varepsilon^2/4G(\mathfrak{E}_0)$ ist. Nach (63) gilt $\|\psi - \omega\|_1 \leq \varepsilon/2$, und wenn wir (59) benutzen, können wir

$$\|\varphi - \omega\|_1 \leq \|\varphi - \psi\|_1 + \|\psi - \omega\|_1 \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

schreiben. Damit ist der Satz bewiesen.

Die Behauptung von Satz 1 läuft darauf hinaus, daß in V_2 bei der Norm $\|\varphi\|_2$ die stückweise konstanten Funktionen überall dicht liegen, und Satz 2 besagt, daß in dem Raum der absolut stetigen Funktionen aus V_1 bei der Norm $\|\varphi\|_1$ die stückweise konstanten Funktionen überall dicht liegen.

80. Hauptsatz. Wir beweisen jetzt den grundlegenden Satz, den wir schon in [73] formuliert haben. Wie schon oben, genügt es, den Satz für den Fall einer nichtnegativen Funktion $\varphi(\mathfrak{E})$ aus V_1 zu beweisen. Wir bezeichnen wieder mit $\varphi^{(ac)}(\mathfrak{E})$ den Grenzwert von $\varphi_n(\mathfrak{E})$, der durch die Formel (89) definiert ist. Die Beziehung $\varphi^{(ac)}(\mathfrak{E}) = \varphi(\mathfrak{E})$ gilt auf Grund der Hilfssätze 1 und 2 genau dann, wenn $\varphi(\mathfrak{E})$ auf $\mathfrak{E}_0$ absolut stetig ist. Es sei nun $\varphi(\mathfrak{E})$ nicht absolut stetig. Wir bilden eine auf $\mathfrak{E}_0$ nichtnegative volladditive Funktion

$$\varphi^{(s)}(\mathfrak{E}) = \varphi(\mathfrak{E}) - \varphi^{(ac)}(\mathfrak{E}). \quad (100)$$

Wir wollen zeigen, daß $\varphi^{(s)}(\mathfrak{E})$ der singuläre Summand in Formel (14) aus [73] ist. Dazu bilden wir

$$\varphi_n^{(s)}(\mathfrak{E}) = \inf [\varphi^{(s)}, nG] = \inf_{\mathfrak{E} = \mathfrak{E}' + \mathfrak{E}''} [\varphi(\mathfrak{E}') - \varphi^{(ac)}(\mathfrak{E}') + nG(\mathfrak{E}'')] . \quad (101)$$

Wir erinnern an die Zerlegung $\mathfrak{E} = \mathfrak{E}'_n + \mathfrak{E}''_n$, die der Bedingung (90) genügt, und an die Ungleichung (91), nach der $G(\mathfrak{E}''_n) \rightarrow 0$ für $n \rightarrow \infty$ gilt. Wenn wir (101) berücksichtigen, können wir

$$\varphi_n^{(s)}(\mathfrak{E}) \leq \varphi(\mathfrak{E}'_n) + nG(\mathfrak{E}''_n) - \varphi^{(ac)}(\mathfrak{E}'_n) = \varphi(\mathfrak{E}'_n) + nG(\mathfrak{E}''_n) - [\varphi^{(ac)}(\mathfrak{E}) - \varphi^{(ac)}(\mathfrak{E}''_n)]$$

schreiben, d. h., nach (90) gilt

$$\varphi_n^{(s)}(\mathfrak{E}) \leq \varphi_n(\mathfrak{E}) + \varepsilon - [\varphi^{(ac)}(\mathfrak{E}) - \varphi^{(ac)}(\mathfrak{E}''_n)] .$$

Unter Berücksichtigung der absoluten Stetigkeit von $\varphi^{(ac)}(\mathfrak{E})$ folgt $0 \leq \varphi^{(ac)}(\mathfrak{E}''_n) \leq \varepsilon$ für alle genügend großen n und folglich

$$\varphi_n^{(s)}(\mathfrak{E}) \leq \varphi_n(\mathfrak{E}) - \varphi^{(ac)}(\mathfrak{E}) + 2\varepsilon .$$

Wenn wir $\varphi_n(\mathfrak{E}) \rightarrow \varphi^{(ac)}(\mathfrak{E})$ beachten, erhalten wir $\lim_{n \rightarrow \infty} \varphi_n^{(s)}(\mathfrak{E}) \leq 2\varepsilon$ oder, weil ε beliebig ist, $\lim_{n \rightarrow \infty} \varphi_n^{(s)}(\mathfrak{E}) = 0$ für jedes $\mathfrak{E}$. Da $\varphi_n^{(s)}(\mathfrak{E})$ nichtnegativ und für wachsendes n nichtfallend ist, gilt für jedes n

$$\varphi_n^{(s)}(\mathfrak{E}) = \inf_{\mathfrak{E} = \mathfrak{E}' + \mathfrak{E}''} [\varphi^{(s)}(\mathfrak{E}') + nG(\mathfrak{E}'')] = 0.$$

Wenn wir dies auf $\mathfrak{E}_0$ für $n=1$ anwenden, so können wir behaupten, daß für ein beliebig gegebenes $\varepsilon > 0$ solche Mengen $\mathfrak{E}_n$ aus $\mathfrak{E}_0$ und B existieren, daß

$$\varphi^{(s)}(\mathfrak{E}_n) + G(\mathfrak{E}_0 - \mathfrak{E}_n) \leq \frac{\varepsilon}{2^n}$$

gilt. Da $\varphi^{(s)}(\mathfrak{E})$ und $G(\mathfrak{E})$ nichtnegativ sind, folgt hieraus

$$\varphi^{(s)}(\mathfrak{E}_n) \leq \frac{\varepsilon}{2^n}, \quad G(\mathfrak{E}_0 - \mathfrak{E}_n) \leq \frac{\varepsilon}{2^n}. \quad (102)$$

Wir bilden die Menge $\mathfrak{E}'_e = \mathfrak{E}_1 + \mathfrak{E}_2 + \dots$, die zu $\mathfrak{E}_0$ und B gehört. Wenn wir beachten, daß $\mathfrak{E}_n \subset \mathfrak{E}'_e$ für beliebiges n ist, können wir $G(\mathfrak{E}_0 - \mathfrak{E}'_e) \leq \varepsilon/2^n$ schreiben, und wenn n gegen ∞ strebt, erhalten wir $G(\mathfrak{E}_0 - \mathfrak{E}'_e) = 0$. Andererseits folgt $\varphi^{(s)}(\mathfrak{E}'_e) \leq \varepsilon$, wenn wir beachten, daß

$$\mathfrak{E}'_e = \mathfrak{E}_1 + (\mathfrak{E}_2 - \mathfrak{E}_1) + (\mathfrak{E}_3 - \mathfrak{E}_1 - \mathfrak{E}_2) + \dots$$

ist, die erste der Ungleichungen (102) gilt und $\varphi^{(s)}(\mathfrak{E})$ nichtnegativ ist. Also gilt

$$\varphi^{(s)}(\mathfrak{E}'_e) \leq \varepsilon \quad \text{und} \quad G(\mathfrak{E}_0 - \mathfrak{E}'_e) = 0.$$

Wenn wir $\mathfrak{E}_0 - \mathfrak{E}'_e = \mathfrak{E}_e$ setzen, können wir behaupten, daß für ein beliebiges $\varepsilon > 0$ eine solche Menge $\mathfrak{E}_e$ aus $\mathfrak{E}_0$ existiert, daß

$$\varphi^{(s)}(\mathfrak{E}_0 - \mathfrak{E}_e) \leq \varepsilon \quad \text{und} \quad G(\mathfrak{E}_e) = 0$$

ist. Wir nehmen $\varepsilon_n \rightarrow 0$, $\varepsilon_n > 0$. Wir können dann

$$\varphi(\mathfrak{E}_0 - \mathfrak{E}_{e_n}) \leq \varepsilon_n \quad \text{und} \quad G(\mathfrak{E}_{e_n}) = 0$$

schreiben. Führen wir nun die Menge $H = \mathfrak{E}_{e_1} + \mathfrak{E}_{e_2} + \dots$ ein, so ist $G(H) = 0$, und wenn wir beachten, daß $\mathfrak{E}_{e_n} \subset H$ für jedes n ist, erhalten wir $\varphi^{(s)}(\mathfrak{E}_0 - H) \leq \varepsilon_n$ auf Grund der vorletzten Ungleichung; für $n \rightarrow \infty$ folgt hieraus $\varphi^{(s)}(\mathfrak{E}_0 - H) = 0$. Es existiert also ein $H \in B$, für welches

$$\varphi^{(s)}(\mathfrak{E}_0 - H) = 0, \quad G(H) = 0 \quad (103)$$

gilt. Nun ist $\mathfrak{E} = \mathfrak{E}H + (\mathfrak{E} - \mathfrak{E}H)$ und $\mathfrak{E} - \mathfrak{E}H \subset \mathfrak{E}_0 - H$, so daß aus der ersten Formel (103), da $\varphi^{(s)}(\mathfrak{E})$ positiv ist, $\varphi^{(s)}(\mathfrak{E} - \mathfrak{E}H) = 0$ folgt. Also ist $\varphi^{(s)}(\mathfrak{E}) = \varphi^{(s)}(\mathfrak{E}H)$. Somit existiert ein H , für welches

$$\varphi^{(s)}(\mathfrak{E}) = \varphi^{(s)}(\mathfrak{E}H) \quad \text{und} \quad G(H) = 0$$

gilt. Also ist $\varphi^{(s)}(\mathfrak{E})$ der singuläre Teil in (14) aus [73], und wenn wir die Formel (100) und die absolute Stetigkeit von $\varphi^{(ac)}(\mathfrak{E})$ beachten, so bleibt zum Beweis des Satzes aus [73] nur noch der folgende Satz nachzutragen:

Satz 3. Jede absolut stetige Funktion $\varphi(\mathfrak{E})$ aus V_1 kann als Integral

$$\varphi(\mathfrak{E}) = \int_{\mathfrak{E}} f(P) G(d\mathfrak{E}) \quad (104)$$

dargestellt werden, wobei $f(P)$ meßbar und auf $\mathfrak{E}_0$ summierbar ist.

Beweis. Nach Satz 2 existieren solche stückweise konstanten Funktionen $\omega_n(\mathfrak{E})$, daß

$$\|\varphi - \omega_n\|_1 \leq \frac{1}{2^{n+1}} \quad (105)$$

ist. Hieraus folgt nach (59)

$$\|\omega_{n+1} - \omega_n\|_1 \leq \|\varphi - \omega_{n+1}\|_1 + \|\varphi - \omega_n\|_1 < \frac{1}{2^n}.$$

Jedes $\omega_n(\mathfrak{E})$ ist aber Integral einer stückweise konstanten Funktion $g_n(P)$ mit endlich vielen endlichen Werten in $\mathfrak{E}_0$ [78]:

$$\omega_n(\mathfrak{E}) = \int_{\mathfrak{E}} g_n(P) G(d\mathfrak{E})$$

und

$$\omega_{n+1}(\mathfrak{E}) - \omega_n(\mathfrak{E}) = \int_{\mathfrak{E}} [g_{n+1}(P) - g_n(P)] G(d\mathfrak{E}).$$

Die Totalvariation dieser durch das Integral dargestellten Mengenfunktion läßt sich durch

$$\|\omega_{n+1} - \omega_n\|_1 = \int_{\mathfrak{E}_0} |g_{n+1}(P) - g_n(P)| G(d\mathfrak{E})$$

ausdrücken [73]. Hieraus folgt

$$\int_{\mathfrak{E}_0} |g_{n+1}(P) - g_n(P)| G(d\mathfrak{E}) \leq \frac{1}{2^n}, \quad (106)$$

und die Reihe aus diesen Integralen ist konvergent. Dann konvergiert die Reihe [54]

$$|g_1(P)| + |g_2(P) - g_1(P)| + |g_3(P) - g_2(P)| + \dots \quad (107)$$

fast überall in $\mathfrak{E}_0$, und somit konvergiert die Reihe

$$f(P) = g_1(P) + [g_2(P) - g_1(P)] + [g_3(P) - g_2(P)] + \dots$$

erst recht fast überall, d. h., fast überall auf $\mathfrak{E}_0$ gilt $g_n(P) \rightarrow g(P)$. Die Summe von (107) ist nach der Abschätzung (106) eine auf $\mathfrak{E}_0$ summierbare Funktion; da $|f(P)|$ aber höchstens gleich dieser Summe ist, muß $f(P)$ ebenfalls summierbar sein. Wir können

$$f(P) = g_n(P) + [g_{n+1}(P) - g_n(P)] + [g_{n+2}(P) - g_{n+1}(P)] + \dots$$

schreiben, und beachten wir (106), so erhalten wir für jedes $\mathfrak{E}$ aus $\mathfrak{E}_0$

$$\int_{\mathfrak{E}} |f(P) - g_n(P)| G(d\mathfrak{E}) \leq \frac{1}{2^n} + \frac{1}{2^{n+1}} + \dots = \frac{1}{2^{n-1}}.$$

Hieraus folgt

$$\lim_{n \rightarrow \infty} \omega_n(\mathfrak{E}) = \lim_{n \rightarrow \infty} \int_{\mathfrak{E}} g_n(P) G(d\mathfrak{E}) = \int_{\mathfrak{E}} f(P) G(d\mathfrak{E})$$

für jedes $\mathfrak{E}$. Andererseits folgt aus der Definition der Totalvariation über $\mathfrak{E}_0$ und (105)

$$|\varphi(\mathfrak{E}) - \omega_n(\mathfrak{E})| \leq \|\varphi - \omega_n\|_1 \leq \frac{1}{2^{n+1}}$$

und hieraus für jedes $\mathfrak{E}$

$$\omega_n(\mathfrak{E}) \rightarrow \varphi(\mathfrak{E}), \quad \text{d. h.} \quad \varphi(\mathfrak{E}) = \int_{\mathfrak{E}} f(P) G(d\mathfrak{E}).$$

Damit ist der Satz bewiesen. Beim Beweis haben wir angenommen, daß die absolut stetige Funktion $\varphi(\mathfrak{E})$ auf dem Körper B definiert ist. Aber die Formel (104) gibt die Möglichkeit, sie auf die Menge L_G aus $\mathfrak{E}_0$ fortzusetzen.

Bisher haben wir vorausgesetzt, daß $G(\mathfrak{E}_0)$ endlich ist. Ist $G(\mathfrak{E}_0) = \infty$, so erhält man das Ergebnis durch Grenzübergang bei den Mengen $\mathfrak{E}_n$ mit endlichem $G(\mathfrak{E}_n)$, für die der Satz schon bewiesen ist, wobei $f(P)$ nicht von n abhängt.

81. Das Hellingersche Integral. Wir untersuchen jetzt ausführlicher die Familie V_2 . Wir bemerken zuerst, daß aus der Bedingung (60) die absolute Stetigkeit von $\varphi(\mathfrak{E})$ folgt, wenn $\varphi(\mathfrak{E})$ zu V_2 gehört. Dabei erhalten wir nach (98)

$$\|\varphi - \varphi_{\delta}\|_2 = \|\varphi\|_2 - \|\varphi_{\delta}\|_2 \quad (108)$$

und außerdem [78]

$$\varphi_{\delta}(\mathfrak{E}) = \int_{\mathfrak{E}} g_{\delta}(P) G(d\mathfrak{E}) \quad \text{und} \quad \|\varphi_{\delta}\|_2 = \int_{\mathfrak{E}} g_{\delta}^2(P) G(d\mathfrak{E}). \quad (109)$$

Nach (76) können wir eine Folge von Zerlegungen von $\mathfrak{E}_0$ so wählen, daß die Ungleichung

$$\|\varphi - \varphi_{\delta_n}\|_2 = \|\varphi\|_2 - \|\varphi_{\delta_n}\|_2 \leq \frac{1}{4^{n+1}G(\mathfrak{E}_0)} \quad (110)$$

gilt. Dabei ist wegen (62) $\|\varphi - \varphi_{\delta_n}\|_1 \leq \frac{1}{2^{n+1}}$, und nach dem Beweis von Satz 3 aus [80] folgt $g_{\delta_n}(P) \rightarrow f(P)$ fast überall auf $\mathfrak{E}_0$ und

$$\varphi(\mathfrak{E}) = \int_{\mathfrak{E}} f(P) G(d\mathfrak{E}). \quad (111)$$

Aus (100) und (110) ergibt sich dann

$$\int_{\mathfrak{E}_0} g_{\delta_n}^2(P) G(d\mathfrak{E}) \leq \|\varphi\|_2 \quad \text{und} \quad \int_{\mathfrak{E}_0} g_{\delta_n}^2(P) G(d\mathfrak{E}) \rightarrow \|\varphi\|_2.$$

Wenn wir den Satz 4 aus [54] benutzen, so können wir

$$\int_{\mathfrak{E}_0} f^2(P) G(d\mathfrak{E}) \leq \|\varphi\|_2 \quad (112)$$

schreiben. Hieraus folgt $f(P) \in L_2(\mathfrak{E}_0)$. Wir beweisen jetzt die zu (112) entgegengesetzte Ungleichung.

Wenden wir auf (111) die Schwarzsche Ungleichung an, so gilt

$$\varphi^2(\mathfrak{E}_k) \leq \int_{\mathfrak{E}_k} f^2(P) G(d\mathfrak{E}) \int_{\mathfrak{E}_k} G(d\mathfrak{E}) = G(\mathfrak{E}_k) \int_{\mathfrak{E}_k} f^2(P) G(d\mathfrak{E}).$$

Wenn wir durch $G(\mathfrak{E}_k)$ dividieren und über k summieren, erhalten wir

$$S_{\delta}(\varphi) = \sum_k \frac{\varphi^2(\mathfrak{E}_k)}{G(\mathfrak{E}_k)} \leq \int_{\mathfrak{E}_0} f^2(P) G(d\mathfrak{E}); \quad (113)$$

daher gilt für die obere Grenze dieser Summen

$$\|\varphi\|_2 \leq \int_{\mathfrak{E}_0} f^2(P) G(d\mathfrak{E}).$$

Der Vergleich mit (112) führt auf

$$\|\varphi\|_2 = \int_{\mathfrak{E}_0} f^2(P) G(d\mathfrak{E}). \quad (114)$$

Wir zeigten, daß für jedes $\varphi(\mathfrak{E})$ aus V_2 die Funktion $f(P)$, die in (111) auftritt, zu L_2 gehört und daß (114) gilt. Wenn umgekehrt bekannt ist, daß sich $\varphi(\mathfrak{E})$ durch (111) mit $f(P) \in L_2(\mathfrak{E}_0)$ darstellen läßt, folgt aus der Abschätzung (113), die auf Grund von (111) gilt, $\varphi(\mathfrak{E}) \in V_2$. Wenn wir die Eindeutigkeit der Darstellung (111) beachten, so gilt auch (114). Wir kommen somit zu dem folgenden wichtigen Satz:

Satz 1. *Eine Funktion $\varphi(\mathfrak{E})$ gehört dann und nur dann auf $\mathfrak{E}_0$ zu V_2 , wenn sie durch (111) mit $f(P) \in L_2(\mathfrak{E}_0)$ darstellbar ist. Ist diese Bedingung erfüllt, so gilt (114).*

Wir geben noch eine andere notwendige und hinreichende Bedingung dafür an, daß $\varphi(\mathfrak{E})$ zu V_2 gehört.

Satz 2. *Eine Funktion $\varphi(\mathfrak{E})$ gehört dann und nur dann auf $\mathfrak{E}_0$ zu V_2 , wenn eine auf $\mathfrak{E}$ volladditive und nichtnegative Funktion $H(\mathfrak{E})$ existiert, für welche die Ungleichung*

$$\varphi^2(\mathfrak{E}) \leq G(\mathfrak{E}) H(\mathfrak{E}) \quad (115)$$

erfüllt ist.

Beweis. Ist diese Bedingung erfüllt, so sind die Summen $S_\delta(\varphi)$ beschränkt:

$$S_\delta(\varphi) = \sum_k \frac{\varphi^2(\mathfrak{E}_k)}{G(\mathfrak{E}_k)} \leq \sum_k H(\mathfrak{E}_k) = H(\mathfrak{E}_0).$$

Gilt umgekehrt $\varphi(\mathfrak{E}) \in V_2$, so gilt (111), und $f(P)$ gehört zu $L_2(\mathfrak{E}_0)$ und folglich auch zu L_2 auf einer beliebigen, bezüglich $G(\mathfrak{E})$ meßbaren Teilmenge von $\mathfrak{E}_0$. Wir setzen

$$H(\mathfrak{E}) = \int_{\mathfrak{E}} f^2(P) g(d\mathfrak{E}). \quad (116)$$

Wenn wir auf (111) die Schwarzsche Ungleichung anwenden, erhalten wir (115). Damit ist der Satz bewiesen.

Wenn $\varphi(\mathfrak{E}) \in V_2$ auf $\mathfrak{E}_0$ ist, nennt man die obere Grenze der Summen $S_\delta(\varphi)$ das *Hellingersche Integral* und bezeichnet es mit

$$\|\varphi\|_2 = \sup_{\delta} S_\delta(\varphi) = \int_{\mathfrak{E}_0} \frac{\varphi^2(d\mathfrak{E})}{G(d\mathfrak{E})}. \quad (117)$$

Die Formel (114) liefert dabei die Transformation des Hellingerschen Integrals in ein Lebesguesches Integral:

$$\int_{\mathfrak{E}_0} \frac{\varphi^2(d\mathfrak{E})}{G(d\mathfrak{E})} = \int_{\mathfrak{E}_0} f^2(P) G(d\mathfrak{E}). \quad (118)$$

Wenn die Zerlegung δ' eine Fortsetzung von δ und i das Integral (117) ist, d. h. die obere Grenze der Summen $S_\delta(\varphi)$, so gilt, wie wir wissen, $S_\delta(\varphi) \leq S_{\delta'}(\varphi) \leq i$. Wenn wir dies beachten, können wir behaupten, daß das Integral (117) die folgende Eigenschaft besitzt: *Für jedes gegebene $\varepsilon > 0$ existiert eine Zerlegung δ_ε derart, daß für jede Fortsetzung δ' von δ_ε die Ungleichung*

$$|i - S_{\delta'}(\varphi)| \leq \varepsilon \quad (119)$$

gilt.

Wir zeigen, daß es nur eine Zahl i mit dieser Eigenschaft gibt. Es möge noch die Zahl i' mit dieser Eigenschaft geben. Neben (119) gilt dann $|i' - S_{\delta'}(\varphi)| \leq \varepsilon$, wobei δ'_ε die Rolle von δ_ε in (119) spielt. Wenn wir das Produkt $\delta''_\varepsilon = \delta_\varepsilon \delta'_\varepsilon$ bilden, können wir die beiden Ungleichungen

$$|i - S_\delta(\varphi)| \leq \varepsilon \quad \text{und} \quad |i' - S_\delta(\varphi)| \leq \varepsilon$$

bilden (δ ist eine Fortsetzung von δ'_i). Wegen $i' - i = (i' - S_\delta(\varphi)) + (S_\delta(\varphi) - i)$ erhalten wir für δ

$$|i' - i| \leq |i' - S_\delta(\varphi)| + |i - S_\delta(\varphi)| \leq 2\varepsilon;$$

hieraus folgt aber $i' = i$, da ε beliebig war.

Es sei nun $\varphi(\mathfrak{E}) \in V_2$ und $\varphi_1(\mathfrak{E}) \in V_2$; wir betrachten die Summe

$$S_\delta(\varphi, \varphi_1) = \sum_k \frac{\varphi(\mathfrak{E}_k) \varphi_1(\mathfrak{E}_k)}{G(\mathfrak{E}_k)}, \quad (120)$$

die wir offenbar in der Form

$$S_\delta(\varphi, \varphi_1) = \frac{1}{2} \sum_k \frac{[\varphi(\mathfrak{E}_k) + \varphi_1(\mathfrak{E}_k)]^2}{G(\mathfrak{E}_k)} - \frac{1}{2} \sum_k \frac{\varphi^2(\mathfrak{E}_k)}{G(\mathfrak{E}_k)} - \frac{1}{2} \sum_k \frac{\varphi_1^2(\mathfrak{E}_k)}{G(\mathfrak{E}_k)} \quad (121)$$

darstellen können, d. h.

$$S_\delta(\varphi, \varphi_1) = \frac{1}{2} S_\delta(\varphi + \varphi_1) - \frac{1}{2} S_\delta(\varphi) - \frac{1}{2} S_\delta(\varphi_1).$$

Für jede der rechts stehenden Summen gilt (119), wobei wir ein und dasselbe δ nehmen können, da verschiedene δ durch ihr Produkt ersetzt werden können. Somit gilt auch für die Summen $S_\delta(\varphi, \varphi_1)$ die Eigenschaft (119); die entsprechende Zahl i für die Summen (120) lautet

$$i = \int_{\mathfrak{E}_0} \frac{\varphi(d\mathfrak{E}) \varphi_1(d\mathfrak{E})}{G(d\mathfrak{E})}. \quad (122)$$

Unter Beachtung von (121) können wir

$$\int_{\mathfrak{E}_0} \frac{\varphi(d\mathfrak{E}) \varphi_1(d\mathfrak{E})}{G(d\mathfrak{E})} = \frac{1}{2} \int_{\mathfrak{E}_0} \frac{[\varphi(d\mathfrak{E}) + \varphi_1(d\mathfrak{E})]^2}{G(d\mathfrak{E})} - \frac{1}{2} \int_{\mathfrak{E}_0} \frac{\varphi^2(d\mathfrak{E})}{G(d\mathfrak{E})} - \frac{1}{2} \int_{\mathfrak{E}_0} \frac{\varphi_1^2(d\mathfrak{E})}{G(d\mathfrak{E})}$$

schreiben oder, wenn wir (118) beachten,

$$\int_{\mathfrak{E}_0} \frac{\varphi(d\mathfrak{E}) \varphi_1(d\mathfrak{E})}{G(d\mathfrak{E})} = \int_{\mathfrak{E}_0} f(P) f_1(P) G(d\mathfrak{E}), \quad (123)$$

wobei $f_1(P)$ eine Punktfunktion aus L_2 ist, die der Funktion $\varphi_1(\mathfrak{E})$ entspricht:

$$\varphi_1(\mathfrak{E}) = \int_{\mathfrak{E}} f_1(P) G(d\mathfrak{E}). \quad (124)$$

Die allgemeineren Summen der Form

$$\sigma_\delta = \sum_k u(P_k) \frac{\varphi(\mathfrak{E}_k) \varphi_1(\mathfrak{E}_k)}{G(\mathfrak{E}_k)}, \quad (125)$$

wobei $u(P)$ beschränkt und bezüglich $G(\mathfrak{E})$ meßbar und P_k ein beliebiger Punkt aus $\mathfrak{E}_k$ ist, lassen sich genauso untersuchen. Wir können zeigen, daß auch für diese Summen eine eindeutig bestimmte Zahl i existiert, die die Eigenschaft (119) besitzt, in der S_δ durch σ_δ ersetzt werden muß, wobei die Ungleichung (119) für jedes P_k erfüllt ist. Diese Zahl i kann durch ein Lebesgue-Stieltjesches Integral ausgedrückt werden:

$$i = \int_{\mathfrak{E}_0} u(P) f(P) f_1(P) G(d\mathfrak{E}). \quad (126)$$

Die Eigenschaft (119) liegt der allgemeinen Definition des Integrals zugrunde. In dem folgenden Abschnitt werden wir den Fall einer Veränderlichen ausführlicher betrachten.

82. Der Fall einer Veränderlichen. Bei der Untersuchung des Falles einer Veränderlichen gehen wir von Punktfunktionen aus und betrachten den einfachsten Fall stetiger Funktionen. Zur Abkürzung werden wir folgende Bezeichnung benutzen: Ist Δ das Intervall $[\alpha, \beta]$, so bezeichne das Symbol $\Delta\tau(x)$ die Differenz $\tau(\beta) - \tau(\alpha)$. Es sei $g(x)$ eine nichtfallende und in dem endlichen Intervall $[a, b]$ stetige Funktion, und $F(x)$ sei eine in diesem Intervall reellwertige und stetige Funktion, die die Eigenschaft besitzt, daß $\Delta F(x) = 0$ für $\Delta g(x) = 0$ ist. Es sei δ eine Zerlegung des Intervalls $[a, b]$ in endlich viele Teilintervalle Δ_k , und es sei

$$S_\delta = \sum_k \frac{(\Delta_k F)^2}{\Delta_k g}. \quad (127)$$

Summanden der Form $\frac{0}{0}$ setzen wir gleich 0. Diese Summe nimmt nicht ab, wenn neue Teilpunkte hinzugefügt werden [78]. Wir benutzen nur Zerlegungen in Intervalle, und wir können die Beweise der Sätze analog denjenigen durchführen, die wir bei Zerlegungen in bezüglich $G(\xi)$ meßbare Mengen angegeben hatten.

Satz 1. *Die Menge der Werte der Summen S_δ ist genau dann beschränkt, wenn eine solche nichtfallende, in $[a, b]$ beschränkte Funktion $h(x)$ existiert, daß für jedes Intervall aus $[a, b]$ die Ungleichung*

$$(\Delta F)^2 \leq \Delta g \cdot \Delta h \quad (128)$$

erfüllt ist.

Beweis. Wenn diese Bedingung erfüllt ist, überschreiten die Summanden der Summe S_δ den Wert $\Delta_k h$ nicht, und für jede Zerlegung gilt $S_\delta \leq h(b) - h(a)$. Wir beweisen jetzt die Notwendigkeit von (128). Wenn die S_δ für das ganze Intervall $[a, b]$ beschränkt sind, so sind sie erst recht auch für jedes Teilintervall beschränkt. Wir bezeichnen mit $h(x)$ die obere Grenze von S_δ im Intervall $[a, x]$: $h(x) = \sup_\delta S_\delta$ für das Intervall $[a, x]$.

Genauso, wie wir die Additivität der Totalvariation [8] bewiesen haben, können wir zeigen, daß die obere Grenze von S_δ für ein beliebiges Teilintervall $[\alpha, \beta]$ gleich $h(\beta) - h(\alpha) = \Delta h$ ist; folglich ist $h(x)$ eine nichtfallende und beschränkte Funktion. Wenn wir Δ nicht in Teilintervalle zerlegen, so reduziert sich die Summe S_δ für das Intervall Δ auf den einen Summanden $(\Delta F)^2 : \Delta g$, und dieser ist nicht größer als die obere Grenze Δh der Summen S_δ für Δ ; folglich ist $(\Delta F)^2 : \Delta g \leq \Delta h$, woraus sich (128) ergibt.

Es sei nun $G(\xi)$ eine Mengenfunktion auf $[a, b]$, die durch die Punktfunktion $g(x)$ erzeugt wird, und wir setzen voraus, daß $F(x)$ die Form

$$F(x) = \int_a^x f(x) G(d\xi) + c \quad (129)$$

hat, wobei $f(x) \in L_2$ ist. Es gilt

$$\Delta F = \int_\Delta f(x) G(d\xi) = \int_\Delta f(x) dg(x),$$

und wenn wir die Schwarzsche Ungleichung anwenden, erhalten wir

$$(\Delta F)^2 \leq \int_\Delta f^2(x) G(d\xi) \int_\Delta G(d\xi) = \Delta g(x) \int_\Delta f^2(x) G(d\xi),$$

d. h., die Bedingung (128) ist erfüllt, und es gilt

$$h(x) = \int_a^x f^2(x) G(d\xi), \quad (130)$$

wobei es wegen der Stetigkeit von $g(x)$ gleichgültig ist, ob das Intervall $[a, x]$ abgeschlossen oder nicht abgeschlossen ist.

Die Formel (129) liefert eine volladditive Mengenfunktion

$$\varphi(\mathfrak{E}) = \int_{\mathfrak{E}} f(x) G(d\mathfrak{E}),$$

die für die Mengen $\mathfrak{E}$ definiert ist, die bezüglich $G(\mathfrak{E})$ meßbar sind und zu $[a, b]$ gehören, wobei $\Delta F = \varphi(\Delta)$ ist. Die Summen

$$\sum_k \frac{\varphi^2(\mathfrak{E}_k)}{G(\mathfrak{E}_k)} \quad (131)$$

haben die obere Grenze i , die durch die Formel (118) ausgedrückt werden kann:

$$i = \int_a^b f^2(x) G(d\mathfrak{E}). \quad (132)$$

Wir zeigen, daß die Summen (127), die der Zerlegung von $[a, b]$ in Intervalle entsprechen, dieselbe obere Grenze haben. Sie kann jedenfalls nicht größer als i sein. Wenn wir die absolute Stetigkeit von $\varphi(\mathfrak{E})$ benutzen, können wir zeigen, daß das Integral (132) auch die obere Grenze der Summen (127) ist, die wir bei einer Zerlegung von $[a, b]$ in Intervalle erhalten. Nach dem oben Gesagten existiert für jedes gegebene positive ε eine solche Zerlegung δ des Intervalls $[a, b]$ in meßbare Mengen $\mathfrak{E}_k$ ($k = 1, 2, \dots, n$), daß

$$\sum_{k=1}^n \frac{\varphi^2(\mathfrak{E}_k)}{G(\mathfrak{E}_k)} \cong i - \varepsilon \quad (133)$$

ist, wobei wir mit i das Integral (132) bezeichnen. Wenn wir die Notwendigkeit der Voraussetzung in Satz 1 aus [37] beachten, so können wir behaupten, daß für jedes $\mathfrak{E}_k$ ein Elementarbereich R_k , d. h. eine endliche Summe von punktfremden halboffenen Intervallen existiert derart, daß

$$R_k + e'_k = \mathfrak{E}_k + e''_k \quad (k = 1, 2, \dots, n) \quad (134)$$

gilt, wobei die Maße von e'_k und e''_k beliebig klein sind. Die $\mathfrak{E}_k$ sind paarweise punktfremde Mengen, die R_k können aber gemeinsame Punkte wegen e''_k haben, d. h. $R_k R_l \subset e''_k e''_l$, und die Maße dieser gemeinsamen Teile sind ebenfalls beliebig klein. In jeder der Gleichungen (134) können wir den Durchschnitt von R_k mit den restlichen R_l , der eine Summe von endlich vielen halboffenen Intervallen ist, zu e'_k rechnen. Wenn η das größte der Maße von e'_k und e''_k ist, gilt bei einer solchen Verschiebung für die neuen e'_k

$$G(e'_k) \leq (n+1) \eta,$$

denn es ist $G(R_k R_l) \leq G(e''_k e''_l) \leq \eta$. Somit können wir annehmen, daß die R_k in den Gleichungen (134) paarweise punktfremd und die Maße von e'_k und e''_k beliebig klein sind.

Wenn wir (133) und die absolute Stetigkeit von $\varphi(\mathfrak{E})$ beachten und die Maße von e'_k und e''_k genügend klein wählen, können wir

$$\sum_{k=1}^n \frac{\varphi^2(R_k)}{G(R_k)} \cong i - 2\varepsilon$$

schreiben. Sind die Δ_s ($s = 1, 2, \dots, m$) Intervalle, die in der Zusammensetzung aller R_k enthalten sind, so erhalten wir, wenn wir (62) beachten,

$$\sum_{s=1}^m \frac{\varphi^2(\Delta_s)}{G(\Delta_s)} \cong i - 2\varepsilon. \quad (135)$$

Da $g(x)$ und $F(x)$ stetig sind, können wir die Δ_k als abgeschlossen oder offen annehmen. Diese Intervalle können $[a, b]$ nicht überdecken. Wenn wir nichtnegative Summanden hinzufügen, die den restlichen Intervallen entsprechen, gilt für die vollständige Summe erst recht (135), und da ε beliebig war, folgt, daß das Integral (132) die obere Grenze der Summen (127) unter der Voraussetzung (129) ist; dabei ist $f(x) \in L_2$. Da $g(x)$ als stetig vorausgesetzt wurde, ist auch die durch (130) definierte Funktion $h(x)$ stetig.

Wir zeigen jetzt, daß $F(x)$ durch (129) mit $f(x) \in L_2$ dargestellt werden kann, wenn die Bedingung (128) erfüllt ist, d. h. die Summen (127) beschränkt sind. Es sei Δ ein beliebiges Intervall aus $[a, b]$ und Δ'_k seien Teilintervalle bei einer Zerlegung. Aus (128) folgt auf Grund der Schwarzschen Ungleichung

$$\left(\sum_k |\Delta'_k F|\right)^2 \leq \left(\sum_k \sqrt{\Delta'_k h} \sqrt{\Delta'_k g}\right)^2 \leq \sum_k \Delta'_k h \sum_k \Delta'_k g, \quad (136)$$

d. h.

$$\left(\sum_k |\Delta'_k F|\right)^2 \leq \Delta g \cdot \Delta h.$$

Für die obere Grenze der links stehenden Summen gilt bei beliebigen Zerlegungen dieselbe Ungleichung; folglich ist $F(x)$ eine Funktion von beschränkter Variation, und für ihre Totalvariation $v(x)$ gilt

$$(\Delta v(x))^2 \leq \Delta g \cdot \Delta h. \quad (137)$$

Sind die Δ'_k beliebige, sich nichtüberdeckende Intervalle aus $[a, b]$, so folgt aus (136)

$$\left(\sum_k |\Delta'_k F|\right)^2 \leq \sum_k \Delta'_k g [h(b) - h(a)].$$

Ist die Summe $\sum_k \Delta'_k g$ höchstens gleich ε , so gilt $\sum_k |\Delta'_k F| \leq \sqrt{\varepsilon} \sqrt{h(b) - h(a)}$, und hieraus folgt, daß $F(x)$ bezüglich $g(x)$ absolut stetig ist. Also ist

$$F(x) = \int_a^x f(x) dg(x) + c. \quad (138)$$

Es muß noch bewiesen werden, daß $f(x)$ zu L_2 gehört. Wir konstruieren eine beschränkte Funktion $f_n(x)$, und zwar sei

$$f_n(x) = \begin{cases} f(x) & \text{für } |f(x)| \leq n, \\ n & \text{für } f(x) > n, \\ -n & \text{für } f(x) < -n, \end{cases}$$

und setzen

$$F_n(x) = \int_a^x f_n(x) dg(x). \quad (139)$$

Die Funktion $f_n(x)$ gehört zu L_2 ; folglich gilt nach dem oben Bewiesenen

$$\sup_b \frac{(\Delta_k F_n)^2}{\Delta_k g} = \int_a^b f_n^2(x) dg(x). \quad (140)$$

Wird das Integral (138) über verschiedene Mengen $\mathcal{E}$ aus $[a, b]$ gebildet, die bezüglich $g(x)$ meßbar sind, so erhalten wir eine Mengenfunktion, deren Totalvariation

in $[a, x]$ durch ein Integral ausgedrückt werden kann [73]:

$$v(x) = \int_a^x |f(x)| \, dg(x). \quad (141)$$

Wenn wir $[a, x]$ nur in Intervalle zerlegen, erhalten wir für die Funktion (138) dieselbe Totalvariation [74].

Unter Berücksichtigung von (137) finden wir

$$\sup_{\delta} \sum_k \frac{(\Delta_k v)^2}{\Delta_k g} = M, \quad (142)$$

wobei M eine endliche Zahl ist. Andererseits gilt nach (139) und (141) die Beziehung $|\Delta_k F_n| \leq \Delta_k v$, und beachten wir (140) und (142), so erhalten wir für jedes n

$$\int_a^b f_n^2(x) \, dg(x) \leq M.$$

Hieraus folgt $f(x) \in L_2$. Diese Überlegungen führen uns auf den folgenden Satz:

Satz 2. *Die Bedingung (128) ist gleichbedeutend damit, daß $F(x)$ in Form von (138) mit $f(x) \in L_2$ darstellbar ist; ist die Bedingung (128) erfüllt, so wird die obere Grenze der Summen (127) durch das Integral (132) ausgedrückt.*

Hierbei brauchen $g(x)$ und $F(x)$ nicht stetig zu sein, denn dann wählen wir ein halboffenes Grundintervall und zerlegen es in halboffene Teilintervalle, für die $\Delta g = g(\beta + 0) - g(\alpha + 0)$ gilt. Alle Resultate, außer der Stetigkeit von $h(x)$, bleiben erhalten. Wir bemerken noch: Aus der Voraussetzung von (128) und der Stetigkeit von $g(x)$ folgt, daß in Bedingung (128) die Funktion $h(x)$ sogar als stetig angenommen werden kann.

83. Eigenschaften des Hellingerschen Integrals. Die obere Grenze der Summen (127) ist ein Hellingersches Integral, das analog zu (117) mit

$$\int_a^b \frac{[dF(x)]^2}{dg(x)} \quad (143)$$

bezeichnet wird; dabei gilt

$$\int_a^b \frac{[dF(x)]^2}{dg(x)} = \int_a^b f^2(x) \, dg(x). \quad (144)$$

Wir wollen beweisen, daß dieses Integral der Grenzwert der Summen (127) bei unbeschränkter Verkleinerung der Intervalle Δ_k ist.

Satz. *Genügt $F(x)$ der Bedingung (128) und $F_1(x)$ der analogen Bedingung*

$$(\Delta F_1)^2 \leq \Delta g \Delta h_1 \quad (145)$$

($g(x)$ sei stetig), so haben die Summe (127) und die Summe

$$\sum_k \frac{\Delta_k F \Delta_k F_1}{\Delta_k g} \quad (146)$$

bei unbeschränkter Verkleinerung der Δ_k einen bestimmten Grenzwert, wobei der Grenzwert von (127) gleich der oberen Grenze der entsprechenden Summen, d. h. gleich dem Integral (143) ist.

Beweis. Es sei, wie schon oben, i die obere Grenze der Summen (127). Für jedes $\varepsilon > 0$ existiert eine solche feste Zerlegung δ_0 in Teilintervalle, daß $S_{\delta_0} \geq i - \varepsilon$ ist. Es sei δ eine so feine Zerlegung, daß jedes Teilintervall von δ nicht mehr als einen Punkt der Zerlegung δ_0 enthält und daß der Zuwachs der stetigen Funktion $h(x)$ in jedem Teilintervall von δ nicht größer als ε ist. Für die Zerlegung $\delta\delta_0$ gilt

$$S_{\delta\delta_0} \geq S_{\delta_0} \geq i - \varepsilon. \quad (147)$$

Ist p die Anzahl der Teilpunkte in der Zerlegung δ_0 , so werden beim Übergang von δ zu $\delta\delta_0$ nicht mehr als p Teilintervalle in zwei Intervalle zerlegt. Dabei wird der entsprechende nichtnegative Summand der Summe S_δ durch zwei nichtnegative Summanden der Summe $S_{\delta\delta_0}$ ersetzt. Jeder dieser drei Summanden ist nach dem oben über den Zuwachs von $h(x)$ und der Eigenschaft (128) Gesagten nicht größer als ε ; folglich ist

$$0 \leq S_{\delta\delta_0} - S_\delta \leq 2p\varepsilon.$$

Ein Vergleich mit (147) ergibt $S_\delta \geq i - (2p+1)\varepsilon$, und hieraus folgt, weil ε beliebig war, daß die Summe (127) bei unbeschränkter Verkleinerung der Δ_k gegen i strebt. Zur Untersuchung der Summen (146) bemerken wir, daß $F(x) + F_1(x)$ durch (138) mit $(f(x) + f_1(x)) \in L_2$ dargestellt werden kann, und die Summen

$$\sum_k \frac{[\Delta_k(F + F_1)]^2}{\Delta_k g}$$

haben wie die analogen Summen für $F_1(x)$ bei unbeschränkter Verkleinerung der Δ_k einen Grenzwert. Hieraus folgt, daß auch die Summe

$$\sum_k \frac{\Delta_k F \Delta_k F_1}{\Delta_k g} = \frac{1}{2} \sum_k \frac{[\Delta_k(F + F_1)]^2}{\Delta_k g} - \frac{1}{2} \sum_k \frac{(\Delta_k F)^2}{\Delta_k g} - \frac{1}{2} \sum_k \frac{(\Delta_k F_1)^2}{\Delta_k g}$$

einen Grenzwert hat. Er stimmt mit dem Integral (122) überein. Somit erhalten wir die Hellingerschen Integrale

$$\int_a^b \frac{(dF)^2}{dg} = \lim \sum_k \frac{(\Delta_k F)^2}{\Delta_k g}, \quad \int_a^b \frac{dF dF_1}{dg} = \lim \sum_k \frac{\Delta_k F \Delta_k F_1}{\Delta_k g}. \quad (148)$$

Nach dem in [81] Gesagten ist

$$\int_a^b \frac{dF dF_1}{dg} = \int_a^b f(x) f_1(x) dg(x),$$

mit

$$F_1(x) = \int_a^b f_1(x) dg(x). \quad (149)$$

Man kann auch die allgemeineren Summen

$$\sum_k u(\xi_k) \frac{(\Delta_k F)^2}{\Delta_k g} \quad (150)$$

und

$$\sum_k u(\xi_k) \frac{\Delta_k F \Delta_k F_1}{\Delta_k g} \quad (151)$$

betrachten, wo $u(x)$ in $[a, b]$ stetig ist, die Integrale (148) existieren und ξ_k beliebige Werte aus Δ_k sind. Diese Summen haben ebenfalls wohlbestimmte Grenzwerte bei unbeschränkter Verkleinerung der Δ_k . Es genügt, dies für die Summen (150) zu beweisen. Wir betrachten die Summen

$$\sum_k m_k \frac{(\Delta_k F)^2}{\Delta_k g}, \quad (152)$$

wobei m_k der kleinste Wert der stetigen Funktion $u(x)$ in dem abgeschlossenen Intervall Δ_k ist. Genauso wie oben kann man beweisen, daß die Summen, wenn neue Teilpunkte hinzugefügt werden, nichtfallend und beschränkt sind und einen wohlbestimmten Grenzwert bei unbeschränkter Verkleinerung der Δ_k haben. Auf Grund der gleichmäßigen Stetigkeit von $u(x)$ und der Bedingung (128) strebt die Differenz der Summen (150) und (152) bei unbeschränkter Verkleinerung der Δ_k gegen 0; folglich haben die Summen (150) einen wohlbestimmten Grenzwert. Wir erhalten somit die Hellingerschen Integrale

$$\int_a^b u(x) \frac{(dF)^2}{dg} = \lim \sum_k u(\xi_k) \frac{(\Delta_k F)^2}{\Delta_k g}, \quad \int_a^b u(x) \frac{dF dF_1}{dg} = \lim \sum_k u(\xi_k) \frac{\Delta_k F \Delta_k F_1}{\Delta_k g}. \quad (153)$$

Die hier dargelegte Theorie kann auch dann angewendet werden, wenn $g(x)$ unstetig ist. Die erwähnten Summen haben dann einen wohlbestimmten Grenzwert für eine im Sinne des allgemeinen Stieltjesschen Integrals reguläre Folge von Zerlegungen δ_n .

Die behandelte Theorie gilt offenbar auch für den Fall, daß $F(x)$, $F_1(x)$ und $u(x)$ komplexe Funktionen sind, wobei $F(x)$ und $F_1(x)$ der Bedingung $|\Delta F| \leq \Delta g \Delta h$ bzw. $|\Delta F_1| \leq \Delta g \Delta h_1$ genügen müssen. Überall muß dann das Quadrat durch das Quadrat des Absolutbetrages ersetzt werden, also z. B. $(\Delta_k F)^2$ durch $|\Delta_k F|^2$.

Wir geben noch einige einfache Eigenschaften des Hellingerschen Integrals an. Es sei $\Phi(x)$ eine der Bedingung (128) genügende, nichtabnehmende Funktion. Mit ihr bilden wir die Funktion

$$F(x) = \int_a^x u(x) d\Phi(x), \quad (154)$$

wobei $u(x)$ stetig sei, und betrachten die Summen

$$\sum_k \frac{(\Delta_k F)^2}{\Delta_k g}. \quad (155)$$

Wenn wir den Mittelwertsatz anwenden, finden wir

$$\Delta_k F = \int_{\Delta_k} u(x) d\Phi(x) = u(\xi_k) \Delta_k \Phi, \quad \xi_k \in \Delta_k.$$

Die Summe (155) lautet nun

$$\sum_k \frac{(\Delta_k F)^2}{\Delta_k g} = \sum_k u^2(\xi_k) \frac{(\Delta_k \Phi)^2}{\Delta_k g},$$

und in der Grenze erhalten wir

$$\int_a^b \frac{(dF)^2}{dg} = \int_a^b u^2(x) \frac{(d\Phi)^2}{dg}.$$

Ganz analog ergibt sich

$$\int_a^b \frac{dF dF_1}{dg} = \int_a^b u(x) u_1(x) \frac{d\Phi d\Phi_1}{dg}, \quad (156)$$

wenn außer (154) noch $F_1(x) = \int_a^x u_1(x) d\Phi_1(x)$ gilt, wobei $\Phi_1(x)$ der Bedingung (128) genügt, eine nichtabnehmende Funktion ist und $u_1(x)$ stetig ist.

Genügen $\Phi(x)$ und $\Phi_1(x)$ der Bedingung (128), sind sie aber nicht monoton, so erhalten wir ebenfalls (156), wenn wir die kanonische Darstellung von $\Phi(x)$ und $\Phi_1(x)$ als Differenz zweier nichtabnehmender Funktionen benutzen. Wir weisen noch darauf hin, daß $F(x)$ und $F_1(x)$ offenbar ebenfalls der Bedingung (128) genügen.

IV. Metrische und normierte Räume

84. Der metrische Raum. Den Beginn dieses Kapitels widmen wir der Theorie der abstrakten Räume und wenden dann diese Theorie auf konkrete Räume an, und zwar in der Hauptsache auf Funktionenräume, d. h. auf Mengen von Funktionen einer bestimmten Klasse. Da ein und derselbe abstrakte Raum verschiedene Konkretisierungen zulassen kann, ist die Theorie der abstrakten Räume äußerst vielseitig anwendbar.

Jeder abstrakte Raum ist eine nichtleere Menge von Elementen, die gewissen Axiomen unterworfen ist. Die Natur dieser Elemente ist unbestimmt, und die Theorie eines beliebigen abstrakten Raumes ist eine Folgerung aus den Axiomen, die diesen Raum definieren. Wir werden bei unseren Darlegungen zunächst vollständig die Theorie eines abstrakten Raumes entwickeln und dann diese Theorie auf verschiedene konkrete Realisierungen dieses Raumes anwenden.

Eine Menge X von Elementen $x, y, z, \dots$ heißt *metrischer Raum*, wenn jedem Paar von Elementen x, y aus X eine nichtnegative Zahl $\varrho(x, y)$, der Abstand zwischen x und y , zugeordnet ist, wobei folgende Bedingungen gelten:

a) $\varrho(x, y) \geq 0$, wobei das Gleichheitszeichen nur für $x = y$ eintritt, d. h., wenn x und y dasselbe Element bezeichnen;

b) es gilt das Symmetrieaxiom

$$\varrho(y, x) = \varrho(x, y); \quad (1)$$

c) es gilt die Dreiecksungleichung

$$\varrho(x, z) \leq \varrho(x, y) + \varrho(y, z). \quad (2)$$

Diese Bedingungen müssen für alle $x, y, z \in X$ gelten. Sind $y_1, y_2, \dots, y_m$ beliebige Elemente aus X , so erhalten wir, wenn wir die Beziehung (2) mehrmals anwenden,

$$\varrho(y_1, y_m) \leq \varrho(y_1, y_2) + \varrho(y_2, y_3) + \dots + \varrho(y_{m-1}, y_m). \quad (2_1)$$

Gegeben sei eine unendliche Folge von Elementen x_n ($n = 1, 2, \dots$), und es existiere ein Element x_0 derart, daß $\varrho(x_0, x_n)$ für $n \rightarrow \infty$ gegen 0 strebt. Dann sagt man, x_0 sei der *Grenzwert* der Folge, der x_n , und schreibt $x_n \rightarrow x_0$ oder $\lim_{n \rightarrow \infty} x_n = x_0$. Es ist leicht zu sehen, daß eine Folge höchstens einen Grenzwert besitzen kann. Wäre nämlich $x_n \rightarrow x_0$, $x_n \rightarrow y_0$, so würde $\varrho(x_0, y_0) \leq \varrho(x_0, x_n) + \varrho(x_n, y_0)$ auf Grund von (2) sein; für $n \rightarrow \infty$ strebt aber die rechte Seite dieser Ungleichung gegen 0, so daß wir $\varrho(x_0, y_0) \leq 0$ erhalten. Es ist jedoch stets $\varrho(x_0, y_0) \geq 0$ und infolgedessen $\varrho(x_0, y_0) = 0$, also $x_0 = y_0$.

Ist $x_n \rightarrow x_0$, so gilt auch für jede unendliche Teilfolge die Beziehung $x_{n_k} \rightarrow x_0$.

Wir zeigen nun, daß $\varrho(x, y)$ eine stetige Funktion von x und y ist, d. h., ist $x_n \rightarrow x_0$ und $y_n \rightarrow y_0$, so strebt $\varrho(x_n, y_n)$ gegen $\varrho(x_0, y_0)$.

Wegen (2₁) können wir

$$\varrho(x_n, y_n) \leq \varrho(x_n, y_0) + \varrho(x_0, y_0) + \varrho(y_0, y_n),$$

$$\varrho(x_0, y_0) \leq \varrho(x_0, x_n) + \varrho(x_n, y_n) + \varrho(y_n, y_0)$$

schreiben; daraus ergibt sich

$$\varrho(x_n, y_n) - \varrho(x_0, y_0) \leq \varrho(x_n, x_0) + \varrho(y_0, y_n),$$

$$\varrho(x_0, y_0) - \varrho(x_n, y_n) \leq \varrho(x_0, x_n) + \varrho(y_n, y_0),$$

also

$$|\varrho(x_0, y_0) - \varrho(x_n, y_n)| \leq \varrho(x_0, x_n) + \varrho(y_0, y_n).$$

Für $n \rightarrow \infty$ strebt die rechte Seite dieser Ungleichung gegen 0. Daraus folgt $\varrho(x_n, y_n) \rightarrow \varrho(x_0, y_0)$.

Hat die Folge der x_n einen Grenzwert ($x_n \rightarrow x_0$), so existiert für jedes gegebene $\varepsilon > 0$ ein N derart, daß

$$\varrho(x_m, x_n) \leq \varepsilon \quad \text{für } m, n \geq N \quad (3)$$

ist. Dies erhalten wir sofort aus der Ungleichung

$$\varrho(x_m, x_n) \leq \varrho(x_m, x_0) + \varrho(x_0, x_n),$$

deren rechte Seite für $m, n \rightarrow \infty$ gegen 0 strebt. Aus (3) folgt aber auf Grund der oben angegebenen Axiome noch nicht, daß die Folge der x_n tatsächlich einen Grenzwert besitzt (die Hinlänglichkeit des Cauchyschen Kriteriums für die Existenz eines Grenzwertes gilt nicht). Führen wir noch die zusätzliche Forderung ein, daß sich aus (3) die Existenz eines Grenzwertes für die Folge der x_n ergeben soll, so nennen wir einen solchen metrischen Raum einen *vollständigen metrischen Raum*.

Es sei U eine Menge von Elementen eines metrischen Raumes. Die Menge U heißt *beschränkt*, wenn es ein Element x_0 und eine positive Zahl A gibt, für die $\varrho(x_0, x) \leq A$ für alle $x \in U$ gilt. Es sei $x_1 \neq x_0$ ein festes Element aus U . Dann gilt

$$\varrho(x_1, x) \leq \varrho(x_1, x_0) + \varrho(x_0, x)$$

und für Elemente aus U

$$\varrho(x_1, x) \leq \varrho(x_1, x_0) + A;$$

auf der rechten Seite dieser Ungleichung steht eine gewisse positive Zahl. Somit ist bei unserer Definition der Beschränktheit von U die Wahl des Elements x_0 gleichgültig. Es läßt sich leicht zeigen, daß die Menge der Elemente x_n beschränkt ist, wenn die Folge der x_n einen Grenzwert hat.

Das Element x_0 heißt *Grenzelement* der Menge $U \subset X$, wenn es eine Folge von Elementen $x_n \in U$ gibt, für die $x_n \rightarrow x_0$ ist. Eine Menge U , die alle ihre Grenzelemente enthält, nennen wir *abgeschlossen*. Ist die Menge U nicht abgeschlossen und fügen wir zu ihr alle ihre Grenzelemente hinzu, so ist die dadurch erhaltene neue Menge $\bar{U}$, die *abgeschlossene Hülle* von U , abgeschlossen [31]. Gehen wir von U zu $\bar{U}$ über, so sagen wir, wir schließen U ab. Ist U schon abgeschlossen, so ist $U = \bar{U}$. Ist U leer, d. h., enthält U kein Element, so ist auch $\bar{U}$ leer. Die Menge der Elemente x , die der Bedingung $\varrho(x_0, x) < R$ genügen, wobei x_0 ein festes Element und R eine positive Zahl ist, nennen wir die *offene Kugel mit dem Mittelpunkt x_0 und dem Radius R* , im Fall $\varrho(x_0, x) \leq R$ die *abgeschlossene Kugel*. Wegen der Stetigkeit von $\varrho(x_0, x)$ als Funktion von x ist die abgeschlossene Kugel eine im Sinne der oben gegebenen Definition abgeschlossene Menge.

Jede nichtleere Menge $U \subset X$ ist ebenfalls ein metrischer Raum, wenn $\varrho(x, y)$ für die Elemente aus U wie in X definiert ist. Der Raum X kann auch endlich viele Elemente enthalten.

Gegeben seien zwei metrische Räume X und X' , und zwischen ihren Elementen bestehe eine solche eindeutige Zuordnung, daß $\varrho(x, y) = \varrho(x', y')$ gilt, wobei x und x' , y und y' beliebige entsprechende Elemente aus X bzw. X' sind. In diesem Fall heißen die Räume X und X' *isometrisch*. Vom Standpunkt der abstrakten Theorie werden isometrische Räume nicht als verschieden angesehen.

85. Vervollständigung eines metrischen Raumes. Wir nennen eine Folge von Elementen $x_n \in X$ eine *Fundamentalfolge* (oder *in sich konvergente Folge*¹⁾), wenn für sie die Bedingung (3) erfüllt ist. Ist der Raum X nicht vollständig, so hat nicht jede Fundamentalfolge einen Grenzwert. Wir werden zeigen, daß in diesem Fall zu dem Raum neue Elemente (auch manchmal „ideale Elemente“ genannt) bei entsprechender Erweiterung des Abstandsbegriffs hinzugefügt werden können, so daß der so erhaltene metrische Raum bereits vollständig ist.

Wir beweisen zunächst den folgenden

Hilfssatz. *Sind $\{x_n\}$ und $\{y_n\}$ zwei Fundamentalfolgen, so existiert der Grenzwert der Zahlenfolge $\{\varrho(x_n, y_n)\}$.*

Beweis. Wegen (2₁) ist

$$\varrho(x_n, y_n) \leq \varrho(x_n, x_m) + \varrho(x_m, y_m) + \varrho(y_m, y_n),$$

woraus

$$\varrho(x_n, y_n) - \varrho(x_m, y_m) \leq \varrho(x_n, x_m) + \varrho(y_m, y_n)$$

folgt. Vertauschen wir m und n und benutzen wir (1), so erhalten wir

$$\varrho(x_m, y_m) - \varrho(x_n, y_n) \leq \varrho(x_n, x_m) + \varrho(y_m, y_n).$$

Aus den letzten beiden Ungleichungen folgt

$$|\varrho(x_n, y_n) - \varrho(x_m, y_m)| \leq \varrho(x_n, x_m) + \varrho(y_m, y_n).$$

Hier strebt die rechte Seite für $m, n \rightarrow \infty$ gegen 0; folglich ist bei der Zahlenfolge $\{\varrho(x_n, y_n)\}$ das Cauchysche Kriterium für die Existenz eines Grenzwerts erfüllt, was zu beweisen war.

Wir verteilen nun alle Fundamentalfolgen auf Klassen, und zwar ordnen wir einer Klasse diejenigen Fundamentalfolgen $\{x_n\}$ und $\{x'_n\}$ zu, für welche $\varrho(x_n, x'_n) \rightarrow 0$ gilt. Liegen $\{x_n\}$ und $\{x'_n\}$ sowie $\{x_n\}$ und $\{x''_n\}$ in einer Klasse, so sind auch $\{x'_n\}$ und $\{x''_n\}$ in einer Klasse enthalten, denn aus $\varrho(x_n, x'_n) \rightarrow 0$ und $\varrho(x_n, x''_n) \rightarrow 0$ folgt wegen (2₁), daß auch $\varrho(x'_n, x''_n)$ gegen 0 konvergiert. Für zwei Folgen $\{x_n\}$ und $\{y_n\}$, die zu verschiedenen Klassen gehören, ist der Grenzwert von $\varrho(x_n, y_n)$ größer als 0. Aus (2) ergibt sich ferner: Hat die Folge $\{x_n\}$ in X einen Grenzwert x_0 , so konvergiert auch jede andere Folge $\{x'_n\}$ derselben Klasse gegen denselben Grenzwert x_0 . Wegen der Stetigkeit des Abstands können Folgen aus verschiedenen Klassen nicht den gleichen Grenzwert besitzen. Die erwähnten Klassen von Fundamentalfolgen zerfallen somit in zwei Arten. Wir beschreiben zunächst die Klassen erster Art. Es sei x_0 ein beliebiges Element aus X . Wir haben die Klasse der Folgen, die x_0 als Grenzwert besitzen. Zu ihr gehört z. B. die Folge $\{x_n\}$, bei der alle Elemente gleich x_0 sind. Für

¹⁾ Vgl. S. 145 (Anm. d. Red.)

jedes x_0 gibt es eine Klasse von Folgen. Die Klassen zweiter Art bestehen aus denjenigen $\{x_n\}$, die in X keinen Grenzwert haben. Ist X vollständig, so gibt es keine Klassen zweiter Art. Ist X nicht vollständig, so konstruieren wir einen neuen metrischen Raum $\tilde{X}$, als dessen Elemente wir die oben erwähnten Klassen von Fundamentalfolgen aus X nehmen. In $\tilde{X}$ müssen wir einen neuen Abstandsbegriff einführen und die Gültigkeit der drei Grundeigenschaften nachprüfen. Es seien $\tilde{x}$ und $\tilde{y}$ zwei Elemente aus $\tilde{X}$. Aus den ihnen entsprechenden Klassen von Folgen wählen wir zwei beliebige Folgen $\{x_n\}$ und $\{y_n\}$ und definieren $\varrho(\tilde{x}, \tilde{y})$ durch

$$\varrho(\tilde{x}, \tilde{y}) = \lim_{n \rightarrow \infty} \varrho(x_n, y_n). \quad (4)$$

Wir zeigen, daß die nichtnegative Zahl $\varrho(x, \tilde{y})$ nicht von der Wahl von $\{x_n\}$ und $\{y_n\}$ abhängt. Es seien $\{x'_n\}$ und $\{y'_n\}$ beliebige Folgen aus den Klassen, welche $\tilde{x}$ bzw. $\tilde{y}$ entsprechen, und es sei

$$\varrho'(\tilde{x}, \tilde{y}) = \lim_{n \rightarrow \infty} \varrho(x'_n, y'_n).$$

Wir müssen also beweisen, daß $\varrho'(\tilde{x}, \tilde{y}) = \varrho(\tilde{x}, \tilde{y})$ ist. Wegen (2₁) gilt

$$\varrho(x_n, y_n) \leq \varrho(x_n, x'_n) + \varrho(x'_n, y'_n) + \varrho(y'_n, y_n).$$

Gehen wir hier zur Grenze über und berücksichtigen wir, daß $\varrho(x_n, x'_n)$ und $\varrho(y'_n, y_n)$ gegen 0 konvergieren, so erhalten wir $\varrho(\tilde{x}, \tilde{y}) \leq \varrho'(\tilde{x}, \tilde{y})$. Analog finden wir $\varrho'(\tilde{x}, \tilde{y}) \leq \varrho(\tilde{x}, \tilde{y})$ und demzufolge $\varrho'(\tilde{x}, \tilde{y}) = \varrho(\tilde{x}, \tilde{y})$. Durch (4) wird also die Zahl $\varrho(\tilde{x}, \tilde{y})$ eindeutig bestimmt.

Wir wollen nun die drei Grundeigenschaften nachprüfen: Offenbar ist $\varrho(\tilde{x}, \tilde{y}) \geq 0$.

a) Es sei $\varrho(\tilde{x}, \tilde{y}) = 0$, also $\varrho(x_n, y_n) \rightarrow 0$. Das besagt, daß die Folgen $\{x_n\}$ und $\{y_n\}$ zu ein und derselben Klasse gehören; also ist $\tilde{x} = \tilde{y}$.

b) Die Eigenschaft $\varrho(\tilde{x}, \tilde{y}) = \varrho(\tilde{y}, \tilde{x})$ folgt unmittelbar aus $\varrho(x_n, y_n) = \varrho(y_n, x_n)$.

c) Wählen wir aus der Klasse der Folgen die den Elementen $\tilde{x}$, $\tilde{y}$ bzw. $\tilde{z}$ entsprechenden Folgen $\{x_n\}$, $\{y_n\}$ und $\{z_n\}$, so gilt

$$\varrho(\tilde{x}, \tilde{z}) = \lim_{n \rightarrow \infty} \varrho(x_n, z_n) \leq \lim_{n \rightarrow \infty} [\varrho(x_n, y_n) + \varrho(y_n, z_n)] \leq \varrho(\tilde{x}, \tilde{y}) + \varrho(\tilde{y}, \tilde{z}).$$

Das Element $\tilde{x}$ gehöre zu einer Klasse erster Art, und x_0 sei der Grenzwert der Folgen dieser Klasse. Wir können ein solches Element $\tilde{x} \in \tilde{X}$ mit dem Element $x_0 \in X$ identifizieren. Die Elemente $\tilde{x}$, die den Klassen zweiter Art zugeordnet sind, sind die Elemente von $\tilde{X}$, die nicht in X liegen. Sind $\tilde{x}$ und $\tilde{y}$ Elemente entsprechender Klassen erster Art und sind x_0 und y_0 die mit ihnen identifizierten Elemente aus X , so können wir $x_n = x_0$ und $y_n = y_0$ in (4) für jedes n setzen, und wir erhalten

$$\varrho(\tilde{x}, \tilde{y}) = \lim_{n \rightarrow \infty} \varrho(x_0, y_0) = \varrho(x_0, y_0),$$

d. h., für die Elemente aus X stimmt die neue Abstandsdefinition mit der alten überein. Ist $\tilde{x}$ einer Klasse erster Art (x_0 sei das entsprechende Element aus X) und $\tilde{y}$ einer Klasse zweiter Art zugeordnet, so ergibt sich aus (4)

$$\varrho(\tilde{x}, \tilde{y}) = \lim_{n \rightarrow \infty} \varrho(x_0, y_n).$$

Wir zeigen ferner, daß $\varrho(\tilde{x}, x_n) \rightarrow 0$ für $n \rightarrow \infty$ ist, wenn die Folge $\{x_n\}$ in der durch das Element $\tilde{x}$ definierten Klasse liegt. Nach Definition ist $\varrho(\tilde{x}, x_n) = \lim_{m \rightarrow \infty} \varrho(x_m, x_n)$. Da $\{x_n\}$ eine Fundamentalfolge ist, gelten die Ungleichungen (3) und $\varrho(\tilde{x}, x_n) \leq \varepsilon$ für $n \geq N$. Also ist $\varrho(\tilde{x}, x_n) \rightarrow 0$ für $n \rightarrow \infty$.

Der Raum X ist in $\tilde{X}$ *dicht*, d. h., ist $\tilde{x}$ ein beliebiges Element aus $\tilde{X}$ und ist ε eine beliebig gegebene positive Zahl, so existiert ein Element $x \in X$, für welches $\varrho(\tilde{x}, x) \leq \varepsilon$ ist. Wenn das Element $\tilde{x}$ einer Klasse erster Art zugeordnet ist und mit dem Element x_0 aus X identifiziert wird, können wir $x = x_0$ für jedes $\varepsilon > 0$ setzen, denn es ist

$$\varrho(\tilde{x}, x_0) = \varrho(x_0, x_0) = 0.$$

Nun sei das Element $\tilde{x}$ nicht in X enthalten, und ihm entspreche eine Fundamentalfolge $\{x_n\}$. Ferner wählen wir ein m mit $\varrho(x_n, x_m) \leq \varepsilon$ für $n \geq m$. Dann können wir $x = x_m$ setzen; denn es ist

$$\varrho(\tilde{x}, x_m) = \lim_{n \rightarrow \infty} \varrho(x_n, x_m),$$

und wegen $\varrho(x_n, x_m) \leq \varepsilon$ für $n \geq m$ erhalten wir $\varrho(\tilde{x}, x_m) \leq \varepsilon$.

Wir beweisen nun, daß $\tilde{X}$ ein vollständiger Raum ist. Es sei $\{\tilde{x}_n\}$ eine Fundamentalfolge in $\tilde{X}$, es gelte also $\varrho(\tilde{x}_n, \tilde{x}_m) \leq \varepsilon$ für $n, m \geq N$. Zu zeigen ist, daß in $\tilde{X}$ ein Element $\tilde{x}$ existiert, für welches $\lim_{n \rightarrow \infty} \varrho(\tilde{x}, \tilde{x}_n) = 0$ ist. Auf Grund des oben Bewiesenen gibt es für jedes n ein solches Element $x_n \in X$, daß $\varrho(\tilde{x}_n, x_n) \leq \frac{1}{n}$ ist. Die Folge $\{x_n\}$ von Elementen aus X ist eine Fundamentalfolge, denn es gilt

$$\varrho(x_n, x_m) \leq \varrho(x_n, \tilde{x}_n) + \varrho(\tilde{x}_n, x_m) + \varrho(\tilde{x}_m, x_m) \leq \frac{1}{n} + \frac{1}{m} + \varrho(\tilde{x}_n, \tilde{x}_m).$$

Die Folge der x_n liegt in einer Klasse, die von einem gewissen Element $\tilde{x} \in \tilde{X}$ definiert ist. Man sieht leicht, daß $\varrho(\tilde{x}, \tilde{x}_n) \rightarrow 0$ ist; dies folgt aus der Ungleichung

$$\varrho(\tilde{x}, \tilde{x}_n) \leq \varrho(\tilde{x}, x_n) + \varrho(x_n, \tilde{x}_n) \leq \varrho(\tilde{x}, x_n) + \frac{1}{n}$$

und aus $\varrho(\tilde{x}, x_n) \rightarrow 0$, wie wir oben sahen. Damit ist die Vollständigkeit von $\tilde{X}$ bewiesen.

Wir beweisen noch einen Satz über die Eindeutigkeit der Vervollständigung eines metrischen Raumes X .

Satz. Die Vervollständigung eines Raumes X , bei welcher X in dem neuen Raum *dicht* ist, ist bis auf Isometrie eindeutig.

Beweis. Es sei Y ein vollständiger metrischer Raum, der X enthält und in welchem X dicht ist. Wir müssen dann zeigen, daß Y dem Raum $\tilde{X}$ isometrisch ist. Dazu nehmen wir natürlich an, daß der Abstand zwischen zwei Elementen aus Y , die auch zu X gehören, der gleiche wie in X ist. Es sei y ein beliebiges Element aus Y . Da X in Y dicht ist, existiert eine Folge von Elementen $x_n \in X$, für welche $\varrho(y, x_n)$ in Y gegen 0 strebt. Somit ist $\{x_n\}$ sowohl in Y als auch in X eine Fundamentalfolge. Dieser Folge entspricht ein bestimmtes Element $\tilde{x} \in \tilde{X}$. Es ist leicht zu sehen, daß $\tilde{x}$ nicht von der Wahl der x_n abhängt, wichtig ist nur, daß $\varrho(y, x_n)$ gegen 0 strebt. Wir ordnen $\tilde{x}$ dem Element $y \in Y$ zu, und es sei $\tilde{x}'$ ein bestimmtes Element aus $\tilde{X}$. Ferner wählen wir eine durch dieses Element bestimmte Folge von Elementen $x'_n \in X$. Sie ist eine Fundamentalfolge im vollständigen Raum Y und definiert somit ein Element $y' \in Y$. Wir sehen sofort, daß y' nicht von der Wahl der Folge $\{x'_n\}$ abhängt, wichtig ist nur, daß diese Folge das Element $\tilde{x}'$ definiert. Wir ordnen y' das Element $\tilde{x}'$ zu. So erkennen wir, daß damit eine eineindeutige Zuordnung zwischen den Elementen von Y und denen von $\tilde{X}$ hergestellt wird. Daß

$$\varrho(\tilde{x}, \tilde{x}') = \varrho(y, y')$$

ist, folgt aus der Definition von $\varrho(\tilde{x}, \tilde{x}')$ in $\tilde{X}$ und aus der Stetigkeit des Abstandes in Y , d. h. aus

$$\varrho(\tilde{x}, \tilde{x}') = \lim_{n \rightarrow \infty} \varrho(x_n, x'_n) = \varrho(y, y').$$

Damit ist der Satz bewiesen.

Wir gingen deshalb so ausführlich auf die Vervollständigung eines metrischen Raumes ein, weil dieser Prozeß bei den Anwendungen der Theorie der metrischen Räume eine wichtige Rolle spielt und eine Beschränkung auf vollständige Räume erlaubt.

Wir bringen nun drei einfache Beispiele.

1. Es sei X der Raum aller reellen rationalen Zahlen $x, y, z, \dots$, wobei der Abstand zweier Zahlen x und y durch $\varrho(x, y) = |x - y|$ definiert ist. Offenbar genügt $\varrho(x, y)$ den drei Bedingungen für metrische Räume. Wir wählen eine Fundamentalfolge reeller rationaler Zahlen x_n , die nach dem Cauchyschen Konvergenzkriterium sicherlich konvergiert; ist aber dieser Grenzwert eine irrationale Zahl, so hat die Folge der x_n keinen Grenzwert in X , so daß der Raum X nicht vollständig ist. Seine Vervollständigung umfaßt schon alle irrationalen Zahlen. Der Raum X aller reellen Zahlen ist bereits vollständig.

2. Wir betrachten den Raum C aller reellen Funktionen $x(t), y(t), z(t), \dots$, die auf einem endlichen Intervall $[a, b]$ stetig sind, und definieren ihren Abstand $\varrho(x, y)$ durch

$$\varrho(x, y) = \max_{a \leq t \leq b} |x(t) - y(t)|$$

(vgl. [14]). Diese Definition von $\varrho(x, y)$ ist erlaubt. Der Konvergenz $\varrho(x, x_n) \rightarrow 0$ entspricht nämlich in diesem Fall die gleichmäßige Konvergenz $x_n(t) \rightarrow x(t)$ auf dem Intervall $a \leq t \leq b$, und ist $\max_{a \leq t \leq b} |x_n(t) - x_m(t)| \rightarrow 0$ für $n, m \rightarrow \infty$, so existiert eine stetige Funktion $x(t)$, gegen welche $x_n(t)$ gleichmäßig konvergiert [I, 144], d. h., der Raum C ist vollständig.

3. Wir untersuchen nun den Raum F der auf dem Intervall $[a, b]$ reellen stetigen Funktionen $x(t), y(t), z(t), \dots$, definieren jetzt aber den Abstand durch

$$\varrho(x, y) = \left[\int_a^b [x(t) - y(t)]^2 dt \right]^{1/2}. \quad (5)$$

Diese Definition von $\varrho(x, y)$ ist ebenfalls zulässig. Wählen wir eine Fundamentalfolge $\{x_n(t)\}$ aus F , für welche

$$\int_a^b [x_n(t) - x_m(t)]^2 dt \rightarrow 0 \quad \text{für } n, m \rightarrow \infty$$

gilt, so hat sie einen Grenzwert im Sinne der Metrik (5) [56], jedoch kann die Grenzfunktion eine beliebige Funktion aus L_2 sein, da die stetigen Funktionen in L_2 überall dicht sind [60]. Ist die Grenzfunktion nicht einer stetigen Funktion äquivalent, so hat eine solche Fundamentalfolge aus F keinen Grenzwert in F , d. h., der Raum F ist nicht vollständig. Seine Vervollständigung ergibt Funktionen aus L_2 , welche nicht äquivalent stetigen Funktionen sind, und führt F in L_2 über.

Statt Funktionen einer Veränderlichen können wir auch Mengen von Funktionen $x(t_1, t_2, \dots, t_n)$ von n Veränderlichen betrachten, die auf einer beschränkten abgeschlossenen Menge des n -dimensionalen Raumes stetig sind.

Wir weisen noch einmal darauf hin, daß es bei der Vervollständigung eines metrischen Raumes wichtig ist, den Sinn der bei der Vervollständigung erhaltenen neuen Elemente deuten zu können. In Beispiel 3 sind dies die Funktionen aus L_2 , welche nicht stetigen Funktionen äquivalent sind. Wie wir oben sahen, läßt sich der Raum L_2 auf einer beliebigen meßbaren Menge betrachten. Wir untersuchten den Raum L_2 auf einer beschränkten abgeschlossenen Menge, insofern wir vom Raum F der stetigen Funktionen ausgingen.

Wir erwähnen noch einen Satz, der in vollständigen metrischen Räumen gilt. Die offene bzw. die abgeschlossene Sphäre in X mit dem Mittelpunkt x und dem Radius r wollen wir mit $S(x, r)$ bzw. $\bar{S}(x, r)$ bezeichnen.

Satz. *In einem vollständigen metrischen Raum X sei eine Folge abgeschlossener Sphären $\bar{S}(x_n, r_n)$, $n = 1, 2, \dots$, derart gegeben, daß jede Sphäre in der vorhergehenden enthalten ist und die Radien r_n für $n \rightarrow \infty$ gegen 0 streben. Dann existiert ein eindeutig bestimmter Punkt, der allen $\bar{S}(x_n, r_n)$ angehört.*

Beweis. Nach Voraussetzung ist $\bar{S}(x_{n+p}, r_{n+p}) \subset \bar{S}(x_n, r_n)$ für $p > 0$; folglich ist $\varrho(x_{n+p}, x_n) \leq r_n$ für jedes $p > 0$, d. h., die Folge der x_n bildet eine Fundamentalfolge, deren Grenzwert wir mit x_0 bezeichnen wollen. Dieser Grenzwert existiert, da X vollständig ist. Dann wählen wir eine bestimmte Sphäre $\bar{S}(x_n, r_n)$ und zeigen, daß x_0 in $\bar{S}(x_n, r_n)$ liegt. Alle Elemente der Folge $x_n, x_{n+1}, \dots$, die x_0 als Grenzwert hat, gehören nach Voraussetzung zur abgeschlossenen Menge $\bar{S}(x_n, r_n)$; also gehört x_0 zu $\bar{S}(x_n, r_n)$. Nun nehmen wir an, es existiere ein Element x'_0 , das zu allen $\bar{S}(x_n, r_n)$ gehört, und werden zeigen, daß $x'_0 = x_0$ ist. Da x_0 und x'_0 in allen $\bar{S}(x_n, r_n)$ liegen, gilt

$$\varrho(x_0, x'_0) \leq \varrho(x_0, x_n) + \varrho(x_n, x'_0) \leq 2r_n.$$

Für $n \rightarrow \infty$ folgt hieraus $\varrho(x_0, x'_0) \leq 0$, also $\varrho(x_0, x'_0) = 0$. Die Elemente x_0 und x'_0 stimmen somit überein. Damit ist der Satz bewiesen.

Jede abgeschlossene Menge U eines vollständigen metrischen Raumes X ist ebenfalls ein vollständiger metrischer Raum (dabei wird vorausgesetzt, daß der Abstand $\varrho(x, y)$ in U gleich dem Abstand zwischen x und y in X ist). Dies folgt daraus, daß jede Fundamentalfolge x_n in U einen Grenzwert in X hat und dieser Grenzwert zu U gehören muß, da U eine abgeschlossene Menge ist.

86. Operatoren und Funktionale. Das Prinzip der kontrahierenden Abbildungen. Gegeben seien zwei metrische Räume X und X' . Unter einem *Operator aus X in X'* verstehen wir eine Vorschrift A , die mit Hilfe der Beziehung $x' = Ax$ den Elementen $x \in X$ bestimmte Elemente $x' \in X'$ zuordnet. Der Operator braucht nicht auf ganz X definiert zu sein. Die Menge der Elemente $x \in X$, auf welchen der Operator A definiert ist, nennen wir den *Definitionsbereich* von A und bezeichnen ihn mit $D(A)$. Den *Wertebereich* von A , d. h. die Menge der Werte Ax , bezeichnen wir mit $R(A)$ ¹⁾; dies ist eine Menge von Elementen aus X' . Stimmt $R(A)$ mit X' überein, so hat die Gleichung

$$x' = Ax \tag{6}$$

für jedes $x' \in X'$ wenigstens eine Lösung. Wir nehmen an, A stelle eine eindeutige Zuordnung zwischen $D(A)$ und $R(A)$ her, d. h., für verschiedene $x \in D(A)$ erhalten wir mit Hilfe von (6) verschiedene $x' \in R(A)$. In diesem Fall hat (6) für jedes $x' \in R(A)$ genau eine Lösung aus $D(A)$.

¹⁾ R ist der Anfangsbuchstabe des englischen Wortes *range* (Anm. d. Red.).

Ein überaus wichtiger Spezialfall der Operatoren sind die *Funktionale*. So nennen wir die Operatoren in dem Fall, daß X' der Raum der reellen Zahlen mit dem in [85] definierten Abstand $\varrho(x', y') = |x' - y'|$ ist. Manchmal wird auch der Raum der komplexen Zahlen mit der gleichen Abstandsdefinition benutzt.

Wenn X' mit X übereinstimmt und $D(A) = X$ ist, sagt man, der Operator A bilde X in sich ab. Wir geben nun in diesem Fall ein Kriterium für die Lösbarkeit der Gleichung $x - Ax = 0$.

Satz (Prinzip der kontrahierenden Abbildungen). *Bildet der Operator A den gesamten vollständigen metrischen Raum X in sich ab und gilt für alle x und y aus X*

$$\varrho(Ax, Ay) \leq \alpha \varrho(x, y) \quad (0 < \alpha < 1), \quad (7)$$

so hat die Gleichung $x = Ax$ genau eine Lösung. Diese Lösung ist der Grenzwert der Folge

$$x_2 = Ax_1, \quad x_3 = Ax_2, \quad x_4 = Ax_3, \dots \quad (8)$$

bei einem beliebigen Anfangselement x_1 .

Beweis. In dem betrachteten Fall ist $D(A) = X$ und $R(A) \subset X$. Ferner gilt

$$\varrho(x_n, x_{n+1}) = \varrho(Ax_{n-1}, Ax_n) \leq \alpha \varrho(x_{n-1}, x_n).$$

Wenden wir diese Abschätzung auf $\varrho(x_{n-1}, x_n)$ usw. an, so erhalten wir

$$\varrho(x_n, x_{n+1}) \leq \alpha^{n-1} \varrho(x_1, x_2) \quad (n = 1, 2, \dots);$$

daraus folgt für $m > n$

$$\begin{aligned} \varrho(x_n, x_m) &\leq \varrho(x_n, x_{n+1}) + \varrho(x_{n+1}, x_{n+2}) + \dots + \varrho(x_{m-1}, x_m) \\ &\leq \alpha^{n-1} (1 + \alpha + \dots + \alpha^{m-n-1}) \varrho(x_1, x_2) \leq \frac{\alpha^{n-1}}{1-\alpha} \varrho(x_1, x_2). \end{aligned}$$

Wegen $\varrho(x_n, x_n) = 0$ und $\varrho(x_n, x_m) = \varrho(x_m, x_n)$ gilt $\varrho(x_n, x_m) \rightarrow 0$ für $n, m \rightarrow \infty$. Da X vollständig ist, hat die Folge x_n einen Grenzwert, den wir mit x_0 bezeichnen ($x_n \rightarrow x_0$). Es ist $Ax_n \rightarrow Ax_0$, denn es ist

$$\varrho(Ax_n, Ax_0) \leq \alpha \varrho(x_{n-1}, x_0) \rightarrow 0$$

wegen $\varrho(x_{n-1}, x_0) \rightarrow 0$. Nehmen wir in der Gleichung $x_n = Ax_{n-1}$ den Grenzübergang vor, so erhalten wir $x_0 = Ax_0$. Um zu zeigen, daß die Lösung von $x = Ax$ eindeutig ist, nehmen wir an, es existiere noch eine andere Lösung x' , d. h., es sei $x' = Ax'$. Für den Abstand zwischen x_0 und x' gilt $\varrho(x_0, x') = \varrho(Ax_0, Ax') \leq \alpha \varrho(x_0, x')$, also $(1 - \alpha) \varrho(x_0, x') = 0$. Daraus folgt $\varrho(x_0, x') = 0$, so daß x_0 mit x' übereinstimmt. Damit ist der Satz bewiesen.

Bemerkung. Es sei U eine abgeschlossene Menge des Raumes X . Stimmt $D(A)$ mit U überein, ist $R(A) \subset U$ und ist die Bedingung (7) erfüllt, so gilt der eben bewiesene Satz, wobei x_0 zu U gehört und jedes x' , das der Gleichung $x' = Ax'$ genügt, gleich x_0 ist. Dabei wird $x' \in U$ angenommen, da A auf U definiert ist.

87. Beispiele. Bevor wir zu Beispielen für das Prinzip der kontrahierenden Abbildungen übergehen, wollen wir einige Beispiele für vollständige metrische Räume betrachten.

1. Der Raum R_n aller Folgen von n reellen Zahlen. Der Abstand zwischen den Elementen $x(a_1, a_2, \dots, a_n)$ und $y(b_1, b_2, \dots, b_n)$ aus R_n wird durch

$$\varrho(x, y) = \left[\sum_{k=1}^n (a_k - b_k)^2 \right]^{1/2} \quad (9)$$

bestimmt. Es läßt sich auch der komplexe Raum R_n der Folgen von n komplexen Zahlen definieren. Dann muß in (9) das Quadrat $(a_k - b_k)^2$ durch $|a_k - b_k|^2$ ersetzt werden. Diese Bemerkung bezieht sich auch auf die folgenden Beispiele für Räume von Folgen oder Funktionen.

2. Der Raum m der unendlichen Folgen x von Zahlen $(a_1, a_2, \dots)$, die in ihrer Gesamtheit beschränkt sind; d. h., für jedes Element $x \in m$ existiert eine positive Zahl m_x derart, daß $|a_i| \leq m_x$ für jedes i gilt. Die Formel für den Abstand lautet hier

$$\varrho(x, y) = \sup_k |a_k - b_k|. \quad (10)$$

Die Konvergenz $x^{(n)} \rightarrow x$ in m ist mit der koordinatenweisen Konvergenz gleichwertig ($a_k^{(n)} \rightarrow a_k$ für $n \rightarrow \infty$; $k = 1, 2, \dots$), die bezüglich der Nummer k der Komponente gleichmäßig ist.

3. Der Raum s der unendlichen Zahlenfolgen, wobei

$$\varrho(x, y) = \sum_{k=1}^{\infty} \frac{|a_k - b_k|}{2^k (1 + |a_k - b_k|)} \quad (11)$$

gilt. Die Zulässigkeit dieser Definition von $\varrho(x, y)$ kann genau so geprüft werden wie später für den analogen Funktionenraum S . Wie im Raum R_n ist auch im Raum s die Konvergenz mit der koordinatenweisen Konvergenz äquivalent.

4. Der Raum l_p ($p \geq 1$) der Folgen von komplexen Zahlen a_k , für die

$$\sum_{k=1}^{\infty} |a_k|^p < \infty \quad (12)$$

gilt. Der Abstand ist durch

$$\varrho(x, y) = \left[\sum_{k=1}^{\infty} |a_k - b_k|^p \right]^{1/p} \quad (13)$$

gegeben. Die Dreiecksungleichung ergibt sich aus der Minkowskischen Ungleichung für Summen für $p > 1$ [62] und gilt offenbar für $p = 1$.

5. Der Raum C der Funktionen $\varphi(x)$, wobei x ein Punkt des n -dimensionalen Raumes R_n ist und die Funktionen auf einer beschränkten abgeschlossenen Menge $\mathcal{E}$ stetig sind. Dabei ist

$$\varrho(\varphi, \psi) = \max_{x \in \mathcal{E}} |\varphi(x) - \psi(x)|. \quad (14)$$

6. Der Raum M der Funktionen $\varphi(x)$, die auf einer Lebesgue-meßbaren Menge $\mathcal{E}$ des R_n definiert sind. Einander äquivalente Funktionen werden miteinander identifiziert, und jede Funktion aus M ist beschränkt (oder einer beschränkten Funktion äquivalent). Die Definition für den Abstand lautet

$$\varrho(\varphi, \psi) = \inf_{m(\mathcal{E}_0)=0} \sup_{\mathcal{E}-\mathcal{E}_0} |\varphi(x) - \psi(x)|. \quad (15)$$

Der Sinn dieser Definition besteht in folgendem: Wir entfernen aus $\mathcal{E}$ eine Menge $\mathcal{E}_0$ vom Maß 0, bestimmen die obere Grenze von $|\varphi(x) - \psi(x)|$ auf der Restmenge $\mathcal{E} - \mathcal{E}_0$, wählen die Menge $\mathcal{E}_0$ auf alle möglichen Arten und bestimmen dann die untere Grenze der Menge der nichtnegativen oberen Grenzen $\sup_{\mathcal{E}-\mathcal{E}_0} |\varphi(x) - \psi(x)|$. Statt (15) schreibt man auch

$$\varrho(\varphi, \psi) = \text{vrai max } |\varphi(x) - \psi(x)|. \quad (16)$$

Ist $\mathcal{E}$ eine beschränkte abgeschlossene Menge, so ist C ein Teil von M , und $\varrho(\varphi, \psi)$ ist in C genauso wie in M definiert, d. h., C ist einem Teil von M isometrisch.

7. Der Raum S aller Funktionen $\varphi(x)$, die auf einer meßbaren Punktmenge $\mathcal{E}$ endlichen Maßes aus R_n meßbar sind, wobei

$$\varrho(\varphi, \psi) = \int_{\mathcal{E}} \frac{|\varphi(x) - \psi(x)|}{1 + |\varphi(x) - \psi(x)|} dx \quad (17)$$

ist.

Hier und im folgenden ist das Maß und das Lebesguesche Integral jeweils im entsprechenden R_n zu verstehen.

8. Der Raum $L_p(\mathfrak{S})$, $p \geq 1$, der Funktionen $\varphi(x)$, die auf einer meßbaren Menge $\mathfrak{S}$ meßbar sind und für welche

$$\int_{\mathfrak{S}} |\varphi(x)|^p dx < \infty$$

gilt. Dabei ist

$$\varrho(\varphi, \psi) = \left[\int_{\mathfrak{S}} |\varphi(x) - \psi(x)|^p dx \right]^{1/p}. \quad (18)$$

Es gelten die in den Axiomen aufgezählten Eigenschaften von $\varrho(\varphi, \psi)$ [62].

9. Der Raum $V[a, b]$ der Funktionen $\varphi(x)$ von beschränkter Variation auf einem abgeschlossenen Intervall $a \leq x \leq b$, die in den inneren Punkten von $[a, b]$ von rechts stetig und für $x = a$ gleich 0 sind. Die Formel für den Abstand lautet

$$\varrho(\varphi, \psi) = V_a^b |\varphi(x) - \psi(x)|. \quad (19)$$

Streicht man die Forderung $\varphi(a) = 0$, so wird $\varrho(\varphi, \psi)$ durch

$$\varrho(\varphi, \psi) = |\varphi(a) - \psi(a)| + V_a^b |\varphi(x) - \psi(x)| \quad (20)$$

definiert. Dabei wird der Raum erweitert, und der ursprüngliche Raum ist einem Teil des erweiterten Raums isometrisch.

Alle obenerwähnten Räume sind vollständig.

Für L_p und l_p wurde die Vollständigkeit schon nachgewiesen. In den anderen Fällen bereitet der Beweis keine Mühe, so daß wir ihn übergehen können.

Wir wollen jetzt den Raum S genauer untersuchen. Da

$$\omega(t) = \frac{t}{1+t} = 1 - \frac{1}{1+t}$$

für $t \geq 0$ wächst, können wir

$$\frac{|t+\tau|}{1+|t+\tau|} \leq \frac{|t|+|\tau|}{1+|t|+|\tau|} \leq \frac{|t|}{1+|t|} + \frac{|\tau|}{1+|\tau|}$$

schreiben. Hieraus folgt die Dreiecksungleichung für S . Wir zeigen nun, daß die Konvergenz in S und die Konvergenz dem Maße nach einander gleichwertig sind. Auf $\mathfrak{S}$ konvergiere $\varphi_n(x)$ dem Maß nach gegen $\varphi(x)$. Wir werden zeigen, daß $\varrho(\varphi, \varphi_n)$ gegen 0 strebt. Zum Beweis führen wir die Menge $\mathfrak{S}_n(\delta) = \mathfrak{S} [|\varphi(x) - \varphi_n(x)| \geq \delta]$ ein. Nach Voraussetzung ist $m[\mathfrak{S}_n(\delta)] \rightarrow 0$ für $n \rightarrow \infty$ und jedes feste $\delta > 0$. Dann gilt

$$\varrho(\varphi, \varphi_n) = \int_{\mathfrak{S}} \frac{|\varphi(x) - \varphi_n(x)|}{1 + |\varphi(x) - \varphi_n(x)|} dx \leq \int_{\mathfrak{S}_n(\delta)} 1 \cdot dx + \int_{\mathfrak{S} - \mathfrak{S}_n(\delta)} \frac{|\varphi(x) - \varphi_n(x)|}{1 + |\varphi(x) - \varphi_n(x)|} dx,$$

woraus sich

$$\varrho(\varphi, \varphi_n) \leq m[\mathfrak{S}_n(\delta)] + \frac{\delta}{1+\delta} m(\mathfrak{S})$$

ergibt, wenn wir berücksichtigen, daß die Funktion $\omega(t)$ wächst und $|\varphi(x) - \varphi_n(x)| < \delta$ auf der Menge $\mathfrak{S} - \mathfrak{S}_n(\delta)$ ist. Gegeben sei nun eine positive Zahl ε . Zu ihr gibt es ein $\delta > 0$

derart, daß $\frac{\delta}{1+\delta} m(\mathfrak{S}) \leq \frac{\varepsilon}{2}$ ist. Ferner existiert ein N , für welches

$$m[\mathfrak{S}_n(\delta)] \leq \frac{\varepsilon}{2} \quad \text{für } n \geq N$$

und folglich

$$\varrho(\varphi, \varphi_n) \leq \varepsilon \quad \text{für } n \geq N$$

ist; also gilt $\varrho(\varphi, \varphi_n) \rightarrow 0$. Umgekehrt nehmen wir nun $\varrho(\varphi, \varphi_n) \rightarrow 0$ an und werden zeigen, daß $\varphi_n(x)$ auf $\mathfrak{S}$ dem Maß nach gegen $\varphi(x)$ konvergiert. Da $\omega(t)$ eine wachsende Funktion ist, gilt

$$\frac{|\varphi(x) - \varphi_n(x)|}{1 + |\varphi(x) - \varphi_n(x)|} \leq \frac{\delta}{1 + \delta} \quad \text{für } x \in \mathfrak{S}_n(\delta).$$

Somit ist

$$\varrho(\varphi, \varphi_n) \leq \int_{\mathfrak{S}_n(\delta)} \frac{\delta}{1 + \delta} dx = \frac{\delta}{1 + \delta} m[\mathfrak{S}_n(\delta)]$$

mit einem festen $\delta > 0$. Nach Voraussetzung ist $\varrho(\varphi, \varphi_n) \rightarrow 0$, so daß wir aus der letzten Ungleichung sehen, daß $m[\mathfrak{S}_n(\delta)]$ gegen 0 strebt, was zu beweisen war.

Mit Hilfe von Satz 7 aus [44] und dem oben Gesagten können wir behaupten, daß für $\varrho(\varphi, \varphi_n) \rightarrow 0$ in S eine Teilfolge $\varphi_{n_k}(x)$ existiert derart, daß $\varphi_{n_k}(x)$ fast überall auf $\mathfrak{S}$ gegen $\varphi(x)$ konvergiert. Die Vollständigkeit von S läßt sich genau wie die von L_2 nachweisen.

Bei der Konstruktion der Funktionenräume L_p , M und S hätten wir auch das Lebesgues-Stieltjessche Integral und das entsprechende Maß verwenden können.

88. Beispiele für die Anwendung des Prinzips der kontrahierenden Abbildungen.

1. Wir untersuchen ein System von n Gleichungen mit n Unbekannten:

$$\xi_i = \lambda \sum_{k=1}^n a_{ik} \xi_k + b_i \quad (i = 1, 2, \dots, n); \quad (21)$$

λ ist ein Zahlenparameter. Die rechten Seiten wollen wir als Operator Ax aus dem R_n in den R_n auffassen, der auf das Element $x(\xi_1, \xi_2, \dots, \xi_n)$ angewendet und im ganzen R_n definiert ist. Aus der Cauchyschen Ungleichung erhalten wir

$$\varrho(Ax, Ay) \leq |\lambda| \left[\sum_{i,k=1}^n |a_{ik}|^2 \right]^{1/2} \varrho(x, y).$$

Somit ist das Prinzip der kontrahierenden Abbildungen nur dann im R_n anwendbar, wenn

$$|\lambda| < \left[\sum_{i,k=1}^n |a_{ik}|^2 \right]^{-1/2}$$

ist.

2. Wir betrachten das unendliche Gleichungssystem

$$\xi_i = \lambda \sum_{k=1}^{\infty} a_{ik} \xi_k + b_i \quad (i = 1, 2, 3, \dots), \quad (22)$$

wobei wir annehmen, daß die Folge $(b_1, b_2, \dots)$ ein Element aus m ist. Bezeichnet

$$\sup_i \sum_{k=1}^{\infty} |a_{ik}| = c$$

eine endliche positive Zahl, so ist durch die rechten Seiten von (22) ein Operator A aus m in m gegeben, der auf ganz m definiert ist. Das Prinzip ist also dann anwendbar, wenn $|\lambda|c < 1$ ist. Ist $(b_1, b_2, \dots)$ ein Element aus l_2 und gilt

$$\sum_{i,k=1}^{\infty} |a_{ik}|^2 = d < \infty,$$

so ergeben die rechten Seiten von (22) einen in l_2 definierten Operator aus l_2 in l_2 ; das Prinzip kann in diesem Fall angewendet werden, wenn $|\lambda| d < 1$ ist. Die Lösbarkeit ist in diesen Räumen eindeutig bestimmt; es können aber noch Lösungen existieren, die nicht zu diesen Räumen gehören.

3. Wir betrachten die Integralgleichung

$$\varphi(x) = \lambda \int_a^b K(x, t) \varphi(t) dt + f(x) \quad (23)$$

(eindimensionaler Fall), wobei $[a, b]$ ein endliches Intervall und $K(x, t)$ im Quadrat Q ($a \leq x \leq b, a \leq t \leq b$) stetig ist. Ist $f(x)$ auf $[a, b]$ stetig, so ist die rechte Seite von (23) ein auf ganz C definierter Operator von $C[a, b]$ in $C[a, b]$. Das Prinzip ist auf (23) anwendbar, wenn

$$|\lambda| \max_{a \leq x \leq b} \int_a^b |K(x, t)| dt < 1$$

ist. Gehört $K(x, t)$ zu $L_2(Q)$ und $f(x)$ zu $L_2[a, b]$ (in diesem Fall kann das Intervall auch unendlich sein), so ist die rechte Seite von (23) ein auf ganz $L_2[a, b]$ definierter Operator von $L_2[a, b]$ in $L_2[a, b]$, und das Prinzip ist für

$$|\lambda| \left[\int_a^b \int_a^b |K(x, t)|^2 dx dt \right]^{1/2} < 1$$

anwendbar.

Das hier Gesagte gilt auch für mehrdimensionale Integralgleichungen.

4. Wir betrachten nun die nichtlineare Integralgleichung

$$\varphi(x) = \lambda \int_a^b K(x, t, \varphi(t)) dt, \quad (24)$$

wobei $[a, b]$ ein endliches Intervall, $K(x, t, z)$ eine für $a \leq x \leq b, a \leq t \leq b$ und $|z| \leq c$ stetige Funktion ihrer Argumente ist (c eine gegebene positive Zahl). Für jede Funktion $\varphi(t)$, die für $a \leq t \leq b$ stetig ist und der Bedingung $|\varphi(t)| \leq c$ genügt, ist $K(x, t, \varphi(t))$ eine stetige Funktion von x und t im Quadrat Q . Es sei $|K(x, t, z)| \leq d$ für $(x, t) \in Q$ und $|z| \leq c$. Ist $|\lambda| d (b-a) \leq c$, so definiert die rechte Seite von (24) einen Operator $A\varphi$ in $C[a, b]$, für den $D(A)$ mit der Sphäre $\varrho(0, \varphi) \leq c$ übereinstimmt. Dabei steht 0 für eine stetige Funktion, die auf $[a, b]$ verschwindet, und $R(A)$ gehört zu der Sphäre $\varrho(0, \varphi) \leq c$. Diese Ungleichung kann auch in der Form $|\varphi(x)| \leq c$ geschrieben werden. Außerdem nehmen wir an, der Kern $K(x, t, z)$ genüge bezüglich z einer Lipschitz-Bedingung, d. h., es sei

$$|K(x, t, z_1) - K(x, t, z_2)| \leq N |z_1 - z_2|$$

für $(x, t) \in Q$ und $|z_1| \leq c, |z_2| \leq c$. Dann ist

$$\varrho(A\varphi, A\psi) \leq |\lambda| N (b-a) \varrho(\varphi, \psi);$$

folglich läßt sich bei Berücksichtigung der Bedingungen

$$|\lambda| d (b-a) \leq c \quad \text{und} \quad |\lambda| N (b-a) < 1$$

auf die Gleichung (24) das Prinzip der kontrahierenden Abbildungen in der Sphäre $\varrho(0, \varphi) \leq c$ anwenden. Diese Gleichung hat in dieser Sphäre eine eindeutig bestimmte Lösung, die auch mit Hilfe sukzessiver Approximationen mit einer beliebigen Anfangsnäherung $\varphi_1(x)$ aus dieser Sphäre konstruiert werden kann. Diese Methode sichert die gleichmäßige Konvergenz der Näherungen gegen die Lösung auf dem Intervall $[a, b]$.

5. Es sei D ein von einer Ljapunoffschen Fläche¹⁾ S berandeter Bereich des dreidimensionalen Raumes. Wir betrachten das Randwertproblem für eine elliptische Differentialgleichung:

$$\Delta u - \lambda f(x, y, z, u) = 0 \quad \text{im Innern von } D, \quad (25)$$

$$u|_S = 0; \quad (26)$$

¹⁾ Vgl. [IV/2, 94] (Anm. d. Red.).

dabei ist Δ der Laplacesche Operator. Wir nehmen an, $f(x, y, z, u)$ sei in demjenigen vierdimensionalen abgeschlossenen Bereich des x, y, z, u -Raumes stetig, der dem Wertebereich von (x, y, z) im abgeschlossenen Bereich $\bar{D}$ für $|u| \leq c$ entspricht. Ferner habe die Funktion $f(x, y, z, u)$ stetige Ableitungen nach ihren Argumenten im Innern von D , wobei diese Ableitungen bis zu der Berandung von D stetig sind. Außerdem setzen wir $|f(x, y, z, u)| \leq d$ für $(x, y, z) \in D$ und $|u| \leq c$ sowie

$$|f(x, y, z, u_1) - f(x, y, z, u_2)| \leq N |u_1 - u_2|$$

für $(x, y, z) \in D$ und $|u_1| \leq c, |u_2| \leq c$. Es sei $G(x, y, z; \xi, \eta, \zeta)$ die Greensche Funktion des Laplaceschen Operators für den Bereich D unter der Randbedingung (26) [IV/2, 122]. Führen wir noch die Punkte $P(x, y, z)$ und $Q(\xi, \eta, \zeta)$ aus D ein, so ist die Lösung des Problems (25), (26) äquivalent mit der Lösung der Integralgleichung

$$u(P) = \lambda \int_D G(P; Q) f(Q; u(Q)) d\tau_Q \quad (27)$$

im Raum $C(\bar{D})$ der in D stetigen Funktionen $u(Q)$ [IV/2, 126]. Bekanntlich [IV/2, 123] ist $G(P; Q) \geq 0$ in D , und es existiert ein endliches

$$\max_{P \in \bar{D}} \int_D G(P; Q) d\tau_Q = G_0.$$

Ist $|\lambda| G_0 d \leq c$, so ist die rechte Seite von (27) ein Operator A in $C(\bar{D})$, für welchen $D(A)$ die Sphäre $\varrho(0, u) \leq c$ in $C(\bar{D})$ (also $|u(P)| \leq c$ in D) ist und $R(A)$ dieser Sphäre angehört. Das Prinzip der kontrahierenden Abbildungen ist auf (27) anwendbar, wenn $|\lambda| N G_0 < 1$ ist. Somit hat das Problem (25), (26) in der Sphäre $|u(P)| \leq c$ eine einzige Lösung, wenn die Bedingungen $|\lambda| G_0 d \leq c$ und $|\lambda| N G_0 < 1$ erfüllt sind. Diese Lösung kann auch mit Hilfe sukzessiver Approximationen aus (27) ermittelt werden, wenn wir eine beliebige Anfangsnäherung aus der erwähnten Sphäre wählen; die Näherungen konvergieren gleichmäßig gegen die Lösung in D .

89. Kompaktheit. Wir führten schon früher den Begriff der Kompaktheit für einen Spezialfall ein [IV/1, 36]. Jetzt wollen wir diesen Begriff für einen allgemeinen metrischen Raum X untersuchen. Eine Menge U von Elementen aus X heißt *im Raum X kompakt* oder einfach *kompakt*, wenn jede Folge von Elementen x_n aus U eine konvergente Teilfolge enthält. Ist die Menge U außerdem abgeschlossen, so heißt sie *in sich kompakt*.

Wir sehen leicht, daß die Beschränktheit der Menge U notwendig für ihre Kompaktheit ist. Wäre nämlich U nicht beschränkt, so würde eine Folge von Elementen $x_n \in U$ existieren, für die $\varrho(a, x_n) \rightarrow \infty$ mit einem beliebigen, aber festen Element a gilt. Aus dieser Folge läßt sich jedoch keine konvergente Teilfolge auswählen, denn jede konvergente Teilfolge ist beschränkt. Die Beschränktheit von U ist auch hinreichend für die Kompaktheit im R_n [IV/1, 15]. Im allgemeinen Fall des metrischen Raumes gilt dies nicht; deshalb geben wir jetzt notwendige und hinreichende Bedingungen für die Kompaktheit an.

Zunächst führen wir noch einen neuen Begriff ein. Wir sagen, die Menge U habe ein *endliches ε -Netz* (ε beliebig positiv), wenn eine endliche Menge von Elementen $x_k \in X$ ($k=1, 2, \dots, l$) derart existiert, daß sich für jedes $x \in U$ ein Element x_s mit $\varrho(x, x_s) \leq \varepsilon$ finden läßt. Die Elemente x_k brauchen nicht zu U zu gehören.

Satz. *Eine Menge U von Elementen eines vollständigen metrischen Raumes ist genau dann kompakt, wenn sie für jedes $\varepsilon > 0$ ein endliches ε -Netz besitzt.*

Beweis. Um die Notwendigkeit zu beweisen, nehmen wir an, U habe für ein gewisses $\varepsilon_0 > 0$ kein endliches ε_0 -Netz, und zeigen, daß U nicht kompakt ist. Es sei x_1 ein Element aus U ; dann gibt es ein solches Element $x_2 \in U$, für das $\varrho(x_1, x_2) > \varepsilon_0$ ist. Anderenfalls wäre nämlich $\varrho(x_1, x) \leq \varepsilon_0$ für jedes $x \in U$, und das eine Element

x_1 würde ein ε_0 -Netz für die Menge U ergeben. Ferner läßt sich ein Element x_3 mit $\varrho(x_i, x_3) > \varepsilon_0$ ($i=1, 2$) angeben; denn anderenfalls würden die Elemente x_1 und x_2 ein ε_0 -Netz für U ergeben, usw. Somit erhalten wir eine unendliche Folge von Elementen $x_n \in U$ mit $\varrho(x_p, x_q) > \varepsilon_0$ für alle p und $q \neq p$. Für eine beliebige Teilfolge x_{n_k} ($k=1, 2, \dots$) würde ebenfalls $\varrho(x_{n_k}, x_{n_l}) > \varepsilon_0$ für $n_k \neq n_l$ gelten; demzufolge ist keine Teilfolge der x_n konvergent, d. h., U ist nicht kompakt.

Um zu zeigen, daß die Bedingung hinreichend ist, setzen wir voraus, U habe ein endliches ε -Netz für jedes $\varepsilon > 0$, und es sei eine Folge von Elementen $x_n \in U$ gegeben. Wir müssen also nachweisen, daß sich aus dieser Folge eine konvergente Teilfolge auswählen läßt. Stimmen für unendlich viele n die Elemente x_n mit ein und demselben Element y überein, so ist die Teilfolge $y, y, y, \dots$ konvergent. Wir nehmen an, dies sei nicht der Fall. Dann erhalten wir eine unendliche Folge von verschiedenen Elementen, wenn wir in der Folge der x_n aus der Gruppe der gleichen Elemente nur ein Element (etwa das mit dem kleinsten Index) übriglassen. Wir können annehmen, daß schon die Grundfolge diese Eigenschaft besitzt. Da für U ein endliches $\frac{\varepsilon}{2}$ -Netz existiert (ε eine beliebige, aber feste Zahl), gibt es endlich viele abgeschlossene Sphären mit dem Radius $\frac{\varepsilon}{2}$ derart, daß alle Elemente von U und damit auch alle x_n in diesen Sphären liegen. Wenigstens eine von ihnen enthält unendlich viele x_n . Diese Sphäre bezeichnen wir mit $S_1\left(\frac{\varepsilon}{2}\right)$. Ferner gibt es endlich viele Sphären mit dem Radius $\frac{\varepsilon}{2^2}$, zu denen alle x_n aus $S_1\left(\frac{\varepsilon}{2}\right)$ gehören. Aus diesen wählen wir die Sphäre $S_2\left(\frac{\varepsilon}{2^2}\right)$, die unendlich viele der erwähnten Elemente enthält. Genauso existiert eine Sphäre $S_3\left(\frac{\varepsilon}{2^3}\right)$ mit dem Radius $\frac{\varepsilon}{2^3}$, in der unendlich viele Elemente x_n enthalten sind, die gleichzeitig zu $S_1\left(\frac{\varepsilon}{2}\right)$ und $S_2\left(\frac{\varepsilon}{2^2}\right)$ gehören. Setzen wir diesen Prozeß fort, so ergibt sich eine unendliche Folge abgeschlossener Sphären $S_k\left(\frac{\varepsilon}{2^k}\right)$ derart, daß $S_k\left(\frac{\varepsilon}{2^k}\right)$ den Radius $\frac{\varepsilon}{2^k}$ hat und unendlich viele Elemente x_n enthält, die gleichzeitig allen Sphären $S_m\left(\frac{\varepsilon}{2^m}\right)$ mit $m < k$ angehören. Aus jedem $S_k\left(\frac{\varepsilon}{2^k}\right)$ wählen wir ein Element x_{n_k} , wobei wir $n_l > n_k$ für $l > k$ annehmen. Auf diese Art ergibt sich eine unendliche Teilfolge von Elementen x_{n_k} . Da für zwei beliebige Elemente x und y , die in der gleichen Sphäre vom Radius r liegen, auf Grund der Dreiecksungleichung die Beziehung $\varrho(x, y) \leq 2r$ gilt, ist

$$\varrho(x_{n_l}, x_{n_k}) \leq \frac{\varepsilon}{2^{k-1}} \quad \text{für } n_l > n_k.$$

Hieraus folgt wegen der Vollständigkeit des Raumes, daß die x_{n_k} eine konvergente Teilfolge bilden. Damit ist der Satz bewiesen.

Bemerkung 1. Für die Kompaktheit ist nicht die Existenz eines endlichen, sondern nur eines kompakten ε -Netzes für jedes $\varepsilon > 0$ hinreichend. Das bedeutet, daß es zu jedem $\varepsilon > 0$ Sphären vom Radius ε gibt, die alle Elemente von U enthalten und deren Mittelpunkte eine kompakte Menge bilden. Diese Menge wollen wir mit U_1 bezeichnen. Für U_1 existiert auf Grund des bewiesenen Satzes (erster Teil des Beweises) ein endliches ε -Netz; aus der Dreiecksungleichung ergibt sich

unmittelbar, daß dieses Netz ein endliches 2ε -Netz für U ist. Daraus folgt, da ε beliebig ist, auf Grund des zweiten Teils des Beweises, daß U eine kompakte Menge ist.

Bemerkung 2. Die Menge U kann mit X übereinstimmen, so daß wir auch von der Kompaktheit des ganzen Raumes X sprechen können. Mit Hilfe der Stetigkeit des Abstands läßt sich leicht zeigen, daß jeder kompakte Raum vollständig ist. Somit bildet jede in sich kompakte Menge U von Elementen aus X einen vollständigen metrischen Raum.

90. Kompaktheit im Raum C . Es sei C der Raum der auf einem endlichen Intervall $[a, b]$ stetigen Funktionen und U eine gewisse Menge von Elementen aus C . Wir sahen [IV/1, 15], daß die Beschränktheit und die gleichgradige Stetigkeit der Funktionen aus U eine hinreichende Bedingung für die Kompaktheit von U sind. Nun wollen wir beweisen, daß diese Bedingungen auch notwendig sind. Es sei U kompakt. Auf Grund des Satzes aus [89] gibt es für jedes $\varepsilon > 0$ endlich viele Funktionen $\varphi_1(t), \varphi_2(t), \dots, \varphi_p(t)$ aus C mit $|\varphi(t) - \varphi_s(t)| \leq \frac{\varepsilon}{3}$ für $a \leq t \leq b$; dabei ist $\varphi(t)$ eine beliebige Funktion aus U und $\varphi_s(t)$ eine der obigen Funktionen. Für alle diese Funktionen gibt es, da von ihnen nur endlich viele existieren, eine nur von ε abhängende positive Zahl η mit

$$|\varphi_k(t+h) - \varphi_k(t)| \leq \frac{\varepsilon}{3} \quad \text{für } |h| \leq \eta; \quad k = 1, 2, \dots, p; \quad t \text{ und } t+h \in [a, b].$$

Daraus erhalten wir

$$|\varphi(t+h) - \varphi(t)| \leq |\varphi(t+h) - \varphi_s(t+h)| + |\varphi_s(t+h) - \varphi_s(t)| + |\varphi_s(t) - \varphi(t)|.$$

Für $|h| \leq \eta$ ist jeder der Summanden auf der rechten Seite dieser Ungleichung höchstens gleich $\frac{\varepsilon}{3}$, also gilt $|\varphi(t+h) - \varphi(t)| \leq \varepsilon$ für $|h| \leq \eta$. Damit ist die gleichgradige Stetigkeit der Funktionen aus U bewiesen. Ihre Beschränktheit folgt unmittelbar daraus, daß die Beschränktheit von U eine notwendige Bedingung für die Kompaktheit ist [89]. Das eben bewiesene Kriterium für die Kompaktheit läßt sich genauso auch für Funktionen mehrerer Veränderlicher nachweisen, die auf einem endlichen abgeschlossenen Bereich des R_n definiert sind. Auch für Funktionen, die auf einer beschränkten abgeschlossenen Menge definiert sind, verläuft der Beweis analog.

91. Kompaktheit im Raum L_p . Wir untersuchen den Raum L_p der Funktionen $\varphi(x, y)$ auf einer meßbaren beschränkten Menge $\mathfrak{E}$ der x, y -Ebene und nehmen im folgenden an, daß die Fortsetzung dieser Funktionen ins Äußere von $\mathfrak{E}$ gleich 0 ist. Das Integralzeichen ohne Hinweis auf das Gebiet bedeutet, daß über die ganze x, y -Ebene integriert wird. Die Integrale sind also tatsächlich Integrale über beschränkte meßbare Mengen.

Satz 1. Eine Menge U von Elementen aus L_p ist dann und nur dann kompakt, wenn alle Funktionen $\varphi(x, y) \in U$ den folgenden Bedingungen genügen:

a) Es existiert ein $c > 0$, für welches

$$\|\varphi\| = \int |\varphi(x, y)|^p dx dy^{1/p} \leq c \quad (28)$$

gilt (Beschränktheit); die Größe $\|\varphi\|$ nennen wir die Norm von $\varphi(x, y)$ in $L_p(\mathfrak{E})$ [62];

b) zu jedem $\varepsilon > 0$ gibt es ein $\eta > 0$ (für alle $\varphi(x, y) \in U$ ein und dasselbe) derart, daß

$$\|\delta_{hk}\varphi\| = \left[\int |\varphi(x+h, y+k) - \varphi(x, y)|^p dx dy \right]^{1/p} \leq \varepsilon \quad \text{für } \sqrt{h^2 + k^2} \leq \eta \quad (29)$$

gilt.

Wir wissen, daß für jede feste Funktion aus L_p und für ein gegebenes $\varepsilon > 0$ ein solches positives η existiert, daß (29) gilt (Stetigkeit im Mittel; vgl. [70]). Das gilt offenbar auch für endlich viele Funktionen $\varphi_s(x, y) \in L_p$ ($s = 1, 2, \dots, n$). Es genügt dann, das kleinste der η zu nehmen, die den $\varphi_s(x, y)$ entsprechen. Die Eigenschaft (29), die für alle $\varphi(x, y) \in U$ gelten muß, können wir *gleichgradige Stetigkeit im Mittel* aller Funktionen aus U nennen. Wir bemerken noch, daß $\varphi(x+h, y+k)$ eine meßbare Funktion und $\varphi(x+h, y+k) - \varphi(x, y) = 0$ außerhalb einer gewissen beschränkten meßbaren Menge ist.

Die Notwendigkeit der Bedingungen (28) und (29) läßt sich wie für den Raum C beweisen. Nur müssen überall die Absolutbeträge der Differenz $\varphi - \psi$ durch $\|\varphi - \psi\| = \varrho(\varphi, \psi)$ in L_p ersetzt werden. Die Beziehung (28), die wir auch in der Form $\varrho(0, \varphi) \leq c$ schreiben können, ist, wie wir wissen, notwendig für die Kompaktheit. Ferner folgt aus der Kompaktheit die Existenz eines endlichen $\frac{\varepsilon}{3}$ -Netzes, d. h. die Existenz endlich vieler Funktionen $\psi_k(x, y) \in L_p$ ($k = 1, 2, \dots, n$), unter denen sich zu jedem $\varphi(x, y) \in U$ eine Funktion $\psi_s(x, y)$ mit $\|\varphi - \psi_s\| \leq \frac{\varepsilon}{3}$ finden läßt. Für alle $\psi_k(x, y)$ gibt es nämlich ein $\eta > 0$ derart, daß die Bedingung

$$\|\delta_{hk}\psi_s\| \leq \frac{\varepsilon}{3} \quad \text{für} \quad \sqrt{h^2 + k^2} \leq \eta \quad (30)$$

erfüllt ist. Weiterhin schreiben wir

$$\begin{aligned} |\varphi(x+h, y+k) - \varphi(x, y)| &\leq |\varphi(x+h, y+k) - \varphi_s(x+h, y+k)| \\ &\quad + |\varphi_s(x+h, y+k) - \varphi_s(x, y)| + |\varphi_s(x, y) - \varphi(x, y)| \end{aligned}$$

und erhalten, wenn wir die Minkowskische Ungleichung für $p > 1$ anwenden, die Bedingung (29) wegen (30) und der Ungleichung $\|\varphi - \psi_s\| \leq \frac{\varepsilon}{3}$. Für $p = 1$ ergibt sich (29) unmittelbar.

Wir wollen nun beweisen, daß die Bedingungen (28) und (29) hinreichend sind. Mit $\varphi_\varrho(x, y)$ bezeichnen wir die Mittelfunktion für $\varphi(x, y)$ [71]. Die Formel (178) aus [71] hat für jedes $p \geq 1$ die Gestalt

$$\|\varphi - \varphi_\varrho\|^p \leq \frac{c_1}{\varrho^2} \int_{u^2+v^2 \leq \varrho^2} \left[\int |\varphi(\xi, \eta) - \varphi(\xi+u, \eta+v)|^p d\xi d\eta \right] dudv, \quad (31)$$

wobei c_1 eine Konstante ist. Auf Grund von (29) existiert für jedes gegebene $\varepsilon > 0$ ein $\eta > 0$ (für alle Funktionen $\varphi(x, y) \in U$ ein und dasselbe), für welches

$$\int |\varphi(\xi, \eta) - \varphi(\xi+u, \eta+v)|^p d\xi d\eta \leq \frac{\varepsilon^p}{c_1 \pi} \quad \text{für} \quad u^2 + v^2 \leq \eta^2$$

gilt. Somit folgt aus (31) für jedes $\varphi(x, y) \in U$

$$\|\varphi - \varphi_\varrho\| \leq \varepsilon \quad \text{für} \quad \varrho \leq \eta. \quad (32)$$

Die Norm $\|\varphi - \varphi_\varrho\|$ ist bezüglich der ganzen Ebene $\mathbb{E}_\infty$ zu verstehen (tatsächlich bezüglich einer beschränkten Menge). Dann ist erst recht $\|\varphi - \varphi_\varrho\| \leq \varepsilon$ für $\varrho \leq \eta$. Wählen wir ein festes $\varrho \leq \eta$, so können wir behaupten, daß die Funktionen $\varphi_\varrho(x, y)$ ein ε -Netz für die Menge U der Funktionen $\varphi(x, y)$ bilden. Es sei Δ ($a \leq x \leq b$; $c \leq y \leq d$) ein die Menge $\mathbb{E}$ enthaltendes Intervall. Wegen (28) und Satz 3 aus [71] ist die Menge der $\varphi_\varrho(x, y)$ kompakt in $C(\Delta)$ und dann erst recht kompakt in $L_p(\mathbb{E})$. Somit bilden die Funktionen $\varphi_\varrho(x, y)$ ein kompaktes ε -Netz für U , und da ε beliebig ist,

folgt die Kompaktheit von U . Damit ist bewiesen, daß die Bedingungen (28) und (29) hinreichend sind.

Wir untersuchen nun den Fall, daß $\mathfrak{E}$ die ganze Ebene $\mathfrak{E}_\infty$ ist. Der vorhergehende Beweis verliert dann seine Gültigkeit, da die Menge der Funktionen $\varphi_\alpha(x, y)$ nicht kompakt zu sein braucht. Die Notwendigkeit von (28) und (29) wird auch wie oben nachgewiesen. Aber diese Bedingungen sind nicht hinreichend. Wir müssen ihnen noch eine Forderung zur Seite stellen, nämlich: Für jedes gegebene $\varepsilon > 0$ existiert ein positives N (für alle Funktionen $\varphi(x, y) \in U$ das gleiche), für welches

$$\int_{\mathfrak{E}_\infty - \mathcal{A}_N} |\varphi(x, y)|^p dx dy \leq \varepsilon^p \quad (33)$$

gilt, wobei $\mathcal{A}_m$ das Intervall $(-m \leq x \leq m; -m \leq y \leq m)$ bezeichnet. Ist die Ungleichung (33) für ein gewisses N erfüllt, so gilt sie auch für größere N .

Wir zeigen nun, daß (33) eine notwendige Bedingung ist. Ist die Menge U kompakt, so existiert für sie ein endliches $\frac{\varepsilon}{2}$ -Netz von Funktionen $\varphi_k(x, y)$, $k = 1, 2, \dots, n$.

Für diese endlich vielen Funktionen existiert nach (33) zu jedem gegebenen $\varepsilon > 0$ ein N derart, daß

$$\int_{\mathfrak{E}_\infty - \mathcal{A}_N} |\varphi_k(x, y)|^p dx dy \leq \frac{\varepsilon^p}{2^p} \quad (k = 1, 2, \dots, n) \quad (34)$$

gilt. Wir wählen dann ein $\varphi(x, y) \in U$. Zu ihm läßt sich eine Funktion $\varphi_s(x, y)$ mit $\|\varphi - \varphi_s\|_{\mathfrak{E}_\infty} \leq \frac{\varepsilon}{2}$ finden. Aus

$$\|\varphi\|_{\mathfrak{E}_\infty - \mathcal{A}_N} \leq \|\varphi - \varphi_s\|_{\mathfrak{E}_\infty - \mathcal{A}_N} + \|\varphi_s\|_{\mathfrak{E}_\infty - \mathcal{A}_N},$$

der Ungleichung (34) und aus

$$\|\varphi - \varphi_s\|_{\mathfrak{E}_\infty - \mathcal{A}_N} \leq \|\varphi - \varphi_s\|_{\mathfrak{E}_\infty} \leq \frac{\varepsilon}{2}$$

folgt

$$\|\varphi\|_{\mathfrak{E}_\infty - \mathcal{A}_N} = \left[\int_{\mathfrak{E}_\infty - \mathcal{A}_N} |\varphi(x, y)|^p dx dy \right]^{1/p} \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

(33) gilt also für jedes $\varphi(x, y) \in U$.

Wir zeigen jetzt, daß die Bedingungen (28), (29) und (33) hinreichend sind. Sie seien für die Funktionen $\varphi(x, y) \in U$ erfüllt, und $\varphi_1(x, y)$, $\varphi_2(x, y)$, ... sei eine beliebige Folge von Funktionen aus U . Wir müssen also aus dieser Folge eine konvergente Teilfolge auswählen können, die in $L_p(\mathfrak{E}_\infty)$ konvergiert. Aus (28) und (29) ergibt sich, daß wir eine Teilfolge $\varphi_{n_1(1)}(x, y)$, $\varphi_{n_2(1)}(x, y)$, ... auswählen können, die in $L_p(\mathcal{A}_1)$ konvergiert. Aus dieser Teilfolge wählen wir eine weitere Teilfolge $\varphi_{n_1(2)}(x, y)$, $\varphi_{n_2(2)}(x, y)$, ... aus, die in $L_p(\mathcal{A}_2)$ konvergiert, usw. Dann konstruieren wir die Folge

$$\varphi_{n_1(1)}(x, y), \quad \varphi_{n_2(2)}(x, y), \quad \varphi_{n_3(3)}(x, y), \dots, \quad (35)$$

die ebenfalls Teilfolge der ursprünglichen Folge der $\varphi_k(x, y)$ ist: Für ein beliebiges natürliches m gehören alle Glieder von (35), angefangen mit $\varphi_{n_m(m)}(x, y)$, zu derjenigen Folge $\varphi_{n_1(m)}(x, y)$, $\varphi_{n_2(m)}(x, y)$, ..., die in $L_p(\mathcal{A}_m)$ konvergiert, d. h., (35) konver-

giert in $L_p(\Delta_m)$ für jedes endliche Intervall Δ_m ($m=1, 2, \dots$). Wir zeigen, daß (35) auch in $L(\mathcal{E}_\infty)$ konvergiert, und betrachten dazu das Integral

$$\begin{aligned} \|\varphi_{n_q} - \varphi_{n_r}\|_{\mathcal{E}_\infty}^p &= \int_{\Delta_m} |\varphi_{n_q}(x, y) - \varphi_{n_r}(x, y)|^p dx dy \\ &\quad + \int_{\mathcal{E}_\infty - \Delta_m} |\varphi_{n_q}(x, y) - \varphi_{n_r}(x, y)|^p dx dy. \end{aligned}$$

Berücksichtigen wir die offenbar gültige Ungleichung $|x+y|^p \leq 2^p|x|^p + 2^p|y|^p$, so erhalten wir

$$\begin{aligned} \|\varphi_{n_q} - \varphi_{n_r}\|_{\mathcal{E}_\infty}^p &\leq \int_{\Delta_m} |\varphi_{n_q}(x, y) - \varphi_{n_r}(x, y)|^p dx dy + 2^p \int_{\mathcal{E}_\infty - \Delta_m} |\varphi_{n_q}(x, y)|^p dx dy \\ &\quad + 2^p \int_{\mathcal{E}_\infty - \Delta_m} |\varphi_{n_r}(x, y)|^p dx dy. \end{aligned}$$

Auf Grund von (33) existiert für jedes gegebene $\varepsilon > 0$ ein solches m , daß auf der rechten Seite dieser Ungleichung die Summe der letzten beiden Glieder kleiner als $\frac{\varepsilon^p}{2}$ ist. Für dieses feste m ergibt sich

$$\|\varphi_{n_q} - \varphi_{n_r}\|_{\mathcal{E}_\infty}^p \leq \int_{\Delta_m} |\varphi_{n_q}(x, y) - \varphi_{n_r}(x, y)|^p dx dy + \frac{\varepsilon^p}{2}.$$

Aus der Konvergenz von (35) in $L_p(\Delta_m)$ folgt, daß das auf der rechten Seite stehende Integral nicht größer ist als $\frac{\varepsilon^p}{2}$ für alle hinreichend großen q und r ; somit gibt es ein M , für welches

$$\|\varphi_{n_q} - \varphi_{n_r}\|_{\mathcal{E}_\infty}^p \leq \varepsilon^p \quad \text{für } q, r \geq M$$

ist, d. h., die Folge (35) ist in $L_p(\mathcal{E}_\infty)$ eine in sich konvergente Folge und hat wegen der Vollständigkeit von $L_p(\mathcal{E}_\infty)$ einen Grenzwert in $L_p(\mathcal{E}_\infty)$. Damit ist bewiesen, daß die Bedingungen (28), (29) und (33) hinreichend sind. Man kann sich leicht davon überzeugen, daß die letzte Forderung nicht eine Folge der ersten beiden ist.

Der Bestimmtheit wegen betrachteten wir die Ebene. Alles oben Gesagte gilt jedoch auch im R_n ($n > 2$). Bezeichnen wir mit $x(x_1, x_2, \dots, x_n)$ einen Punkt des R_n und setzen wir $dx = dx_1 dx_2 \dots dx_n$, so können wir den Forderungen (28) und (29) die Gestalt

$$[\int |\varphi(x)|^p dx]^{1/p} \leq c, \quad (36)$$

$$[\int |\varphi(x+y) - \varphi(x)|^p dx]^{1/2} \leq \varepsilon \quad \text{für } |y| \leq \eta \quad (37)$$

geben, wobei y die Komponenten $(y_1, y_2, \dots, y_n)$ hat und $|y| = \sqrt{y_1^2 + y_2^2 + \dots + y_n^2}$ ist.

92. Kompaktheit im Raum l_p . Wir beweisen den folgenden

Satz. Eine Menge U von Elementen aus l_p ($p \geq 1$) ist dann und nur dann kompakt, wenn alle Elemente $x(\xi_1, \xi_2, \dots)$ aus U den beiden folgenden Bedingungen genügen:

a) Es existiert eine Zahl $c > 0$, für welche

$$\|x\| = \left(\sum_{l=1}^{\infty} |\xi_l|^p \right)^{1/p} \leq c \quad (38)$$

ist (Beschränktheit);

b) zu jedem gegebenen $\varepsilon > 0$ existiert eine ganze positive Zahl n_ε (für alle $x \in U$ die gleiche) mit

$$\left(\sum_{i=n_\varepsilon}^{\infty} |\xi_i|^p \right)^{1/p} \leq \varepsilon. \quad (39)$$

Beweis. Zunächst beweisen wir die Notwendigkeit. Wie wir wissen, ist Beschränktheit (also die Beziehung (38)) notwendig für die Kompaktheit. Ferner folgt aus der Kompaktheit von U die Existenz endlich vieler Elemente x_k ($k=1, 2, \dots, m$) aus l_p derart, daß $\varrho(x, x_s) = \|x - x_s\| \leq \frac{\varepsilon}{2}$ für ein beliebiges $x \in U$ ist; das Element x_s ist eines der Elemente x_k . Für die endlich vielen Elemente $x_k(\xi_1^{(k)}, \xi_2^{(k)}, \dots)$ gibt es eine ganze positive Zahl n_ε , so daß die Ungleichung

$$\left(\sum_{i=n_\varepsilon}^{\infty} |\xi_i^{(k)}|^p \right)^{1/p} \leq \frac{\varepsilon}{2} \quad (k=1, 2, \dots, m) \quad (40)$$

erfüllt ist. Aus $\|x - x_s\| \leq \frac{\varepsilon}{2}$ ergibt sich

$$\left(\sum_{i=n_\varepsilon}^{\infty} |\xi_i - \xi_i^{(s)}|^p \right)^{1/p} \leq \left(\sum_{i=1}^{\infty} |\xi_i - \xi_i^{(s)}|^p \right)^{1/p} \leq \frac{\varepsilon}{2},$$

und wenden wir die Minkowskische Ungleichung für Summen an ($p > 1$), so erhalten wir

$$\left(\sum_{i=n_\varepsilon}^{\infty} |\xi_i|^p \right)^{1/p} \leq \left(\sum_{i=n_\varepsilon}^{\infty} |\xi_i - \xi_i^{(s)}|^p \right)^{1/p} + \left(\sum_{i=n_\varepsilon}^{\infty} |\xi_i^{(s)}|^p \right)^{1/p} \leq \varepsilon.$$

Damit ist die Notwendigkeit von (39) gezeigt. Ist $p=1$, so brauchen wir die Minkowskische Ungleichung nicht.

Um zu beweisen, daß die Bedingungen hinreichend sind, setzen wir voraus, die Elemente von U genügen den Bedingungen (38) und (39), und weisen die Kompaktheit von U nach. Es sei $\varepsilon > 0$. Jedem Element $(\xi_1, \xi_2, \dots)$ aus U ordnen wir das Abschnittselement $(\xi_1, \xi_2, \dots, \xi_{n_\varepsilon-1}, 0, \dots)$ zu, und es sei U_ε die Menge dieser Abschnittselemente. Aus (39) folgt, daß zu jedem $x \in U$ ein Element $y \in U_\varepsilon$ mit $\|x - y\| \leq \varepsilon$ existiert, d. h., die Menge U_ε ist ein ε -Netz für U . Es genügt zu zeigen, daß U_ε eine kompakte Menge ist [89]. Der Beweis verläuft analog zu dem, daß jede im R_n beschränkte Menge kompakt ist.

Auf Grund von (38) ist $|\xi_s| \leq c$ für eine beliebige Komponente aus U_ε der Elemente von U . Aus jeder Folge von Elementen aus U_ε können wir eine Teilfolge auswählen, in der die ersten $n_\varepsilon - 1$ Komponenten einen endlichen Grenzwert besitzen. Die übrigen Komponenten dieser Elemente sind gleich 0. Daraus folgt, daß die Teilfolge in l_p gegen ein Element konvergiert, dessen sämtliche Komponenten ξ_s für $s \geq n_\varepsilon$ gleich 0 sind. Damit ist die Kompaktheit von U_ε bewiesen.

Aus diesem Satz folgt, daß die Sphäre $\|x\| \leq r$ in l_p nicht kompakt ist.

93. Funktionale auf in sich kompakten Mengen. Wir nehmen an, daß ein reellwertiges Funktional $l(x)$ auf einer in sich kompakten Menge U des metrischen Raumes X definiert ist. Wir nennen dieses Funktional *stetig*, wenn $l(x_n) \rightarrow l(x_0)$ aus $x_n \rightarrow x_0$ folgt.

Für solche Funktionale gilt ein Satz, der dem über stetige Funktionen auf beschränkten abgeschlossenen Mengen des R_n analog ist:

Satz 1. Ist U eine in sich kompakte Menge des Raumes X und $l(x)$ ein reelles stetiges Funktional auf U , so ist $l(x)$ beschränkt und erreicht auf U seine untere sowie seine obere Grenze.

Beweis. Wir werden nur zeigen, daß das Funktional von unten beschränkt ist und seine untere Grenze erreicht; den Beweis führen wir indirekt. Ist die Menge der Werte von $l(x)$ nicht von unten beschränkt, so existiert eine Folge von Elementen $x_n \in U$ derart, daß $l(x_n)$ gegen $-\infty$ strebt. Da U kompakt ist, können wir aus dieser Folge eine konvergente Teilfolge $x_{n_k} \rightarrow x_0$ auswählen; wegen der Kompaktheit in sich gehört x_0 zu U . Dann gilt $l(x_{n_k}) \rightarrow l(x_0)$ wegen der Stetigkeit von $l(x)$; diesem widerspricht aber $l(x_{n_k}) \rightarrow -\infty$; denn $l(x_0)$ ist eine endliche Zahl.

Es sei a die untere Grenze für die Menge der Werte von $l(x)$ auf U . Dann existiert eine Folge von Elementen $x_n \in U$ derart, daß $a \leq l(x_n) \leq a + \frac{1}{n}$ gilt. Wie oben können wir $x_{n_k} \rightarrow x_0$ mit $x_0 \in U$ annehmen; folglich ist $l(x_{n_k}) \rightarrow l(x_0)$. Aus $a \leq l(x_{n_k}) \leq a + \frac{1}{n_k}$ ergibt sich aber $l(x_{n_k}) \rightarrow a$. Also ist $l(x_0) = a$, was zu beweisen war.

Früher führten wir den Begriff des unteren und des oberen Limes einer Folge reeller Zahlen a_n ($n = 1, 2, \dots$) ein. Dafür können wir auch

$$S = \underline{\lim} a_n, \quad T = \overline{\lim} a_n$$

schreiben. Diese Ausdrücke können auch gleich ∞ oder $-\infty$ sein.

Hat die Folge der a_n einen Grenzwert, so stimmen S und T mit ihm überein. Außerdem besagt die Definition von S und T , daß keine Teilfolge a_{n_k} der Folge a_n einen Grenzwert haben kann, der kleiner als S oder größer als T ist, aber es gibt eine Teilfolge, die den Grenzwert S , und eine, die den Grenzwert T hat. Das Funktional $l(x)$ heißt *von unten halbstetig* auf U , wenn $\underline{\lim} l(x_n) \geq l(x_0)$ aus $x_n \rightarrow x_0$, und *von oben halbstetig*, wenn $\overline{\lim} l(x_n) \leq l(x_0)$ aus $x_n \rightarrow x_0$ folgt.

Wir beweisen nun eine für die Anwendungen wichtige Verallgemeinerung von Satz 1.

Satz 2. Ein Funktional $l(x)$, das auf einer in sich kompakten Menge U eines metrischen Raumes definiert und von unten (von oben) halbstetig ist, ist von unten (von oben) beschränkt und erreicht auf U seine untere (obere) Grenze.

Beweis. Wir nehmen ein von unten halbstetiges Funktional und zeigen wie in Satz 1 seine Beschränktheit von unten. Die Annahme $l(x_n) \rightarrow -\infty$ führt auf eine Teilfolge $x_{n_k} \rightarrow x_0$, wobei $x_0 \in U$ und $l(x_{n_k}) \rightarrow -\infty$ ist. Auf Grund der Halbstetigkeit von unten gilt $\underline{\lim} l(x_{n_k}) \geq l(x_0)$; diesem widerspricht aber $l(x_{n_k}) \rightarrow -\infty$, da $l(x_0)$ eine endliche Zahl ist.

Es sei a die untere Grenze für die Menge der Werte von $l(x)$ auf U . Wie beim Beweis von Satz 1 erhalten wir $x_{n_k} \rightarrow x_0$ und $a \leq l(x_{n_k}) \leq a + \frac{1}{n_k}$. Aus dem ersten Teil dieser Ungleichung ergibt sich $\underline{\lim} l(x_{n_k}) \geq l(x_0)$, aus dem zweiten Teil $\lim l(x_{n_k}) = a$, also $l(x_0) \leq a$. Da a aber die untere Grenze der Werte von $l(x)$ sein soll, muß $l(x_0) = a$ gelten, was zu beweisen war.

94. Separabilität. Ein metrischer Raum X mit unendlich vielen Elementen heißt *separabel*, wenn eine abzählbare und in X dichte Menge von Elementen $x_1, x_2, \dots$

aus X existiert, d. h., wenn es für jedes $x \in X$ und jedes $\varepsilon > 0$ dieser Menge ein Element x_s gibt, für welches $\varrho(x, x_s) \leq \varepsilon$ gilt.

Schon früher zeigten wir die Separabilität von l_p und L_p ($p \geq 1$) [59, 60]. Im Raum C ist die eben erwähnte Menge z. B. die Menge aller Polynome mit rationalen Koeffizienten. Im Raum R_n ist diese Menge die Menge der Elemente $(a_1, a_2, \dots, a_n)$, bei denen alle Zahlen a_k rational sind oder (im Fall eines komplexen Raumes) die Form $a_k = \alpha_k + \beta_k i$ haben, wobei α_k und β_k reelle rationale Zahlen sind.

Im Raum s ist diese Menge gleich der Menge der Elemente $(a_1, a_2, \dots, a_n, 0, \dots)$ mit rationalen a_k .

Wir zeigen nun, daß der Raum m nicht separabel ist. Dazu betrachten wir eine Menge U von verschiedenen Elementen $x(a_1, a_2, \dots)$ aus m , bei denen die Zahlen a_k entweder gleich 0 oder 1 sind. Aus der Betrachtung im Dualsystem geschriebener Zahlen $0, a_1 a_2 a_3 \dots$ folgt, daß die Menge U nicht abzählbar ist. Berücksichtigen wir das in [1] Gesagte, so sehen wir, daß die Menge U die Mächtigkeit des Kontinuums besitzt. Für je zwei verschiedene Elemente x und y aus U gilt $\varrho(x, y) = 1$. Wir nehmen nun an, der Raum m sei separabel, d. h., es gebe eine abzählbare und in m dichte Menge von Elementen x_k ($k = 1, 2, \dots$) aus m , und mit S_k bezeichnen wir die Sphären mit dem Mittelpunkt x_k und dem Radius $\frac{1}{3}$. Die Menge dieser Sphären ist abzählbar, und wenigstens eine von ihnen enthält mehr als ein Element aus U . Sind y und z zwei verschiedene Elemente aus U , die sich in einer Sphäre befinden, so ist $\varrho(y, z) \leq 2 \cdot \frac{1}{3}$. Dies widerspricht aber der Beziehung $\varrho(x, y) = 1$. Damit ist bewiesen, daß m nicht separabel ist.

Satz. Jede Menge U von Elementen eines separablen Raumes X ist ebenfalls separabel.

Beweis. Wir müssen also die Existenz einer endlichen oder abzählbaren Menge von Elementen aus U nachweisen, die in U dicht ist. Wegen der Separabilität von X gibt es eine abzählbare Menge von Elementen x_n ($n = 1, 2, \dots$) aus X , die in X dicht ist. Mit $S(x_n, r)$ bezeichnen wir eine Sphäre mit dem Mittelpunkt x_n und dem Radius r . Wir betrachten die Sphären $S\left(x_n, \frac{1}{2^k}\right)$, $k = 1, 2, \dots$; enthält irgendeine dieser Sphären Elemente von U , so greifen wir eines dieser Elemente heraus. So erhalten wir eine endliche oder abzählbare Menge von Elementen u_m ($m = 1, 2, \dots$) aus U . Nun sei u ein beliebiges Element aus U und ε eine gegebene positive Zahl. Wir werden beweisen, daß wenigstens für eines der Elemente u_m die Ungleichung $\|u - u_m\| \leq \varepsilon$ erfüllt ist. Dabei können wir $\varepsilon < 1$ annehmen, so daß eine ganze positive Zahl l existiert, für welche

$$\frac{1}{2^{l-1}} \leq \varepsilon < \frac{1}{2^{l-2}}$$

gilt. Da die Menge der x_n in X dicht ist, gibt es ein $n = n_0$ mit $\|u - x_{n_0}\| \leq \frac{\varepsilon}{4} < \frac{1}{2^l}$; die Sphäre $S\left(x_{n_0}, \frac{1}{2^l}\right)$ enthält also Elemente aus U . Ist $u_n \in U$ dasjenige Element, das wir dieser Sphäre entnehmen (u_n braucht mit u nicht übereinzustimmen), so gilt wegen $u, u_n \in S\left(x_{n_0}, \frac{1}{2^l}\right)$

$$\|u - u_n\| < \frac{1}{2^{n-1}} \leq \varepsilon,$$

was zu beweisen war.

95. Lineare normierte Räume. Wir führen nun abstrakte Räume ein, die metrisch sind, jedoch noch andere Eigenschaften besitzen. Die Elemente eines solchen Raumes bezeichnen wir wie oben mit $x, y, z, \dots$ und Zahlen mit $a, b, c, \dots$. Diese Zahlen können reell oder komplex sein. Im ersten Fall haben wir dann einen reellen, im zweiten Fall einen komplexen Raum.

Wir werden stets komplexe Räume betrachten, falls nichts anderes ausdrücklich vermerkt wird.

Eine Menge X von Elementen $x, y, z, \dots$ heißt ein *linearer Raum*, wenn ihre Elemente den folgenden Axiomen genügen:

Axiom A. Die Elemente von X können mit Zahlen multipliziert und addiert werden, d. h., sind x, y Elemente aus X und ist a eine Zahl, so sind ax und $x + y$ ebenfalls wohlbestimmte Elemente aus X .

Diese Operationen sind den folgenden Gesetzen unterworfen:

1. $x + y = y + x$;
2. $x + (y + z) = (x + y) + z$;
3. $a(x + y) = ax + ay$;
4. $(a + b)x = ax + bx$;
5. $a(bx) = (ab)x$;
6. $1x = x$;
7. aus $x + y = x + z$ folgt $y = z$.

Wir führen nun den Begriff des Nullelements ein. Es seien x und y zwei beliebige Elemente aus X . Für sie gilt $0x = 0y$; denn setzen wir $0x = \theta$, $0y = \theta_1$, so erhalten wir mit Hilfe der Gesetze 4 und 6

$$x + \theta = 1x + 0x = (1 + 0)x = 1x = x$$

(und analog $y + \theta_1 = y$), ferner mit Hilfe der Gesetze 1 und 2

$$(x + y) + \theta = (x + \theta) + y = x + y$$

(und analog $(x + y) + \theta_1 = x + y$), so daß sich also $(x + y) + \theta = (x + y) + \theta_1$ und wegen Gesetz 7

$$\theta = \theta_1$$

ergibt. Infolgedessen entsteht bei Multiplikation eines Elements mit der Zahl 0 stets ein und dasselbe Element, das wir das *Nullelement* nennen und mit θ bezeichnen. Für θ lassen sich einfache Folgerungen aus den obigen Gesetzen nachprüfen. Das Produkt $a\theta$ ist für jedes komplexe a gleich θ . Ist $ax = \theta$ und $a \neq 0$, so ist $x = \theta$. Gilt $ax = bx$ und $x \neq \theta$, so ist $a = b$. Aus $ax = ay$ und $a \neq 0$ folgt $x = y$. Mit $-x$ bezeichnen wir das Produkt $(-1)x$. Die Differenz $x - y$ ist durch

$$x - y = x + (-y)$$

definiert. Auch für sie gelten die üblichen Regeln aus der Algebra.

Im folgenden werden wir oft das Nullelement einfach mit 0 bezeichnen. Das gibt keine Verwechslung mit der Zahl 0, wenn wir die folgenden Gleichungen aufmerksam betrachten; denn steht auf einer Seite einer Gleichung ein Element aus X und auf der anderen das Zeichen 0, so bedeutet 0 das Nullelement von X .

Definition. Die Elemente $x_1, x_2, \dots, x_m$ heißen *linear unabhängig*, wenn die Gleichung

$$c_1x_1 + c_2x_2 + \dots + c_mx_m = 0$$

nur erfüllt ist, sobald alle Zahlen c_k ($k = 1, 2, \dots, m$) gleich 0 sind.

Für den n -dimensionalen komplexen Raum, der schon in Teil III/1 untersucht wurde, ist die maximale Anzahl linear unabhängiger Elemente gleich n . Es wird manchmal auch ein Axiom eingeführt, das die Möglichkeit eines endlichdimensionalen Raumes ausschließt.

Axiom B. Für jedes ganze positive n existieren n linear unabhängige Elemente.

Im folgenden wird dieses Axiom keine wesentliche Rolle spielen.

Axiom C. Jedem Element x läßt sich eine bestimmte reelle nichtnegative Zahl $\|x\|$, die Norm von x , zuordnen; sie muß den drei folgenden Bedingungen genügen:

1. $\|\theta\| = 0$ und $\|x\| > 0$ für $x \neq \theta$;
2. $\|x + y\| \leq \|x\| + \|y\|$;
3. $\|ax\| = |a| \cdot \|x\|$.

Dabei ist a eine beliebige Zahl und $|a|$ der Absolutbetrag von a . Aus der zweiten und der dritten Eigenschaft folgt $\|-x\| = \|x\|$ und

$$\|x - y\| \geq \|x\| - \|y\|, \quad \|x - y\| \geq \|y\| - \|x\|,$$

d. h.

$$\|x - y\| \geq \left| \|x\| - \|y\| \right|. \quad (41)$$

Der Abstand zwischen zwei Elementen x, y wird durch $\varrho(x, y) = \|x - y\|$ definiert. Wir sehen sofort, daß $\varrho(x, y)$ den drei Forderungen in der Definition des metrischen Raumes [84] genügt, d. h., jeder lineare normierte Raum mit $\varrho(x, y) = \|x - y\|$ ist gleichzeitig ein metrischer Raum, so daß alles das über metrische Räume Gesagte gilt. Die Norm eines Elements x läßt sich vermöge der Formel $\|x\| = \varrho(x, \theta)$ durch den Abstand ausdrücken.

Wenn wir noch Vollständigkeit fordern, nennen wir den linearen normierten Raum einen *Banachraum* (oder einfach einen *Raum B*).¹⁾ Das Folgende wird sich auf Banachräume beziehen.

Ein nicht vollständiger linearer normierter Raum läßt sich vervollständigen [85]. Die Norm der neu hinzukommenden Elemente wird durch $\|x\| = \varrho(x, \theta)$ definiert. Bei der Vervollständigung bleiben alle Axiome erhalten, insbesondere das Axiom A. Dies folgt aus der Stetigkeit der Summe $x + y$ und des Produkts ax , wovon wir noch sprechen werden.

Im folgenden benutzen wir konvergente Folgen von Zahlen und Elementen. Wie oben werden wir für konvergente Folgen von Zahlen die Bezeichnung $a_n \rightarrow a_0$, für Folgen von Elementen die Bezeichnung $x_n \rightarrow x_0$ verwenden. Die Konvergenz $x_n \rightarrow x_0$ ist gleichwertig mit $\|x_0 - x_n\| \rightarrow 0$.

Wir können in einem Banachraum B unendliche Reihen $u_1 + u_2 + \dots$ mit $u_k \in B$ ($k = 1, 2, \dots$) betrachten. Hat eine Folge von Elementen $x_n = u_1 + u_2 + \dots + u_n$ aus einem Banachraum den Grenzwert x_0 , so sagt man, die Reihe konvergiert und hat die Summe x_0 .

Wir zeigen nun, daß die Ausdrücke $x + y$ und ax stetig sind, d. h., aus $x_n \rightarrow x_0$,

¹⁾ B von BANACH; Räume dieser Art wurden zuerst von dem polnischen Mathematiker S. BANACH systematisch untersucht (Ann. d. Red.).

$y_n \Rightarrow y_0$ und $a_n \rightarrow a_0$ folgt

$$x_n + y_n \Rightarrow x_0 + y_0 \quad \text{und} \quad a_n x_n \Rightarrow a_0 x_0.$$

Es gilt

$$|(x_0 + y_0) - (x_n + y_n)| \leq |x_0 - x_n| + |y_0 - y_n|;$$

hier strebt die rechte Seite und demzufolge auch die linke Seite gegen 0. Also ist $x_n + y_n \Rightarrow x_0 + y_0$. Für die Differenz $a_0 x_0 - a_n x_n$ schreiben wir $a_0 x_0 - a_n x_0 + a_n x_0 - a_n x_n$, und es gilt

$$|a_0 x_0 - a_n x_n| \leq |a_0 x_0 - a_n x_0| + |a_n x_0 - a_n x_n| = |a_0 - a_n| \cdot |x_0| + |a_n| \cdot |x_0 - x_n|;$$

hieraus ist ersichtlich, daß die rechte Seite gegen 0 strebt, denn aus $a_n \rightarrow a_0$ folgt die Beschränktheit von $|a_n|$. Ferner ergibt sich $|x_n| \rightarrow |x_0|$ aus $x_n \Rightarrow x_0$ (auf Grund der Formel $|x_n| = \rho(x_n, \theta)$ und der Stetigkeit des Abstands).

Wir definieren nun in Banachräumen eine lineare Mannigfaltigkeit: Eine Menge von Elementen aus U heißt *lineare Mannigfaltigkeit*, wenn mit $x_p \in U$ ($k = 1, 2, \dots, m$) auch jede Linearkombination $c_1 x_1 + c_2 x_2 + \dots + c_m x_m$ dieser Elemente zu U gehört. Wir brauchen uns nur davon zu überzeugen, daß mit $x, y \in U$ auch $x + y \in U$ und $ax \in U$ für jedes a gilt. Setzen wir $a = 0$, so sehen wir, daß das Nullelement in jeder nichtleeren linearen Mannigfaltigkeit enthalten ist. Eine abgeschlossene lineare Mannigfaltigkeit wollen wir *Unterraum* nennen. Ist U eine nicht abgeschlossene lineare Mannigfaltigkeit, so ist die abgeschlossene Menge $\bar{U}$ ein Unterraum, d. h., die Abschließung einer linearen Mannigfaltigkeit führt auf einen Unterraum. Dies folgt aus der oben bewiesenen Stetigkeit von $x + y$ und ax . Ist U keine lineare Mannigfaltigkeit, so erhalten wir, wenn wir alle möglichen Linearkombinationen $c_1 x_1 + c_2 x_2 + \dots + c_m x_m$ von Elementen $x_k \in U$ bilden, eine neue Menge V von Elementen aus U , die bereits eine lineare Mannigfaltigkeit ist. Sie wird im allgemeinen die *lineare Hülle* von U genannt; sie ist die kleinste, U enthaltende lineare Mannigfaltigkeit.

Sind $x_1, x_2, \dots, x_k$ linear unabhängige Elemente, so ist die Menge U von Elementen aus einem Banachraum, die in der Form $x = c_1 x_1 + c_2 x_2 + \dots + c_k x_k$ mit beliebigen Zahlen c_s darstellbar sind, offenbar eine lineare Mannigfaltigkeit. Wir können leicht zeigen, daß sie eine abgeschlossene Menge (ein Unterraum) ist. Auf Grund der linearen Unabhängigkeit der x_s ist die obige Darstellung von x eindeutig; eine solche lineare Mannigfaltigkeit heißt im allgemeinen *endlichdimensional*. Alle Elemente aus U lassen sich in der Gestalt $x = c_1 y_1 + c_2 y_2 + \dots + c_k y_k$ ausdrücken, wobei die y_s ($s = 1, 2, \dots, k$) beliebige linear unabhängige Elemente aus U und die c_s willkürlich gewählte Zahlen sind, und jede Formel, die alle Elemente von U in der obigen Form darstellt, weist k Summanden auf. Die Zahl k gibt die *Dimension* von U an.

Wir erwähnen noch, daß jeder Unterraum eines Banachraumes ebenfalls ein Banachraum ist.

Schon früher definierten wir den Begriff der Isometrie für metrische Räume. Nun wollen wir die Isometrie für Banachräume definieren. Zwei solche Räume X und X' heißen *isometrisch*, wenn zwischen ihren Elementen eine eindeutige Zuordnung hergestellt werden kann, für welche die beiden folgenden Bedingungen erfüllt sind: 1. Sind x und x' sowie y und y' zwei beliebige Paare von einander entsprechenden Elementen aus X bzw. X' , so sind auch die Elemente $ax + by$ und $ax' + by'$ für beliebige Zahlen a und b einander zugeordnet; 2. einander entsprechende Elemente haben die gleiche Norm.

Daraus folgt, daß die Nullelemente von X und X' einander zugeordnet und die Abstände zwischen entsprechenden Elementen in X und in X' gleich sein müssen.

Vom Standpunkt der abstrakten Theorie werden isometrische Räume nicht voneinander unterschieden; wir können also $X = X'$ schreiben.

96. Beispiele für normierte Räume. 1. Die in [87] erwähnten Räume sind außer s und S Banachräume, wenn für sie $\|x\| = \varrho(0, x)$ gesetzt wird. Dann führt in Folgenräumen die Multiplikation eines Elementes mit einer Zahl a nach Definition auf die Multiplikation jeder Zahl einer Folge mit a , und die Addition von Elementen auf die Addition von Zahlen dieser Folgen mit gleichem Index:

$$a(\xi_1, \xi_2, \dots) = (a\xi_1, a\xi_2, \dots), \quad (\xi_1, \xi_2, \dots) + (\eta_1, \eta_2, \dots) = (\xi_1 + \eta_1, \xi_2 + \eta_2, \dots).$$

In Funktionenräumen führt die Multiplikation eines Elements mit einer Zahl a nach Definition auf die Multiplikation einer Funktion mit a und die Addition von Elementen auf die Addition der entsprechenden Funktionen. Das Nullelement ist im Folgenraum eine Folge, die nur aus Nullen besteht, und im Funktionenraum diejenigen Funktionen, die entweder identisch 0 (in C und V) oder äquivalent 0 (in M , S und L_p) ist.

2. Wir untersuchen im R_n ein beschränktes Gebiet D und eine Menge $C^{(l)}$ von Funktionen $\varphi(x)$, die innerhalb von D stetige partielle Ableitungen bis einschließlich l -ter Ordnung haben, wobei diese Ableitungen Grenzwerte auf der Berandung von D besitzen und Funktionen darstellen, die in dem abgeschlossenen Bereich $\bar{D}$ stetig sind. In diesem Fall sagen wir kurz, die Funktionen haben in $\bar{D}$ stetige Ableitungen. Diese Menge von Funktionen ist ein linearer Raum. In ihm führen wir die Norm

$$\|\varphi\|_{C^{(l)}} = \max_{\substack{x \in \bar{D} \\ 0 \leq k \leq l}} |D^k \varphi| \quad (42)$$

ein, wobei $D^k \varphi$ eine beliebige Ableitung k -ter Ordnung ist. Das Maximum wird über alle $x \in \bar{D}$ von der Funktion $\varphi(x)$ und allen ihren Ableitungen bis einschließlich l -ter Ordnung genommen. Es ist leicht zu sehen, daß die Norm den drei Grundforderungen genügt [95].

Die Konvergenz in $C^{(l)}$ ist eine gleichmäßige Konvergenz in $\bar{D}$ der Funktionen und aller ihrer Ableitungen bis einschließlich l -ter Ordnung. Auf Grund des Cauchyschen Konvergenzkriteriums und des bekannten Satzes über die gliedweise Differentiation von Funktionenfolgen können wir behaupten: Konvergiert eine Folge von Elementen $\varphi_n(x) \in C^{(l)}$ in sich, so konvergiert sie gegen ein Element $\varphi(x) \in C^{(l)}$, d. h., der Raum $C^{(l)}$ ist ein Banachraum.

97. Operatoren in normierten Räumen. Früher definierten wir Operatoren in einem metrischen Raum X . In linearen normierten Räumen gibt es neue Gesichtspunkte zu berücksichtigen. Wir setzen voraus, ein Operator A sei auf einer linearen Mannigfaltigkeit $D(A)$ eines metrischen Banachraumes X gegeben und sein Wertebereich $R(A)$ gehöre zu einem Banachraum X' . Der Operator heißt *distributiv*, wenn für x_k ($k = 1, 2, \dots, m$) aus $D(A)$ und beliebige Zahlen c_k die Bedingung

$$A(c_1 x_1 + c_2 x_2 + \dots + c_m x_m) = c_1 A x_1 + c_2 A x_2 + \dots + c_m A x_m \quad (43)$$

erfüllt ist. Es genügt, die Gültigkeit der Beziehungen $A(c_1 x_1) = c_1 A x$ und $A(x + y) = A x + A y$ nachzuweisen. Aus (43) folgt, daß $R(A)$ eine lineare Mannigfaltigkeit in X' ist und $A\theta = \theta'$ gilt, wenn θ bzw. θ' das Nullelement von X bzw. X' ist. Es ist nämlich $A\theta = A(0x)$, $x \in D(A)$, und $A(0x) = 0Ax = \theta'$. Im folgenden werden wir nur von distributiven Operatoren auf linearen Mannigfaltigkeiten sprechen.

Wir erinnern noch an die Definition der Stetigkeit. Ein Operator A heißt *stetig auf einem Element* x_0 , wenn die folgende Bedingung erfüllt ist: Gehören die Elemente x_n ($n = 1, 2, \dots$) zu $D(A)$ und ist $x_n \rightarrow x_0$ in X , so ist $Ax_n \rightarrow Ax_0$ in X' . Ist ein distributiver Operator A auf einem Element $y_0 \in D(A)$ stetig, so ist er auch auf jedem Element $z_0 \in D(A)$ stetig. Um also nachzuweisen, daß $Az \rightarrow Az_0$ gilt, nehmen wir $z_n, z_0 \in D(A)$ und $z_n \rightarrow z_0$ an und konstruieren Elemente $y_n = (z_n - z_0) + y_0$ aus $D(A)$,

für welche $y_n \rightarrow y_0$ gilt. Dann ist $Az_n = Az_0 + Ay_n - Ay_0$, und wegen $Ay_n \rightarrow Ay_0$ gilt $Az_n \rightarrow Az_0$. Es ist also nicht zweckmäßig, von Stetigkeit auf einem Element von $D(A)$ zu sprechen, sondern nur von der *Stetigkeit auf ganz $D(A)$* .

Ein distributiver Operator A heißt *beschränkt*, wenn für eine positive Zahl c und jedes $x \in D(A)$

$$\|Ax\| \leq c\|x\| \quad (44)$$

gilt. Hier bezieht sich die Norm rechts auf X , links auf X' .

Wir zeigen nun, daß für einen distributiven Operator Beschränktheit und Stetigkeit auf $D(A)$ gleichwertig sind. Auf Grund des oben Gesagten genügt es, die Stetigkeit auf dem Nullelement θ zu untersuchen. Es gelte (44). Wir müssen also zeigen, daß $Ax_n \rightarrow \theta'$ aus $x_n \in D(A)$ und $x_n \rightarrow \theta$ folgt. Ist $x_n \rightarrow \theta$, so existiert für jedes gegebene $\varepsilon > 0$ ein solches N , daß $\|x_n\| \leq \varepsilon$ für $n \geq N$ und wegen (44) dann $\|Ax_n\| \leq c\varepsilon$ für $n \geq N$ gilt. Also ergibt sich $Ax_n \rightarrow \theta'$, da ε beliebig ist. Wir nehmen nun an, es sei $Ax_n \rightarrow \theta'$ für $x_n \rightarrow \theta$, und zeigen, daß (44) gilt. Ist $x = \theta$, so geht (44) in die Ungleichung $\|\theta'\| \leq c\|\theta\|$ über, d. h., es ist $0 \leq c\|\theta\|$; dies gilt für jedes c . Es genügt also, (44) für $x \neq \theta$ zu beweisen. Den Beweis führen wir indirekt. Gilt (44) nicht, so gibt es eine Folge von Elementen $x_n \in D(A)$ mit $\|x_n\| > 0$, so daß $\|Ax_n\| > c_n\|x_n\|$ mit $c_n \rightarrow \infty$ für $n \rightarrow \infty$ gilt.

Führen wir Elemente $z_n = \frac{1}{c_n\|x_n\|} x_n \in D(A)$ ein, für welche $\|z_n\| \rightarrow 0$ gilt, so erhalten wir $\|Az_n\| = 1$. Dies widerspricht aber der Stetigkeit von Ax auf dem Nullelement.

Ist A der Nulloperator, d. h., gilt $Ax = \theta'$ für jedes $x \in D(A)$, so können wir in (44) $c = 0$ setzen. Für jeden anderen Operator ist notwendig $c > 0$, und es existiert ein kleinstes positives c , für welches (44) erfüllt ist. Es heißt die *Norm* des Operators A und ergibt sich vermöge der Formel

$$n_A = \sup_{\substack{\|x\|=1 \\ x \in D(A)}} \|Ax\|. \quad (45)$$

Die Norm von A wird auch mit $\|A\|$ bezeichnet, so daß wir

$$\|Ax\| \leq n_A\|x\| \quad \text{oder} \quad \|Ax\| \leq \|A\| \cdot \|x\| \quad (46)$$

schreiben können.

Diese Überlegungen sind denen aus Teil IV/1 für den dort behandelten Spezialfall vollkommen analog.

Satz 1. *Ist für einen distributiven beschränkten Operator A die lineare Mannigfaltigkeit $D(A)$ in X dicht, so läßt sich A auf ganz X unter Beibehaltung der Distributivität und der Norm erweitern.*

Beweis. Wegen $\overline{D(A)} = X$ können wir jedes Element $x_0 \in X$ als Grenzwert darstellen: $x_n \rightarrow x_0$ mit $x_n \in D(A)$. Wir wollen zeigen, daß in X' ein Grenzwert von Ax_n existiert, der nicht von der Wahl der Folge x_n abhängt. Es ist

$$\|Ax_n - Ax_m\| = \|A(x_n - x_m)\| \leq \|A\| \cdot \|x_n - x_m\|;$$

wegen $x_n \rightarrow x_0$ strebt die rechte Seite für $n, m \rightarrow \infty$ gegen 0, also ist auch $\|Ax_n - Ax_m\| \rightarrow 0$, und da X' vollständig ist, hat Ax_n einen Grenzwert. Wir müssen noch zeigen, daß der Grenzwert von Ax_n von der Wahl der Folge x_n unabhängig ist. Es sei $x_n, x'_n \in D(A)$ und $x_n \rightarrow x_0, x'_n \rightarrow x_0$; dann muß die Übereinstimmung der Grenzwerte von Ax_n und Ax'_n nachgewiesen werden. Die Existenz der Grenzwerte folgt aus dem Vorhergehenden.

Wir erkennen leicht, daß die Folge

$$x_1, x'_1, x_2, x'_2, x_3, x'_3, \dots$$

ebenfalls den Grenzwert x_0 hat. Dann hat die Folge

$$Ax_1, Ax'_1, Ax_2, Ax'_2, Ax_3, Ax'_3, \dots$$

einen Grenzwert $y \in X'$, so daß die Folgen Ax_n und Ax'_n den gleichen Grenzwert, nämlich y , haben.

Ist $x_0 \in X$, aber $x_0 \notin D(A)$, so wählen wir eine beliebige Folge $x_n \in D(A)$ mit $x_n \rightarrow x_0$ und setzen $Ax_0 = \lim_{n \rightarrow \infty} Ax_n$. Der so in X definierte Operator ist distributiv, und seine Norm wird beim Übergang von $D(A)$ zu X nicht größer. Sind x'_n und x''_n zwei Folgen aus $D(A)$ mit den Grenzwerten x'_0 bzw. x''_0 aus X (ist etwa $x'_0 \in D(A)$, so können wir alle x'_n gleich x'_0 setzen), so erhalten wir, da A in $D(A)$ distributiv ist und Addition und Multiplikation mit einer Zahl stetig sind,

$$\begin{aligned} A(c_1 x'_0 + c_2 x''_0) &= \lim_{n \rightarrow \infty} A(c_1 x'_n + c_2 x''_n) \\ &= c_1 \lim_{n \rightarrow \infty} Ax'_n + c_2 \lim_{n \rightarrow \infty} Ax''_n = c_1 Ax'_0 + c_2 Ax''_0. \end{aligned}$$

Die Invarianz der Norm folgt aus der Ungleichung $\|Ax'_n\| \leq \|A\| \cdot \|x'_n\|$, in welcher $\|A\|$ die Norm von A in $D(A)$ nach dem Grenzübergang ist. Die Norm kann offenbar nicht kleiner werden. Damit ist der Satz bewiesen.

Die eben gezeigte Methode der Erweiterung von A heißt im allgemeinen *Erweiterung bezüglich der Stetigkeit*.

Die Erweiterung des Operators A von $D(A)$ auf den Raum X ist in gewissem Sinne eindeutig: Ist B ein in X distributiver beschränkter Operator, der auf $D(A)$ mit A übereinstimmt, so stimmt B überall mit der Erweiterung von A bezüglich der Stetigkeit überein; denn ist $x_0 \notin D(A)$, $x_n \in D(A)$ und $x_n \rightarrow x_0$, so gilt wegen der Stetigkeit von B und der Übereinstimmung von B und A auf $D(A)$

$$Bx_0 = \lim_{n \rightarrow \infty} Bx_n = \lim_{n \rightarrow \infty} Ax_n = Ax_0.$$

Ist der Operator A auf dem ganzen Raum X definiert, distributiv und beschränkt, so nennen wir ihn *linear* (wir sprechen in diesem Fall auch manchmal von einem beschränkten linearen Operator). Entsprechen verschiedenen x aus der linearen Mannigfaltigkeit verschiedene $Ax \in R(A)$, so existiert ein *inverser Operator*, der auf $R(A)$ definiert ist und jedem $x' \in R(A)$ ein einziges Element $x \in D(A)$ zuordnet, das mit x' durch die Beziehung $x' = Ax$ verknüpft ist. Aus der Distributivität von A folgt unmittelbar die von A^{-1} ; jedoch braucht A^{-1} nicht beschränkt zu sein, wenn A beschränkt ist.

Als Beispiel betrachten wir im Raum $C[0, 1]$ den Operator $\varphi = A f$,

$$\varphi(x) = \int_0^x f(t) dt; \quad (47)$$

dabei nehmen wir an, X' sei der gleiche Raum C . Dies ist möglich wegen $\varphi(x) \in C[0, 1]$. Der Operator (47) führt ganz C in eine lineare Mannigfaltigkeit über, die aus denjenigen Funktionen $\varphi(x)$ besteht, welche stetige Ableitungen besitzen und für $x=0$ verschwinden. Auf dieser linearen Mannigfaltigkeit existiert ein inverser distributiver Operator $f(x) = \varphi'(x)$, aber er ist nicht beschränkt; die Funktionen $\varphi_n(x) = \sin n\pi x$, die zu dieser linearen Mannigfaltigkeit gehören, haben nämlich für jedes n die Norm 1, und die Funktionen $\varphi'_n(x) = n\pi \cos n\pi x$ haben die Norm $n\pi$, die für $n \rightarrow \infty$ unbegrenzt wächst.

Satz 2. *Ist B ein linearer Operator in X , $R(B) \subset X$ und $\|B\| = \alpha < 1$, so hat der Operator $E - B$, wobei E der Einheitsoperator (also $Ex = x$) ist, ein auf ganz X definiertes, distributives und beschränktes Inverses $(E - B)^{-1}$.*

Beweis. Wir betrachten die Gleichung

$$y = x - Bx \quad (48)$$

oder

$$x = y + Bx, \quad (49)$$

wobei y ein gegebenes und x das gesuchte Element aus X ist. Der Operator $Ax = y + Bx$ (A ist die Bezeichnung aus [86]) genügt der Bedingung für die Anwendbarkeit des Prinzips der kontrahierenden Abbildungen, denn es ist

$$\|Ax_2 - Ax_1\| = \|B(x_2 - x_1)\| \leq \alpha \|x_2 - x_1\| \quad (0 < \alpha < 1).$$

Somit hat (48) oder (49) für jedes $y \in X$ eine einzige Lösung, d. h., es existiert ein auf ganz X definierter inverser Operator mit $x = (E - B)^{-1}y$. Er ist offenbar distributiv, aber auch beschränkt; es ist nämlich

$$\|x - y\| = \|Bx\| \leq \alpha \|x\|,$$

also erst recht

$$\|x\| - \|y\| \leq \alpha \|x\|, \quad \text{also} \quad \|x\| < \frac{1}{1 - \alpha} \|y\|;$$

d. h., die Norm von $(E - B)^{-1}$ ist nicht größer als $\frac{1}{1 - \alpha}$.

Satz 3. *Ein linearer Operator bildet eine kompakte Menge wieder in eine kompakte Menge ab.*

Beweis. Es sei U eine kompakte Menge von Elementen aus X , x_n ($n = 1, 2, \dots$) eine beliebige Folge von Elementen aus U und A ein linearer Operator. Wir müssen zeigen, daß wir aus der Folge Ax_n ($n = 1, 2, \dots$) eine in X' konvergente Teilfolge auswählen können. Aus der Kompaktheit von U folgt die Existenz einer Teilfolge x_{n_k} , die in X den Grenzwert x_0 hat. Dann ist auf Grund der Stetigkeit von A

$$Ax_{n_k} \Rightarrow Ax_0 \quad \text{in} \quad X',$$

was zu beweisen war.

Wir geben ohne Beweis noch den folgenden Satz an:

Satz 4. *Ist im Banachraum X ein linearer Operator definiert, der den Raum X eineindeutig in den ganzen Banachraum X' abbildet, so ist der (auf ganz X' definierte) inverse Operator A^{-1} ebenfalls linear.*

98. Lineare Funktionale. Wir betrachten einen reellen Banachraum X (die Elemente aus X werden nur mit reellen Zahlen multipliziert). Ein Funktional in X ist ein Operator, dessen Wertebereich im Raum der reellen Zahlen liegt. Dieser Raum ist ein reeller Banachraum bei der gewöhnlichen Definition der Addition reeller Zahlen und ihrer Multiplikation mit einer reellen Zahl. Die Norm ist hier der Absolutbetrag der reellen Zahl [87].

Die Überlegungen aus [97] für Operatoren gelten auch für Funktionale. Die Beschränktheit eines Funktional wird durch die Ungleichung

$$|l(x)| \leq \|l\| \cdot \|x\| \quad (50)$$

definiert, wobei $|l(x)|$ der Absolutbetrag der reellen Zahl $l(x)$ und $||l||$ die Norm von $l(x)$ ist. Das lineare Funktional ist ein Spezialfall des linearen Operators. Dabei stimmt $D(l)$ mit ganz X überein.

Satz 1 (HAHN-BANACH). *Ist auf einer linearen Mannigfaltigkeit U ein distributives beschränktes Funktional $l(x)$ gegeben, so kann es auf den ganzen Raum X derart erweitert werden, daß $l(x)$ in X ein lineares Funktional mit der gleichen Norm wie in U ist.*

Beweis. Nach Voraussetzung gilt außer der Distributivität noch

$$|l(x)| \leq ||l||_U ||x|| \quad (x \in U), \quad (51)$$

wobei $||l||_U$ die Norm von $l(x)$ in U ist. Zum Beweis nehmen wir an, daß X separabel ist, was die Überlegungen vereinfacht. Jedoch gilt der Satz auch für nicht separable Räume.

Auf Grund der Separabilität existiert eine abzählbare Menge von Elementen, die in X dicht ist. Wir erhalten in dieser Menge nur Elemente, die nicht zu U gehören. Gibt es solche Elemente nicht, so ist U in X überall dicht, und $l(x)$ läßt sich stetig auf den ganzen Raum X fortsetzen [97]. Anderenfalls können die übrigen Elemente der abzählbaren Menge numeriert werden: $x_1, x_2, \dots$

Wir betrachten die Menge U_1 von Elementen z der Form $z = y + tx_1$, wobei y ein beliebiges Element von U und t eine beliebige reelle Zahl ist. Wir sehen leicht, daß U_1 genauso wie U eine lineare Mannigfaltigkeit ist. Die Darstellung von z ist eindeutig. Hätte nämlich z zwei verschiedene Darstellungen

$$z = y + tx_1 = y' + t'x_1, \quad (52)$$

so wäre $t \neq t'$ (weil aus $t = t'$ die Beziehung $y = y'$ folgt); das kann aber nicht sein, denn mit $t \neq t'$ ergibt sich aus (52)

$$x_1 = \frac{1}{t-t'} (y' - y),$$

also $x_1 \in U$, was dem oben Gesagten widerspricht. Wir werden nun die für das Folgende wesentliche Ungleichung (54) herleiten. Für zwei beliebige Elemente x' und x'' aus U gilt

$$l(x') - l(x'') = l(x' - x'') \leq ||l||_U ||x' - x''||. \quad (53)$$

Steht links eine negative Zahl, so ist (53) trivial. Unter Berücksichtigung von

$$||x' - x''|| = ||(x' + x_1) - (x'' + x_1)|| \leq ||x' + x_1|| + ||x'' + x_1||$$

folgt wegen (53)

$$l(x') - ||l||_U ||x' + x_1|| \leq l(x'') + ||l||_U ||x'' + x_1||.$$

Nehmen wir die obere Grenze der links stehenden Zahlenmenge und die untere Grenze für die rechte Seite, so erhalten wir, wenn x' und x'' unabhängig voneinander ganz U durchlaufen,

$$\sup_{x' \in U} (l(x') - ||l||_U ||x' + x_1||) \leq \inf_{x'' \in U} (l(x'') + ||l||_U ||x'' + x_1||);$$

dabei ist offenbar die rechte Seite und infolgedessen auch die linke Seite endlich. Somit existiert eine reelle Zahl a , die der Ungleichung

$$\sup_{x' \in U} (l(x') - ||l||_U ||x' + x_1||) \leq a \leq \inf_{x'' \in U} (l(x'') + ||l||_U ||x'' + x_1||) \quad (54)$$

genügt.

Wir erweitern nun $l(x)$ von U auf U_1 . Es sei $z = y + tx_1$ ein beliebiges Element aus U_1 , und wir setzen

$$l(z) = l(y) - ta, \quad (55)$$

wobei a eine feste reelle Zahl ist, die der Ungleichung (54) genügt. Ist $z \in U$, so gilt $t = 0$, und $l(z)$ stimmt also mit $l(y)$ überein, d. h., die Beziehung (55) definiert auf U_1 ein Funktional, das auf U_1 mit dem vorhergehenden übereinstimmt. Daher benutzen wir für das erweiterte Funktional die gleiche Bezeichnung. Die Distributivität von $l(z)$ folgt unmittelbar aus (55), aus der Distributivität von $l(y)$ auf U und aus den Formeln

$$\begin{aligned} z &= y + tx_1, & cz &= cy + ct x_1, \\ z' &= y' + t' x_1, & z'' &= y'' + t'' x_1, \\ z' + z'' &= (y' + y'') + (t' + t'') x_1. \end{aligned}$$

Wir zeigen schließlich, daß die Norm von $l(z)$ in U_1 nicht größer als $\|l\|_U$ ist (kleiner kann sie nicht werden). Zunächst sei $t > 0$. Wegen $l(y) = t l\left(\frac{1}{t} y\right)$ mit $\frac{1}{t} y \in U$ erhalten wir

$$l(z) = t \left[l\left(\frac{1}{t} y\right) - a \right]. \quad (56)$$

Nun folgt aus (54)

$$a \leq l\left(\frac{1}{t} y\right) - \|l\|_U \left\| \frac{1}{t} y + x_1 \right\|,$$

und ersetzen wir a in (56) durch eine kleinere Zahl, so ergibt sich

$$l(z) \leq t \|l\|_U \left\| \frac{1}{t} y + x_1 \right\| = \|l\|_U \|y - tx_1\| = \|l\|_U \|z\| \quad (t > 0).$$

Wir gehen nun zum Fall $t < 0$ über. Aus (54) folgt

$$a \leq l\left(\frac{1}{t} y\right) + \|l\|_U \left\| \frac{1}{t} y + x_1 \right\|,$$

und ersetzen wir a in (56) durch eine größere Zahl, so finden wir ($|t| = -t$)

$$l\left(\frac{1}{t} y\right) - a \leq -\|l\|_U \left\| \frac{1}{t} y + x_1 \right\| = -\frac{1}{|t|} \|l\|_U \|y + tx_1\| = \frac{1}{t} \|l\|_U \|z\|.$$

Multiplizieren wir beide Seiten dieser Ungleichung mit negativem t , so ergibt sich mit Hilfe von (56)

$$l(z) \leq \|l\|_U \|z\|.$$

Ist $t = 0$, so gehört z zu U , und die Ungleichung ist klar. Somit gilt für jedes $z \in U_1$ die Relation $l(z) \leq \|l\|_U \|z\|$. Ersetzen wir in dieser Ungleichung z durch $-z$, so erhalten wir wegen $l(-z) = -l(z)$ und $\| -z \| = \|z\|$

$$-l(z) \leq \|l\|_U \|z\|.$$

Die letzten beiden Ungleichungen führen schließlich auf die Ungleichung

$$|l(z)| \leq \|l\|_U \|z\|, \quad (57)$$

aus der folgt, daß bei der erwähnten Erweiterung des Funktionals $l(x)$ von U auf U_1 die Norm invariant bleibt.

Wir kommen jetzt zur weiteren Fortsetzung eines Funktional. Gehört das Element x_2 aus der obengenannten Folge zu U_1 , so schließen wir es aus. Gehört x_2 nicht zu U_1 , so erweitern wir $l(x)$ wie oben von der linearen Mannigfaltigkeit U_1 auf eine lineare Mannigfaltigkeit U_2 , die aus den Elementen $z = y + tx_2$ besteht, wobei y ein beliebiges Element aus U_1 und t eine beliebige reelle Zahl ist. Wir setzen dies weiter fort und behalten die vorige Bezeichnung $x_1, x_2, \dots$ für die Elemente, die bei der oben beschriebenen Konstruktion nicht entfernt wurden, bei (es können auch endlich viele Elemente sein). Auf diese Weise setzen wir $l(x)$ auf eine lineare Mannigfaltigkeit V fort, deren Elemente die Gestalt

$$y = c_1 x_1 + c_2 x_2 + \dots + c_n x_n$$

haben, wobei y ein beliebiges Element aus U und n eine ganze positive Zahl ist (die nicht größer als die Anzahl der Elemente x_k ist, wenn diese Anzahl endlich ist) und ferner die c_k beliebige reelle Zahlen sind. Diese lineare Mannigfaltigkeit V ist in unserem Banachraum dicht, und auf ihr ist das Funktional $l(x)$ distributiv und beschränkt, und es hat die Norm $\|l\|_V$. Man braucht also jetzt nur noch $l(x)$ in den ganzen Raum X bezüglich der Stetigkeit fortzusetzen. Damit ist der Satz bewiesen.

Für den Fall eines komplexen Banachraumes ist der Beweis von Satz 1 z. B. bei SUCHOMLINOW¹⁾ sowie bei RIESZ und SZ.-NAGY²⁾ zu finden.

Der bewiesene Satz läßt sich nicht auf Operatoren übertragen.

Satz 2. Ist $x_0 \neq 0$ ein beliebiges, aber festes Element aus X , so existiert ein lineares Funktional $l(x)$ mit der Norm 1 ($\|l\| = 1$), für welches $l(x_0) = \|x_0\|$ gilt.

Beweis. Wir betrachten die lineare Mannigfaltigkeit U von Elementen der Gestalt $x = tx_0$ (t beliebig reell) und definieren auf U ein distributives Funktional $l(x)$ mit Hilfe der Formel $l(tx_0) = t\|x_0\|$. Für $t=1$ gilt $l(x_0) = \|x_0\|$ und $\|l\|_U = 1$. Auf Grund von Satz 1 können wir $l(x)$ auf den ganzen Raum X unter Beibehaltung der Norm erweitern. Damit ist der Satz bewiesen. Aus ihm folgt unter anderem, daß in jedem Banachraum Funktionale mit positiver Norm existieren.

Der Satz 2 gilt auch dann, wenn $x_0 = \theta$ das Nullelement ist. In diesem Fall braucht man nur ein von θ verschiedenes Element x_1 zu nehmen und ein lineares Funktional $l(x)$ mit $\|l\| = 1$ und $l(x_1) = \|x_1\|$ zu bilden. Für dieses Funktional gilt $l(x_0) = l(\theta) = 0 = \|x_0\|$.

99. Konjugierte Räume. Wir betrachten den Raum X^* , dessen Elemente alle möglichen linearen Funktionale im Banachraum X seien. Die Funktionale in X , d. h., die den Elementen x zugeordneten reellen Zahlen, bezeichnen wir mit $l(x)$, $m(x)$, $n(x)$, ... Als Elemente von X^* bezeichnen wir sie nur mit $l, m, n, \dots$. Der Raum X^* ist linear; die Addition und die Multiplikation von Funktionalen mit einer Zahl wird nämlich durch

$$(l + m)(x) = l(x) + m(x), \quad (al)(x) = al(x)$$

definiert, wobei alle Eigenschaften aus Axiom A erfüllt sind.

Das Nullelement von X^* ist das Nullfunktional, für welches also $l(x) = 0$, $x \in X$, gilt. Die Norm eines Elements aus X^* ist gleich der Norm des entsprechenden Funk-

¹⁾ G. A. SUCHOMLINOW, Über Fortsetzung von linearen Funktionalen in linearen komplexen Räumen und linearen Quaternionenräumen (russ.), Mat. sbornik (2) **8**, 353–358 (1938).

²⁾ F. RIESZ und B. SZ.-NAGY, Vorlesungen über Funktionalanalysis (vgl. das Literaturverzeichnis).

tionals. Diese Norm ist nichtnegativ; gleich 0 ist sie nur für das Nullfunktional. Auch die beiden anderen Eigenschaften aus Axiom C gelten. Die zweite Eigenschaft folgt aus den Ungleichungen

$$|l_1(x) + l_2(x)| \leq |l_1(x)| + |l_2(x)| \leq \|l_1\| \cdot \|x\| + \|l_2\| \cdot \|x\| = (\|l_1\| + \|l_2\|) \|x\|,$$

die dritte ist klar.

Wir zeigen nun, daß X^* ein vollständiger Raum ist. Es existiere eine in sich konvergente Folge von Elementen aus X^* :

$$\|l_n - l_m\| \rightarrow 0 \quad \text{für } n, m \rightarrow \infty. \quad (58)$$

Wir müssen beweisen, daß es ein Element $l \in X^*$ gibt derart, daß $\|l - l_n\| \rightarrow 0$ für $n \rightarrow \infty$ gilt. Wir setzen $l_n - l_m = l_{nm} \in X^*$, und es ist $l_n = l_m + l_{nm}$ sowie $\|l_n\| \leq \|l_m\| + \|l_{nm}\|$. Wegen (58) gibt es ein solches N , daß $\|l_{nm}\| \leq 1$ für $n, m \geq N$ ist. Wählen wir ein festes $m = m_0 \geq N$, so erhalten wir $\|l_n\| \leq \|l_{m_0}\| + 1$ für $n \geq N$. Unter den endlich vielen nichtnegativen Zahlen $\|l_n\|$ ($n = 1, 2, \dots, N-1$) gibt es eine größte; somit folgt aus (58), daß ein $c > 0$ mit $\|l_n\| \leq c$ für alle n existiert. Ferner ist

$$|l_n(x) - l_m(x)| \leq \|l_n - l_m\| \cdot \|x\|,$$

und aus (58) folgt

$$|l_n(x) - l_m(x)| \rightarrow 0 \quad \text{für } n, m \rightarrow \infty \quad \text{und jedes } x \in X.$$

Auf Grund des Cauchyschen Konvergenzkriteriums (für Zahlen) hat die Folge der Zahlen $l_n(x)$ einen Grenzwert. Dieser Grenzwert, den wir mit $l(x)$ bezeichnen wollen ($l_n(x) \rightarrow l(x)$), ist ein in ganz X definiertes Funktional. Es ist linear; seine Distributivität folgt nämlich aus den Gleichungen $l_n(x+y) = l_n(x) + l_n(y)$ und $l_n(ax) = al_n(x)$, seine Beschränktheit aus der Ungleichung $\|l_n\| \leq c$, d. h. aus $|l_n(x)| \leq c\|x\|$ durch Grenzübergang für $n \rightarrow \infty$. Damit ist $l(x)$ ein lineares Funktional in X , also $l \in X^*$. Wir müssen nun noch $\|l - l_n\| \rightarrow 0$ nachweisen. Zu jedem $\varepsilon > 0$ gibt es wegen (58) ein solches N , daß $|l_n(x) - l_m(x)| \leq \varepsilon\|x\|$ für $n, m \geq N$ ist. Lassen wir m unbegrenzt wachsen, so erhalten wir $|l_n(x) - l(x)| \leq \varepsilon\|x\|$ für $n \geq N$, also $\|l - l_n\| \leq \varepsilon$ für $n \geq N$, woraus sich $\|l - l_n\| \rightarrow 0$ ergibt. Damit ist die Vollständigkeit von X^* bewiesen. Der Raum X^* ist also ein Banachraum; er heißt der zu X konjugierte Raum.

Wir führen noch den *bikonjugierten Raum* $X^{**} = (X^*)^*$ ein, dessen Elemente alle möglichen linearen Funktionale auf X^* sind. Der Raum X^{**} ergibt sich aus X^* genauso wie X^* aus X , und auch X^{**} ist ein Banachraum.

Wählen wir ein festes $x \in X$, so entspricht jedem Element $l \in X^*$ eine bestimmte reelle Zahl $l(x)$, d. h., $l(x)$ ist für ein festes x ein Funktional in X^* , wenn l den Raum X^* durchläuft. Dieses Funktional wollen wir mit $L_x(l)$ bezeichnen. Aus

$$(l_1 + l_2)(x) = l_1(x) + l_2(x) \quad \text{und} \quad (al)(x) = al(x)$$

folgt

$$L_x(l_1 + l_2) = L_x(l_1) + L_x(l_2) \quad \text{und} \quad L_x(al) = aL_x(l);$$

$L_x(l)$ ist also ein distributives Funktional in X^* . Aus der Ungleichung

$$|l(x)| \leq \|l\| \cdot \|x\| \quad (59)$$

folgt wegen $L_x(l) = l(x)$

$$|L_x(l)| \leq \|x\| \cdot \|l\|, \quad (60)$$

wobei $\|x\|$ die Norm von x in X und $\|l\|$ die Norm von l in X^* ist. Die Ungleichung (60) besagt, daß die Norm des Funktionals $L_x(l)$ in X^* nicht größer als $\|x\|$ ist, d. h., $L_x(l)$ ist in X^* beschränkt. Also ist $L_x(l)$ ein lineares Funktional in X^* . Ist $x = \theta$ das

Nullelement von X , so gilt $L_\theta(l) = l(\theta) = 0$ für jedes $l \in X^*$, d. h., $L_\theta(l)$ ist das Nullelement von X^{**} . Aus (60) folgt, daß die Norm von $L_x(l)$ nicht größer als $\|x\|$ ist; auf Grund von Satz 2 aus [98] gibt es nun für jedes $x \neq \theta$ ein Funktional $l(x)$ mit $l(x) = \|x\|$ und $\|l\| = 1$. Für ein solches l sind beide Seiten von (60) gleich $\|x\|$, und es gilt das Gleichheitszeichen. Daraus folgt, daß die Norm von $L_x(l)$ gleich $\|x\|$ ist. Ferner folgen aus $l(x_1 + x_2) = l(x_1) + l(x_2)$ und $l(x_1 c_1) = c_1 l(x_1)$ die Beziehungen $L_{x_1 + x_2}(l) = L_{x_1}(l) + L_{x_2}(l)$ und $L_{c_1 x_1}(l) = c_1 L_{x_1}(l)$, allgemein $L_{c_1 x_1 + c_2 x_2}(l) = c_1 L_{x_1}(l) + c_2 L_{x_2}(l)$. Insbesondere ist $L_{x_1 - x_2}(l) = L_{x_1}(l) - L_{x_2}(l)$, woraus wir auf Grund des oben Gesagten $\|L_{x_1} - L_{x_2}\| = \|x_1 - x_2\|$ erhalten; verschiedenen x entsprechen also verschiedene Elemente L_x des Raumes X^{**} . Diese Ergebnisse lassen sich in dem folgenden Satz zusammenfassen:

Satz. *Jedem Element $x \in X$ kann man ein Element $L_x \in X^{**}$ zuordnen. Bei dieser Zuordnung entsprechen verschiedenen x verschiedene L_x , der Addition und der Multiplikation mit einer Zahl in X die gleichen Operationen für die zugeordneten Elemente in X^{**} , und die Normen der einander zugeordneten Elemente in X und in X^{**} sind gleich.*

Dieser Satz gibt die Möglichkeit, L_x und x zu identifizieren, d. h. X in X^{**} einzubetten, was wir mit $X \subseteq X^{**}$ bezeichnen. Mit anderen Worten, X ist einem Teil von X^{**} isometrisch.

Wir sehen ferner, daß X in gewissen Fällen dem ganzen Raum X^{**} isometrisch ist, $X^{**} = X$. Einen solchen Raum nennen wir *regulär*. Statt $l(x) = L_x(l)$ wird häufig die Schreibweise (l, x) oder (x, l) benutzt,

$$(x, l) = (l, x) = l(x); \quad (61)$$

(x, l) heißt das *innere Produkt* eines Elements $l \in X^*$ mit einem Element $x \in X$. Die Elemente l und x sind *orthogonal*, wenn $(x, l) = 0$ ist.

Die Ungleichung (59) hat in dieser Schreibweise die Gestalt

$$|(x, l)| \leq \|l\| \cdot \|x\|, \quad (62)$$

und es ist

$$(ax, bl) = ab(x, l),$$

$$(x_1 + x_2, l) = (x_1, l) + (x_2, l), \quad (x, l_1 + l_2) = (x, l_1) + (x, l_2),$$

wobei a und b beliebige reelle Zahlen sind.

Bis jetzt untersuchten wir reelle Banachräume. Für komplexe Banachräume gilt das gleiche (die Multiplikation eines Elements x wird nun aber mit einer komplexen Zahl vorgenommen). In diesem Fall können auch die Funktionale beliebige komplexe Werte annehmen. Im folgenden bezeichnen wir wie früher mit $\bar{\alpha}$ die zu α konjugiert komplexe Zahl ($\alpha = a + bi$, $\bar{\alpha} = a - bi$). Ist $l(x)$ ein lineares Funktional in B , so ist $\overline{l(x)}$ ebenfalls beschränkt in X , aber kein lineares Funktional, denn $\overline{l(x)}$ wird mit $\bar{c}$ multipliziert, wenn x mit einer komplexen Zahl c multipliziert wird. Der konjugierte Raum X^* ist nicht als Menge der $l(x)$, sondern als Menge der $\overline{l(x)}$ definiert.

Die Addition von Elementen aus X^* und die Multiplikation mit einer komplexen Zahl ist die gewöhnliche Addition und Multiplikation komplexer Zahlen. Es gilt $|\overline{l(x)}| \leq \|l\| \cdot \|x\|$, d. h. $\|\bar{l}\| = \|l\|$: Die Zahl $\|l\|$ kann nicht durch eine kleinere ersetzt werden. Der Raum X^* ist ein Banachraum. Das innere Produkt wird durch

$$(l, x) = \overline{l(x)}, \quad (x, l) = \overline{(l, x)} = l(x) \quad (63)$$

definiert. Für alle komplexen a und b gilt dann

$$(al, bx) = a\bar{b}(l, x). \quad (64)$$

Der Raum $X^{**} = (X^*)^*$ hängt mit X genauso wie im Fall eines reellen Raumes zusammen.

100. Schwache Konvergenz von Funktionalen. Wir untersuchten in [99] die folgende Konvergenz von Folgen linearer Funktional $l_n(x)$ gegen ein Funktional $l(x)$: Es ist $\|l - l_n\| \rightarrow 0$ und damit erst recht $l_n(x) \rightarrow l(x)$ für jedes $x \in X$. Diese Konvergenz heißt im allgemeinen *Konvergenz der Norm nach*. Dabei folgt aus [99], daß die Normen der $l_n(x)$ beschränkt sind, also für jedes n eine gewisse positive Zahl nicht überschreiten. Wir führen nun einen neuen Konvergenzbegriff ein. Wir sagen, eine Folge linearer Funktional $l_n(x)$ *konvergiere schwach*, wenn für jedes $x \in X$ die Zahlenfolge $l_n(x)$ einen (endlichen) Grenzwert $l(x)$ besitzt. Dieser Grenzwert ist ein in ganz X definiertes Funktional. Seine Distributivität folgt aus der von $l_n(x)$, und wir könnten auch auf seine Beschränktheit (und damit auf seine Linearität) schließen, wenn wir wüßten, ob die Folge der $\|l_n\|$ beschränkt ist. Dies ist aber tatsächlich der Fall.

Satz 1. *Es sei L eine Menge linearer Funktional $l(x)$, wobei zu jedem Element x eine positive Zahl m_x existiert, für welche $|l(x)| \leq m_x$ gilt, wenn l zu L gehört, d. h., für ein festes x ist die Menge der Zahlen $|l(x)|$ beschränkt. Dabei sind die Normen der Funktional $l(x)$, $l \in L$, beschränkt.*

Beweis. Hinreichend für den Beweis des Satzes ist die Feststellung, daß $|l(x)|$ in irgendeiner Sphäre beschränkt ist. In der Tat, gegeben sei eine positive Zahl b , für welche

$$|l(x)| \leq b \quad (l \in L) \quad (65)$$

gilt, wenn x zu einer abgeschlossenen Sphäre $\bar{S}(x_0, a)$ ($\|x - x_0\| \leq a$) gehört, und es sei y ein vom Nullelement verschiedenes Element aus X . Das Element $x = \frac{a}{\|y\|}y + x_0$ liegt in $\bar{S}(x_0, a)$, und wegen (65) gilt

$$\left| \frac{a}{\|y\|} l(y) + l(x_0) \right| \leq b \quad \text{und somit} \quad \frac{a}{\|y\|} |l(y)| - |l(x_0)| \leq b,$$

woraus sich

$$|l(y)| \leq \frac{b + |l(x_0)|}{a} \|y\| \leq \frac{2b}{a} \|y\| \quad (66)$$

für jedes $y \in X$, also $\|l\| \leq \frac{2b}{a}$ ergibt. Dies ist genau die Behauptung des Satzes.

Können wir nachweisen, daß die Menge der Zahlen $|l(x)|$ in irgendeiner Sphäre beschränkt ist, so ist der Satz bewiesen. Wir nehmen dazu an, die erwähnte Menge sei nicht beschränkt in einer beliebigen abgeschlossenen Sphäre, und führen den Beweis auf einen Widerspruch. Wir wählen eine beliebige abgeschlossene Sphäre $\bar{S}_1$. Auf Grund des Bewiesenen gibt es ein Element $x_1 \in \bar{S}_1$ und ein Funktional $l_1 \in L$ mit $|l_1(x_1)| > 1$. Da $l_1(x)$ stetig ist, können wir annehmen, daß x_1 innerhalb von $\bar{S}_1$ liegt und die Ungleichung $|l_1(x)| > 1$ in der ganzen Sphäre $\bar{S}(x_1, r_1)$ erfüllt ist, wobei r_1 eine hinreichend kleine positive Zahl ist. Wie oben läßt sich hier ein Element $x_2 \in \bar{S}(x_1, r_1)$, ein Funktional $l_2 \in L$ und ein so kleines positives r_2 finden, daß $\bar{S}(x_2, r_2)$ zu $\bar{S}(x_1, r_1)$ gehört und in der ganzen Sphäre $\bar{S}(x_2, r_2)$ die Ungleichung $|l_2(x)| > 2$

gilt. Führen wir diesen Prozeß weiter fort, so erhalten wir eine Folge von Sphären

$$S(x_1, r_1) \supset \bar{S}(x_2, r_2) \supset \bar{S}(x_3, r_3) \supset \dots$$

und von Funktionalen $l_k \in L$ derart, daß in ganz $\bar{S}(x_k, r_k)$ die Ungleichung $|l_k(x)| > k$ erfüllt ist. Dabei kann $r_k \rightarrow 0$ für $k \rightarrow \infty$ angenommen werden. Im Punkt x_0 , der allen diesen Sphären angehört [85], gilt somit $|l_k(x_0)| > k$. Das widerspricht aber der Tatsache, daß die Menge der Zahlen $|l_k(x_0)|$ beschränkt sein sollte. Damit ist der Satz bewiesen.

Hat eine Folge von Funktionalen $l_n(x)$ für jedes x einen endlichen Grenzwert, so ist die Folge der Zahlen $|l_n(x)|$ für jedes x und auf Grund des Satzes auch die Folge der $\|l_n\|$ beschränkt. Daraus folgt, wie wir schon oben bemerkten, daß der Grenzwert $l(x)$ einer schwach konvergenten Folge linearer Funktione ebenfalls ein lineares Funktional ist.

Ist die Folge der $\|l_n\|$ beschränkt, so braucht für die schwache Konvergenz von Funktionalen die Existenz eines Grenzwertes für die $l_n(x)$ nicht auf ganz X , sondern nur auf einer in X dichten linearen Mannigfaltigkeit gefordert zu werden.

Satz 2. *Eine Folge von Funktionalen $l_n(x)$ konvergiert schwach genau dann, wenn die Folge der $\|l_n\|$ beschränkt ist und der Grenzwert der $l_n(x)$ auf einer in X dichten linearen Mannigfaltigkeit U existiert.*

Beweis. Die Notwendigkeit der ersten Bedingung folgt aus Satz 1, die der zweiten ist klar. Wir zeigen also, daß die Bedingungen hinreichend sind. Auf Grund der ersten Bedingung ist $\|l_n\| \leq c$, wobei c eine positive Zahl ist. Bezeichnen wir mit $l(x)$ den Grenzwert der $l_n(x)$ für $x \in U$, so können wir behaupten, daß $l(x)$ ein auf U beschränktes distributives Funktional ist (dessen Norm höchstens gleich c ist). Wir können $l(x)$ auf den ganzen Raum X stetig fortsetzen. Das so erhaltene Funktional bezeichnen wir ebenfalls mit $l(x)$ und zeigen, daß $l_n(x) \rightarrow l(x)$ für jedes $x \in X$ gilt. Dies ist der Fall für $x \in U$. Es sei $x_0 \in U$. Zu einem gegebenen $\varepsilon > 0$ existiert ein Element $x'_0 \in U$ mit $\|x_0 - x'_0\| \leq \frac{\varepsilon}{4c}$. Berücksichtigen wir, daß die Normen von $l_n(x)$ und $l(x)$ höchstens gleich c sind, so können wir

$$\begin{aligned} |l(x_0) - l_n(x_0)| &\leq |l(x_0) - l(x'_0)| + |l(x'_0) - l_n(x'_0)| + |l_n(x'_0) - l_n(x_0)| \\ &\leq \|l\| \cdot \|x_0 - x'_0\| + |l(x'_0) - l_n(x'_0)| + \|l_n\| \cdot \|x'_0 - x_0\| \\ &\leq \frac{\varepsilon}{2} + |l(x'_0) - l_n(x'_0)| \end{aligned}$$

schreiben. Nun ist $l_n(x'_0) \rightarrow l(x'_0)$; folglich existiert ein N derart, daß $|l(x'_0) - l_n(x'_0)| \leq \frac{\varepsilon}{2}$ für $n \geq N$ ist, und aus der obigen Ungleichung ergibt sich $|l(x_0) - l_n(x_0)| \leq \varepsilon$ für $n \geq N$, also $l_n(x_0) \rightarrow l(x_0)$.

Bemerkung. Die zweite Forderung des Satzes kann durch die folgende ersetzt werden: Es existiert ein Grenzwert der $l_n(x)$ auf einer Menge V von Elementen, deren lineare Hülle U (eine lineare Mannigfaltigkeit) in X dicht ist. Aus der Konvergenz der $l_n(x)$ auf V folgt nämlich wegen der Distributivität der $l_n(x)$ die Konvergenz der $l_n(x)$ auf U .

Der Begriff der schwachen Konvergenz von Funktionalen führt natürlich auf den Begriff der *schwachen Kompaktheit*. Eine Menge W von Elementen aus X^* heißt schwach kompakt, wenn man aus jeder Folge von Funktionalen $l_n \in W$ eine schwach konvergente Teilfolge auswählen kann.

Satz 3. *Ist X separabel, so ist jede beschränkte Menge von Funktionalen ($\|l\| \leq r$, $r > 0$) schwach kompakt.*

Beweis. Wir müssen zeigen: Sind bei einer Folge linearer Funktionale $l_n(x)$ die Normen beschränkt, $\|l_n\| \leq r$, so können wir eine für jedes $x \in X$ konvergente Teilfolge $l_{n_k}(x)$ auswählen. Die Elemente $x_1, x_2, \dots$ aus X mögen eine in X dichte abzählbare Menge V bilden. Für jedes m gilt $|l_n(x_m)| \leq r \|x_m\|$, d. h., die Folge der Zahlen $l_n(x_m)$, $n = 1, 2, \dots$, ist beschränkt. Mit Hilfe des üblichen Diagonalverfahrens [IV/1, 15] bilden wir die Teilfolge $l_{n_k}(x)$, die auf allen Elementen von V konvergiert. Aus der Bemerkung zu Satz 2 folgt, daß diese Folge auf ganz X konvergiert. Damit ist der Satz bewiesen.

Bemerkung. Schwach kompakt ist offenbar auch jede beschränkte Menge von Elementen aus X^* , da man sie in eine Sphäre $\|l\| \leq r$ einbetten kann.

101. Schwache Konvergenz von Elementen. Wir führen nun den Begriff der schwachen Konvergenz von Elementen aus einem Banachraum X ein. Man sagt, eine Folge von Elementen $x_n \in X$ *konvergiere schwach* gegen ein Element x_0 (in Zeichen $x_n \rightarrow x_0$), wenn $l(x_n) \rightarrow l(x)$ für jedes lineare Funktional $l(x)$ gilt. Das Element x_0 heißt *schwacher Grenzwert* von x_n . Wir wollen zeigen, daß eine Folge x_n nicht gegen verschiedene Grenzwerte schwach konvergieren kann. Ist $x_n \rightarrow x_0$ und $x_n \rightarrow y_0$, so gilt nach Definition $l(x_n) \rightarrow l(x_0)$ und $l(x_n) \rightarrow l(y_0)$ für jedes $l \in X^*$; daraus folgt $l(y_0) = l(x_0)$ oder $l(y_0 - x_0) = 0$ für jedes $l \in X^*$. Ist hierbei $y_0 \neq x_0$, d. h., ist $y_0 - x_0$ nicht das Nullelement, so existiert ein Element $l \in X^*$, für welches $l(y_0 - x_0) = \|y_0 - x_0\| > 0$ gilt. Das widerspricht jedoch dem oben Gesagten, so daß also die Behauptung bewiesen ist.

Strebt x_n schwach gegen x_0 , so ist offenbar $x_{n_k} \rightarrow x_0$ für jede Teilfolge. Aus $x_n \rightarrow x_0$, $y_n \rightarrow y_0$ und $a_n \rightarrow a_0$ folgt $a_n x_n \rightarrow a_0 x_0$ und $x_n + y_n \rightarrow x_0 + y_0$. Dies ergibt sich unmittelbar aus der Distributivität der entsprechenden Funktionale.

Die Konvergenz der Norm nach ($\|x_0 - x_n\| \rightarrow 0$), die wir früher auch in der Form $x_n \Rightarrow x_0$ schrieben, wird manchmal auch *starke Konvergenz* oder einfach *Konvergenz* genannt. Aus $x_n \Rightarrow x_0$ folgt auf Grund der Stetigkeit eines linearen Funktionals die Beziehung $l(x_n) \rightarrow l(x_0)$ für jedes $l \in X^*$, d. h., die starke Konvergenz zieht die schwache nach sich.

Aus der schwachen Konvergenz folgt im allgemeinen die starke Konvergenz nicht. Dazu geben wir ein Beispiel an. Wir betrachten den Raum $L_2[0, 1]$ und wählen als Folge der x_n die Funktionen $\sin n\pi t$ ($n = 1, 2, \dots$). Wie wir in [102] zeigen werden, wird die allgemeine Gestalt der Funktionale in $L_2[0, 1]$ durch

$$l(x) = \int_0^1 f(t) x(t) dt$$

angegeben, wobei $f(t)$ eine feste Funktion aus $L_2[0, 1]$ und $x(t)$ ein beliebiges Element dieses Raumes ist. Insbesondere gilt für die Elemente $x_n(t) = \sin n\pi t$

$$l(x_n) = \int_0^1 f(t) \sin n\pi t dt ;$$

daraus ist ersichtlich, daß die $l(x_n)$ bis auf den Faktor $\sqrt{2}$ die Fourierkoeffizienten von $f(t)$ in bezug auf das System der Funktionen $\sin n\pi t$ auf dem Intervall $[0, 1]$ ist. Wir wissen, daß $l(x_n)$ für jedes $f(t) \in L_2[0, 1]$ gegen 0 konvergiert, also $l(x_n) \rightarrow l(\theta)$ ist; dabei ist θ das Nullelement von $L_2[0, 1]$, d. h. irgendeine Funktion, die äqui-

valent 0 ist. Somit ist $\sin n\pi t \rightarrow \theta$ für $n \rightarrow \infty$ in $L_2[0, 1]$. Jedoch liegt keine starke Konvergenz vor, denn es ist

$$\|\theta - x_n\|^2 = \int_0^1 \sin^2 n\pi t dt = \frac{1}{2}.$$

Satz 1. *Ist $x_n \rightarrow x_0$, so ist die Folge der $\|x_n\|$ beschränkt.*

Beweis. Wir betrachten x_n und x_0 als Elemente von X^{**} . Dabei folgt aus $x_n \rightarrow x_0$, daß die den x_n entsprechenden Funktionale in X^* gegen das dem Element x_0 entsprechende Funktional schwach konvergieren. Dann bilden auf Grund von Satz 1 aus [100] die Normen dieser Funktionale, die gleich $\|x_n\|$ sind, eine beschränkte Menge. Damit ist der Satz bewiesen.

Die *schwache Kompaktheit einer Menge von Elementen* aus X wird genauso definiert wie die schwache Kompaktheit einer Menge von Funktionalen. Jede beschränkte Menge von Elementen x ($\|x\| \leq c$) ist eine beschränkte Menge in X^{**} ; diese ist als Menge von Funktionalen in X^* , wenn X^* separabel ist, schwach kompakt. Ist X ein regulärer Raum, d. h., ist $X^{**} = X$, so folgt hieraus der

Satz 2. *Ist X regulär und sind X und X^* separabel, so ist jede beschränkte Menge von Elementen aus X schwach kompakt.*

Ist X kein regulärer Raum, ist also $X^{**} \supset X$, so kann der Grenzwert einer Folge von Elementen aus X^{**} dasjenige Element aus X^{**} sein, dem kein Element von X entspricht. Ist X separabel und regulär, so ist auch X^* separabel.

Aus der obenerwähnten Zuordnung von Elementen $x \in X$ und Elementen aus X^{**} sowie aus Satz 2 aus [100] ergibt sich unmittelbar:

Satz 3. *Eine Folge von Elementen x_n eines regulären Raumes X ist genau dann schwach konvergent, wenn die Folge der $\|x_n\|$ beschränkt ist und wenn der Grenzwert der $l(x_n)$ auf einer in X^* dichten linearen Mannigfaltigkeit U von Elementen $l \in X^*$ existiert.*

Wie in [100] kann man U durch eine Menge V von Elementen aus X^* ersetzen, deren lineare Hülle in X^* dicht ist.

Satz 4. *Es sei A ein linearer Operator im Banachraum X , und $R(A)$ gehöre zum Banachraum X' . Ist dann $x_n \rightarrow x_0$ in X , so gilt $Ax_n \rightarrow Ax_0$ in X' .*

Beweis. Es sei $m(y)$ ein beliebiges lineares Funktional in X' . Wir sehen sofort, daß $m(Ax)$ ein lineares Funktional in X ist, und aus $x_n \rightarrow x_0$ in X folgt $m(Ax_n) \rightarrow m(Ax_0)$. Dies gilt für jedes Funktional in X' ; folglich ist $Ax_n \rightarrow Ax_0$ in X' , was zu beweisen war.

Wir wissen, daß ein linearer Operator im Sinne der starken Konvergenz stetig ist. Satz 4 besagt, daß er auch im Sinne der schwachen Konvergenz stetig ist.

Wir sahen, daß $\|x_n\| \rightarrow \|x_0\|$ aus $x_n \rightarrow x_0$ folgt [95]. Bei schwacher Konvergenz gilt diese Eigenschaft nicht. Für die oben betrachtete Folge der Funktionen $\sin n\pi x$ aus $L_2[0, 1]$ war $\sin n\pi x \rightarrow \theta$, aber $\|\sin n\pi x\| = \frac{1}{\sqrt{2}}$.

Satz 5. *Ist $x_n \rightarrow x_0$, so gilt $\|x_0\| \leq \overline{\lim} \|x_n\|$.*

Beweis. Wir bemerken vor allem, daß $\overline{\lim} \|x_n\|$ auf Grund von Satz 1 endlich ist. Den Beweis führen wir indirekt. Es sei $\|x_0\| > \overline{\lim} \|x_n\|$, und wir wählen eine

Zahl m , die der Ungleichung

$$||x_0|| > m > \overline{\lim} ||x_n|| \quad (67)$$

genügt. Daraus folgt, daß ein N existiert, so daß $||x_n|| < m$ für $n \geq N$ ist. Ferner gibt es ein lineares Funktional $l(x)$ mit $l(x_0) = ||x_0||$ und $||l|| = 1$ [98], und es gilt $|l(x_n)| \leq ||l|| \cdot ||x_n||$, also $|l(x_n)| \leq ||x_n|| < m$ für $n \geq N$ oder wegen (67)

$$l(x_0) = ||x_0|| > m.$$

Somit strebt $l(x_n)$ nicht gegen $l(x_0)$; das widerspricht jedoch den Voraussetzungen des Satzes.

Wir nehmen nun an, daß der Raum X die folgenden Eigenschaften besitzt: Zu jedem gegebenen $\delta > 0$ existiert eine Zahl $\eta < 1$ derart, daß die Beziehungen $||x|| = ||y|| = 1$ und $||x - y|| = \delta$ die Ungleichung $\frac{||x + y||}{2} < \eta$ nach sich ziehen. In diesem Fall nennt man X einen *gleichmäßig konvexen* Raum. Für solche Räume gilt: Aus $x_n \rightarrow x_0$ und $||x_n|| \rightarrow ||x_0||$ folgt $x_n \Rightarrow x_0$. In Kapitel V werden wir diese Behauptung für einen Spezialfall, und zwar für den Hilbertraum, beweisen. Auch die Räume L_p ($p > 1$) sind gleichmäßig konvex.¹⁾

Satz 6. Ist $x_n \rightarrow x_0$, so gehört x_0 zu der (bezüglich der Norm) abgeschlossenen Hülle der linearen Hülle der Menge der Elemente x_n ($n = 1, 2, \dots$).

Beweis. Wir führen den Beweis indirekt. Es sei U die lineare Hülle der Elemente x_n , und wir setzen voraus, x_0 gehöre nicht zu U , d. h., es sei

$$\inf_{y \in U} ||x_0 - y|| = d > 0. \quad (68)$$

Für jede von 0 verschiedene Zahl t gilt

$$||tx_0 + y|| \geq |t| d, \quad (69)$$

denn es ist

$$||tx_0 + y|| = |t| \cdot \left\| x_0 - \frac{1}{t} y \right\|.$$

Mit y gehört auch $\frac{1}{t} y$ zu U , und (69) folgt sofort aus (68). Wir betrachten nun die Menge von Elementen der Gestalt

$$x = tx_0 + y, \quad (70)$$

wobei y zu U gehört und t beliebig ist. Wie in [98] können wir leicht erkennen, daß die Darstellung von x in der Form (70) eindeutig bestimmt ist und die Elemente x eine lineare Mannigfaltigkeit bilden, die wir mit V bezeichnen. Auf V definieren wir ein distributives Funktional durch die Formel $l(x) = t$, so daß $l(y) = 0$ für $y \in U$ gilt. Dieses Funktional ist auf V beschränkt; denn auf Grund von (69) erhalten wir für $t \neq 0$ die Ungleichung

$$|l(tx_0 + y)| = |t| \leq \frac{1}{d} ||tx_0 + y||,$$

die für $t = 0$ evident ist. Somit ist $||l|| \leq \frac{1}{d}$ auf V . Wir können l auf den ganzen Raum X unter Beibehaltung der Norm fortsetzen und auf diese Art ein gewisses Funktio-

¹⁾ Vgl. etwa S. L. SOBOLEW, Einige Anwendungen der Funktionalanalysis in der mathematischen Physik (vgl. das Literaturverzeichnis).

nal $l(x)$ erhalten. Nach Voraussetzung ist $l(x_0) = 1$ und $l(x_n) = 0$, denn alle x_n liegen in U . Also strebt $l(x_n)$ nicht gegen $l(x_0)$, was der Voraussetzung $x_n \rightarrow x_0$ des Satzes widerspricht. Damit ist der Satz bewiesen.

Den Satz 6 können wir auch folgendermaßen formulieren: *Ist $x_n \rightarrow x_0$, so existiert eine Folge von Linearkombinationen der Elemente x_n ,*

$$c_1^k x_1 + c_2^k x_2 + \dots + c_{n_k}^k x_{n_k} \quad (k = 1, 2, \dots),$$

die stark gegen x_0 strebt,

$$c_1^k x_1 + c_2^k x_2 + \dots + c_{n_k}^k x_{n_k} \rightarrow x_0 \quad \text{für } k \rightarrow \infty.$$

Es läßt sich auch noch eine stärkere Behauptung beweisen: Ist $x_n \rightarrow x_0$, so existiert eine Teilfolge x_{n_k} ($k = 1, 2, \dots$) mit

$$\frac{1}{k} (x_{n_1} + x_{n_2} + \dots + x_{n_k}) \rightarrow x_0 \quad \text{für } k \rightarrow \infty.$$

Satz 7. *Ist der Raum X regulär und ist eine Folge von Elementen $x_n \in X$ eine in sich schwach konvergente Folge, d. h., ist $l(x_n) - l(x_m) = l(x_n - x_m) \rightarrow 0$ für $n, m \rightarrow \infty$ und jedes Element $l \in X^*$, so ist die Folge der x_n schwach konvergent.*

Beweis. Aus der Voraussetzung und aus dem Cauchyschen Konvergenzkriterium für Zahlenfolgen ergibt sich, daß die $l(x_n)$ für jedes $l \in X^*$ einen Grenzwert besitzen, d. h., die linearen Funktionale $L_{x_n}(l) = l(x_n)$ haben in X^* für jedes $l \in X^*$ einen Grenzwert, der ebenfalls ein lineares Funktional in X^* ist. Da X regulär ist, muß dieser Grenzwert die Gestalt $L_{x_0}(l) = l(x_0)$ haben, d. h., für jedes $l \in X^*$ gilt $l(x_n) \rightarrow l(x_0)$, also $x_n \rightarrow x_0$, was zu beweisen war.

Auch Satz 7 läßt sich anders formulieren: *Ein regulärer Raum X ist schwach vollständig.*

102. Lineare Funktionale in den Räumen C , L_p und l_p .

1. Wir wissen, daß ein lineares Funktional im Raum $C[a, b]$ durch die Formel

$$l(f) = \int_a^b f(x) dg(x) \quad (71)$$

allgemein dargestellt wird [15], wobei $g(x)$ eine Funktion von beschränkter Variation und von rechts stetig ist sowie der Bedingung $g(a) = 0$ genügt; verschiedene Funktionen $g(x)$ erzeugen verschiedene Funktionale $l(f)$. Wir wissen auch, daß $\|l\| = \bigvee_a^b g(x)$ gilt. Somit können wir jedem $l(f)$ in C eine Funktion $g(x)$ mit den erwähnten Eigenschaften und mit der Norm $\|g\| = \bigvee_a^b g(x)$ zuordnen und den Raum C^* mit dem Raum V dieser Funktionen identifizieren. Der Raum V ist ein Banachraum [96].

Wir betrachten nun Funktionale in C^* , d. h. in V . Ein solches Funktional ist durch (71) für jede feste und auf $[a, b]$ stetige Funktion $f(x)$ gegeben. Der Raum C ist also in C^{**} enthalten.

Wir wollen zeigen, daß nicht alle Funktionale aus V durch (71) darstellbar sind. Als Funktional $l_0(g)$ aus V wählen wir die Summe der Sprünge von $g(x)$ und weisen nach, daß es für keine stetige Funktion $f(x)$ durch (71) dargestellt werden kann.

Dazu konstruieren wir das folgende Element von V :

$$g_0(x) = \begin{cases} 0 & \text{für } a \leq x < c, \\ 1 & \text{für } c \leq x \leq b, \quad c > a. \end{cases}$$

Würde (71) gelten, so erhielten wir $l_0(g_0) = f(c)$. Da die Summe der Sprünge von $g_0(x)$ gleich 1 ist, muß $f(c) = 1$ sein, d. h., die stetige Funktion $f(x)$ muß identisch 1 sein. Dann folgt aus (71) $l_0(g) = g(b) - g(a)$; diese Differenz ist aber nicht für jedes $g(x) \in V$ eine Summe von Sprüngen. Somit ist C^{**} umfassender als C , d. h., C ist kein regulärer Raum. Alles oben Gesagte bezieht sich auf reelle Funktionen, läßt sich aber auch auf komplexe Funktionen übertragen.

2. Wir geben jetzt die allgemeine Form eines linearen Funktional im Raum L_p ($p > 1$) der reellen Funktionen auf einer beschränkten meßbaren Menge $\mathfrak{E}_0$ des R_n an. Es sei $l(f)$ ein solches Funktional und $\omega_{\mathfrak{E}}(x)$ die charakteristische Funktion einer beliebigen, in $\mathfrak{E}_0$ enthaltenen meßbaren Menge $\mathfrak{E}$. Offenbar ist $\omega_{\mathfrak{E}}(x) \in L_p(\mathfrak{E}_0)$, und wir setzen

$$l(\omega_{\mathfrak{E}}) = F(\mathfrak{E}). \quad (72)$$

Wir zeigen zunächst, daß diese Mengenfunktion, die auf allen meßbaren Mengen $\mathfrak{E}$ aus $\mathfrak{E}_0$ definiert ist, volladditiv ist. Es sei $\mathfrak{E} = \mathfrak{E}_1 + \mathfrak{E}_2 + \dots$, wobei die meßbaren Mengen $\mathfrak{E}_k$ paarweise punktfremd sein mögen. Die Reihe

$$\sum_{k=1}^{\infty} \omega_{\mathfrak{E}_k}(x) \quad (73)$$

konvergiert in $L_p(\mathfrak{E}_0)$; denn es ist

$$\left\| \sum_{k=q}^r \omega_{\mathfrak{E}_k}(x) \right\|_{L_p(\mathfrak{E}_0)} = \left[\sum_{k=q}^r \int_{\mathfrak{E}_k} dx \right]^{1/p} = \left[\sum_{k=q}^r m(\mathfrak{E}_k) \right]^{1/p},$$

und die letzte Summe strebt für $q, r \rightarrow \infty$ auf Grund der Volladditivität des Maßes gegen 0. In jedem Punkt $x \in \mathfrak{E}$ konvergiert (73) gegen $\omega_{\mathfrak{E}}(x)$. Folglich strebt (73) auch in $L_p(\mathfrak{E}_0)$ gegen $\omega_{\mathfrak{E}}(x)$ [62] (oder gegen eine ihr äquivalente Funktion), d. h.

$$\omega_{\mathfrak{E}}(x) = \sum_{k=1}^{\infty} \omega_{\mathfrak{E}_k}(x), \quad (74)$$

wobei die Konvergenz als Konvergenz in $L_p(\mathfrak{E}_0)$ aufgefaßt werden kann. Auf Grund der Stetigkeit des Funktional $l(f)$ in $L_p(\mathfrak{E}_0)$ erhalten wir aus (74)

$$F(\mathfrak{E}) = F(\mathfrak{E}_1) + F(\mathfrak{E}_2) + \dots,$$

also eine volladditive Funktion $F(\mathfrak{E})$. Ist $\mathfrak{E}'$ eine Menge vom Maß 0, so gilt

$$|l(\omega_{\mathfrak{E}'}(x))| \leq \|l\| \cdot \|\omega_{\mathfrak{E}'}(x)\|_{L_p(\mathfrak{E}_0)} \leq \|l\| \left[\int_{\mathfrak{E}'} dx \right]^{1/p} = 0$$

und somit $l(\omega_{\mathfrak{E}'}(x)) = 0$ für $m(\mathfrak{E}') = 0$; der Satz aus [73] liefert für $F(\mathfrak{E})$ die Darstellung

$$F(\mathfrak{E}) = \int_{\mathfrak{E}} \psi(x) dx,$$

wobei die Funktion $\psi(x)$ auf $\mathfrak{E}_0$ summierbar ist. Damit ist

$$l(\omega_{\mathfrak{E}}(x)) = \int_{\mathfrak{E}_0} \psi(x) \omega_{\mathfrak{E}}(x) dx$$

bewiesen.

Auf Grund der Distributivität von $l(f)$ gilt

$$l(\varphi) = \int_{\mathfrak{E}_0} \psi(x) \varphi(x) dx \quad (75)$$

für jede beschränkte Funktion $\varphi(x)$ mit endlich vielen Werten. Wir zeigen jetzt, daß (75) auch für jede beschränkte und auf $\mathfrak{E}_0$ meßbare Funktion $\varphi(x)$ gilt (jede solche Funktion gehört zu $L_p(\mathfrak{E}_0)$), und nehmen zunächst $\varphi(x) \geq 0$ an. Infolge von Satz 1 aus [46] existiert eine Folge von Funktionen $\varphi_n(x)$ mit endlich vielen Werten, die auf $\mathfrak{E}_0$ gleichmäßig gegen $\varphi(x)$ konvergiert. Somit sind die $\varphi_n(x)$ auf $\mathfrak{E}_0$ durch ein und dieselbe Zahl beschränkt. Für $\varphi_n(x)$ gilt

$$l(\varphi_n) = \int_{\mathfrak{E}_0} \varphi(x) \varphi_n(x) dx. \quad (76)$$

Offenbar ist $\varphi_n(x) \Rightarrow \varphi(x)$ in $L_p(\mathfrak{E}_0)$, so daß in (76) der Grenzübergang unter dem Integralzeichen vorgenommen werden kann [54]. Die Stetigkeit des Funktional $l(\varphi)$ führt auf die Formel (75), die somit für jede nichtnegative, beschränkte und auf $\mathfrak{E}_0$ meßbare Funktion $\varphi(x)$ gilt. Der Fall einer beschränkten Funktion beliebigen Vorzeichens reduziert sich mit Hilfe der Darstellung $\varphi(x) = \varphi^+(x) - \varphi^-(x)$, wobei $\varphi^+(x)$ und $\varphi^-(x)$ der positive bzw. der negative Teil von $\varphi(x)$ ist, auf den betrachteten Fall. Wir zeigen nun, daß $\varphi(x)$ zu $L_{p'}(\mathfrak{E}_0)$ gehört; dabei ist $\frac{1}{p} + \frac{1}{p'} = 1$. Setzen wir in (75) die durch

$$\varphi(x) = \begin{cases} |\varphi(x)|^{p'-1} \operatorname{sgn} \varphi(x) & \text{für } |\varphi(x)| \leq N, \\ N^{p'-1} \operatorname{sgn} \varphi(x) & \text{für } |\varphi(x)| > N \end{cases} \quad (77)$$

mit

$$\operatorname{sgn} \alpha = \begin{cases} -1 & \text{für } \alpha < 0, \\ 0 & \text{für } \alpha = 0, \\ 1 & \text{für } \alpha > 0 \end{cases}$$

definierte beschränkte meßbare Funktion $\varphi(x)$ ein, so erhalten wir

$$l(\varphi) \equiv \int_{\mathfrak{E}_0} |\varphi(x)|^p dx, \quad (78)$$

denn es ist

$$|\varphi(x)| \equiv |\varphi(x)|^{\frac{1}{p'-1}} \quad \text{und} \quad \frac{p'}{p'-1} = p.$$

Andererseits gilt

$$l(\varphi) \equiv ||l|| \cdot ||\varphi|| = ||l|| \left[\int_{\mathfrak{E}_0} |\varphi(x)|^p dx \right]^{1/p}$$

und wegen (78)

$$\int_{\mathfrak{E}_0} |\varphi(x)|^p dx \equiv ||l|| \left[\int_{\mathfrak{E}_0} |\varphi(x)|^p dx \right]^{1/p};$$

daraus folgt

$$\left[\int_{\mathfrak{E}_0} |\varphi(x)|^p dx \right]^{1/p'} \equiv ||l||.$$

Aus (77) ergibt sich aber

$$|\varphi(x)|^p = \begin{cases} |\varphi(x)|^{p'} & \text{für } |\varphi(x)| \leq N, \\ N^{p'} & \text{für } |\varphi(x)| > N \end{cases}$$

und infolgedessen

$$\left[\int_{\mathfrak{E}_0} |\varphi_N(x)|^{p'} dx \right]^{1/p'} \equiv ||l||, \quad (79)$$

wobei $\varphi_N(x)$ die Abschnittsfunktion von $\varphi(x)$ ist. Daraus erhalten wir $\varphi(x) \in L_{p'}(\mathfrak{E}_0)$ und

$$||l|| \cong ||\varphi||_{L_{p'}(\mathfrak{E}_0)}. \quad (80)$$

Es sei nun $\varphi(x)$ eine beliebige Funktion aus $L_p(\mathfrak{E}_0)$. Es existiert eine Folge meßbarer beschränkter Funktionen $\varphi_n(x)$, die in $L_p(\mathfrak{E}_0)$ gegen $\varphi(x)$ strebt. Infolge der Stetigkeit des Funktionals ist $l(\varphi_n) \rightarrow l(\varphi)$ und auf Grund von [62]

$$\int_{\mathfrak{E}_0} \varphi(x) \varphi_n(x) dx \rightarrow \int_{\mathfrak{E}_0} \varphi(x) \varphi(x) dx.$$

Für $\varphi_n(x)$ gilt die Beziehung (75), und aus dem oben Gesagten folgt, daß sie auch für jede Funktion $\varphi(x) \in L_p(\mathfrak{E}_0)$ richtig ist. Mit Hilfe der Hölderschen Ungleichung finden wir

$$|l(\varphi)| = \left[\int_{\mathfrak{E}_0} |\varphi(x)|^{p'} dx \right]^{1/p'} ||\varphi||_{L_p(\mathfrak{E}_0)}$$

und

$$||l|| \cong ||\varphi||_{L_{p'}(\mathfrak{E}_0)}; \quad (81)$$

dies ergibt zusammen mit (80)

$$||l|| = ||\varphi||_{L_{p'}(\mathfrak{E}_0)}. \quad (82)$$

Somit läßt sich jedes lineare Funktional in $L_p(\mathfrak{E}_0)$ durch (75) mit $\varphi(x) \in L_{p'}(\mathfrak{E}_0)$ darstellen, und es gilt (82).

Nun sei $\varphi(x)$ eine beliebige feste Funktion aus $L_{p'}(\mathfrak{E}_0)$. Auf Grund der Hölderschen Ungleichung liefert (75) ein lineares Funktional in $L_p(\mathfrak{E}_0)$, dessen Norm der Ungleichung (81) genügt, d. h., (75) ist die allgemeine Form eines linearen Funktionals in $L_p(\mathfrak{E}_0)$. Berücksichtigen wir, daß aus $p(p' - 1) = p'$ die Zugehörigkeit der Funktionen $\varphi(x) = |\varphi(x)|^{p'-1} \operatorname{sgn} \varphi(x)$ zu $L_p(\mathfrak{E}_0)$ folgt, so können wir diese Funktion in (75) einsetzen:

$$l(|\varphi(x)|^{p'-1} \operatorname{sgn} \varphi(x)) = \int_{\mathfrak{E}_0} |\varphi(x)|^{p'} dx. \quad (83)$$

Die Norm von $\varphi(x) = |\varphi(x)|^{p'-1} \operatorname{sgn} \varphi(x)$ in $L_p(\mathfrak{E}_0)$ ist gleich

$$\left[\int_{\mathfrak{E}_0} |\varphi(x)|^p dx \right]^{1/p} = \left[\int_{\mathfrak{E}_0} |\varphi(x)|^{p'} dx \right]^{1/p},$$

und aus (83) erhalten wir

$$\int_{\mathfrak{E}_0} |\varphi(x)|^{p'} dx \cong ||l|| \left[\int_{\mathfrak{E}_0} |\varphi(x)|^{p'} dx \right]^{1/p},$$

also

$$||l|| \cong ||\varphi(x)||_{L_{p'}(\mathfrak{E}_0)};$$

diese Ungleichung und (81) ergeben wiederum die Beziehung (82). Somit stellt die Formel (75), in der $\varphi(x)$ eine beliebige Funktion aus $L_{p'}(\mathfrak{E}_0)$ ist, die allgemeine Form eines linearen Funktionals in $L_p(\mathfrak{E}_0)$ dar; dabei gilt (82).

Äquivalente Funktionen $\varphi(x)$ ergeben offenbar gleiche Funktionale (die auf allen $\varphi(x) \in L_p(\mathfrak{E}_0)$ übereinstimmen). Wir zeigen, daß nicht äquivalente $\varphi(x)$ verschiedene Funktionale ergeben. Dies läßt sich offenbar auf den Beweis der folgenden Behauptung zurückführen: Ist $\varphi_0(x) \in L_p(\mathfrak{E}_0)$ und gilt für jedes $\varphi(x) \in L_p(\mathfrak{E}_0)$

$$\int_{\mathfrak{E}_0} \varphi_0(x) \varphi(x) dx = 0,$$

so ist $\varphi_0(x)$ äquivalent 0. Setzen wir $\varphi(x) = |\varphi_0(x)|^{p'-1} \operatorname{sgn} \varphi_0(x)$, so erhalten wir

$$\int_{\mathfrak{E}_0} |\varphi_0(x)|^{p'} dx = 0;$$

daraus folgt sofort, daß $\varphi_0(x)$ äquivalent 0 ist [51].

Wir untersuchen nun den Raum $L_p(\mathfrak{E}_\infty)$, wobei $\mathfrak{E}_\infty$ der ganze R_n sei. Wie oben läßt sich zeigen, daß die Formel (75) für $\varphi(x) \in L_{p'}(\mathfrak{E}_\infty)$ ein lineares Funktional in $L_p(\mathfrak{E}_\infty)$ definiert und daß (82) gilt. Wir wollen beweisen, daß jedes Funktional in $L_p(\mathfrak{E}_\infty)$ durch (75) mit $\varphi(x) \in L_{p'}(\mathfrak{E}_\infty)$ dargestellt werden kann. Dazu betrachten wir diejenigen Funktionen $\varphi(x) \in L_p(\mathfrak{E}_\infty)$, welche außerhalb des Intervalls Δ_m ($-m \leq x_k \leq m$; $k=2, 1, \dots, n$) gleich 0 sind; sie bilden den Raum $L_p(\Delta_m)$. Für diese Funktionen ist ein Funktional $l(\varphi)$ auf $L_p(\mathfrak{E}_\infty)$ auch ein Funktional auf $L_p(\Delta_m)$, und seine allgemeine Form ist

$$l_m(\varphi) = \int_{\mathfrak{E}_\infty} \psi_m(x) \varphi(x) dx$$

mit $\psi_m(x) \in L_{p'}(\Delta_m)$, wobei $\|\psi_m(x)\|_{L_{p'}(\Delta_m)} \leq \|l\|$ gilt. Aus dem oben Gesagten folgt unmittelbar, daß $\psi_{m+k}(x)$ und $\psi_m(x)$ auf Δ_m für $k > 0$ äquivalent sind. Somit erhalten wir eine Funktion $\varphi(x) \in L_{p'}(\mathfrak{E}_\infty)$, die auf Δ_m der Funktion $\psi_m(x)$ äquivalent ist, und es gilt

$$l(\varphi) = \int_{\mathfrak{E}_\infty} \varphi(x) \varphi(x) dx.$$

Da die finiten Funktionen in $L_p(\mathfrak{E}_\infty)$ überall dicht sind, gelten alle oben für $L_p(\mathfrak{E}_0)$ angestellten Überlegungen auch für $L_p(\mathfrak{E}_\infty)$. Die erhaltenen Resultate lassen sich auch auf den komplexen Raum $L_p(\mathfrak{E}_0)$ ausdehnen, wobei die Funktionale komplexe Werte annehmen können. Aus den obigen Überlegungen folgt sofort, daß der Raum $L_p^*(\mathfrak{E}_0)$ mit $L_{p'}(\mathfrak{E}_0)$ identifiziert werden kann und somit $L_p^{**}(\mathfrak{E}_0)$, also $L_p^*(\mathfrak{E}_0)$, mit $L_p(\mathfrak{E}_0)$ übereinstimmt. Mit anderen Worten, die rechte Seite von (75) liefert für ein festes $\varphi(x) \in L_p(\mathfrak{E}_0)$ die allgemeine Form eines linearen Funktionals in $L_{p'}(\mathfrak{E}_0)$ mit der Norm $\|\varphi\|_{L_p(\mathfrak{E}_0)}$. Der Raum $L_p(\mathfrak{E}_0)$ ist demzufolge regulär. Da $L_p(\mathfrak{E}_0) = L_p^*(\mathfrak{E}_0)$ gilt und $L_{p'}(\mathfrak{E}_0)$ separabel ist, können wir behaupten, daß jede Sphäre in $L_p(\mathfrak{E}_0)$ (oder jede beschränkte Menge) schwach kompakt ist. Für $p=2$ gilt $p'=2$, d. h., es ist $L_2^*(\mathfrak{E}_0) = L_2(\mathfrak{E}_0)$. Mit diesem Fall werden wir uns in Kapitel V eingehend beschäftigen. Alle diese Überlegungen gelten auch für $L_p(\mathfrak{E}_\infty)$.

Mit Hilfe der Formel für ein lineares Funktional in $L_p(\mathfrak{E}_0)$, $p > 1$, beweisen wir den folgenden

Satz. Ist $\psi(x)$ eine auf der beschränkten meßbaren Menge $\mathfrak{E}_0$ meßbare Funktion und ist das Produkt $\psi(x) \varphi(x)$ auf $\mathfrak{E}_0$ für jedes $\varphi(x) \in L_p(\mathfrak{E}_0)$, $p > 1$, summierbar, so gehört $\psi(x)$ zu $L_p(\mathfrak{E}_0)$.

Beweis. Aus der Bedingung des Satzes ergibt sich unmittelbar, daß $\psi(x)$ nur auf einer Menge vom Maß 0 unendlich sein kann, und wir können voraussetzen, daß $\psi(x)$ nur endlicher Werte fähig ist. Wir bestimmen die Folge der Funktionen

$$\psi_n(x) = \begin{cases} \psi(x) & \text{für } |\psi(x)| \leq n, \\ n & \text{für } |\psi(x)| > n, \end{cases} \quad (84)$$

die in jedem Punkt x gegen $\psi(x)$ strebt. Ist $\varphi(x)$ eine beliebige Funktion aus $L_p(\mathfrak{E}_0)$, so gilt $|\psi_n(x) \varphi(x)| \leq |\psi(x) \varphi(x)|$, wobei das Produkt $\psi(x) \varphi(x)$ nach Voraussetzung auf $\mathfrak{E}_0$ summierbar ist. Daraus folgt

$$\lim_{n \rightarrow \infty} \int_{\mathfrak{E}_0} \psi_n(x) \varphi(x) dx = \int_{\mathfrak{E}_0} \psi(x) \varphi(x) dx.$$

Die $\psi_n(x)$ liegen als beschränkte Funktionen in $L_{p'}(\mathbb{E}_0)$, und die links stehenden Integrale sind lineare Funktionale von $\varphi(x)$ in $L_p(\mathbb{E}_0)$. Da sie auf jedem Element $\varphi(x) \in L_p(\mathbb{E}_0)$ einen Grenzwert besitzen, sind ihre Normen durch eine Zahl A beschränkt,

$$\int_{\mathbb{E}_0} |\psi_n(x)|^{p'} dx \leq A^{p'}$$

(vgl. [100]); daraus folgt für $n \rightarrow \infty$ [54]

$$\int_{\mathbb{E}_0} |\psi(x)|^{p'} dx \leq A^{p'},$$

was zu beweisen war.

Der Fall $p=1$ stellt eine Besonderheit dar. Es läßt sich nämlich zeigen, daß der Raum L_1^* dem Raum M (dem Raum der meßbaren beschränkten Funktionen) isometrisch und der Raum L_1 nicht regulär ist.

3. Wir untersuchen jetzt lineare Funktionale im Raum $l_p(p>1)$. Es sei $l(x)$ ein solches Funktional. Jedem Element $x(\xi_1, \xi_2, \dots)$ aus l_p entspricht ein Abschnittelement $x_n(\xi_1, \xi_2, \dots, \xi_n, 0, 0, \dots)$, und da die Reihe mit dem allgemeinen Glied $|\xi_k|^p$ konvergiert, gilt $x_n \rightarrow x$. Wir führen Elemente $y_k(\xi_1^{(k)}, \xi_2^{(k)}, \dots)$, $k=1, 2, \dots$, ein, so daß $\xi_i^{(k)}=0$ für $i \neq k$ und $\xi_i^{(i)}=1$ ist, und setzen $l(y_k)=a_k$. Auf Grund der Distributivität von $l(x)$ gilt $l(x_n)=a_1\xi_1+a_2\xi_2+\dots+a_n\xi_n$ oder, da $l(x)$ stetig ist,

$$l(x)=a_1\xi_1+a_2\xi_2+\dots \quad (85)$$

Wir untersuchen nun die Zahlen a_k . Aus l_p wählen wir die Elemente $z_N(\eta_1^{(N)}, \eta_2^{(N)}, \dots)$, wobei

$$\eta_k^{(N)} = \begin{cases} |a_k|^{p'-1} \operatorname{sgn} a_k & \text{für } k \leq N, \\ 0 & \text{für } k > N \end{cases}$$

ist. Dann gilt $l(z_N) = \sum_{k=1}^N |a_k|^{p'}$ und

$$\sum_{k=1}^N |a_k|^{p'} = l(z_N) \leq \|l\| \cdot \|z_N\| = \|l\| \left[\sum_{k=1}^N |a_k|^{p'} \right]^{1/p},$$

woraus sich

$$\left[\sum_{k=1}^N |a_k|^{p'} \right]^{1/p'} \leq \|l\|$$

oder für $N \rightarrow \infty$

$$\left[\sum_{k=1}^{\infty} |a_k|^{p'} \right]^{1/p'} \leq \|l\|$$

ergibt, d. h., $v(a_1, a_2, \dots)$ gehört zu $l_{p'}$, und es ist

$$\|v\|_{l_{p'}} \leq \|l\|. \quad (86)$$

Wenden wir die Höldersche Ungleichung auf die Summe (85) an, so ergibt sich $\|l\| \leq \|v\|_{l_{p'}}$ oder wegen (86)

$$\|l\| = \|v\|_{l_{p'}}. \quad (87)$$

Genau wie für den Raum L_p können wir auch hier zeigen, daß die Formel (85), wobei $v(a_1, a_2, \dots)$ ein beliebiges Element aus $l_{p'}$ ist, die allgemeine Form eines

linearen Funktional in l_p angibt; das Element v ist dabei eindeutig durch das Funktional $l(x)$ definiert, und es gilt (87). Daraus ergibt sich, daß l_p^* mit l_p übereinstimmt und l_p ein regulärer Raum ist.

Ein dem obigen Satz analoger Satz lautet: *Konvergiert die Reihe $\sum_{k=1}^{\infty} b_k a_k$, mit festen b_k für jedes $(a_1, a_2, \dots) \in l_p$ ($p > 1$), so gehört $(b_1, b_2, \dots)$ zu l_p .*

103. Schwache Konvergenz in den Räumen C , L_p und l_p .

1. Die schwache Konvergenz von Elementen $f_n(x) \in C$ gegen ein Element $f(x) \in C$ (auf einem endlichen Intervall $[a, b]$) läßt sich durch

$$\lim_{n \rightarrow \infty} \int_a^b f_n(x) dg(x) = \int_a^b f(x) dg(x) \quad (88)$$

für jede Funktion $g(x)$ von beschränkter Variation definieren. Notwendige und hinreichende Bedingungen für diese Konvergenz sind: a) Es existiert eine Zahl $c > 0$ derart, daß $|f_n(x)| \leq c$ ist ($n = 1, 2, \dots$); b) es ist $f_n(x) \rightarrow f(x)$ für jedes $x \in [a, b]$.

Die Bedingung a) folgt unmittelbar aus [101]. Ist ferner $x = x_0$ ein beliebiger fester Wert aus $[a, b]$ und $f(x)$ ein beliebiges Element aus C , so gibt $l_0(f) = f(x_0)$ offenbar ein lineares Funktional über C an; wegen $f_n(x) \rightarrow f(x)$ muß $l_0(f_n) \rightarrow l_0(f)$, also $f_n(x_0) \rightarrow f(x_0)$ gelten. Wir müssen jetzt noch zeigen, daß (88) aus a) und b) folgt für eine beliebige Funktion $g(x)$ von beschränkter Variation. Auf Grund von $|f_n(x)| \leq c$ und $f_n(x) \rightarrow f(x)$ ist der Grenzübergang in (88) zulässig, wenn das Integral als Lebesgue-Stieltjessches Integral aufgefaßt wird [54]. Wegen der Stetigkeit von $f(x)$ und $f_n(x)$ lassen sich diese Integrale auch als gewöhnliche Stieltjessche Integrale auffassen.

Nach einem Satz aus [101] existiert bei Berücksichtigung der Bedingungen a) und b) eine Folge von Linearkombinationen der $f_n(x)$, die auf $[a, b]$ gleichmäßig gegen $f(x)$ strebt. Die Funktionen $f_n(x)$ und $f(x)$ werden wie oben als stetig vorausgesetzt.

Wir weisen noch darauf hin, daß ein nichtregulärer Raum C nicht schwach vollständig sein kann. Dies entspricht der Tatsache, daß der Grenzwert der Folge der stetigen Funktionen $f_n(x)$, die in jedem Punkt von $[a, b]$ konvergieren und in ihrer Gesamtheit beschränkt sind ($|f_n(x)| \leq c$), nicht eine stetige Funktion zu sein braucht.

2. Die schwache Konvergenz von Elementen $\varphi_n(x) \in L_p(\mathfrak{E}_0)$, $p > 1$, gegen ein Element $\varphi(x) \in L_p(\mathfrak{E}_0)$ wird durch die Beziehung

$$\lim_{n \rightarrow \infty} \int_{\mathfrak{E}_0} \psi(x) \varphi_n(x) dx = \int_{\mathfrak{E}_0} \psi(x) \varphi(x) dx \quad (89)$$

für jede Funktion $\psi(x) \in L_p(\mathfrak{E}_0)$ definiert. Auf Grund von Satz 3 aus [101] lassen sich die notwendige und hinreichende Bedingung für die schwache Konvergenz in $L_p(\mathfrak{E}_0)$ folgendermaßen formulieren: a) Die Normen von $\varphi_n(x)$ sind beschränkt, d. h., es ist

$$\left[\int_{\mathfrak{E}_0} |\varphi_n(x)|^p dx \right]^{1/p} \leq c; \quad (90)$$

b) die Formel (89) gilt auf der Menge der Elemente $\psi(x) \in L_p(\mathfrak{E}_0)$, deren lineare Hülle in $L_p(\mathfrak{E}_0)$ überall dicht ist. Beim Erfülltsein von (90) ist z. B. hinreichend, daß (89) für alle charakteristischen Funktionen $\omega_g(x)$ der in $\mathfrak{E}_0$ enthaltenen meßbaren Mengen $\mathfrak{E}$ gilt ($\mathfrak{E}_0$ ist eine beschränkte Menge). Für ein eindimensionales end-

liches oder unendliches Intervall ξ_0 genügt es, daß die Bedingung (90) und die Gleichungen

$$\lim_{n \rightarrow \infty} \int_{\alpha}^{\xi} \varphi_n(x) dx = \int_{\alpha}^{\xi} \varphi(x) dx \quad (91)$$

erfüllt sind; dabei ist α eine beliebige, aber feste Zahl und ξ eine willkürliche Zahl aus diesem Intervall.

3. Die schwache Konvergenz der Elemente $(\xi_1^{(n)}, \xi_2^{(n)}, \dots) \in l_p$ ($p > 1$) gegen ein Element $(\xi_1, \xi_2, \dots) \in l_p$ ist definiert durch

$$\lim_{n \rightarrow \infty} (b_1 \xi_1^{(n)} + b_2 \xi_2^{(n)} + \dots) = b_1 \xi_1 + b_2 \xi_2 + \dots \quad (92)$$

für jedes Element $(b_1, b_2, \dots) \in l_{p'}$. Notwendige und hinreichende Bedingungen für die schwache Konvergenz sind

$$\sum_{k=1}^{\infty} |\xi_k^{(n)}|^p \leq c^p, \quad (93)$$

$$\xi_k^{(n)} \rightarrow \xi_k \quad (k=1, 2, \dots). \quad (94)$$

(93) ist die übliche Bedingung, und die Notwendigkeit von (94) folgt aus (92), wenn $b_i = 0$ für $i \neq k$ und $b_k = 1$ gesetzt wird. Wir nehmen an, die Bedingungen (93) und (94) wären erfüllt. Aus (94) folgt, daß die Bedingung (92) auf den Elementen der Form $(0, \dots, 0, 1, 0, \dots)$, d. h. auf den Einheitsvektoren des Raumes $l_{p'}$, erfüllt ist. Die lineare Hülle der Einheitsvektoren ist in $l_{p'}$ dicht; denn die Abschnittelemente, deren Komponenten von einer gewissen Nummer an gleich 0 sind, sind in $l_{p'}$ dicht [59]. Somit ergibt sich aus [101], daß die Bedingungen (93) und (94) hinreichend sind.

104. Der Raum der linearen Operatoren und die Konvergenz von Operatorfolgen. Früher betrachteten wir den Raum der linearen Funktionale und die Frage nach der Konvergenz von Folgen von Funktionalen (Konvergenz der Norm nach und schwache Konvergenz). Wir wenden uns nun diesen Fragen für lineare Operatoren in einem Banachraum X zu. Es sei Y der Raum aller möglichen linearen Operatoren X , deren Wertebereiche in einem Banachraum X' liegen. Die Addition und die Multiplikation mit einer Zahl ist analog wie für Funktionale durch

$$(A + B)x = Ax + Bx, \quad (cA)x = c(Ax) \quad (95)$$

definiert. Die Norm eines Elements $A \in Y$ führen wir als Norm $\|A\|$ des entsprechenden Operators ein. Wie in [99] läßt sich zeigen, daß Y ein Banachraum ist.

Wir betrachten nun eine Folge von linearen Operatoren A_n ($n=1, 2, \dots$) von X in X' . Auf Grund des oben Gesagten existiert, wenn $\|A_n - A_m\|$ für $n, m \rightarrow \infty$ gegen 0 strebt, ein linearer Operator A , für den $\|A - A_n\| \rightarrow 0$ ist und damit für jedes $x \in X$ die Beziehung $A_n x \rightarrow Ax$ in X' gilt. Die Konvergenz $\|A - A_n\| \rightarrow 0$ heißt *Konvergenz der Norm nach*. Dabei ist $\|A_n\|$ ($n=1, 2, \dots$) beschränkt, was sich aus der Ungleichung $\|A_n\| \leq \|A\| + \|A_n - A\|$ ergibt.

Für die Konvergenz der Norm nach, $\|A - A_n\| \rightarrow 0$, ist die Konvergenz der Norm nach in sich notwendig und hinreichend, d. h. $\|A_n - A_m\| \rightarrow 0$ für $n, m \rightarrow \infty$ (Vollständigkeit des Raumes der Operatoren).

Offenbar gilt die Ungleichung

$$\|Ax - A_n x\| \leq \|A - A_n\| \cdot \|x\|. \quad (96)$$

Gehört x zu einer beschränkten Menge $U \in X$, so gibt es ein d mit $\|x\| \leq d$ für $x \in U$, und aus (96) folgt $\|Ax - A_n x\| \leq \|A - A_n\| d$. Also existiert zu jedem $\varepsilon > 0$ eine (nur von ε und nicht von x abhängige) Zahl N derart, daß $\|Ax - A_n x\| \leq \varepsilon$ für $n \geq N$ und $x \in U$ ist, d. h., die Konvergenz von $A_n x$ gegen Ax ist auf jeder beschränkten Menge U gleichmäßig. Daher spricht man statt von Konvergenz von Operatoren der Norm nach auch von der *gleichmäßigen Konvergenz von Operatoren*. Wir betrachten noch eine andere Konvergenz von Operatoren. Wir nennen eine Folge von Operatoren A_n *stark konvergent* gegen einen linearen Operator A , wenn $A_n x \rightarrow Ax$ in X' für jedes $x \in X$ ist.

Wie in [99] läßt sich beweisen, daß die Folge der Normen $\|A_n\|$ beschränkt ist, wenn $A_n x$ eine in X' für jedes $x \in X$ konvergente Folge bildet. Auch können wir zeigen, daß für die starke Konvergenz einer Folge A_n die Beschränktheit der $\|A_n\|$ ($\|A_n\| \leq c$) und die Konvergenz der $A_n x$ auf einer in X dichten linearen Mannigfaltigkeit hinreichend ist. Wir nehmen nun an, $A_n x$ konvergiere in X' für jedes $x \in X$ gegen einen Grenzwert, den wir mit Ax bezeichnen ($A_n x \rightarrow Ax$ in X'). Der Operator A ist in X distributiv auf Grund der Distributivität der A_n und beschränkt auf Grund der Beschränktheit der $\|A_n\|$, d. h., A ist ein linearer Operator. Ist somit $A_n x$ für jedes $x \in X$ eine in X' konvergente Folge, so konvergiert A_n stark gegen den linearen Operator A . Auf Grund der Vollständigkeit von X' braucht man statt der Konvergenz der Folge $A_n x$ nur ihre Konvergenz in sich zu fordern. Somit ist der Raum der linearen Operatoren nicht nur bezüglich der Konvergenz der Norm nach, sondern auch bezüglich der starken Konvergenz vollständig. Wie wir schon oben erwähnten, zieht die Konvergenz der Norm nach die starke Konvergenz nach sich.

Wir definieren noch eine dritte Art der Konvergenz von Operatoren. Wir sagen, eine Folge linearer Operatoren A_n *konvergiere schwach* gegen einen linearen Operator A , wenn $A_n x \rightarrow Ax$ in X' für jedes $x \in X$ gilt. Die starke Konvergenz von Operatoren zieht offenbar die schwache Konvergenz nach sich. Bei Funktionalen unterscheiden sich starke und schwache Konvergenz nicht voneinander.

Früher definierten wir die Addition linearer Operatoren und die Multiplikation eines linearen Operators mit einer Zahl. Es läßt sich natürlich auch die Multiplikation von Operatoren definieren. Ist A ein linearer Operator von X in X' und B ein linearer Operator von X' in X'' , so ist der durch

$$(BA)x = B(Ax)$$

definierte Operator BA ein linearer Operator von X in X'' . Seine Distributivität folgt aus der von A und B und seine Beschränktheit aus der Ungleichung

$$\|(BA)x\|_{X''} \leq \|B\| \cdot \|Ax\|_{X'} \leq \|B\| \cdot \|A\| \cdot \|x\|_X;$$

also gilt $\|BA\| \leq \|B\| \cdot \|A\|$. Wir können auch das Produkt von mehreren Faktoren bilden. Ist A ein linearer Operator von X in X , so lassen sich ganze positive Potenzen von ihm bilden: $A^2 = A(Ax)$ usw.

Das Produkt zweier Operatoren kann von der Reihenfolge der Faktoren abhängen. Sind etwa A und B lineare Operatoren von X in X , so lassen sich die beiden folgenden Operatoren von X in X bilden:

$$(BA)x = B(Ax) \quad \text{und} \quad (AB)x = A(Bx).$$

Diese Operatoren können voneinander verschieden sein. Analoges gilt auch für mehr als zwei Faktoren.

Die starke Konvergenz von Operatoren wird manchmal einfach Konvergenz genannt; für sie benutzen wir die Schreibweise $A_n \rightarrow A$. Sind A_n und B_n Folgen

linearer Operatoren von X in X' und ist a_n eine Zahlenfolge, so läßt sich zeigen, daß $a_n A_n \rightarrow aA$ und $A_n + B_n \rightarrow A + B$ gilt, wenn $a_n \rightarrow a$, $A_n \rightarrow A$ und $B_n \rightarrow B$ ist; eine analoge Behauptung gilt auch für die Konvergenz der Norm nach. Sind die A_n lineare Operatoren von X in X' und die B_n solche von X' in X'' , so folgt $B_n A_n \rightarrow BA$ aus $A_n \rightarrow A$, $B_n \rightarrow B$ (dasselbe gilt für Konvergenz der Norm nach). Diese letzte Behauptung wollen wir beweisen. Es gilt

$$BAx - B_n A_n x = (B - B_n)(Ax) + B_n(A - A_n)x$$

und

$$\|BAx - B_n A_n x\|_{X''} \leq \|(B - B_n)(Ax)\|_{X''} + \|B_n\| \cdot \|(A - A_n)x\|_{X'}.$$

Der erste Summand strebt wegen $B_n \rightarrow B$, der zweite wegen der Beschränktheit der $\|B_n\|$ und wegen $A_n \rightarrow A$ gegen 0.

105. Adjungierte Operatoren. Es sei A ein linearer Operator von X in X' und $l'(x)$ ein beliebiges Funktional aus X' ($l' \in X'^*$). Die Räume X und X' seien Banachräume.

Wir sehen leicht, daß dann $l'(Ax)$ ein lineares Funktional aus X ist:

$$l'(Ax) = l(x).$$

Für ein gegebenes A ordnet diese Gleichung jedem Element $l' \in X'^*$ ein Element $l \in X^*$ zu; dies läßt sich durch $l = A^* l'$ wiedergeben, wobei der Operator A^* , der den Definitionsbereich X'^* und einen in X^* liegenden Wertebereich hat, als der zu A adjungierte Operator definiert wird. Aus der Distributivität der linearen Funktionale und des Operators A folgt die Distributivität von A^* . Wir zeigen nun, daß A^* beschränkt und $\|A^*\| = \|A\|$ ist; dabei ist die linke Seite die Norm des Operators in X'^* , die rechte in X . Es gilt

$$|l(x)| = |l'(Ax)| \leq \|l'\| \cdot \|Ax\| \leq \|l'\| \cdot \|A\| \cdot \|x\|,$$

woraus sich $\|l\| \leq \|l'\| \cdot \|A\|$ ergibt. Ferner ist $l = A^* l'$ und folglich $\|A^*\| \leq \|A\|$. Ist x_0 ein beliebiges, aber festes Element aus X und l' ein solches Element aus X'^* , daß $\|l'\| = 1$ und $l'(Ax_0) = \|Ax_0\|$ gilt, so erhalten wir

$\|Ax_0\| = l'(Ax_0) = l(x_0) \leq \|l\| \cdot \|x_0\| = \|A^* l'\| \cdot \|x_0\| \leq \|A^*\| \cdot \|l'\| \cdot \|x_0\| = \|A^*\| \cdot \|x_0\|$, also $\|Ax_0\| \leq \|A^*\| \cdot \|x_0\|$ oder $\|A\| \leq \|A^*\|$; daraus ergibt sich zusammen mit $\|A^*\| \leq \|A\|$ die Beziehung $\|A^*\| = \|A\|$. Damit gelangen wir zu dem folgenden

Satz. Der zu einem linearen Operator A von X in X' adjungierte Operator A^* ist ein linearer Operator von X'^* in X^* , und es gilt $\|A^*\| = \|A\|$.

Bemerkung. Stimmt X' mit X überein, so ist A^* ein linearer Operator in X^* , dessen Wertebereich ebenfalls in X^* liegt.

Sind A und B lineare Operatoren von X in X' , so folgt aus der Definition des adjungierten Operators $(A+B)^* = A^* + B^*$. Sind A und B lineare Operatoren von X in X , so gilt $(BA)^* = A^* B^*$, denn es besteht die Gleichung

$$l(BAx) = (B^* l)(Ax) = A^*(B^* l)(x) = (A^* B^* l)(x).$$

In einem reellen Raum gilt $(cA)^* = cA^*$, in einem komplexen Raum $(cA)^* = \bar{c}A^*$.

106. Vollstetige Operatoren. Ein linearer Operator A von X in X' heißt *vollstetig*, wenn er jede in X beschränkte Menge in eine in X' kompakte Menge abbildet. Es ist leicht zu sehen, daß ein distributiver Operator A , der in ganz X definiert ist und jede beschränkte Menge in eine kompakte abbildet, ein beschränkter und dem-

zufolge ein linearer Operator ist. Nach Voraussetzung bildet nämlich A die Menge $\|x\| \leq 1$ in die kompakte Menge Ax ab. Nun ist jede kompakte Menge beschränkt, d. h., es existiert eine positive Zahl c mit $\|Ax\| \leq c$ für $\|x\| \leq 1$, woraus sich die Beschränktheit von A ergibt. Somit läßt sich die Definition eines vollstetigen Operators auch folgendermaßen formulieren: Ein distributiver, auf ganz X definierter Operator A heißt *vollstetig*, wenn er jede beschränkte Menge in eine kompakte abbildet. Ein so definierter vollstetiger Operator ist natürlich auch linear.

Satz 1. *Ist A ein vollstetiger Operator und gilt $x_n \rightarrow x_0$, so ist $Ax_n \rightarrow Ax_0$.*

Beweis. Wir wissen, daß $Ax_n \rightarrow Ax_0$ gilt, wenn A ein linearer Operator ist. Nach Voraussetzung ist $x_n \rightarrow x_0$, so daß die Folge der Zahlen $\|x_n\|$ beschränkt ist. Auf Grund der Vollstetigkeit von A können wir aus der Folge der Ax_n eine Teilfolge auswählen, die gegen ein gewisses Element $y_0 \in X'$ stark konvergiert. Andererseits folgt aus dem oben Gesagten, daß diese Folge schwach gegen Ax_0 konvergiert; also ist $y_0 = Ax_0$. Somit konvergiert jede stark konvergente Teilfolge von Ax_n stark gegen Ax_0 . Den Beweis, daß die Folge der Ax_n stark gegen Ax_0 konvergiert, führen wir indirekt. Wir nehmen an, es existiere eine Zahl $\delta > 0$ und eine solche unendliche Teilfolge Ax_{n_k} , daß $\|Ax_{n_k} - Ax_0\| \geq \delta$ gilt. Aus der Folge der Ax_{n_k} können wir wiederum eine stark konvergente Teilfolge auswählen, die auf Grund des oben Gesagten stark gegen Ax_0 konvergieren muß. Dies widerspricht aber der Ungleichung $\|Ax_{n_k} - Ax_0\| \geq \delta > 0$. Damit ist der Satz bewiesen.

Satz 2. *Es sei A_m ($m=1, 2, \dots$) eine Folge vollstetiger Operatoren, die der Norm nach gegen einen linearen Operator A konvergiert ($\|A - A_m\| \rightarrow 0$). Dann ist auch der Operator A vollstetig.*

Beweis. Wir müssen zeigen, daß A jede beschränkte Folge von Elementen $x_n \in X$ ($n=1, 2, \dots$) in eine kompakte überführt. Die Folge A_mx_n ist für ein beliebiges festes m kompakt. Andererseits folgt aus der Konvergenz der Norm nach die gleichmäßige Konvergenz auf jeder beschränkten Menge. Somit existiert für ein gegebenes $\varepsilon > 0$ ein m , für welches $\|Ax_n - A_mx_n\| \leq \varepsilon$ ($n=1, 2, \dots$) gilt, d. h., Ax_n hat ein kompaktes ε -Netz A_mx_n ; daraus folgt die Kompaktheit von Ax_n .

Es läßt sich zeigen, daß dann, wenn A ein vollstetiger Operator ist, auch der Operator A^* , dessen Definitionsbereich in X'^* und dessen Wertebereich in X^* liegt, vollstetig ist.

Wir werden diesen Satz in Kapitel V für Operatoren im Hilbertraum beweisen. Für diesen Raum werden wir die Theorie der vollstetigen Operatoren eingehender darlegen.

107. Operatorengleichungen. Wir werden jetzt voraussetzen, daß der Wertebereich der im Raum X definierten linearen Operatoren ebenfalls zu X gehört (X' stimme also mit X überein). Dann liegt der Wertebereich des in X^* definierten Operators A^* ebenfalls in X^* . Mit E bezeichnen wir wie früher den Einheitsoperator in einem beliebigen Raum X vom Typ B , d. h., es sei $Ex = x$ für jedes $x \in X$. Wir erinnern noch daran, daß wir unter dem Nulloperator den linearen Operator verstanden, der jedes Element x in das Nullelement überführt. Seine Norm ist gleich 0, während sie für jeden anderen Operator positiv ist.

Wir betrachten die Gleichung

$$(A - E)x = y, \quad (97)$$

in der y das gegebene und x das gesuchte Element aus X ist. Wir können (97) auch

in der Gestalt

$$x = Ax - y \quad (98)$$

schreiben. Die rechte Seite von (98) ist ein Operator von X in X . Setzen wir $Ax - y = Bx$ (B ist kein linearer Operator), so haben wir

$$Bx_1 - Bx_2 = Ax_1 - Ax_2,$$

woraus sich

$$\|Bx_1 - Bx_2\| \leq \|A\| \cdot \|x_1 - x_2\|$$

ergibt. Ist $\|A\| < 1$, so läßt sich auf (98) das Prinzip der kontrahierenden Abbildungen anwenden. Hieraus erhalten wir den folgenden

Satz 1. *Ist $\|A\| < 1$, so hat (97) für jedes vorgegebene $y \in X$ eine eindeutig bestimmte Lösung, die auch mit Hilfe der Methode der sukzessiven Approximationen aus (98) bei einer beliebigen Anfangsnäherung erhalten werden kann.*

Im folgenden werden wir es oft mit einer Gleichung zu tun haben, die einen Parameter enthält:

$$(A - \lambda E)x = y. \quad (99)$$

Ist X ein reeller Banachraum, so ist λ eine reelle Zahl. Für komplexe Räume kann λ auch komplex sein. Nehmen wir $\lambda \neq 0$ an, so läßt sich (99) auch in der Form

$\left(\frac{1}{\lambda}A - E\right)x = \frac{1}{\lambda}y$ schreiben. Aus dem obigen Satz folgt, daß (99) für jedes y eine eindeutig bestimmte Lösung besitzt, wenn $|\lambda| > \|A\|$ ist; diese Lösung läßt sich auch mit Hilfe der Methode der sukzessiven Approximationen aus der Gleichung $x = \frac{1}{\lambda}Ax - \frac{1}{\lambda}y$ errechnen.

Die Gleichung (97) wird manchmal auch in der Gestalt

$$(\lambda A - E)x = y \quad \text{oder} \quad (E - \lambda A)x = y$$

angegeben. Dann muß die Bedingung $|\lambda| > \|A\|$ durch $|\lambda| < \|A\|^{-1}$ ersetzt werden.

Wir nehmen nun an, daß A ein *vollstetiger Operator* ist, und schreiben die beiden Gleichungen

$$(A - E)x = y, \quad (100)$$

$$(A^* - E)x^* = y^* \quad (101)$$

auf, wobei (100) in X und (101) in X^* gilt und y, y^* die gegebenen, x, x^* die gesuchten Elemente sind. Die zu (100) und (101) gehörigen homogenen Gleichungen lauten

$$(A - E)x = \theta, \quad (102)$$

$$(A^* - E)x^* = \theta^*; \quad (103)$$

θ bzw. θ^* ist das Nullelement von X bzw. X^* . Die Lösungen dieser Gleichungen bilden lineare Mannigfaltigkeiten.

Wir wollen nun einige Resultate formulieren, die diese Gleichungen betreffen. Ihr Beweis ist in Kapitel V für den Fall eines Hilbertraumes angegeben.

Hat eine der Gleichungen (100) oder (101) bei beliebiger rechter Seite eine Lösung, so hat auch die andere Gleichung diese Eigenschaft. Dann hat jede der Gleichungen

bei beliebiger rechter Seite eine eindeutig bestimmte Lösung, so daß die Gleichungen (102) und (103) nur die trivialen Lösungen $x = \theta$ bzw. $x^* = \theta^*$ besitzen.

Hat eine der Gleichungen (102) oder (103) eine von 0 verschiedene Lösung, so hat auch die andere Gleichung diese Eigenschaft. Es gibt endlich viele linear unabhängige Lösungen von (102) und (103), und zwar gleich viele [IV/1, 9]. Die linearen Mannigfaltigkeiten der Lösungen seien endlichdimensionale Unterräume. Dann ist (100) bzw. (101) genau dann lösbar, wenn y bzw. y^* zu allen Lösungen von (103) bzw. (102) orthogonal ist, d. h. $(y, x^*) = 0$ bzw. $(y^*, x) = 0$ gilt.

Diese Resultate behalten natürlich ihre Gültigkeit auch für Gleichungen mit einem Parameter. Im Fall eines komplexen Banachraumes lassen sich die Gleichungen (100) bis (103) wie folgt schreiben:

$$(A - \lambda E)x = y, \quad (104)$$

$$(A^* - \bar{\lambda} E)x^* = y^*, \quad (105)$$

$$(A - \lambda E)x = \theta, \quad (106)$$

$$(A^* - \bar{\lambda} E)x^* = \theta^*. \quad (107)$$

Wir geben noch ein Resultat an: Es existieren nur endlich viele Werte λ mit $|\lambda| \leq R$ ($R > 0$ beliebig gegeben), für welche (106) und folglich auch (107) Lösungen besitzen, die von der trivialen Lösung $x = \theta$ (bzw. $x^* = \theta^*$) verschieden sind.

Diese Resultate sind den Sätzen vollständig analog, die wir aus der Theorie der Integralgleichungen kennen.

Ist X ein reeller Banachraum, so muß λ als reell angenommen werden ($\lambda = \bar{\lambda}$).

Die Zahl λ , für welche (106) eine nichttriviale Lösung besitzt, nennen wir einen *Eigenwert* des Operators A ; die Anzahl der linear unabhängigen Lösungen von (106) gibt den *Rang* oder die *Vielfachheit* des entsprechenden Eigenwertes an. Die Lösungen $x_1, x_2, \dots, x_m$ bilden ein Fundamentalsystem linear unabhängiger Lösungen von (106), wenn $x = c_1x_1 + c_2x_2 + \dots + c_mx_m$ mit beliebigen Zahlen c_k die allgemeine Form der Lösungen von (106) ist. Die Darstellung von x ist eindeutig auf Grund der linearen Unabhängigkeit der x_k . Linear unabhängige Lösungen lassen sich auf verschiedene Arten wählen, aber in einem Fundamentalsystem linear unabhängiger Lösungen ist ihre Anzahl stets dieselbe. Hat (106) eine nichttriviale Lösung und genügt das in (104) auftretende y der obenerwähnten Lösbarkeitsbedingung, so lassen sich alle Lösungen von (104) durch

$$x = x_0 + c_1x_1 + c_2x_2 + \dots + c_mx_m \quad (108)$$

darstellen, wobei x_0 eine beliebige Lösung von (104), $x_1, x_2, \dots, x_m$ ein Fundamentalsystem linear unabhängiger Lösungen von (106) ist und die c_k beliebige Zahlen sind. Dies ergibt sich unmittelbar aus der Linearität von (104) und (106) [IV/1, 9, 10].

108. Vollstetige Operatoren in den Räumen C , L_p und l_p .

1. Im Raum C betrachten wir den Integraloperator

$$\varphi(x) = \int_a^b K(x, t) \psi(t) dt; \quad (109)$$

dabei sei $[a, b]$ ein endliches Intervall. Ist der Kern $K(x, t)$ im Quadrat $Q(a \leq x \leq b; a \leq y \leq b)$ stetig, so liefert (109) offenbar einen distributiven Operator von C in C . Seine Beschränktheit folgt unmittelbar aus der Ungleichung

$$\max_{a \leq x \leq b} |\varphi(x)| \leq \max_{a \leq t \leq b} |\psi(t)| \max_{a \leq x \leq b} \int_a^b |K(x, t)| dt. \quad (110)$$

Ist U eine beschränkte Menge von Funktionen $\varphi(t) \in C$, für die also $\max_{a \leq t \leq b} |\varphi(t)| \leq A$ gilt, so ist die Menge der entsprechenden $\varphi(x)$ kompakt. Die Beschränktheit folgt auf Grund von $\max_{a \leq t \leq b} |\varphi(t)| \leq A$ aus (110) und die gleichgradige Stetigkeit aus der Ungleichung

$$|\varphi(x_2) - \varphi(x_1)| \leq A \int_a^b |K(x_2, t) - K(x_1, t)| dt. \quad (111)$$

Bei Stetigkeit des Kerns in Q ist also (109) in C vollstetig.

Die Norm des Operators (109) ist, wie man beweisen kann, gleich

$$\max_{a \leq x \leq b} \int_a^b |K(x, t)| dt.$$

Der Operator (109) ist auch bei schwächeren Voraussetzungen über den Kern in C vollstetig. Nehmen wir z. B. an, $K(x, t)$ sei eine beschränkte meßbare Funktion in Q , und es gelte

$$\lim_{x' \rightarrow x} K(x', t) = K(x, t) \quad (112)$$

für jedes $x \in [a, b]$ und fast alle t , dann ist [54]

$$\lim_{x' \rightarrow x} \int_a^b |K(x', t) - K(x, t)| dt = 0,$$

und für jedes $\varepsilon > 0$ existiert ein $\eta > 0$ derart, daß

$$\int_a^b |K(x_2, t) - K(x_1, t)| dt \leq \varepsilon \quad \text{für} \quad |x_2 - x_1| \leq \eta$$

gilt. Dies läßt sich wie die gleichmäßige Stetigkeit der Funktionen nachweisen, die auf einem endlichen abgeschlossenen Intervall stetig sind [I, 43].

Die Beschränktheit und gleichgradige Stetigkeit von $\varphi(x)$ können wir wie oben beweisen.

Im folgenden werden wir uns mit Integraloperatoren mit polaren Kernen eingehend beschäftigen.

2. Wir betrachten nun den Operator (109) im Raum L_p ($p > 1$) und setzen voraus, daß $K(x, t)$ zu $L_{p'}(Q)$ gehört, d. h., es sei

$$\int_a^b \int_a^b |K(x, t)|^{p'} dx dt = A^{p'} < \infty. \quad (113)$$

Für jede Funktion $\varphi(x) \in L_p[a, b]$ ist das Integral (109) sinnvoll. Es läßt sich leicht zeigen, daß es eine meßbare Funktion $\varphi(x)$ definiert (vgl. [68]). Auf Grund der Hölderschen Ungleichung gilt

$$|\varphi(x)| \leq \left[\int_a^b |K(x, t)|^{p'} dt \right]^{1/p'} \left[\int_a^b |\varphi(x)|^p dx \right]^{1/p}.$$

Erheben wir beide Seiten dieser Ungleichung in die p' -te Potenz und integrieren wir nach x , so erhalten wir

$$\|\varphi\|_{L_p}^{p'} \leq A^{p'} \|\varphi\|_{L_p}^{p'}, \quad \text{also} \quad \|\varphi\|_{L_{p'}} \leq A \|\varphi\|_{L_p}, \quad (114)$$

d. h., (109) ist ein linearer Operator von L_p in $L_{p'}$. Wir könnten auch zeigen, daß A die Norm von (109) ist.

Wir wollen nun nachweisen, daß der Operator (109) vollstetig ist. Es sei U die in L_p beschränkte Menge der Funktionen $\psi(x)$ und V die Menge der entsprechenden Funktionen $\varphi(x) \in L_{p'}$. Zu zeigen ist die Kompaktheit von V . Nach Voraussetzung ist $\|\psi\|_{L_p} \leq c$ für $\psi(x) \in U$, und aus (114) folgt damit $\|\varphi\|_{L_{p'}} \leq Ac$. Wir müssen noch beweisen, daß $\varphi(x)$ gleichgradig stetig im Mittel ist. Wir setzen $\varphi(x)$ ins Äußere von $[a, b]$ und $K(x, t)$ ins Äußere von Q durch die Funktion 0 fort und können damit

$$\varphi(x+h) - \varphi(x) = \int_a^b [K(x+h, t) - K(x, t)] \psi(t) dt$$

schreiben, woraus sich wie oben

$$\|\varphi(x+h) - \varphi(x)\|_{L_{p'}} \leq \left[\int_a^b \int_a^b |K(x+h, t) - K(x, t)|^{p'} dx dt \right]^{1/p'} \|\psi\|_{L_p},$$

also

$$\|\varphi(x+h) - \varphi(x)\|_{L_{p'}} \leq \left[\int_a^b \int_a^b |K(x+h, t) - K(x, t)|^{p'} dx dt \right]^{1/p'} c \quad (115)$$

ergibt. Da $K(x, t)$ in $L_{p'}(Q)$ im Mittel stetig ist, existiert zu jedem $\varepsilon > 0$ ein $\eta > 0$ derart, daß

$$\int_a^b \int_a^b |K(x+h, t) - K(x, t)|^{p'} dx dt \leq \frac{\varepsilon^{p'}}{c^{p'}} \quad \text{für } |h| \leq \eta$$

gilt, und aus (115) folgt

$$\|\varphi(x+h) - \varphi(x)\|_{L_{p'}} \leq \varepsilon \quad \text{für } |h| \leq \eta,$$

wobei η für alle $\varphi(x) \in V$ ein und dieselbe Zahl ist. Damit ist die Behauptung bewiesen.

Der Beweis für den Fall eines unendlichen Intervalls und mehrerer unabhängiger Veränderlicher läßt sich analog führen.

3. Wir betrachten nun den Operator von l_p in $l_{p'} (p > 1)$, der durch die Formeln

$$\eta_i = a_{i1}\xi_1 + a_{i2}\xi_2 + \dots \quad (116)$$

unter der Bedingung

$$\sum_{i,k=1}^{\infty} |a_{ik}|^{p'} = A^{p'} < \infty \quad (117)$$

gegeben ist. Für die Elemente führen wir die Bezeichnungen $x(\xi_1, \xi_2, \dots)$ und $y(\eta_1, \eta_2, \dots)$ ein und wenden die Höldersche Ungleichung für Summen an; dann ergibt sich wie oben

$$\|y\|_{l_{p'}} \leq A \|x\|_{l_p}, \quad (118)$$

so daß der Operator (116) ein linearer Operator von l_p in $l_{p'}$ ist. Er ist auch vollstetig. Zum Beweis bezeichnen wir wieder mit U eine beschränkte Menge von Elementen $x \in l_p$ ($\|x\| \leq c$) und mit V die entsprechende Menge von Elementen $y \in l_{p'}$. Zu zeigen ist die Kompaktheit von V . Daß V beschränkt ist, folgt aus (118); wir brauchen nur noch nachzuweisen, daß es zu jedem $\varepsilon > 0$ eine ganze positive Zahl n_ε gibt, für die

$$\sum_{i=n_\varepsilon}^{\infty} |\eta_i|^{p'} \leq \varepsilon^{p'} \quad (119)$$

gilt. Aus (116) und der Hölderschen Ungleichung ergibt sich

$$\sum_{i=n_e}^{\infty} |\eta_i|^{p'} \leq \sum_{i=n_e}^{\infty} \sum_{k=1}^{\infty} |a_{ik}|^{p'} \|x\|_{l_p}^{p'} \leq \sum_{i=n_e}^{\infty} \sum_{k=1}^{\infty} |a_{ik}|^{p'} c^{p'}. \quad (120)$$

Die Doppelreihe mit dem allgemeinen Glied $|a_{ik}|^{p'}$ konvergiert auf Grund von (117); also gibt es ein n_e , für das

$$\sum_{i=n_e}^{\infty} \sum_{k=1}^{\infty} |a_{ik}|^{p'} \leq \frac{c^{p'}}{c^{p'}} \quad (121)$$

gilt. Hieraus folgt mit Hilfe von (120) die Ungleichung (119).

109. Verallgemeinerte Ableitungen. Wir führen jetzt einen neuen Begriff einer Ableitung ein, der in der modernen mathematischen Physik häufig benutzt wird. Es sei D ein beschränktes Gebiet des n -dimensionalen euklidischen Raumes R_n , dessen Punkte x durch kartesische Koordinaten $(x_1, x_2, \dots, x_n)$ bestimmt werden. Unter einem *Gebiet* wollen wir stets eine offene zusammenhängende Menge verstehen, und wir werden annehmen, daß die Ränder der Gebiete, von denen wir reden werden, das räumliche Maß 0 haben. Mit $\bar{D}$ bezeichnen wir, wie üblich, das Gebiet D nebst seinem Rand (d. h. das *abgeschlossene Gebiet* oder den *Bereich*). Wir sagen, ein Gebiet D' liege echt in D , wenn $D' \subset D$ ist und der Abstand zwischen D' und dem Rand von D positiv ist. Damit gleichwertig ist die Schreibweise $\bar{D}' \subset D$. Wie früher nennen wir eine Funktion *finit* in D , wenn sie außerhalb eines in D echt enthaltenen Gebietes D' gleich 0 ist (D' kann für verschiedene Funktionen verschieden sein). Wir nehmen nun an, die Funktionen $\varphi(x)$ und $\psi(x)$ haben stetige Ableitungen bis einschließlich l -ter Ordnung in D , und $\psi(x)$ sei finit. Wir betrachten dann eine beliebige Ableitung l -ter Ordnung

$$D^l \varphi = \frac{\partial^l \varphi}{\partial x_1^{l_1} \partial x_2^{l_2} \dots \partial x_n^{l_n}} \quad (l_1 + l_2 + \dots + l_n = l). \quad (122)$$

Integrieren wir partiell, so erhalten wir, da $\psi(x)$ finit ist,

$$\int_D D^l \varphi(x) \psi(x) dx = (-1)^l \int_D \varphi(x) D^l \psi(x) dx. \quad (123)$$

Die Formel (123) kann der Definition der verallgemeinerten Ableitung zugrunde gelegt werden.

Definition 1. Die Funktionen $\varphi(x)$ und $\chi(x)$ seien in jedem echt in D enthaltenen Teilgebiet D' summierbar und mögen für jede finite, l -mal stetig differenzierbare Funktion $\psi(x)$ die Beziehung

$$\int_D \chi(x) \psi(x) dx = (-1)^l \int_D \varphi(x) D^l \psi(x) dx \quad (124)$$

erfüllen. Dann heißt die Funktion $\chi(x)$ die *verallgemeinerte Ableitung* der Gestalt (122) der Funktion $\varphi(x)$ in D .

Wir wollen uns davon überzeugen, daß eine *gegebene Funktion* $\varphi(x)$ nur eine *verallgemeinerte Ableitung der Gestalt* (122) *haben kann*. Es seien $\chi(x)$ und $\chi_1(x)$ zwei verallgemeinerte Ableitungen, und außer für $\varphi(x)$ und $\chi(x)$ gelte auch für $\varphi(x)$ und $\chi_1(x)$ die Beziehung (124). Subtrahieren wir gliedweise, so erhalten wir

$$\int_D [\chi(x) - \chi_1(x)] \psi(x) dx = 0; \quad (125)$$

da die finite Funktion $\psi(x)$ beliebig war, folgt aus (125), daß $\chi(x)$ und $\chi_1(x)$ in D äquivalente Funktionen sind [51].

Hat $\varphi(x)$ im Innern von D stetige Ableitungen einschließlich l -ter Ordnung, so gilt (123), und es ist $\chi(x) = D^l \varphi(x)$. Im folgenden werden wir die Bezeichnung (122) auch für verallgemeinerte Ableitungen beibehalten.

Wir erwähnen noch einige Eigenschaften der verallgemeinerten Ableitung, die unmittelbar aus ihrer Definition hervorgehen. Die verallgemeinerte Ableitung $D^l \varphi(x)$ hängt nicht davon ab, in welcher Reihenfolge differenziert wird, denn in (124) können wir die Reihenfolge der Differentiation bei der Funktion $\psi(x)$, die stetige Ableitungen besitzt, beliebig ändern. Haben $\varphi_1(x)$ und $\varphi_2(x)$ verallgemeinerte Ableitungen $\chi_1(x)$ bzw. $\chi_2(x)$ der Form (122), so hat $c_1 \varphi_1(x) + c_2 \varphi_2(x)$ die verallgemeinerte Ableitung $c_1 \chi_1(x) + c_2 \chi_2(x)$ derselben Form (c_1, c_2 konstant). Ist die Funktion $\chi(x)$ die verallgemeinerte Ableitung von $\varphi(x)$ in D , so ist sie auch verallgemeinerte Ableitung derselben Gestalt in jedem in D enthaltenen Gebiet D' .

Hat $\varphi(x)$ die verallgemeinerte Ableitung $\frac{\partial \varphi(x)}{\partial x_1} = \chi(x)$ und $\chi(x)$ die verallgemeinerte Ableitung $\frac{\partial \chi(x)}{\partial x_2}$, so hat $\varphi(x)$ die verallgemeinerte Ableitung $\frac{\partial^2 \varphi(x)}{\partial x_1 \partial x_2} = \frac{\partial \chi(x)}{\partial x_2}$.

Hat ferner $\varphi(x)$ die verallgemeinerten Ableitungen $\frac{\partial \varphi(x)}{\partial x_2}$ und $\frac{\partial^2 \varphi(x)}{\partial x_1 \partial x_2}$, so ist $\frac{\partial^2 \varphi(x)}{\partial x_1 \partial x_2}$ die verallgemeinerte Ableitung von $\frac{\partial \varphi(x)}{\partial x_2}$ nach x_1 . Später werden wir beweisen, daß unter gewissen zusätzlichen Einschränkungen für die Differentiation eines Produkts die übliche Formel

$$\frac{\partial [\varphi_1(x) \varphi_2(x)]}{\partial x_1} = \frac{\partial \varphi_1(x)}{\partial x_1} \varphi_2(x) + \varphi_1(x) \frac{\partial \varphi_2(x)}{\partial x_1} \quad (126)$$

gilt.

Wir stellen nun einen Zusammenhang zwischen der verallgemeinerten Differentiation und der Operation der Mittelung her. Es seien $\omega_1(|x-y|)$ ein beliebiger gemittelter Kern, der von dem Abstand zwischen den Punkten x und y abhängt, und $\varphi_h(x)$ die für $\varphi(x)$ konstruierten Mittelfunktionen:

$$\varphi_h(x) = \frac{1}{h^n} \int \omega_h(|x-y|) \varphi(y) dy. \quad (127)$$

Wir setzen voraus, $\varphi(x)$ habe in D eine verallgemeinerte Ableitung $\chi(x) = D^l \varphi(x)$ der Form (122) und ermitteln die entsprechende (offenbar gewöhnliche) Ableitung der Mittelfunktionen [71]:

$$D_x^l \varphi_h(x) = \frac{1}{h^n} \int \varphi(y) D_x^l \omega_h(|x-y|) dy = \frac{(-1)^l}{h^n} \int \varphi(y) D_y^l \omega_h(|x-y|) dy. \quad (128)$$

Der Punkt $x \in D$ habe vom Rand von D einen Abstand, der größer als h sei. Da die Funktion $\omega_h(|x-y|)$ außerhalb einer Kugel mit dem Radius h und dem Mittelpunkt x verschwindet, können wir sie als die finite Funktion in (124) auffassen. Zusammen mit (128) führt dies auf die Beziehung

$$D_x^l \varphi_h(x) = \frac{1}{h^n} \int \omega_h(|x-y|) D_y^l \varphi(y) dy, \quad (129)$$

die sich folgendermaßen beschreiben läßt: *Die Mittelfunktionen von verallgemeinerten Ableitungen stimmen mit den Ableitungen der Mittelfunktionen in allen Punkten von D überein, deren Abstand vom Rand von D größer ist als der Mittelungsradius h .*

Auf Grund der Eigenschaften der Mittelfunktionen [71] können wir behaupten, daß $\varphi_h(x) \rightarrow \varphi(x)$ und $D^l \varphi_h(x) \rightarrow D^l \varphi(x)$ für $h \rightarrow 0$ in $L(D')$ gilt, wobei D' in D echt enthalten sein soll. Nehmen wir noch zusätzlich an, daß $\varphi(x)$ in einem beliebigen,

in D echt enthaltenen Gebiet D' mit einem Exponenten $p \geq 1$ summierbar ist, desgleichen die verallgemeinerte Ableitung $D^i \varphi(x)$ mit einem Exponenten $q \geq 1$, so konvergieren $\varphi_h(x)$ und $D^i \varphi_h(x)$ in $L_p(D')$ bzw. $L_q(D')$. Wir müssen aber vor folgendem warnen: Ist die Funktion $\varphi(x)$ irgendwie auf ganz R_n definiert (z. B. außerhalb von D durch die Funktion 0), so sind die Funktionen $\varphi_h(x)$ ebenfalls in ganz R_n definiert und konvergieren für $h \rightarrow 0$ gegen $\varphi(x)$ in $L_p(D)$; jedoch konvergiert $D^i \varphi_h(x)$ im allgemeinen nicht gegen $D^i \varphi(x)$ im Raum $L_q(D)$. Dies hängt damit zusammen, daß die so fortgesetzte Funktion $\varphi(x)$ keine entsprechende verallgemeinerte Ableitung im ganzen R_n haben kann.

Wir kehren nun zum Beweis der Formel (126) zurück. Zunächst weisen wir die folgende einfache Behauptung nach: Ist $\varphi_1(x) \in L_p(D')$, $p > 1$, und $\varphi_2(x) \in L_{p'}(D')$, $\frac{1}{p} + \frac{1}{p'} = 1$, in einem beliebigen, in D echt enthaltenen Gebiet D' und ist $\psi(x)$ eine beschränkte und in D finite Funktion, so gilt [62]

$$\int_D \varphi_{1h}(x) \varphi_{2h}(x) \psi(x) dx \rightarrow \int_D \varphi_1(x) \varphi_2(x) \psi(x) dx \quad \text{für } h \rightarrow 0.$$

Benutzen wir nämlich die Höldersche Ungleichung, so finden wir

$$\begin{aligned} & \left| \int_D [\varphi_{1h}(x) \varphi_{2h}(x) - \varphi_1(x) \varphi_2(x)] \psi(x) dx \right| \\ & \leq \int_D |\varphi_{1h}(x)| \cdot |\varphi_{2h}(x) - \varphi_2(x)| \cdot |\psi(x)| dx + \int_D |\varphi_2(x)| \cdot |\varphi_{1h}(x) - \varphi_1(x)| \cdot |\psi(x)| dx \\ & \leq c [\|\varphi_{1h}\|_{L_p(D')} \cdot \|\varphi_{2h} - \varphi_2\|_{L_{p'}(D')} + \|\varphi_2\|_{L_{p'}(D')} \cdot \|\varphi_{1h} - \varphi_1\|_{L_p(D')}]; \end{aligned}$$

dabei ist $c = \sup |\psi(x)|$ und D' ein Teilgebiet von D , in dessen Äußerem $\psi(x)$ verschwindet. Die rechte Seite der letzten Ungleichung strebt für $h \rightarrow 0$ gegen 0, da $\varphi_{2h}(x) \rightarrow \varphi_2(x)$ in $L_{p'}(D')$ und $\varphi_{1h}(x) \rightarrow \varphi_1(x)$ in $L_p(D')$ gilt und die in $L_p(D')$ konvergente Folge $\varphi_{1h}(x)$ in $L_p(D')$ der Norm nach beschränkt ist.

Wir stellen nun die Formel (126) unter der Bedingung auf, daß $\varphi_1(x)$ und $\frac{\partial \varphi_1(x)}{\partial x_1}$ zu $L_p(D')$ und $\varphi_2(x)$ und $\frac{\partial \varphi_2(x)}{\partial x_1}$ zu $L_{p'}(D')$ für jedes echt in D enthaltene Teilgebiet D' gehört. Benutzen wir die vorhergehende Behauptung, so haben wir für jede stetig differenzierbare finite Funktion $\psi(x)$, die außerhalb von D' gleich 0 ist,

$$\int_D \varphi_1(x) \varphi_2(x) \frac{\partial \psi(x)}{\partial x_1} dx = \lim_{h \rightarrow 0} \int_D \varphi_{1h}(x) \varphi_{2h}(x) \frac{\partial \psi(x)}{\partial x_1} dx$$

und folglich

$$\int_D \varphi_1(x) \varphi_2(x) \frac{\partial \psi(x)}{\partial x_1} dx = - \lim_{h \rightarrow 0} \int_D \left[\frac{\partial \varphi_{1h}(x)}{\partial x_1} \varphi_{2h}(x) + \varphi_{1h}(x) \frac{\partial \varphi_{2h}(x)}{\partial x_1} \right] \psi(x) dx. \quad (130)$$

Für hinreichend kleines h läßt sich in D' die Formel (129) anwenden und in (130) $\frac{\partial \varphi_{1h}(x)}{\partial x_1}$ durch $\left(\frac{\partial \varphi_1(x)}{\partial x_1} \right)_h$ und $\frac{\partial \varphi_{2h}(x)}{\partial x_1}$ durch $\left(\frac{\partial \varphi_2(x)}{\partial x_1} \right)_h$ ersetzen. Wenden wir auf die rechte Seite von (130) wieder die vorhergehende Behauptung an, so erhalten wir

$$\int_D \varphi_1(x) \varphi_2(x) \frac{\partial \psi(x)}{\partial x_1} dx = - \int_D \left[\frac{\partial \varphi_1(x)}{\partial x_1} \varphi_2(x) + \varphi_1(x) \frac{\partial \varphi_2(x)}{\partial x_1} \right] \psi(x) dx.$$

Diese Gleichung besagt, daß das Produkt $\varphi_1(x) \varphi_2(x)$ in D eine verallgemeinerte Ableitung nach x_1 besitzt, die mit Hilfe von (126) ermittelt werden kann.

Wir weisen darauf hin, daß (126) auch für $p = 1$ gilt. In diesem Fall brauchen wir nur $p' = \infty$ zu setzen, d. h. $\varphi_2(x)$ und $\frac{\partial \varphi_2(x)}{\partial x_1}$ als in einem beliebigen Teilgebiet D' beschränkt anzunehmen.

Wir zeigen nun, daß es noch eine andere Definition der verallgemeinerten Ableitung gibt und sich mit Hilfe von (129) die Äquivalenz dieser Definition mit der ersten nachweisen läßt.

Definition 2. Eine Funktion $\varphi(x)$ hat in D eine *verallgemeinerte Ableitung* $\chi(x)$ der Gestalt (122), wenn es eine Folge von Funktionen $\varphi_m(x)$ gibt, die im Innern von D l -mal stetig differenzierbar sind, und wenn $\varphi_m(x)$ und $D^l \varphi_m(x)$ gegen $\varphi(x)$ bzw. $\chi(x)$ in $L(D')$ streben, wobei D' ein beliebiges in D echt enthaltenes Teilgebiet ist.

Satz 1. Die Definitionen 1 und 2 sind miteinander äquivalent.

Beweis. Wir nehmen an, $\chi(x)$ sei die verallgemeinerte Ableitung von $\varphi(x)$ gemäß Definition 2. Es gilt (123), wenn wir $\varphi(x)$ durch $\varphi_m(x)$ ersetzen, und da $\varphi_m(x)$ gegen $\varphi(x)$ und $D^l \varphi_m(x)$ gegen $\chi(x)$ in $L(D')$ strebt, können wir bei beliebiger Wahl der finiten Funktion $\psi(x)$ mit den erwähnten Eigenschaften unter dem Integralzeichen den Grenzübergang vornehmen (vgl. [54]), woraus sich (124) ergibt. Wir nehmen nun an, $\chi(x)$ sei die verallgemeinerte Ableitung von $\varphi(x)$ gemäß Definition 1. Dann wird auf Grund von (129) und Satz 4 aus [71] die in Definition 2 verlangte Folge von Funktionen $\varphi_m(x)$ durch die Mittelfunktionen $\varphi_{h_m}(x)$ für eine beliebige, gegen 0 konvergierende Folge h_m angegeben (dabei wird angenommen, daß $\varphi(x)$ durch die Funktion 0 ins Äußere von D fortgesetzt wird). Damit ist Satz 1 bewiesen.

Aus diesem Satz folgt, daß die verallgemeinerte Ableitung, falls sie existiert, auch im Sinne von Definition 2 eindeutig bestimmt ist.

Wir beweisen nun einen Satz, der besagt, daß die verallgemeinerten Ableitungen den schwachen Grenzübergang in $L_p(D')$ erlauben.

Satz 2. Die im Innern von D definierten Funktionen $\varphi_k(x)$, $k = 1, 2, \dots$, mögen gegen eine Funktion $\varphi(x)$ in $L_p(D')$, $p > 1$, schwach konvergieren, wobei D' ein in D echt enthaltenes Teilgebiet ist, und in D verallgemeinerte Ableitungen $D^l \varphi_k(x)$ der Form (122) besitzen; die Normen von $D^l \varphi_k(x)$ mögen in $L_p(D')$ durch eine Zahl $M(D')$ beschränkt sein, die von der Wahl von D' abhängt. Dann hat $\varphi(x)$ in D die verallgemeinerte Ableitung $D^l \varphi(x)$ der Form (122); $D^l \varphi(x)$ ist der schwache Grenzwert von $D^l \varphi_k(x)$ in $L_p(D')$.

Beweis. Auf Grund der schwachen Kompaktheit beschränkter Mengen in L_p ($p > 1$) folgt aus der Ungleichung

$$\|D^l \varphi_k\|_{L_p(D')} \leq M(D') \quad (131)$$

die Existenz einer Teilfolge $\varphi_{n_k}(x)$, für welche $D^l \varphi_{n_k}(x)$ in $L_p(D')$ schwach konvergiert. Wir nehmen eine Folge von Gebieten D'_m , für die $D'_m \subset D'_{m+1} \subset D$ gilt und die gegen D konvergiert, und konstruieren mit Hilfe des Diagonalverfahrens eine Teilfolge $D^l \varphi_{m_k}(x)$, die in $L_p(D')$ schwach gegen eine Funktion $\chi(x)$ in jedem echt in D enthaltenen Teilgebiet D' konvergiert. Es ist klar, daß $\chi(x)$ überall in D definiert ist und für jedes echt in D enthaltene Gebiet D' zu $L_p(D')$ gehört.

Die Gleichung (123) behält ihre Gültigkeit, wenn $\varphi(x)$ durch $\varphi_{m_k}(x)$ ersetzt wird. Gehen wir dann bei einer festen Funktion in (123) zur Grenze über, so gelangen wir, da $\psi(x)$ finit ist, zu der Gleichung (124) (schwache Konvergenz); daraus folgt, daß

$\chi(x)$ die verallgemeinerte Ableitung von $\varphi(x)$ in D ist. Aus den obigen Überlegungen ergibt sich, daß für alle schwach konvergenten Teilfolgen $D^l\varphi_{m_k}(x)$ ein einziger Grenzwert $\chi(x)$ existiert (Eindeutigkeit der verallgemeinerten Ableitung); also konvergiert auch die Folge $D^l\varphi_k(x)$ schwach gegen $\chi(x)$.

Bemerkungen. 1. Aus Satz 2 folgt, daß in D die verallgemeinerte Ableitung $D^l\varphi(x) \in L_p(D')$ existiert, wenn $\varphi(x)$ zu $L_p(D')$ gehört und die Ableitungen der Mittelfunktionen $\varphi_h(x)$ der Abschätzung (131) genügen. Wie wir nämlich wissen, ist in diesem Fall $D^l\varphi_h(x) \rightarrow D^l\varphi(x)$ in $L_p(D')$; folglich genügt auch die Norm von $D^l\varphi(x)$ der Abschätzung (131).

2. In den Voraussetzungen von Satz 2 können die Funktionen $\varphi(x)$ und $D^l\varphi(x)$ zu $L_p(D')$ bzw. zu $L_q(D')$ auch für $p \neq q$ gehören.

3. Satz 2 bleibt für $p=1$ in Kraft, wenn statt (131) die schwache Kompaktheit der Funktionen $D^l\varphi_k(x)$ in $L(D')$ für jedes in D echt enthaltene Gebiet D' gefordert wird.

110. Verallgemeinerte Ableitungen (Fortsetzung). Wir stellen jetzt den Zusammenhang zwischen der Existenz der verallgemeinerten Ableitungen und der absoluten Stetigkeit der Funktionen her und betrachten den Fall einer unabhängigen Veränderlichen; als Grundgebiet D wählen wir das Intervall $0 < x < 1$. Die Funktion $\varphi(x)$ sei auf dem Intervall $[0, 1]$ absolut stetig. Wie wir wissen [74], hat $\varphi(x)$ auf $[0, 1]$ die Ableitung $\varphi'(x)$, die eine auf $[0, 1]$ summierbare Funktion ist. Die Formel für die partielle Integration [74] liefert für jede stetig differenzierbare und finite Funktion $\psi(x)$ die Beziehung

$$\int_0^1 \varphi(x) \psi'(x) dx = - \int_0^1 \varphi'(x) \psi(x) dx ; \quad (132)$$

$\varphi'(x)$ ist also die verallgemeinerte Ableitung von $\varphi(x)$.

Die Funktion $\varphi(x)$ gehöre zu $L[0, 1]$ und habe die verallgemeinerte Ableitung $\frac{d\varphi(x)}{dx}$ in D , die ebenfalls in $L[0, 1]$ liegt. Wir wollen zeigen, daß dann $\varphi(x)$ einer in $[0, 1]$ absolut stetigen Funktion äquivalent ist. Wir setzen

$$\varphi_1(x) = \int_0^x \frac{d\varphi(t)}{dt} dt ;$$

$\varphi_1(x)$ ist absolut stetig, und die Ableitung $\varphi_1'(x)$ ist mit $\frac{d\varphi(x)}{dx}$ äquivalent [74]. Die Differenz $\varphi^*(x) = \varphi(x) - \varphi_1(x)$ hat offenbar eine Ableitung, die äquivalent 0 ist. Wir wählen ein festes $\varepsilon > 0$ und betrachten das Intervall $[\varepsilon, 1 - \varepsilon]$. Für hinreichend kleines h ist die Ableitung der Mittelfunktion $\varphi_h^*(x)$ gleich 0 in $[\varepsilon, 1 - \varepsilon]$, so daß also $\varphi_h^*(x)$ dort eine Konstante ist. Da der Grenzwert von Konstanten nur eine Konstante sein kann und $\varphi_h^*(x) \rightarrow \varphi^*(x)$ in $L[\varepsilon, 1 - \varepsilon]$ gilt, ist $\varphi^*(x)$ mit einer Konstanten im Intervall $[\varepsilon, 1 - \varepsilon]$ äquivalent. Daraus folgt unmittelbar, daß überall in D

$$\varphi(x) = \varphi(0) + \int_0^x \frac{d\varphi(t)}{dt} dt \quad (133)$$

bis auf Äquivalenz genau gilt. Damit haben wir gezeigt, daß die Existenz der verallgemeinerten Ableitung mit der absoluten Stetigkeit von $\varphi(x)$ gleichwertig ist.

Für den Fall mehrerer Veränderlicher läßt sich analog zeigen, daß dann, wenn $\varphi(x_1, x_2, \dots, x_n)$ beispielsweise im Kubus $[0 < x_k < 1; k = 1, 2, \dots, n]$ die verallgemeinerte Ableitung $\frac{\partial \varphi(x)}{\partial x_1}$ hat und dort die Funktionen $\varphi(x)$ und $\frac{\partial \varphi(x)}{\partial x_1}$ zu L_p ($p \geq 1$) gehören, für fast alle Werte $(x_2, x_3, \dots, x_n)$ aus dem Kubus $[0 < x_k < 1; k = 2, 3, \dots, n]$ die Funktion $\varphi(x)$ auf $0 \leq x_1 \leq 1$ absolut stetig und die Gleichung

$$\varphi(x_1, x_2, \dots, x_n) = \varphi(0, x_2, \dots, x_n) + \int_0^{x_1} \frac{\partial \varphi(t, x_2, \dots, x_n)}{\partial t} dt \quad (134)$$

für alle x_1 aus $0 \leq x_1 \leq 1$ erfüllt ist.

Die Gleichungen (133) und (134) bedürfen noch einer Erläuterung. Die Funktion $\varphi(x)$ und ihre verallgemeinerte Ableitung $D\varphi(x)$ sind nämlich bis auf eine Menge vom Maß 0 bestimmt. Daher muß man (133) und (134) in dem Sinne auffassen, daß Funktionen aus der Klasse der zu $\varphi(x)$ äquivalenten Funktionen existieren, für welche diese Gleichungen gelten.

Wir betrachten als Beispiel die Funktion $\varphi(x_1, x_2)$, welche die verallgemeinerte gemischte Ableitung $\frac{\partial^2 \varphi(x_1, x_2)}{\partial x_1 \partial x_2}$, aber keine verallgemeinerten ersten Ableitungen besitzt. Diese Eigenschaft hat die Funktion $\varphi(x_1, x_2) = f(x_1) + f(x_2)$ [$0 \leq x_k \leq 1; k = 1, 2$], wobei $f(x)$ die stetige Funktion aus [76] ist. Die Funktion $\varphi(x_1, x_2)$ hat keine verallgemeinerten ersten Ableitungen, denn $f(x)$ ist nicht absolut stetig. Die verallgemeinerte Ableitung $\frac{\partial^2 \varphi(x_1, x_2)}{\partial x_1 \partial x_2}$ existiert und ist identisch 0; für jede stetig differenzierbare und finite Funktion $\psi(x_1, x_2)$ gilt nämlich

$$\int_0^1 \int_0^1 f(x_1) \frac{\partial^2 \varphi(x_1, x_2)}{\partial x_1 \partial x_2} dx_1 dx_2 = \int_0^1 dx_1 \int_0^1 \frac{\partial \left[f(x_1) \frac{\partial \varphi(x_1, x_2)}{\partial x_1} \right]}{\partial x_2} dx_2 = 0$$

und analog für $f(x_2)$, d. h.

$$\int_0^1 \int_0^1 [f(x_1) + f(x_2)] \frac{\partial^2 \varphi(x_1, x_2)}{\partial x_1 \partial x_2} dx_1 dx_2 = 0,$$

woraus sich auf Grund der Definition 1 aus [109]

$$\frac{\partial^2 \varphi(x_1, x_2)}{\partial x_1 \partial x_2} = 0$$

ergibt.

Ist eine Funktion $\varphi(x_1, x_2, \dots, x_n)$ stetig in D und kann das Gebiet D mit Hilfe endlich vieler glatter Flächen in endlich viele Gebiete D_i ($i = 1, 2, \dots, l$) zerlegt werden, wobei $\varphi(x)$ in allen D_i bis zum Rand nach einem gewissen x_k stetig differenzierbar ist, so hat $\varphi(x)$ in D eine verallgemeinerte Ableitung, die in jedem der D_i gleich $\frac{\partial \varphi(x)}{\partial x_k}$ ist. (Diese Ableitung kann auf den obengenannten Flächen Unstetigkeitsstellen erster Art besitzen.) Die Behauptung folgt sofort aus der Formel für die partielle Integration:

$$\int_{D_i} \varphi(x) \frac{\partial \varphi(x)}{\partial x_k} dx = - \int_{D_i} \frac{\partial \varphi(x)}{\partial x_k} \varphi(x) dx + \int_{\bar{S}_i} \varphi(x) \psi(x) \cos(n, x_k) dS; \quad (135)$$

dabei ist S_i die Berandung von D_i und n die Richtung der äußeren Normalen von D_i auf S_i . Es muß nur darauf hingewiesen werden, daß sich die Integrale über die Flächen S_i bei Summation über i aufheben.

Hat $\varphi(x)$ verschiedene Grenzwerte auf einem $(n-1)$ -dimensionalen Stück Σ einer sich in D befindenden Fläche und liegt die Richtung x_k nicht in der Tangentialebene dieser Fläche, so gibt es in D keine verallgemeinerte Ableitung $\frac{\partial \varphi(x)}{\partial x_k}$. Dies folgt aus dem oben hergestellten Zusammenhang zwischen der absoluten Stetigkeit von $\varphi(x)$ und der Existenz der verallgemeinerten Ableitung.

Bemerkung. Man kann nicht nur den Begriff der einzelnen verallgemeinerten Ableitung einer summierbaren Funktion $\varphi(x)$, sondern auch einen verallgemeinerten linearen Differentialoperator beliebiger Ordnung einführen, nämlich

$$L(\varphi) = \sum_{i,k=1}^n a_{ik} \frac{\partial^2 \varphi(x)}{\partial x_i \partial x_k} + \sum_{k=1}^n b_k \frac{\partial \varphi(x)}{\partial x_k} + c\varphi(x), \quad (136)$$

wobei die Koeffizienten hinreichend oft stetig differenzierbare Funktionen von $(x_1, x_2, \dots, x_n)$ sind.

Ein solcher verallgemeinerter Operator wird durch die zu (124) analoge Gleichung

$$\int_D \varphi(x) M(\psi) dx = \int_D L(\varphi) \psi(x) dx \quad (137)$$

definiert, wobei $M(\psi)$ der zu $L(\varphi)$ adjungierte Differentialoperator und $\psi(x)$ eine beliebige, stetig differenzierbare und in D finite Funktion ist. Die Existenz der in $L(\varphi)$ auftretenden einzelnen verallgemeinerten Ableitungen wird dabei nicht vorausgesetzt.

111. Sternförmige Gebiete. In [109] zeigten wir mit Hilfe der Mittelfunktionen, daß zu jeder Funktion $\varphi(x) \in L_p(D')$, $p \geq 1$, deren verallgemeinerte Ableitung $\chi(x) = D^l \varphi(x)$ ebenfalls in $L_p(D')$ liegt, eine Folge von l -mal in $\bar{D}$ stetig differenzierbaren Funktionen $\varphi_k(x)$ existiert, für welche $\varphi_k(x) \rightarrow \varphi(x)$ und $D^l \varphi_k(x) \rightarrow \chi(x)$ in $L_p(D')$ ist. (Wie früher ist D' ein beliebiges, in D echt enthaltenes Gebiet.) Wir zeigen nun, daß für eine wichtige Klasse von Gebieten eine analoge Approximation der Funktionen $\varphi(x)$ und $D^l \varphi(x)$ auch im Raum $L_p(D)$ möglich ist.

Wir nennen ein Gebiet D von *sternförmigem Typ*, wenn es einen solchen inneren Punkt $x_0 \in D$ gibt, daß jeder von x_0 ausgehende Radiusvektor die Berandung von D in genau einem Punkt schneidet. Ein solches Gebiet heißt auch *sternförmig bezüglich x_0* .

Satz. *Es sei D ein Gebiet von sternförmigem Typ, und $\varphi(x)$ habe in D die verallgemeinerte Ableitung $D^l \varphi(x)$, wobei $\varphi(x)$ und $D^l \varphi(x)$ zu $L_p(D)$, $p \geq 1$, gehören mögen. Dann gibt es eine Folge von l -mal in $\bar{D}$ stetig differenzierbaren Funktionen $\varphi_k(x)$ derart, daß $\varphi_k(x)$ und $D^l \varphi_k(x)$ gegen $\varphi(x)$ bzw. $D^l \varphi(x)$ in $L_p(D)$ konvergieren.*

Beweis. Wir sprechen von l -mal in $\bar{D}$ stetig differenzierbaren Funktionen $\varphi_k(x)$, wenn die $\varphi_k(x)$ stetig in $\bar{D}$ und l -mal stetig differenzierbar im Innern von D sind und ihre Ableitungen so bis an die Berandung von D definiert werden können, daß sich Funktionen ergeben, die in $\bar{D}$ stetig sind.

Wir wählen den Punkt x_0 als Koordinatenursprung und konstruieren eine Folge von Funktionen $\varphi\left(\frac{k-1}{k}x\right)$, $k=2, 3, \dots$; diese seien im Innern von Gebieten D_k definiert, welche das Gebiet D echt in sich enthalten, wobei D_k durch eine Ähnlichkeitstransformation mit dem Ähnlichkeitsfaktor $\frac{k}{k-1}$ aus D hervorgeht.

Wir setzen $\varphi\left(\frac{k-1}{k}x\right) = \varphi^{(k)}(x)$ und zeigen, daß in $L_p(D)$ die Funktionen $\varphi^{(k)}(x)$ gegen $\varphi(x)$ und die verallgemeinerte Ableitung $D^l \varphi^{(k)}(x)$ gegen $\chi(x) = D^l \varphi(x)$ konvergieren. Wir beweisen z. B. die zweite Behauptung. Es gilt

$$\begin{aligned} \|D^l \varphi(x) - D^l \varphi^{(k)}(x)\| &= \left[\int_D \left| \chi(x) - \left(\frac{k-1}{k}\right)^l \chi\left(\frac{k-1}{k}x\right) \right|^p dx \right]^{1/p} \\ &\leq \left[1 - \left(\frac{k-1}{k}\right)^l \right] \left[\int_D \left| \chi\left(\frac{k-1}{k}x\right) \right|^p dx \right]^{1/p} \\ &\quad + \left[\int_D \left| \chi(x) - \chi\left(\frac{k-1}{k}x\right) \right|^p dx \right]^{1/p}. \end{aligned}$$

Der Abstand zwischen den Punkten $\frac{k-1}{k}x$ und x ist nicht größer als $\frac{d}{k}$, wenn d der Durchmesser von D ist, und strebt folglich in D gleichmäßig gegen 0. Wiederholen wir die Überlegungen, die wir bei dem Satz über die Stetigkeit im Mittel [70] angestellt haben, so können wir uns davon überzeugen, daß der zweite Summand auf der rechten Seite der letzten Ungleichung für $k \rightarrow \infty$ gegen 0 strebt. Im ersten Summanden verschwindet der Faktor $1 - \left(\frac{k-1}{k}\right)^l$ für $k \rightarrow \infty$, während der zweite Faktor auf Grund des oben Gesagten gegen $\|\chi(x)\|_{L_p(D)}$ konvergiert (Stetigkeit der Norm). Noch einfacher läßt sich die Konvergenz der $\varphi^{(k)}(x)$ gegen $\varphi(x)$ in $L_p(D)$ nachweisen.

Für ein festes k ist das Gebiet D echt in D_k enthalten, so daß die Mittelfunktionen $\varphi_h^{(k)}(x)$ gegen $\varphi^{(k)}(x)$ und ihre Ableitungen $D^l \varphi_h^{(k)}(x)$ gegen $D^l \varphi^{(k)}(x)$ für $h \rightarrow 0$ in $L_p(D)$ konvergieren. Daraus ergibt sich, daß die Funktionen $\varphi(x)$ und $\chi(x)$ in der Metrik von $L_p(D)$ durch unendlich oft in D differenzierbare Funktionen $\varphi_{h_k}^{(k)}(x)$ bzw. $D^l \varphi_{h_k}^{(k)}(x)$ approximiert werden können, wobei die h_k eine geeignete, gegen 0 konvergierende Folge bilden. Die Funktionen $\varphi_{h_k}^{(k)}(x)$ können als die im Satz genannten $\varphi_k(x)$ genommen werden. Damit ist der Satz bewiesen.

112. Die Räume $\tilde{W}_p^{(l)}$ und $W_p^{(l)}$. Es sei D wieder ein beschränktes Gebiet des n -dimensionalen Raumes. Wir betrachten die Menge der Funktionen $\varphi(x)$, welche alle verallgemeinerte Ableitungen l -ter Ordnung haben; $\varphi(x)$ und alle $D^l \varphi(x)$ gehören zu $L_p(D)$, $p \geq 1$. Diese Funktionenklasse bezeichnen wir mit $\tilde{W}_p^{(l)}(D)$; sie bildet einen normierten Raum, wenn wir mit Hilfe von

$$\|\varphi\|_{\tilde{W}_p^{(l)}(D)}^p = \int_D |\varphi(x)|^p dx + \int_D \sum_{l_1+l_2+\dots+l_n=l} \left| \frac{\partial^l \varphi(x)}{\partial x_1^{l_1} \dots \partial x_n^{l_n}} \right|^p dx \quad (138)$$

in ihr eine Norm einführen. Hier und im folgenden bezeichne $\sum_{l_1+l_2+\dots+l_n=l}$ die Summation über alle möglichen Systeme natürlicher Zahlen $l_1, l_2, \dots, l_n$, deren Summe die Zahl l ergibt. Diese Norm besitzt die in [95] angegebenen Eigenschaften. Wir wollen zeigen, daß der Raum $\tilde{W}_p^{(l)}$ vollständig ist. Es sei $\varphi_k(x)$ eine in sich konvergente Folge in $\tilde{W}_p^{(l)}(D)$, d. h., es gelte

$$\int_D \left[|\varphi_k(x) - \varphi_m(x)|^p + \sum_{l_1+l_2+\dots+l_n=l} \left| \frac{\partial^l \varphi_k(x)}{\partial x_1^{l_1} \dots \partial x_n^{l_n}} - \frac{\partial^l \varphi_m(x)}{\partial x_1^{l_1} \dots \partial x_n^{l_n}} \right|^p \right] dx \rightarrow 0$$

für $k, m \rightarrow \infty$. Daraus ergibt sich, daß die Folgen der $\varphi_k(x)$ und der $D^l \varphi_k(x)$ in $L_p(D)$ in sich konvergieren. Auf Grund der Vollständigkeit von $L_p(D)$ und mit Hilfe von Satz 2 aus [109] finden wir, daß die $\varphi_k(x)$ in $L_p(D)$ gegen eine Funktion $\varphi(x)$ konvergieren, welche alle verallgemeinerten Ableitungen l -ter Ordnung in $L_p(D)$ besitzt, und daß $D^l \varphi_k(x) \rightarrow D^l \varphi(x)$ in $L_p(D)$ ist. Dies ist gleichwertig mit der Konvergenz von $\varphi_k(x)$ gegen $\varphi(x)$ in $\tilde{W}_p^{(l)}(D)$. Der Raum $\tilde{W}_p^{(l)}(D)$ ist somit ein Banachraum (also ein vollständiger linearer normierter Raum). Wir gehen jetzt zum Beweis der Separabilität von $\tilde{W}_p^{(l)}(D)$ über. Zu diesem Zweck stellen wir D als abzählbare Menge sich nicht überlappender, halboffener Intervalle dar [32]. Die ihnen entsprechenden offenen Intervalle numerieren wir, bezeichnen sie mit D_k ($k=1, 2, \dots$) und führen die Menge $V_p^{(l)}(D)$ der Funktionen $\varphi(x)$ ein, die zu $\tilde{W}_p^{(l)}(D_k)$ in jedem Intervall D_k gehören und für die die Reihe

$$\|\varphi\|_{V_p^{(l)}(D)}^p = \sum_{k=1}^{\infty} \|\varphi\|_{\tilde{W}_p^{(l)}(D_k)}^p \quad (139)$$

konvergiert. Die Vorgabe einer Norm vermöge Formel (139) führt $V_p^{(l)}(D)$ in einen linearen normierten Raum über. Es ist leicht zu sehen, daß die Funktionen aus $\tilde{W}_p^{(l)}(D)$ in $V_p^{(l)}(D)$ auftreten und $\|\varphi\|_{\tilde{W}_p^{(l)}(D)} = \|\varphi\|_{V_p^{(l)}(D)}$ ist. $\tilde{W}_p^{(l)}(D)$ ist also ein Unter-
raum von $V_p^{(l)}(D)$, und es genügt, wenn wir jetzt die Separabilität von $V_p^{(l)}(D)$ nachweisen [94].

Die Menge der Funktionen aus $V_p^{(l)}(D)$, die nur auf endlich vielen Intervallen von 0 verschieden sind, ist in $V_p^{(l)}(D)$ dicht. Ist nämlich $\varphi(x) \in V_p^{(l)}(D)$ und eine beliebige Zahl $\varepsilon > 0$ gegeben und setzen wir

$$\varphi(x) = \begin{cases} \varphi_m(x) & \text{für } x \in \sum_{k=1}^m D_k, \\ 0 & \text{im übrigen Teil von } D, \end{cases}$$

so ist offenbar $\varphi_m(x) \in V_p^{(l)}(D)$, und für hinreichend großes m gilt

$$\|\varphi - \varphi_m\|_{V_p^{(l)}(D)}^p = \sum_{k=m+1}^{\infty} \|\varphi\|_{\tilde{W}_p^{(l)}(D_k)}^p \leq \varepsilon^p$$

auf Grund der Konvergenz von (139). Die Funktionen $\varphi_m(x)$ können in jedem der Intervalle D_k ($k \leq m$) durch l -mal stetig differenzierbare Funktionen in der Metrik von $\tilde{W}_p^{(l)}(D_k)$ approximiert werden [111], die sich ihrerseits nebst ihren Ableitungen durch Polynome mit rationalen Koeffizienten in D_k gleichmäßig approximieren lassen. Daraus folgt, daß in $V_p^{(l)}(D)$ die Menge derjenigen Funktionen dicht ist, die nur in endlich vielen Intervallen D_k von 0 verschieden sind und wobei jede dieser Funktionen in jedem der D_k mit einem Polynom mit rationalen Koeffizienten übereinstimmt. Die Menge dieser Funktionen ist offenbar abzählbar; folglich ist $V_p^{(l)}(D)$ separabel. Damit ist auch die Separabilität von $\tilde{W}_p^{(l)}(D)$ bewiesen.

Wir gehen nun zu einem speziellen Problem über. In einem linearen normierten Raum X sei neben der Grundnorm $\|x\|$ noch eine Norm $\|x\|_1$ eingeführt, wobei für alle $x \in X$

$$c_2 \|x\| \leq \|x\|_1 \leq c_1 \|x\| \quad (140)$$

gilt und $c_1 > 0$, $c_2 > 0$ Konstanten sind. Normen, die der Bedingung (140) genügen, heißen *äquivalent*. Offenbar ist eine Folge, die der einen Norm nach konvergiert, auch der anderen Norm nach konvergent. Bei der Untersuchung von Vollständigkeit, Separabilität, Kompaktheit usw. ist es gleichgültig, welche der beiden Normen betrachtet wird. Genauso ist ein distributiver Operator, der bezüglich der einen

Norm beschränkt ist, auch bezüglich der anderen (äquivalenten) Norm beschränkt. Dabei ändert sich die Norm eines beschränkten Operators im allgemeinen beim Übergang zu einer äquivalenten Normierung des Raumes, aber sie bleibt endlich.

Auf die Frage nach äquivalenten Normierungen des n -dimensionalen reellen euklidischen Raumes R_n wollen wir ausführlicher eingehen. Im R_n führten wir eine Norm mit Hilfe der Formel

$$||x|| = \sqrt{x_1^2 + x_2^2 + \dots + x_n^2} \quad (141)$$

ein. Nun besitze eine Funktion $f(x) = f(x_1, x_2, \dots, x_n)$ die Eigenschaften der Norm [95]; außerdem sei sie auf der Oberfläche der Einheitskugel $x_1^2 + x_2^2 + \dots + x_n^2 = 1$ stetig. Wir werden zeigen, daß $||x||_1 = f(x)$ eine zu (141) äquivalente Norm ist. Die Funktion $f(x)$ erreicht auf der Oberfläche der Einheitskugel ihren größten und ihren kleinsten Wert; wir wählen dafür die Bezeichnung $c_1 = \sup f(x)$ und $c_2 = \inf f(x)$ für $||x|| = 1$. Da $f(x)$ positiv und stetig ist, gilt $0 < c_2 \leq c_1 < \infty$. Auf Grund der Eigenschaften von $f(x)$ gilt überall im R_n

$$c_2 ||x|| \leq f(x) \leq c_1 ||x|| ,$$

d. h., die Normen (141) und $f(x)$ sind äquivalent.

Für $f(x)$ kann man z. B. die Funktionen

$$f(x) = \left[\sum_{k=1}^n |x_k|^p \right]^{1/p}, \quad p > 1, \quad \text{oder} \quad f(x) = \max_k |x_k|$$

wählen. Stützen wir uns auf die vorhergehenden Überlegungen, so können wir leicht nachprüfen, daß die durch

$$||\varphi|| = \sum_{l_1 + \dots + l_n = l} |D^l \varphi|_{L_p(D)} + ||\varphi||_{L_p(D)}, \quad (142)$$

$$||\varphi|| = \max_{l_1 + \dots + l_n = l} ||D^l \varphi|_{L_p(D)} + ||\varphi||_{L_p(D)}, \quad (143)$$

$$||\varphi|| = \left\{ \int_D \left[\sum_{l_1 + \dots + l_n = l} (D^l \varphi)^2 \right]^{p/2} dx \right\}^{1/p} + ||\varphi||_{L_p(D)} \quad (144)$$

im Raum $\tilde{W}_p^{(l)}(D)$ definierten Normen mit der Grundnorm (138) äquivalent sind. Diese Bemerkung werden wir im folgenden benutzen.

Wir führen nun einen weiteren Funktionenraum ein. Auf der Menge der Funktionen, welche in D alle verallgemeinerten Ableitungen l -ter Ordnung besitzen und nebst diesen Ableitungen zu $L_p(D)$ gehören, definieren wir eine Norm durch

$$||\varphi||_{W_p^{(l)}(D)} = \sum_{k=0}^l \sum_{k_1 + \dots + k_n = k} ||D^k \varphi||_{L_p(D)}. \quad (145)$$

Den so erhaltenen linearen normierten Raum bezeichnen wir mit $W_p^{(l)}(D)$. Genauso wie für den Raum $\tilde{W}_p^{(l)}(D)$ können wir nachweisen, daß der Raum $W_p^{(l)}(D)$ vollständig und separabel ist. Die obige Bemerkung über die Äquivalenz von Normen in $\tilde{W}_p^{(l)}(D)$ läßt sich in gleichem Maße auch auf den Raum $W_p^{(l)}(D)$ übertragen. In [116] werden wir zeigen, daß für eine hinreichend umfangreiche Klasse von Gebieten die Räume $W_p^{(l)}(D)$ und $\tilde{W}_p^{(l)}(D)$ aus ein und denselben Mengen von Funktionen bestehen und die Normen (138) und (145) äquivalent sind. Für $l=0$ und $l=1$ stimmen $W_p^{(l)}(D)$ und $\tilde{W}_p^{(l)}(D)$ nach Definition überein, wobei offenbar $W_p^{(0)}(D)$ mit $L_p(D)$ übereinstimmt.

113. Eigenschaften der Funktionen aus der Klasse $W_p^{(l)}(D)$. Die Funktionen aus den Klassen $W_p^{(l)}(D)$ und $\tilde{W}_p^{(l)}(D)$ werden wir in den folgenden Abschnitten, die den

sogenannten Einbettungssätzen gewidmet sind, systematisch untersuchen. Die Resultate, die wir jetzt angeben wollen, sind Spezialfälle dieser allgemeinen Sätze. Auf Grund der Wichtigkeit dieser Resultate werden wir hier jedoch unabhängige und bedeutend einfachere Beweise angeben.

Zunächst weisen wir auf eine Eigenschaft des Raumes $W_p^{(l)}(D)$ hin, die sich unmittelbar aus seiner Definition ergibt. Wir ersetzen die Veränderlichen $x(x_1, x_2, \dots, x_n)$ durch $y(y_1, y_2, \dots, y_n)$ derart, daß D eindeutig auf ein gewisses Gebiet $\bar{D}_1$ abgebildet wird, wobei die Abbildung in beiden Richtungen durch Funktionen vermittelt wird, die in abgeschlossenen Gebieten $\bar{D}$ bzw. $\bar{D}_1$ stetige Ableitungen bis einschließlich l -ter Ordnung haben. Werden bei den Funktionen die unabhängigen Veränderlichen wie oben ersetzt, so geht dabei $W_p^{(l)}(D)$ in $W_p^{(l)}(D_1)$ über.

Aus der Definition des Raumes $W_p^{(l)}(D)$ folgt unmittelbar: Ist $\varphi(x) \in W_p^{(l)}(D)$, so gehört $\varphi(x)$ für $q \leq p$ und $m \leq l$ zu $W_q^{(m)}(D)$. Im Zusammenhang damit beweisen wir den folgenden Satz:

Satz 1. *Ist U eine in $W_p^{(l)}(D)$, $l \geq 1$, beschränkte Menge von Elementen $\varphi(x) \in W_p^{(l)}(D)$, so ist sie in $W_p^{(l-1)}(D')$ kompakt, wobei D' ein beliebiges, in D echt enthaltenes Gebiet ist.*

Beweis. Wir wollen den Beweis für $l = 1$ führen. Nach Voraussetzung existiert eine Konstante c derart, daß für $\varphi(x) \in U$

$$\|\varphi\|_{W_p^{(1)}(D)}^p = \int_D \left[|\varphi(x)|^p + \sum_{s=1}^n \left| \frac{\partial \varphi(x)}{\partial x_s} \right|^p \right] dx \leq c^p \quad (146)$$

gilt. Wir müssen zeigen, daß U in $L_p(D')$ kompakt ist, wobei D' ein festes, in D echt enthaltenes Gebiet ist. Die Beschränktheit von U in $L_p(D')$ folgt unmittelbar aus (146); wir brauchen also nur noch zu beweisen, daß die Funktionen $\varphi(x) \in U$ in $L_p(D')$ im Mittel gleichgradig stetig sind [70]. Für hinreichend kleine $|\Delta x| = \sqrt{(\Delta x_1)^2 + (\Delta x_2)^2 + \dots + (\Delta x_n)^2}$ beweisen wir die Ungleichung

$$\int_{D'} |\varphi(x + \Delta x) - \varphi(x)|^p dx \leq c_1 |\Delta x|^p, \quad (147)$$

wobei c_1 eine Konstante ist (und zwar für alle $\varphi(x) \in U$ dieselbe); daraus folgt dann die gleichgradige Stetigkeit.

Durch Drehung der Koordinatenachse können wir erreichen, daß Δx gleich $(\Delta x_1, 0, 0, \dots, 0)$ und $\Delta x_1 > 0$ ist. Es sei D'' ein Gebiet von derselben Art wie D' , und D' sei echt in D'' enthalten. Die Größe Δx_1 soll so klein sein, daß $x + \Delta x_1$ noch in D'' liegt, wenn $x \in D'$ ist. Wir nehmen zunächst an, die Funktion $\varphi(x)$ sein in $\bar{D}$ stetig differenzierbar. Offenbar gilt

$$R = \int_{D'} |\varphi(x + \Delta x_1) - \varphi(x)|^p dx = \int_{D'} \left| \int_0^{\Delta x_1} \frac{\partial \varphi(x_1 + \tau, x_2, \dots, x_n)}{\partial \tau} d\tau \right|^p dx.$$

Für $p > 1$ läßt sich das innere Integral mit Hilfe der Hölderschen Ungleichung abschätzen:

$$\begin{aligned} R &\leq \int_{D'} (\Delta x_1)^{p/p'} \left[\int_0^{\Delta x_1} \left| \frac{\partial \varphi(x_1 + \tau, x_2, \dots, x_n)}{\partial \tau} \right|^p d\tau \right] dx \\ &= (\Delta x_1)^{p/p'} \int_0^{\Delta x_1} \left[\int_{D'} \left| \frac{\partial \varphi(x_1 + \tau, x_2, \dots, x_n)}{\partial \tau} \right|^p dx \right] d\tau \\ &\leq (\Delta x_1)^{p/p'} \int_0^{\Delta x_1} \|\varphi\|_{W_p^{(1)}(D'')}^p d\tau = (\Delta x_1)^p \|\varphi\|_{W_p^{(1)}(D')}^p. \end{aligned}$$

Daraus folgt

$$\int_{D'} |\varphi(x + \Delta x) - \varphi(x)|^p dx \leq (\Delta x_1)^p \|\varphi\|_{W_p^{(1)}(D')}^p. \quad (148)$$

Für $p=1$ ergibt sich diese Abschätzung unmittelbar durch Änderung der Reihenfolge der Integrationen. Die Ungleichung (148) gilt nicht nur für stetig differenzierbare Funktionen, sondern auch für jede Funktion $\varphi(x) \in W_p^{(1)}(D)$. Davon können wir uns leicht überzeugen, wenn wir eine Folge von stetig differenzierbaren Funktionen wählen, die gegen $\varphi(x)$ in $W_p^{(1)}(D)$ konvergieren, und den Grenzübergang in (148) vornehmen. Zusammen mit (146) führt (148) auf (147). Damit ist der Satz für $l=1$ bewiesen. Berücksichtigen wir nun, daß etwa $\frac{\partial^2 \varphi(x)}{\partial x_1^2}$ die verallgemeinerte Ableitung von $\frac{\partial \varphi(x)}{\partial x_1}$ nach x_1 ist, so können wir den Satz auf Grund des Beweises für $l=1$ auch für $l=2$ beweisen. Analog läßt er sich für beliebige l beweisen.

Wir erwähnen noch einen Satz, der unmittelbar aus dem in [IV/2, 59] bewiesenen Satz folgt:

Satz 2. *Ist die Folge der stetigen und bis zur Ordnung $l = \left[\frac{n}{2}\right] + 1$ stetig differenzierbaren Funktionen $\varphi_k(x)$ in $W_2^{(l)}(D')$ konvergent, wobei D' ein beliebiges, in D enthaltenes Gebiet ist, so konvergiert sie in jedem Gebiet D' gleichmäßig.*

Daraus ergibt sich unmittelbar, daß die Grenzfunktion $\varphi(x)$ im Innern von D stetig ist. Wir nehmen nun an, es gebe eine Funktion $\varphi(x) \in W_2^{(l)}(D)$, $l \equiv \left[\frac{n}{2}\right] + 1$, und die $\varphi_k(x)$ seien Mittelfunktionen für $\varphi(x)$, wobei der Mittelungsradius für $k \rightarrow \infty$ gegen 0 strebt. Berücksichtigen wir die Eigenschaft der Mittelfunktionen verallgemeinerter Ableitungen [109] und den Satz 2, so gelangen wir zu der folgenden Behauptung: *Ist $\varphi(x) \in W_2^{(l)}(D)$ und $l \equiv \left[\frac{n}{2}\right] + 1$, so ist $\varphi(x)$ mit einer in $\bar{D}$ stetigen Funktion äquivalent.*

Wir setzen jetzt voraus, daß die Folge der Funktionen $\varphi_k(x)$ in $W_p^{(l)}(D)$ konvergiert. Diese Funktionen untersuchen wir auf einem beliebigen Schnitt von D . Es sei $\bar{D}$ der durch die folgenden Bedingungen definierte Zylinder: $0 \leq x_n \leq a$ (a endlich) und $(x_1, x_2, \dots, x_{n-1})$ gehöre zu der abgeschlossenen Hülle $\bar{\mathcal{E}}$ eines endlichen Gebietes $\mathcal{E}$ der $x_1, x_2, \dots, x_{n-1}$ -Ebene. Den Schnitt von $\bar{D}$ mit den Ebenen $x_n = \text{const}$ bezeichnen wir mit $\mathcal{E}_{x_n}$.

Satz 3. *Es seien $\varphi_k(x)$, $k=1, 2, \dots, n$, stetige Funktionen, die eine stetige Ableitung $\frac{\partial \varphi_k(x)}{\partial x_n}$ in D haben, und sowohl die Folge der $\varphi_k(x)$ als auch die der $\frac{\partial \varphi_k(x)}{\partial x_n}$ konvergiere in $L_p(D)$, $p>1$. Dann konvergiert die Folge der $\varphi_k(x)$ in $L_p(\mathcal{E}_{x_n})$ gleichmäßig bezüglich $x_n \in [0, a]$; dadurch ist die Grenzfunktion $\varphi(x) \in L_p(D)$ auf allen Schnitten $\mathcal{E}_{x_n}$ definiert und als Element von $L_p(\mathcal{E}_{x_n})$ stetig von x_n abhängig.*

Beweis. Wir weisen die Behauptung zuerst für $x_n \in \left[\frac{a}{2}, a\right]$ nach und wählen auf dem Intervall $[0, a]$ die Funktionen

$$\zeta(x_n) = \begin{cases} 0 & \text{für } x_n = 0, \\ 1 & \text{für } x_n \in \left[\frac{a}{2}, a\right]. \end{cases}$$

Es läßt sich leicht nachprüfen, daß die Funktionen $\psi_k(x) = \zeta(x_n) \varphi_k(x)$ und $\frac{\partial \psi_k(x)}{\partial x_n}$

in $L_p(D)$ konvergieren. Mit Hilfe der Formel

$$\psi_k(x_1, x_2, \dots, x_n) = \int_0^{x_n} \frac{\partial \psi_k(x_1, x_2, \dots, x_{n-1}, \tau)}{\partial \tau} d\tau \quad (149)$$

und der Hölderschen Ungleichung erhalten wir

$$\begin{aligned} & \int_{\mathcal{E}_{x_n}} |\psi_l(x) - \psi_k(x)|^p dx_1 \dots dx_{n-1} \\ &= \int_{\mathcal{E}_{x_n}} \left| \int_0^{x_n} \left[\frac{\partial \psi_l(x_1, \dots, x_{n-1}, \tau)}{\partial \tau} - \frac{\partial \psi_k(x_1, \dots, x_{n-1}, \tau)}{\partial \tau} \right] d\tau \right|^p dx_1 \dots dx_{n-1} \\ &\leq x_n^{p/p'} \int_{\mathcal{E}_{x_n}} \int_0^{x_n} \left| \frac{\partial \psi_l(x_1, \dots, x_{n-1}, \tau)}{\partial \tau} - \frac{\partial \psi_k(x_1, \dots, x_{n-1}, \tau)}{\partial \tau} \right|^p dx_1 \dots dx_{n-1} d\tau \\ &\leq a^{p/p'} \left\| \frac{\partial \psi_l}{\partial x_n} - \frac{\partial \psi_k}{\partial x_n} \right\|_{L_p(D)}^p. \end{aligned}$$

Daraus folgt, daß die $\psi_k(x)$ gleichmäßig bezüglich x_n gegen eine Funktion $\varphi(x)$ der Norm von $L_p(\mathcal{E}_{x_n})$ nach konvergieren. Somit erhalten wir wegen $\zeta(x_n) = 1$ für $x_n \in \left[\frac{a}{2}, a\right]$ eine Grenzfunktion $\varphi(x)$ in $L_p(\mathcal{E}_{x_n})$ für die $\varphi_k(x)$; $\varphi(x)$ ist auf jedem Schnitt $\mathcal{E}_{x_n} \left(x_n \in \left[\frac{a}{2}, a\right]\right)$ definiert. Diese Behauptung läßt sich analog für $x_n \in \left[0, \frac{a}{2}\right]$ nachweisen. Die Grenzfunktion $\varphi(x)$, die auf allen Schnitten $\mathcal{E}_{x_n}$ definiert ist, ist auch in $L_p(D)$ Grenzfunktion der $\varphi_k(x)$. Wir müssen noch beweisen, daß $\varphi(x)$ als Element aus $L_p(\mathcal{E}_{x_n})$ stetig von x_n abhängt, d. h., daß

$$\lim_{\delta \rightarrow 0} \int_{\mathcal{E}_{x_n}} |\varphi(x_1, \dots, x_{n-1}, x_n + \delta) - \varphi(x_1, \dots, x_{n-1}, x_n)|^p dx_1 \dots dx_{n-1} = 0$$

gilt. Auf Grund der Hölderschen Ungleichung ist

$$\begin{aligned} & \int_{\mathcal{E}_{x_n}} |\psi_k(x_1, \dots, x_{n-1}, x_n + \delta) - \psi_k(x_1, \dots, x_{n-1}, x_n)|^p dx_1 \dots dx_{n-1} \\ &= \int_{\mathcal{E}_{x_n}} \left| \int_{x_n}^{x_n + \delta} \frac{\partial \psi_k(x_1, \dots, x_{n-1}, x_n)}{\partial x_n} dx_n \right|^p dx_1 \dots dx_{n-1} \\ &\leq \delta^{p/p'} \int_{\mathcal{E}_{x_n}} \int_{x_n}^{x_n + \delta} \left| \frac{\partial \psi_k}{\partial x_n} \right|^p dx \leq \delta^{p/p'} \left\| \frac{\partial \psi_k}{\partial x_n} \right\|_{L_p(D)}^p \end{aligned}$$

und nach Grenzübergang

$$\int_{\mathcal{E}_{x_n}} |\psi(x_1, \dots, x_{n-1}, x_n + \delta) - \psi(x_1, \dots, x_{n-1}, x_n)|^p dx_1 \dots dx_{n-1} \leq \delta^{p/p'} \left\| \frac{\partial \psi}{\partial x_n} \right\|_{L_p(D)}^p;$$

daraus ergibt sich die geforderte Beziehung für $x_n \in \left[\frac{a}{2}, a\right]$. Das gleiche gilt auch für $x_n \in \left[0, \frac{a}{2}\right]$.

Bemerkung. Die Voraussetzung der Stetigkeit der Ableitungen $\frac{\partial \varphi_m(x)}{\partial x_n}$ läßt sich abschwächen, indem man nur die Existenz verallgemeinerter Ableitungen aus $L_p(D)$ fordert. In diesem Fall [110] gilt (149) für alle $x_n \in [0, a]$ und fast alle $(x_1, x_2, \dots, x_{n-1})$ aus $\mathfrak{E}$, und alle weiteren Überlegungen bleiben gültig. Der Grenzwert der $\frac{\partial \varphi_m(x)}{\partial x_n}$ in $L_p(D)$ ist offenbar die verallgemeinerte Ableitung $\frac{\partial \varphi(x)}{\partial x_n}$ in D . Aus dem bewiesenen Satz und den Eigenschaften der verallgemeinerten Ableitungen ergibt sich: Ist die Funktion $\varphi(x)$ etwa im Kubus Q $[0 \leq x_i \leq 1; i = 1, 2, \dots, n]$ gegeben und gehört $\varphi(x)$ nebst der verallgemeinerten Ableitung $\frac{\partial \varphi(x)}{\partial x_n}$ zu $L_p(Q)$, $p > 1$, so ist sie mit einer Funktion äquivalent, die auf jedem Schnitt des Kubus Q mit der Ebene $x_n = b$ ($0 \leq b \leq 1$) definiert ist, auf diesen Schnitten zu $L_p(\mathfrak{E}_{x_n})$ gehört und von x_n bezüglich der Norm von $L_p(\mathfrak{E})$ stetig abhängt. Dies folgt daraus, daß eine solche Funktion $\varphi(x)$ durch in $\bar{Q}$ stetig differenzierbare Funktionen $\varphi_k(x)$ approximiert werden kann [111], so wie dies in den Voraussetzungen des Satzes erwähnt wurde. Insbesondere existieren in diesem Sinn auch die Randwerte von $\varphi(x)$ auf den Rändern $x_n = 0$ und $x_n = 1$ des Kubus.

Wir nehmen nun an, die Funktion $\varphi(x)$ sei auf einem beschränkten Gebiet D gegeben und gehöre zu $W_p^{(1)}(D)$. Ferner enthalte die Grenze des Gebietes D ein $(n-1)$ -dimensionales glattes Teilgebiet S . Wir bilden den Teil von D , der sich an S anschließt, mit Hilfe einer bis an die Berandung stetig differenzierbaren Transformation $y = y(x)$ in ein Parallelepiped Q ab (S soll so gewählt sein, daß diese Transformation möglich ist). Ist S durch

$$x_n = F(x_1, \dots, x_{n-1}) \quad (\alpha \leq x_k \leq \beta; k = 1, 2, \dots, n-1)$$

definiert und gehören die Punkte $(x_1, x_2, \dots, x_n)$, die den Bedingungen

$$\alpha \leq x_k \leq \beta \quad (k = 1, 2, \dots, n-1), \quad 0 \leq x_n - F(x_1, \dots, x_{n-1}) \leq \gamma$$

genügen, zu $\bar{D}$, so können wir

$$y_1 = x_1, \quad y_2 = x_2, \dots, y_n = x_n - F(x_1, \dots, x_{n-1})$$

als neue Veränderliche y_k nehmen. Das Parallelepiped Q wird dann durch die Ungleichungen

$$\alpha \leq y_k \leq \beta \quad (k = 1, 2, \dots, n-1), \quad 0 \leq y_n \leq \gamma$$

definiert. In den neuen Veränderlichen gehört $\varphi(y)$ zu $W_p^{(1)}(Q)$; für $\varphi(y)$ gilt das oben Gesagte.

Insbesondere kann sich bei Annäherung an die Berandung T von Q , die das Bild von S ist, der Wert der Funktion $\varphi(y)$ bezüglich der Norm von $L_p(T)$ gegen die Werte von $\varphi(y)$ auf T nähern. In alten Koordinaten bedeutet dies, daß die Werte von $\varphi(x)$ auf S und die Werte von $\varphi(x)$ auf den „auf entsprechende Weise verschobenen Flächen“ im Sinne der Norm von $L_p(S)$ einander benachbart sind. In diesem Fall kann man bei den Funktionen $\varphi(x) \in W_p^{(1)}(D)$ von ihren Werten auf $(n-1)$ -dimensionalen glatten Flächen sprechen (insbesondere von ihren Werten auf $(n-1)$ -dimensionalen glatten Stücken der Berandung) und davon, daß diese Werte stetig angenommen werden.

Wir zeigen jetzt, daß für Funktionen des Raumes $W_p^{(1)}(D)$ bei den folgenden Voraussetzungen die gewöhnliche Formel für die partielle Integration gilt:

$$\int_D \frac{\partial \varphi(x)}{\partial x_k} \psi(x) dx = - \int_D \varphi(x) \frac{\partial \psi(x)}{\partial x_k} dx + \int_S \varphi(x) \psi(x) \cos(n, x_k) dS; \quad (150)$$

dabei ist n die äußere Normale auf der Berandung S von D . Wir setzen voraus, das Gebiet D lasse sich in endlich viele Gebiete D_k zerlegen, von denen jedes in bezug auf einen seiner Punkte sternförmig ist und eine stückweise glatte Berandung besitzt. Können wir die Gültigkeit von (150) für jedes der D_k beweisen, so können wir uns, wenn wir diese Gleichungen für alle D_k addieren, auch davon überzeugen, daß (150) für das ganze Gebiet D gilt. Es sei also D sternförmig, und es gehöre $\varphi(x)$ zu $W_p^{(1)}(D)$ und $\psi(x)$ zu $W_{p'}^{(1)}(D)$, $\frac{1}{p} + \frac{1}{p'} = 1$. Auf Grund des Satzes aus [111] existieren Folgen von in $\bar{D}$ stetig differenzierbaren Funktionen $\varphi_k(x)$ und $\psi_k(x)$, die gegen $\varphi(x)$ in $W_p^{(1)}(D)$ bzw. $\psi(x)$ in $W_{p'}^{(1)}(D)$ konvergieren. Wegen Satz 3 ist $\varphi_k(x) \rightarrow \varphi(x)$ in $L_p(S)$ und $\psi_k(x) \rightarrow \psi(x)$ in $L_{p'}(S)$. Für die Funktionen $\varphi_k(x)$ und $\psi_k(x)$ ist (150) richtig. Gehen wir in (150) mit k zur Grenze über, so gilt (150) auch für $\varphi(x)$ und $\psi(x)$.

Bei der Lösung von Randwertproblemen der mathematischen Physik werden Unterräume der Räume $W_p^{(l)}(D)$ betrachtet, die aus Elementen bestehen, welche homogenen Randbedingungen genügen. Zuerst wurden sie von K. FRIEDRICHS eingeführt.¹⁾ Es sei D wie oben ein beschränktes Gebiet des R_n . Mit $\dot{C}^{(l)}(D)$ bezeichnen wir die Menge der finiten, stetigen und l -mal in D stetig differenzierbaren Funktionen. In dieser linearen Mannigfaltigkeit führen wir die Norm von $W_p^{(l)}(D)$ ein, und die abgeschlossene Hülle von $\dot{C}^{(l)}(D)$ bezüglich dieser Norm bezeichnen wir mit $\dot{W}_p^{(l)}(D)$. Ist $\varphi(x) \in \dot{C}^{(l)}(D)$ und $\psi(x) \in W_{p'}^{(l)}(D)$, $\frac{1}{p} + \frac{1}{p'} = 1$, so gilt nach Definition der verallgemeinerten Ableitung

$$\int_D D^l \varphi(x) \psi(x) dx = (-1)^l \int_D \varphi(x) D^l \psi(x) dx. \quad (151)$$

Da die Elemente von $\dot{W}_p^{(l)}(D)$ die Grenzwerte der Elemente aus $\dot{C}^{(l)}(D)$ bezüglich der Norm von $W_p^{(l)}(D)$ sind, können wir behaupten, daß die Formel (151) für alle $\varphi(x) \in \dot{W}_p^{(l)}(D)$ und $\psi(x) \in W_{p'}^{(l)}(D)$ gilt. Offenbar ist $\dot{W}_p^{(l)}(D)$ in $W_p^{(l)}(D)$ enthalten. Wir sehen leicht, daß $\dot{W}_p^{(l)}(D)$ ein echter Teil von $W_p^{(l)}(D)$, $l \geq 1$, ist. Um dies zu zeigen, nehmen wir etwa $l=1$ an. Die Formel für die partielle Integration lautet in diesem Fall

$$\int_D \frac{\partial \varphi(x)}{\partial x_1} \psi(x) dx = - \int_D \varphi(x) \frac{\partial \psi(x)}{\partial x_1} dx + \int_S \varphi(x) \psi(x) \cos(n, x_1) dS,$$

wobei $\varphi(x)$ und $\psi(x)$ in $\bar{D}$ stetig differenzierbare Funktionen sind. In $W_p^{(1)}(D)$ gibt es Funktionen $\varphi(x)$ dieser Art, die auf S von 0 verschieden sind und für die das Integral über S bei entsprechender Wahl von $\psi(x)$ ebenfalls von 0 verschieden ist. Für jede Funktion $\varphi(x)$ aus $\dot{W}_p^{(1)}(D)$ gilt aber (151) mit $k=1$, wobei das Integral über S fehlt.

Die Beziehung (151) gibt uns, da $\psi(x)$ beliebig aus $W_{p'}^{(l)}(D)$ ist, die Möglichkeit zu sagen, daß die Funktionen $\varphi(x)$ aus $\dot{W}_p^{(l)}(D)$ „nebst ihren Ableitungen bis zur $(l-1)$ -ten Ordnung auf der Berandung verschwinden“.

Ist die Berandung S hinreichend glatt, so streben die Funktion $\varphi(x)$ und ihre Ableitungen bis zur $(l-1)$ -ten Ordnung bei Annäherung an S der Norm von $L_p(S)$ nach gegen 0, wie wir schon oben erwähnten.

Im allgemeinen gilt diese Randbedingung nur, wenn für alle $\varphi(x) \in \dot{W}_p^{(l)}(D)$ und $\psi(x) \in W_{p'}^{(l)}(D)$ die Formel (151) erfüllt ist. Bei der Untersuchung des Raumes $\dot{W}_p^{(l)}(D)$

¹⁾ Vgl. R. COURANT und D. HILBERT, Methoden der mathematischen Physik, Band 2, 2. Aufl., Springer-Verlag, Berlin-Heidelberg-New York 1968.

spielt die Glattheit der Berandung keine wesentliche Rolle, denn schließen wir D in das Innere einer Kugel D_1 ein und setzen wir $\varphi(x)$ in $D_1 - D$ gleich 0, so gehört $\varphi(x)$ zu $\dot{W}_p^{(l)}(D_1)$. Eine solche Fortsetzung der Funktionen aus $\dot{W}_p^{(l)}(D)$ gibt uns die Möglichkeit, stärkere Schlußfolgerungen zu ziehen, ohne das in D echt enthaltene Gebiet D' einführen zu müssen. Insbesondere:

1. Die Funktionen $\varphi(x) \in \dot{W}_p^{(l)}(D)$ können durch unendlich oft differenzierbare und in D finite Funktionen bezüglich der Norm von $\dot{W}_p^{(l)}(D)$ approximiert werden.

2. Ist $\varphi(x) \in \dot{W}_2^{(l)}(D)$ mit $l \geq \left\lfloor \frac{n}{2} \right\rfloor + 1$, so ist $\varphi(x)$ mit einer in $\bar{D}$ stetigen Funktion äquivalent, die auf der Berandung von D gleich 0 ist.

Abschließend zeigen wir, daß die abgeschlossene Hülle der Funktionen aus $\dot{C}^{(\infty)}(D)$ in $\dot{W}_p^{(l)}(D)$, $p \geq 1$, d. h. in $L_p(D)$, den ganzen Raum $L_p(D)$ ergibt. Mit anderen Worten, *im Raum $L_p(D)$ bilden die stetig differenzierbaren und finiten Funktionen eine dichte Menge*. Um dies nachzuweisen, nehmen wir an, D_δ sei die Menge derjenigen Punkte, deren Abstand von der Berandung von D nicht kleiner als δ ist und für jede Funktion $\varphi(x) \in L_p(D)$ gelte

$$\varphi^{(\delta)}(x) = \begin{cases} \varphi(x) & \text{für } x \in D_\delta, \\ 0 & \text{für } x \notin D_\delta. \end{cases}$$

Offenbar sind die Funktionen $\varphi^{(\delta)}(x)$ in $L_p(D)$ dicht, denn für $\delta \rightarrow 0$ gilt

$$\|\varphi - \varphi^{(\delta)}\|_{L_p(D)}^p = \int_{D-D_\delta} |\varphi(x)|^p dx \rightarrow 0.$$

Zu $\varphi^{(\delta)}(x)$ konstruieren wir die Mittelfunktionen $\varphi_h^{(\delta)}(x)$ für $h \leq \delta/2$. Dies sind stetig differenzierbare, in D finite Funktionen, die in $L_p(D)$ für $h \rightarrow 0$ gegen $\varphi^{(\delta)}(x)$ konvergieren; sie bilden eine in $L_p(D)$ dichte Menge. Damit haben wir gezeigt, daß für $l=0$ die Räume $\dot{W}_p^{(l)}(D)$ und $\dot{W}_p^{(l)}(D)$ übereinstimmen.

114. Einbettungssätze. Wir wollen nun die Funktionen aus $\dot{W}_p^{(l)}(D)$ eingehend betrachten und den Zusammenhang zwischen ihrem Verhalten und dem ihrer Ableitungen auf dem Gebiet D und auf seinen Schnitten verschiedener Dimensionen herstellen. Dabei erhalten wir eine Reihe wichtiger Ungleichungen, mit deren Hilfe sich die äquivalenten Normierungen des Raumes $\dot{W}_p^{(l)}(D)$ untersuchen lassen. Die Gesamtheit dieser Resultate werden im allgemeinen Sobolewsche „Einbettungssätze“ genannt. In diesem Abschnitt wollen wir diese Sätze formulieren.

Wir nehmen an, daß D ein bezüglich aller Punkte einer in D enthaltenen Kugel K sternförmiges Gebiet ist oder daß D mit Hilfe glatter Flächen in endlich viele Gebiete von sternförmigem Typ zerlegt werden kann.

Für Gebiete dieses Typs gilt die folgende Behauptung: *Hat $\varphi(x)$ in D verallgemeinerte Ableitungen der Ordnung $l \geq 1$, wobei $\varphi(x)$ und diese Ableitungen zu $L_p(D)$, $p > 1$, gehören, so hat $\varphi(x)$ verallgemeinerte Ableitungen bis zur Ordnung l , die ebenfalls zu $L_p(D)$ gehören.* Mit anderen Worten, die Funktionenklasse $\dot{W}_p^{(l)}(D)$ wird für $m < l$ sowohl in die Klasse $\dot{W}_p^{(m)}(D)$ als auch in die Klasse $\dot{W}_p^{(m)}(D)$ eingebettet. Hieraus folgt unter anderem, daß für Gebiete des obenerwähnten Typs die Klassen $\dot{W}_p^{(l)}(D)$ und $\dot{W}_p^{(l)}(D)$ ein und dieselben Funktionen aufweisen. Zur Abkürzung setzen wir

$$\|\varphi\|_{p,l} = \|\varphi\|_{\dot{W}_p^{(l)}(D)}.$$

Die Ungleichung (wir werden sie später beweisen)

$$\sum_{k=1}^{l-1} \|\varphi\|_{p,k} \leq A \|\varphi\|_{p,l}, \quad (152)$$

wobei A eine nicht von der Wahl von $\varphi(x) \in \tilde{W}_p^{(l)}(D)$ abhängige Konstante ist, besagt, daß die Normen der Räume $W_p^{(l)}(D)$ und $\tilde{W}_p^{(l)}(D)$ äquivalent sind. Im folgenden werden diese beiden Räume nicht besonders unterschieden. Die obigen Behauptungen sind Spezialfälle der allgemeineren Sätze, zu deren Formulierung wir jetzt gelangen, die wir jedoch erst in [115] bis [118] beweisen werden.

Vorbereitend führen wir den Raum $C^{(l)}(D)$ der in $\bar{D}$ stetigen Funktionen ein, welche Ableitungen bis zur l -ten Ordnung besitzen, die ebenfalls in $\bar{D}$ stetig sind; dabei wird die Norm in diesem Raum durch

$$\|\varphi\|_{C^{(l)}(D)} = \max_{\substack{x \in \bar{D} \\ 0 \leq k \leq l}} |D^k \varphi(x)|$$

definiert, wobei $D^k \varphi(x)$ eine beliebige Ableitung k -ter Ordnung ist und das Maximum für alle $x \in \bar{D}$ und alle möglichen Ableitungen bis zur l -ten Ordnung genommen wird. Unter der Stetigkeit einer Ableitung $D^k \varphi(x)$ in $\bar{D}$ verstehen wir, wie wir uns erinnern, folgendes: $\varphi(x)$ hat im Innern von D die stetige Ableitung $D^k \varphi(x)$, die auf der Berandung von D so definiert sein kann, daß sich eine in $\bar{D}$ stetige Funktion ergibt. Statt der obigen Norm in $C^{(l)}(D)$ können wir auch vermöge der Formel

$$\|\varphi\| = \sum_{k=0}^l \sum_{k_1+\dots+k_n=k} \max_{x \in \bar{D}} \left| \frac{\partial^k \varphi(x)}{\partial x_1^{k_1} \dots \partial x_n^{k_n}} \right|$$

eine äquivalente Norm einführen. Der Raum $C^{(l)}(D)$ ist offenbar ein vollständiger Banachraum. Den Raum $C^{(0)}(D)$ bezeichnen wir wie früher mit $C(D)$.

Satz 1. Ist $p > 1$ und $pl > n$, so ist jede Funktion $\varphi(x) \in W_p^{(l)}(D)$ mit einer Funktion aus $C(D)$ äquivalent, und es gilt

$$\|\varphi\|_{C(D)} \leq M \|\varphi\|_{W_p^{(l)}(D)}, \quad (153)$$

wobei M eine nur vom Gebiet D abhängige Konstante ist. Jede in $W_p^{(l)}(D)$ beschränkte Menge U ist in $C(D)$ kompakt.

Satz 2. Ist $p > 1$ und $pl \leq n$, so ist jede Funktion $\varphi(x) \in W_p^{(l)}(D)$ mit einer Funktion $\varphi(x)$ äquivalent, die fast überall auf dem Durchschnitt D_s des Gebiets mit einer beliebigen Ebene definiert ist (die Dimension von D_s sei $s > n - pl$). Ferner ist $\varphi(x)$ auf D_s mit beliebigem Exponenten q summierbar, der der Ungleichung

$$q < \frac{ps}{n - pl} \quad (154)$$

genügt, und es gilt

$$\|\varphi\|_{L_q(D_s)} \leq M_1 \|\varphi\|_{W_p^{(l)}(D)}, \quad (155)$$

wobei M_1 eine nur von D und D_s abhängige Konstante ist. Außerdem existiert zu einem beliebig vorgegebenen $\varepsilon > 0$ ein $\eta > 0$ derart, daß

$$\int_{D_s} |\varphi(x + \Delta x) - \varphi(x)|^q ds \leq \varepsilon \quad (156)$$

gilt, wenn $|\Delta x| \leq \eta$ ist und die Punkte x und $x + \tau \Delta x$ ($0 \leq \tau \leq 1$) zu D gehören; diese Zahl η ist ein und dieselbe für alle Funktionen $\varphi(x)$, deren Normen in $W_p^{(l)}(D)$ nicht größer als eine beliebige feste Zahl sind.

Hieraus folgt, daß eine in $W_p^{(l)}(D)$ beschränkte Menge in $L_q(D_s)$ kompakt ist.

Als D_s können wir sowohl einen vollständigen als auch einen nicht vollständigen s -dimensionalen Durchschnitt von D mit einer Ebene, aber auch ein in D enthalte-

nes n -dimensionales Gebiet wählen. Ist $pl = n$, so kann in Satz 2 die Zahl q beliebig, aber größer als 1 sein. Für $pl < n$ ist die rechte Seite von (154) größer als 1. Der Satz 2 bleibt auch dann gültig, wenn die Ebene durch eine glatte Fläche ersetzt wird.

Bemerkung. Auf Grund von Satz 1 ist jede Funktion $\varphi(x) \in W_p^{(l)}(D)$ für $pl > n$ auch eine Funktion aus $C(D)$, d. h., aus $W_p^{(l)}(D)$ wird sie in $C(D)$ eingebettet. Wegen (153) ist der Einbettungsoperator, der jeder Funktion aus $W_p^{(l)}(D)$ die gleiche Funktion, aber als Element aus $C(D)$, gegenüberstellt, beschränkt. Aus der letzten Behauptung von Satz 1 folgt, daß der Einbettungsoperator auch *vollstetig* ist. Eine analoge Bemerkung können wir auch für Satz 2 machen.

Wir gehen nun noch auf einige Folgerungen aus den Einbettungssätzen ein. Ist $pl > n$ und genügt eine ganze Zahl m der Ungleichung $0 < m < l - \frac{n}{p}$, so ist jede Funktion $\varphi(x) \in W_p^{(l)}(D)$ in $\bar{D}$ stetig differenzierbar bis zur Ordnung m , wobei die Funktionen $D^k \varphi(x)$ für $k \leq m$ den entsprechenden, in $\bar{D}$ stetigen Ableitungen von $\varphi(x)$ äquivalent sind, und es existiert eine nur von D abhängige Zahl A , für die

$$|D^k \varphi(x)| \leq A \|\varphi\|_{W_p^{(l)}(D)} \quad (157)$$

gilt. Daraus folgt, daß $W_p^{(l)}(D)$ für $pl > n$ ein Teil des Raumes derjenigen Funktionen ist, die in D stetig differenzierbar bis zur Ordnung $l - \left[\frac{n}{p}\right] - 1$ sind (also ein Teil des Raumes $C^{l - \left[\frac{n}{p}\right] - 1}$). Ist $m \geq 0$, $m \leq l - \frac{n}{p}$ und $s > n - (l - m)p$, so gilt auf jeder hinreichend glatten s -dimensionalen Mannigfaltigkeit, F_s aus D

$$D^m \varphi(x) \in L_q(F_s) \quad \text{für} \quad q < \frac{sp}{n - (l - m)p}, \quad (158)$$

und es existiert eine nur von D und F_s abhängige Zahl A_1 , für die

$$\|D^m \varphi(x)\|_{L_q(D_s)} \leq A_1 \|\varphi\|_{W_p^{(l)}(D)} \quad (159)$$

gilt. In allen diesen Fällen sind die Einbettungsoperatoren *vollstetig*.

Die Sätze 1 und 2 erlauben, verschiedene Normen in $W_p^{(l)}(D)$ zu konstruieren, die der Grundnorm (145) oder (138) äquivalent sind. Ein allgemeineres Resultat in dieser Richtung liefert der

Satz 3. *Gegeben seien lineare und in $W_p^{(l)}(D)$ beschränkte Funktionale $l_k(u)$, $k = 1, 2, \dots, N$, derart, daß sie auf keinem nicht identisch verschwindenden Polynom von höchstens $(l - 1)$ -tem Grad alle gleichzeitig gleich 0 sind. Dann definiert die Formel*

$$\|\varphi\|^p = \sum_{l_1 + l_2 + \dots + l_N = l} \|D^l \varphi\|_{L_p(D)}^p + \sum_{k=1}^N |l_k(\varphi)|^p \quad (160)$$

eine zur Grundnorm (144) oder (138) äquivalente Norm.

Die Norm (160) ist auf Grund des in [112] Gesagten z. B. mit der Norm

$$\|\varphi\| = \sum_{l_1 + l_2 + \dots + l_N = l} \|D^l \varphi\|_{L_p(D)} + \sum_{k=1}^N |l_k(\varphi)| \quad (161)$$

und anderen analogen Ausdrücken äquivalent.

Wir geben nun einige Beispiele für äquivalente Normen im Raum $W_p^{(l)}(D)$ an.

Aus Satz 3 folgt, daß wir in $W_p^{(1)}(D)$ eine Norm mit Hilfe von

$$\|\varphi\| = \sum_{k=1}^n \left\| \frac{\partial \varphi}{\partial x_k} \right\|_{L_p(D)} + \left| \int_D \varphi(x) dx \right| \quad (162)$$

einführen können; das Funktional $\int_D \varphi(x) dx$ ist nämlich in $W_p^{(1)}(D)$ auf Grund der Ungleichung

$$\left| \int_D \varphi(x) dx \right| \leq \|\varphi\|_{L_p(D)} (mD)^{1/p'} \leq \|\varphi\|_{W_p^{(1)}(D)} (mD)^{1/p'}$$

linear, und $\int c dx$ ist für jede Konstante $c \neq 0$ von 0 verschieden (mD bezeichnet das Maß von D).

Aus Satz 2 folgt, daß die Norm

$$\|\varphi\| = \sum_{k=1}^n \left\| \frac{\partial \varphi}{\partial x_k} \right\|_{L_p(D)} + \|\varphi\|_{L_p(D)}, \quad (163)$$

mit der Norm (138) für jedes q , das der Ungleichung $1 \leq q < \frac{pn}{n-p}$ ($p < n$) genügt, äquivalent ist. Für $p=n$ kann man ein beliebiges $q \geq 1$ wählen. Ist $p > n$, so kann man ebenfalls ein beliebiges $q \geq 1$ wählen oder sogar den zweiten Summanden von (163) durch $\|\varphi\|_{C(D)}$ ersetzen. Diese Bemerkung gilt auch für die folgende Beziehung (164). Der Leser kann leicht nachprüfen, daß der Ausdruck

$$\|\varphi\| = \sum_{k=1}^n \left\| \frac{\partial \varphi}{\partial x_k} \right\|_{L_q(D)} + \|\varphi\|_{L_q(D)} \quad (164)$$

ebenfalls eine zu (138) äquivalente Norm definiert; dabei ist S eine beliebige, in D enthaltene glatte $(n-1)$ -dimensionale Mannigfaltigkeit, und der Exponent q genügt der Bedingung $1 \leq q < \frac{p(n-1)}{n-p}$ ($p < n$). Analoge Überlegungen führen zu verschiedenartigen äquivalenten Normierungen des Raumes $W_p^{(l)}(D)$ für $l \geq 1$.

115. Integraloperatoren mit polaren¹⁾ Kernen. Ehe wir zu der Darlegung der Beweise der in [114] angegebenen Einbettungssätze kommen, müssen wir zunächst noch die Klasse der Integraloperatoren mit polaren Kernen untersuchen. Wie üblich bezeichnen wir mit D ein beschränktes Gebiet des n -dimensionalen Raumes R_n . Es sei σ_n die $(n-1)$ -dimensionale Einheitskugel im R_n und $|\sigma_n|$ ihr Flächeninhalt. Für ein Volumenelement des R_n gilt nach [II, 186]

$$dx = dx_1 dx_2 \dots dx_n = r^{n-1} dr d\sigma_n$$

mit

$$d\sigma_n = \sin^{n-2} \theta_1 \sin^{n-3} \theta_2 \dots \sin \theta_{n-2} d\theta_1 d\theta_2 \dots d\theta_{n-2} d\varphi.$$

Satz 1. Es sei $B(x, y)$ ein für $x, y \in D$ beschränkter und für $x \neq y$ stetiger Kern. Der Integraloperator

$$u(x) = \int_D \frac{B(x, y)}{|x-y|^\lambda} f(y) dy \quad (165)$$

ist dann für $\lambda < \frac{n}{p} \left(\frac{1}{p} + \frac{1}{p'} = 1 \right)^2$ als Operator von $L_p(D)$ in $C(D)$ vollstetig (und damit beschränkt).

¹⁾ Vgl. [IV/1, 16] (Anm. d. Red.).

²⁾ Hier und im folgenden wird $p > 1$ vorausgesetzt.

Beweis. Nach Voraussetzung ist $|B(x, y)| \leq B = \text{const.}$ Auf Grund der Hölderschen Ungleichung ist

$$|u(x)| \leq B \left[\int_D |f(y)|^p dy \right]^{1/p} \left[\int_{K_R} r^{-\lambda p'} dy \right]^{1/p'},$$

wobei K_R die das Gebiet D enthaltende Kugel mit dem Radius R und dem Koordinatenursprung als Mittelpunkt und $r = |x - y|$ ist. Gehen wir zu Kugelkoordinaten mit dem Koordinatenursprung im Punkt x über, so erhalten wir

$$\left[\int_{K_R} r^{-\lambda p'} dy \right]^{1/p'} \leq \left[\int_0^{2R} r^{n-1-\lambda p'} dr \int_{\sigma_n} d\sigma_n \right]^{1/p'} \quad (166)$$

und folglich

$$|u(x)| \leq B \left(\frac{|\sigma_n|}{n - \lambda p'} \right)^{1/p'} (2R)^{(n/p') - \lambda} \|f\|_{L_p(D)}. \quad (167)$$

Diese Abschätzung besagt: Ist $\|f\|_{L_p(D)} \leq c$, wobei c eine gewisse Konstante ist, so sind die entsprechenden Funktionen $u(x)$ in D gleichmäßig beschränkt. Zum Beweis des Satzes genügt es also zu zeigen, daß die Funktionen $u(x)$ gleichgradig stetig sind. Es sei $\delta > 0$ eine hinreichend kleine Zahl und $D^{(\delta)}$ die Menge derjenigen Punkte $y \in D$, für die $|y - x| \leq \delta$ gilt. Für $x \in D$ und $x + \Delta x \in D$ ergibt sich

$$\begin{aligned} |u(x + \Delta x) - u(x)| &\leq \int_{D^{(\delta)}} \left| \frac{B(x, y)}{|x - y|^\lambda} - \frac{B(x + \Delta x, y)}{|x + \Delta x - y|^\lambda} \right| \cdot |f(y)| dy \\ &\quad + B \int_{D - D^{(\delta)}} \frac{|f(y)|}{|x - y|^\lambda} dy + B \int_{D - D^{(\delta)}} \frac{|f(y)|}{|x + \Delta x - y|^\lambda} dy, \end{aligned}$$

wobei wir $|\Delta x| \leq \delta/2$ annehmen. Für ein beliebig vorgegebenes $\varepsilon > 0$ gibt es dann ein $\eta > 0$ derart, daß der Absolutbetrag der im ersten Integranden stehenden Differenz höchstens gleich ε ist, wenn $|\Delta x| \leq \eta$ gilt. Mit Hilfe der Hölderschen Ungleichung erhalten wir für das erste Integral die Abschätzung $\varepsilon |D|^{1/p'} c$, wobei $|D|$ das Maß von D ist. Das zweite Integral läßt sich mit Hilfe von (167) für $2R = \delta$ abschätzen, und das dritte Integral ist nicht größer als das entsprechende Integral über den Teil der Kugel

$$|y - (x + \Delta x)| \leq \frac{3\delta}{2},$$

der in D liegt; die Abschätzung dieses Integrals läßt sich ebenfalls mit Hilfe von (167) für $2R = \frac{3\delta}{2}$ durchführen. Schließlich erhalten wir

$$|u(x + \Delta x) - u(x)| \leq \varepsilon |D|^{1/p'} c + B \left(\frac{|\sigma_n|}{n - \lambda p'} \right)^{1/p'} \left[\delta^{(n/p') - \lambda} + \left(\frac{3\delta}{2} \right)^{(n/p') - \lambda} \right].$$

Daraus ergibt sich, da δ und ε beliebig sind, die gleichgradige Stetigkeit von $u(x)$ für $\|f\|_{L_p(D)} \leq c$, was zu beweisen war.

Satz 2. Es sei $n > \lambda \geq \frac{n}{p'}$, $s > n - (n - \lambda)p$ (oder, was dasselbe ist, $\frac{s}{p} > \lambda - \frac{n}{p'}$) und $s \leq n$; dabei sei s eine ganze Zahl. Der Integraloperator (165) ist dann als Operator von $L_p(D)$ in $L_q(D_s)$ vollstetig, wobei D_s ein beliebiger s -dimensionaler Schnitt von D mit einer Ebene und q eine beliebige Zahl ist, die der Ungleichung

$$q < q^* = \frac{sp}{n - (n - \lambda)p}$$

genügt. Setzen wir voraus, $B(x, y)$ sei in einem das Gebiet D im Innern enthaltenden Gebiet D_1 definiert und besitze in D_1 die oben erwähnten Eigenschaften, dann erhalten wir außerdem

$$\|u(x + \alpha x^0) - u(x)\|_{L_q(D_s)} \leq \varepsilon(\alpha) \|f\|_{L_p(D)}, \quad (168)$$

wobei x^0 einen festen Vektor bezeichnet und $\varepsilon(\alpha)$ für $0 \leq \alpha \leq \delta$ ($\delta > 0$) stetig, für $\alpha = 0$ gleich 0 ist und durch die Konstante B sowie durch die Dimensionen von D und D_s bestimmt ist. Bei $s = n$ kann für D_s ein beliebiges Teilgebiet von D genommen werden.

Bemerkung. Wir brauchen nicht vorauszusetzen, daß $B(x, y)$ in einem größeren Gebiet D_1 definiert ist, aber dann müssen wir in (168) die Verschiebungen αx^0 als zulässig annehmen, d. h., die Punkte $x + \alpha x^0$ sollen in D liegen. Wir weisen noch darauf hin, daß die Zahl q für $n - (n - \lambda)p = 0$ beliebig gewählt werden kann.

Den Beweis von Satz 2 führen wir in zwei Schritten.

Hilfssatz 1. Unter den erwähnten Bedingungen ist der Operator (165) als Operator von $L_p(D)$ in $L_q(D_s)$ beschränkt.

Auf Grund der Voraussetzungen von Satz 2 ist $q^* > p$, und wir nehmen vorläufig $q > p$ an. Dann gilt

$$\frac{s}{q} > \frac{s}{q^*} = \lambda - \frac{n}{p'}, \quad \text{also} \quad \lambda = \frac{s}{q} + \frac{n}{p'} - 2\beta \quad (\beta > 0),$$

und wir setzen

$$\alpha_1 = \frac{1}{q}, \quad \alpha_2 = \frac{1}{p} - \frac{1}{q}, \quad \alpha_3 = \frac{1}{p'} \quad (\alpha_1 + \alpha_2 + \alpha_3 = 1).$$

Wenden wir zweimal die Höldersche Ungleichung auf ein Produkt von drei Faktoren an, so gelangen wir zu der Ungleichung

$$\int_D |f_1 f_2 f_3| dx \leq \left[\int_D |f_1|^{\alpha_1} dx \right]^{\alpha_1} \left[\int_D |f_2|^{\alpha_2} dx \right]^{\alpha_2} \left[\int_D |f|^{\alpha_3} dx \right]^{\alpha_3}.$$

Wenden wir dies auf das Integral in der offenbar richtigen Abschätzung

$$|u(x)| \leq B \int_{K_R} (|f(y)|^{\frac{p}{q}} r^{-\frac{s}{q} + \beta}) \left(|f(y)|^{p(\frac{1}{p} - \frac{1}{q})} \right) (r^{-\frac{n}{p'} + \beta}) dy$$

an, wobei $f(y)$ auf den Teil der Kugel, der außerhalb von D liegt, durch die Funktion 0 fortgesetzt ist, so erhalten wir

$$|u(x)| \leq B \left[\int_{K_R} |f(y)|^p r^{-s+q\beta} dy \right]^{1/q} \left[\int_{K_R} |f(y)|^p dy \right]^{(1/p) - (1/q)} \left[\int_{K_R} r^{-n+p'\beta} dy \right]^{1/p'}.$$

Der zweite Faktor auf der rechten Seite ist gleich $\|f\|_{L_p(D)}^{1 - (p/q)}$, der dritte läßt sich bekanntlich durch

$$\int_{K_R} r^{-n+p'\beta} dy \leq \frac{|\sigma_n| (2R)^{p'\beta}}{p'\beta}$$

abschätzen, so daß wir

$$|u(x)| \leq B \left(\frac{|\sigma_n|}{p'\beta} \right)^{1/p'} (2R)^\beta \|f\|_{L_p(D)}^{1 - (p/q)} \left[\int_{K_R} |f(y)|^p r^{-s+q\beta} dy \right]^{1/q}$$

und

$$\int_{D_s} |u(x)|^q dx^{(s)} \leq B^q \left(\frac{|\sigma_n|}{p'\beta} \right)^{q/p'} (2R)^{q\beta} \|f\|_{L_p(D)}^{q - \frac{p}{q}} \int_{D_s} dx^{(s)} \int_{K_R} |f(y)|^p r^{-s+q\beta} dy \quad (169)$$

erhalten; das Differential $dx^{(s)}$ bezieht sich auf D_s . Wir ändern die Reihenfolge der Integration und wollen das innere Integral von $r^{-s+q\beta}$ über D_s für konstantes y ab-

schätzen. Zunächst bemerken wir, daß

$$-s + q\beta = -q \left(\frac{s}{q} - \beta \right) = -q \left(\lambda - \frac{n}{p'} + \beta \right) < 0$$

gilt. Führen wir in D_s Kugelkoordinaten ein, wobei der Koordinatenursprung die Projektion von y auf D_s sei, so ergibt sich $dx^{(s)} = \varrho^{s-1} d\varrho d\sigma_s$ und $r^{-s+q\beta} \leq \varrho^{-s+q\beta}$, denn es ist $\varrho \leq r$ und $-s + q\beta < 0$. Daraus folgt

$$\int_{D_s} r^{-s+q\beta} dx^{(s)} \leq \int_{D_s} \varrho^{-s+q\beta} \varrho^{s-1} d\varrho d\sigma_s = \frac{|\sigma_s| (2R)^{q\beta}}{q\beta}$$

und auf Grund von (169)

$$\|u\|_{L_q(D_s)} \leq (2R)^{2\beta} c \|f\|_{L_p(D)}, \quad (170)$$

wobei

$$c = B \left(\frac{|\sigma_n|}{p'\beta} \right)^{1/p'} \left(\frac{|\sigma_s|}{q\beta} \right)^{1/q} \quad (171)$$

nicht von den Abmessungen von D abhängt.

Nimmt man nicht wie bisher $q > p$, sondern $q \leq p$ an, so braucht man nur die Ungleichung

$$\|u\|_{L_q(D_s)} \leq \|u\|_{L_{q_1}(D_s)} |D_s|^{1-(1/q_1)} \quad (p < q_1 < q^*) \quad (172)$$

zu benutzen ($|D_s|$ ist das Maß von D_s), die unmittelbar aus der Hölderschen Ungleichung folgt; dann tritt in dem Ausdruck für c noch der Faktor $|D_s|^{1-(q/q_1)}$ auf.

Hilfssatz 2. *Der Operator (165) ist bezüglich einer Verschiebung in dem in Satz 2 genannten Sinne stetig.*

Wie in Satz 1 ist

$$\begin{aligned} |u(x + \Delta x) - u(x)| &\leq \int_{D(s)} \left| \frac{B(x, y)}{|y - x|^\lambda} - \frac{B(x + \Delta x, y)}{|y - (x + \Delta x)|^\lambda} \right| \cdot |f(y)| dy \\ &\quad + B \int_{|y-x| \leq \delta} \frac{|f(y)|}{|y-x|^\lambda} dy + B \int_{|y-x| \leq \delta} \frac{|f(y)|}{|y-(x+\Delta x)|^\lambda} dy, \quad (173) \end{aligned}$$

wobei auch jetzt $|\Delta x| \leq \delta/2$ sei. Zu einem vorgegebenen $\varepsilon > 0$ existiert ein $\eta > 0$ derart, daß wir für $|\Delta x| \leq \eta$

$$\|u(x + \Delta x) - u(x)\|_{L_q(D_s)} \leq \varepsilon |D|^{1/p'} |D_s|^{1/q} \|f\|_{L_p(D)} + \dots$$

erhalten; die Punkte bedeuten die Summe der Normen des zweiten und des dritten Summanden auf der rechten Seite von (173). Für diese Normen gilt die Abschätzung (170) mit $2R = \delta$ und $2R = 3\delta/2$, so daß also

$$\|u(x + \Delta x) - u(x)\|_{L_q(D_s)} \leq (c_1 \varepsilon + c_2 \delta^{2\beta}) \|f\|_{L_p(D)} \quad (174)$$

mit

$$c_1 = |D|^{1/p'} |D_s|^{1/q}, \quad c_2 = B \left(\frac{|\sigma_n|}{p'\beta} \right)^{1/p'} \left(\frac{|\sigma_s|}{q\beta} \right)^{1/q} \left[1 + \left(\frac{3}{2} \right)^{2\beta} \right] \quad (q > p)$$

folgt. Ist $q \leq p$, so tritt in dem Ausdruck für c_2 noch ein zusätzlicher Faktor auf.

Für den Beweis von Satz 2 genügt es zu zeigen, daß sich für $\|f\|_{L_p(D)} \leq A = \text{const}$ eine in $L_q(D_s)$ kompakte Menge von Funktionen $u(x)$ findet. Die Beschränktheit von $u(x)$

in $L_q(D_s)$ folgt aus (170), die gleichgradige Stetigkeit in $L_q(D_s)$ aus (174), wenn wir annehmen, daß Δx in D_s liegt, und beachten, daß die für ε definierte Zahl $\eta > 0$ nur vom Kern $B(x, y)$ und nicht von $f(y)$ abhängt. Damit ist Satz 2 bewiesen.

Satz 3. Es seien $B_1(x, y)$ und $B_2(x, y)$ für $x, y \in \bar{D}$ beschränkte und für $x \neq y$ stetige Kerne. Dann läßt sich das Integral

$$I(x, z) = \int_D \frac{B_1(x, y) B_2(x, y)}{|x-y|^\lambda |y-z|^\mu} dy \quad (\lambda < n, \mu < n) \quad (175)$$

für $x, z \in \bar{D}$ in der Gestalt

$$I(x, z) = B(x, z) \varphi(|x-z|) \quad (176)$$

darstellen, wobei $B(x, z)$ die gleichen Eigenschaften wie $B_i(x, y)$, $i = 1, 2$, hat und

$$\varphi(\xi) = \begin{cases} \xi^{-(\lambda+\mu-n)} & \text{für } \lambda + \mu > n, \\ 1 + |\lg \xi| & \text{für } \lambda + \mu = n, \\ 1 & \text{für } \lambda + \mu < n \end{cases} \quad (177)$$

gilt.

Beweis. Es sei $x \neq z$. Das Gebiet D zerlegen wir in zwei Teile, $D = D_1 + D_2$, und zwar so, daß von den beiden Punkten x und z der Punkt x in D_1 und der Punkt z in D_2 liegt. Wählen wir $p' > 1$ so, daß $\lambda < n/p'$ ist, und setzen wir $f(y) = B_2(y, z)/|y-z|^\mu$ in (165), so sehen wir mit Hilfe von Satz 1, daß das Integral (175), über D_1 genommen, eine stetige Funktion von x und z ist. Analoges gilt für das Integral über D_2 . Damit ist die Stetigkeit von $I(x, z)$ für $x \neq z$ bewiesen.

Für den Beweis von (176) brauchen wir nur die Gültigkeit der Ungleichung

$$\int_D \frac{dy}{|x-y|^\lambda |y-z|^\mu} \leq c \varphi(|x-z|) \quad (178)$$

nachzuweisen, wobei c eine positive Konstante ist. Dazu müssen wir drei Fälle unterscheiden.

1. Es sei $\lambda + \mu > n$. Wir setzen $|x-z| = \delta$ und führen neue Koordinaten

$$x' = \frac{x}{\delta}, \quad y' = \frac{y}{\delta}, \quad z' = \frac{z}{\delta}$$

ein; damit ist $|x'-z'| = 1$. Bezeichnen wir wie üblich mit R_n den n -dimensionalen Raum, so erhalten wir

$$\int_D \frac{dy}{|x-y|^\lambda |y-z|^\mu} \leq \int_{R_n} \frac{dy}{|x-y|^\lambda |y-z|^\mu} = \frac{\delta^n}{\delta^\lambda \delta^\mu} \int_{R_n} \frac{dy'}{|x'-y'|^\lambda |y'-z'|^\mu}.$$

Zur Berechnung des Integrals über R_n legen wir den Koordinatenursprung in den Punkt x' und wählen die y'_1 -Achse (die erste Komponente von y') in Richtung von x' nach z' . Dann erhalten wir

$$\int_D \frac{dy}{|x-y|^\lambda |y-z|^\mu} \leq \delta^{-(\lambda+\mu-n)} \int_{R_n} \frac{dy'}{|y'|^\lambda |y'-z'_0|^\mu},$$

wobei z'_0 die Koordinaten $(1, 0, 0, \dots, 0)$ hat. Das letzte Integral konvergiert wegen $\lambda < n$, $\mu < n$ und $\lambda + \mu > n$ und hängt offenbar nicht von x und z ab. Damit ergibt sich

$$\int_D \frac{dy}{|x-y|^\lambda |y-z|^\mu} \leq c |x-z|^{-(\lambda+\mu-n)}.$$

2. Der Fall $\lambda + \mu = n$. Wir führen eine D enthaltende Kugel K_R mit dem Radius R und dem Mittelpunkt im Koordinatenursprung ein und benutzen dieselben Koordinaten wie im Fall 1. Dann erhalten wir

$$\int_D \frac{dy}{|x-y|^\lambda |y-z|^\mu} \cong \int_{K_R} \frac{dy}{|x-y|^\lambda |y-z|^\mu} = \int_{K_{R/\delta}} \frac{dy'}{|x'-y'|^\lambda |y'-z'|^\mu} \cong \int_{K_{2R/\delta}} \frac{dy'}{|y'|^\lambda |y'-z'_0|^\mu}.$$

Im letzten Integral verdoppelten wir den Radius, aber dafür können wir wie früher über die Kugel mit dem Zentrum im Koordinatenursprung integrieren. Wir können ein hinreichend kleines $\delta > 0$ nehmen, und es sei $2R/\delta > 2$. Das Integral über $K_{2R/\delta}$ zerlegen wir in zwei Integrale, und zwar in ein Integral über K_2 und ein Integral über $K_{2R/\delta} - K_2$. Die Integration über K_2 ergibt eine gewisse positive Konstante c_1 . Also bleibt noch das Integral

$$I_1 = \int_{K_{2R/\delta} - K_2} \frac{dy'}{|y'|^\lambda |y'-z'_0|^\mu}$$

abzuschätzen. Auf Grund von $|y'-z'_0| \geq |y'| - 1$ ergibt sich, wenn wir $|y'| = r$ setzen,

$$I_1 \leq |\sigma_n| \int_2^{2R/\delta} \frac{r^{n-1}}{r^\lambda (r-1)^\mu} dr = |\sigma_n| \int_2^{2R/\delta} \frac{dr}{r \left(1 - \frac{1}{r}\right)^\mu} \quad (\lambda + \mu = n)$$

oder wegen $r \geq 2$

$$I_1 \leq |\sigma_n| 2^\mu \int_2^{2R/\delta} \frac{dr}{r} < |\sigma_n| 2^\mu \lg \frac{2R}{\delta}.$$

Also ist

$$\int_D \frac{dy}{|y-x|^\lambda |z-y|^\mu} \leq c_1 + c_2 |\lg \delta| \leq c (1 + \lg |x-z|), \quad c \geq c_1, \quad c \geq c_2,$$

woraus (178) für $\lambda + \mu = n$ folgt.

3. Die Untersuchung des Falls $\lambda + \mu < n$ verläuft analog den vorhergehenden Fällen. In diesem Fall konvergiert das Integral (178) auch für $x = z$. Ähnlich wie bei Fall 2 gilt die Abschätzung

$$\begin{aligned} \int_D \frac{dy}{|x-y|^\lambda |y-z|^\mu} &\leq \delta^{n-\lambda-\mu} \int_{K_{2R/\delta}} \frac{dy'}{|y'|^\lambda |y'-z'_0|^\mu} \leq \delta^{n-\lambda-\mu} \left[c_1 + |\sigma_n| \int_2^{2R/\delta} \frac{r^{n-1} dr}{r^{\lambda+\mu} \left(1 - \frac{1}{r}\right)^\mu} \right] \\ &\leq \delta^{n-\lambda-\mu} \left[c_1 + |\sigma_n| 2^\mu \int_2^{2R/\delta} r^{n-\lambda-\mu-1} dr \right] < \delta^{n-\lambda-\mu} \left[c_1 + \frac{|\sigma_n| 2^\mu}{n-\lambda-\mu} \left(\frac{2R}{\delta} \right)^{n-\lambda-\mu} \right]. \end{aligned}$$

Also gilt

$$\int_D \frac{dy}{|x-y|^\lambda |y-z|^\mu} \leq c.$$

Damit ist Satz 3 bewiesen.

116. Sobolewsche Integraldarstellungen. Wir setzen jetzt voraus, daß das Gebiet D bezüglich aller Punkte einer im Innern von D enthaltenen Kugel K sternförmig ist. Der Mittelpunkt dieser Kugel liege im Koordinatenursprung, ihr Radius sei R . Wir führen

die unendlich oft differenzierbare Funktion

$$p(y) = \begin{cases} c \exp \left[-\frac{R^2}{R^2 - |y|^2} \right] & \text{für } |y| < R, \\ 0 & \text{für } |y| \geq R \end{cases} \quad (179)$$

ein [71], wobei die Konstante c so gewählt sei, daß das Integral von $p(y)$ über K gleich 1 wird.

Es sei $u(y)$ eine beliebige, stetige und in $\bar{D}$ stetig differenzierbare Funktion. Führen wir Kugelkoordinaten mit dem Anfangspunkt in einem gewissen Punkt x aus D ein, so können wir $u(y)$ als Funktion von x , r und ω_n betrachten, wobei mit ω_n die Gesamtheit der Winkelkoordinaten auf der Einheitskugel bezeichnet ist [IV/2, 59], d. h., es ist $u(y) = u(x, r, \omega_n)$ mit $u(x, 0, \omega_n) = u(x)$.

Wir wollen nun das Produkt $u(y) p(y)$ über D integrieren. Praktisch bleibt die Integration auf die Kugel K beschränkt. Schreiben wir $p(y)$ in der Form $p(x, r, \omega_n)$ und integrieren wir partiell nach r , so erhalten wir

$$\begin{aligned} \int_D u(y) p(y) dy &= \int_{\omega_n} \left[\int_0^\infty u(x, r, \omega_n) p(x, r, \omega_n) r^{n-1} dr \right] d\omega_n \\ &= - \int_{\omega_n} \left\{ \int_0^\infty u(x, r, \omega_n) dr \left[\int_r^\infty p(x, \varrho, \omega_n) \varrho^{n-1} d\varrho \right] \right\} d\omega_n \\ &= - \int_{\omega_n} \left[u(x, r, \omega_n) \int_r^\infty p(x, \varrho, \omega_n) \varrho^{n-1} d\varrho \right]_{r=0}^{r=\infty} d\omega_n \\ &\quad + \int_{\omega_n} \left\{ \int_0^\infty \frac{\partial u(x, r, \omega_n)}{\partial r} \left[\int_r^\infty p(x, \varrho, \omega_n) \varrho^{n-1} d\varrho \right] dr \right\} d\omega_n \\ &= u(x) + \int_{\omega_n} \left[\int_0^\infty \frac{1}{r^{n-1}} \frac{\partial u(x, r, \omega_n)}{\partial r} \left(\int_r^\infty p(x, \varrho, \omega_n) \varrho^{n-1} d\varrho \right) r^{n-1} dr \right] d\omega_n \end{aligned}$$

und schließlich

$$\int_D u(y) p(y) dy = u(x) - \int_D \frac{\partial u(x, r, \omega_n)}{\partial r} \frac{1}{r^{n-1}} B(x, y) dy \quad (180)$$

mit

$$B(x, y) = - \int_r^\infty p(x, \varrho, \omega_n) \varrho^{n-1} d\varrho; \quad (181)$$

dabei hat der Punkt y die Kugelkoordinaten (r, ω_n) mit dem Koordinatenursprung in x . Die Funktion $B(x, y)$ ist für $x, y \in \bar{D}$ beschränkt und für $x \neq y$ stetig. Strebt y auf einem Strahl gegen x , so hat $B(x, y)$ den Grenzwert

$$- \int_r^\infty p(x, \varrho, \omega_n) \varrho^{n-1} d\varrho,$$

der nur von den Winkelkoordinaten des Strahls abhängt. Aus der Definition von $B(x, y)$ folgt unmittelbar: Gehört x zur Kugel K , so ist $B(x, y) = 0$ für diejenigen y , die in D , aber nicht in K liegen; befindet sich x im Äußeren von K , so ist $B(x, y) = 0$ für diejenigen y , die zu D gehören und außerhalb des Gebietes liegen, das gebildet wird von K und dem Teil des Tangentialkegels mit der Spitze in x an die Kugel, die sich zwischen x und K befindet. Beachten wir die Formel

$$\frac{\partial u(y)}{\partial r} = \sum_{i=1}^n \frac{\partial u(y)}{\partial y_i} \cos(r, y_i) = \sum_{i=1}^n \frac{\partial u(y)}{\partial y_i} \frac{y_i - x_i}{|y - x|},$$

so erhalten wir wegen (180)

$$u(x) = U + \sum_{i=1}^n \int_D \frac{\partial u(y)}{\partial y_i} \frac{B_i(x, y)}{|x-y|^{n-1}} dy \quad (182)$$

mit

$$U = \int_D u(y) p(y) dy, \quad B_i(x, y) = B(x, y) \frac{y_i - x_i}{|y-x|}. \quad (183)$$

Die Kerne $B_i(x, y)$ haben offenbar die gleichen Eigenschaften wie der Kern $B(x, y)$: Sie sind für $x, y \in \bar{D}$ beschränkt, für $x \neq y$ stetig und in dem oben erwähnten Teil von D gleich 0.

Wir nehmen nun an, daß $u(y)$ stetig ist und stetige Ableitungen bis zur Ordnung l in D hat, und wollen eine Formel herleiten, durch welche die Funktion $u(x)$ mit Hilfe ihrer Ableitungen l -ter Ordnung ausgedrückt wird. Benutzen wir (182), so können wir

$$\frac{\partial u(y)}{\partial y_i} = U_i + \sum_{k=1}^n \int_D \frac{\partial^2 u(z)}{\partial z_i \partial z_k} \frac{B_k(y, z)}{|y-z|^{n-1}} dz \quad (184)$$

mit

$$U_i = \int_D \frac{\partial u(y)}{\partial y_i} p(y) dy \quad (185)$$

schreiben. Setzen wir (184) in (182) ein und benutzen wir den Satz 3 aus [115], so finden wir

$$u(x) = U + \sum_{i=1}^n U_i b_i(x) + \sum_{i,k=1}^n \frac{\partial^2 u(z)}{\partial z_i \partial z_k} \frac{B_{ik}(x, z)}{|x-z|^{n-2}} dz, \quad (186)$$

wobei

$$b_i(x) = \int_D \frac{B_i(x, y)}{|x-y|^{n-1}} dy, \quad \frac{B_{ik}(x, z)}{|x-z|^{n-2}} = \int_D \frac{B_i(x, y) B_k(y, z)}{|x-y|^{n-1} |y-z|^{n-1}} dy \quad (187)$$

ist, die Kerne $B_{ik}(x, z)$ dieselben Eigenschaften wie der Kern $B(x, y)$ haben und die $b_i(x)$ auf Grund von Satz 1 aus [115] ($f(y) \equiv 1$) in D stetig sind. Analog ergibt sich die Darstellung der Funktion $u(x)$ mit Hilfe ihrer l -ten Ableitungen:

$$u(x) = U + \sum_{k=1}^{l-1} \sum_{i_1, \dots, i_k=1}^n U_{i_1, \dots, i_k} b_{i_1, \dots, i_k}(x) + \sum_{i_1, \dots, i_l=1}^n \int_D \frac{\partial^l u(y)}{\partial y_{i_1} \dots \partial y_{i_l}} \frac{B_{i_1, \dots, i_l}(x, y)}{|x-y|^{n-l}} dy; \quad (188)$$

die $b_{i_1, \dots, i_k}(x)$ sind in $\bar{D}$ stetig, die Kerne $B_{i_1, \dots, i_l}(x, y)$ besitzen dieselben Eigenschaften wie der Kern $B(x, y)$, es ist

$$U_{i_1, \dots, i_k} = \int_D \frac{\partial^k u(y)}{\partial y_{i_1} \dots \partial y_{i_k}} p(y) dy, \quad (189)$$

und die Summation über $i_1, \dots, i_l$ ist die Summation über alle Arten von Ableitungen l -ter Ordnung. Im Fall $n < l$ sind die Kerne in der Darstellung (188) beschränkt [115].

Wir weisen noch darauf hin, daß die Größen $U_{i_1, \dots, i_k}$ lineare Funktionale in $L_p(D)$ sind. Davon kann man sich überzeugen, wenn man sie in der Gestalt

$$U_{i_1, \dots, i_k} = (-1)^k \int_D u(y) \frac{\partial^k p(y)}{\partial y_{i_1} \dots \partial y_{i_k}} dy \quad (190)$$

darstellt.

Wir gehen nun zum Studium des Raumes $W_p^{(l)}(D)$ über. Die Norm in $W_p^{(l)}(D)$ ist, wie wir uns erinnern, durch

$$\|u\|_{W_p^{(l)}(D)}^p = \|u\|_{L_p(D)}^p + \int_D \sum_{i_1, \dots, i_l} \left| \frac{\partial^l u(y)}{\partial y_{i_1} \dots \partial y_{i_l}} \right|^p dy \quad (191)$$

gegeben (eine der äquivalenten Normen [112]). Wir wollen zeigen, daß die Integraldarstellung (188) für jede Funktion $u(x) \in W_p^{(l)}(D)$ gilt. Gegeben sei eine Folge von Funktionen $u_m(x)$ aus $C^{(l)}(D)$, die in $W_p^{(l)}(D)$ gegen $u(x)$ konvergiert [112]. Wir schreiben (188) für die Funktionen $u_m(x)$ auf, geben den Größen $U_{i_1, \dots, i_l}$ die Form (190) und gehen für $m \rightarrow \infty$ zur Grenze über. Auf Grund der Sätze 1 und 2 aus [115] sind die Integraloperatoren mit den Kernen $B_{i_1, \dots, i_l}(x, y) |x - y|^{-(n-l)}$ stetig in $L_p(D)$; dies erlaubt uns, unter dem Integralzeichen den Grenzübergang vorzunehmen. Somit gilt die Integralsdarstellung (188) für alle Funktionen aus $W_p^{(l)}(D)$.

Wir zeigen nun, daß die Funktionen aus $W_p^{(l)}(D)$ alle verallgemeinerte Ableitungen von beliebiger Ordnung $k < l$ aus $L_p(D)$ besitzen. Wenden wir (182) auf die $(l-1)$ -te Ableitung von $u_m(x)$ an, so erhalten wir

$$\frac{\partial^{l-1} u_m(x)}{\partial x_{i_1} \dots \partial x_{i_{l-1}}} = U_{i_1, \dots, i_{l-1}}^{(m)} + \int_D \sum_{k=1}^n \frac{\partial^l u_m(y)}{\partial y_{i_1} \dots \partial y_{i_{l-1}} \partial y_k} \frac{B_k(x, y)}{|x - y|^{n-1}} dy \quad (192)$$

mit

$$U_{i_1, \dots, i_{l-1}}^{(m)} = \int_D \frac{\partial^{l-1} u_m(x)}{\partial y_{i_1} \dots \partial y_{i_{l-1}}} p(y) dy = (-1)^{l-1} \int_D u_m(y) \frac{\partial^{l-1} p(y)}{\partial y_{i_1} \dots \partial y_{i_{l-1}}} dy.$$

Auf Grund von Satz 1 und 2 aus [115] können wir, je nachdem, ob $n-1 < \frac{n}{p'}$ oder $n-1 \geq \frac{n}{p'}$ ist, behaupten, daß der auf der rechten Seite von (192) stehende Integraloperator als Operator von $L_p(D)$ in $C(D)$ oder in $L_p(D)$, d. h. in jedem Fall in $L_p(D)$, stetig ist. Nach Voraussetzung konvergiert die Folge der $u_m(x)$ der Norm von $W_p^{(l)}(D)$ nach gegen $u(x)$; daraus folgt, daß für $m \rightarrow \infty$ die rechte Seite von (192) in $L_p(D)$ einen Grenzwert und infolgedessen die Funktion $u(x)$ in D verallgemeinerte $(l-1)$ -te Ableitungen besitzt, für welche (192) gilt, wenn $u_m(x)$ durch $u(x)$ sowie $U_{i_1, \dots, i_{l-1}}^{(m)}$ durch

$$U_{i_1, \dots, i_{l-1}} = (-1)^{l-1} \int_D u(y) \frac{\partial^{l-1} p(y)}{\partial y_{i_1} \dots \partial y_{i_{l-1}}} dy$$

ersetzt wird.

Benutzen wir die zu (188) und (192) analogen Integralsdarstellungen der Ableitungen von beliebiger Ordnung $k < l$ durch die Ableitungen l -ter Ordnung, so können wir auch die Existenz von verallgemeinerten Ableitungen niedrigerer Ordnung aus $L_p(D)$ nachweisen. Dabei lassen sich die erwähnten Integralsdarstellungen automatisch auf ganz $W_p^{(l)}(D)$ übertragen. Die entsprechenden Integraloperatoren haben polare Kerne $[n-(l-k)]$ -ter Ordnung. Die Beschränktheit dieser Integraloperatoren in $L_p(D)$ führt unmittelbar auf die Abschätzung

$$\left\| \frac{\partial^k u(x)}{\partial x_{i_1} \dots \partial x_{i_k}} \right\|_{L_p(D)} \leq c \left[\|u\|_{L_p(D)} + \sum_{i_1, \dots, i_l} \left\| \frac{\partial^l u(x)}{\partial x_{i_1} \dots \partial x_{i_l}} \right\|_{L_p(D)} \right] \quad (k = 1, 2, \dots, l-1). \quad (193)$$

Daraus folgt (152).

Den Größen (190) kann jetzt wieder die Form (189) gegeben werden. Damit ist bewiesen, daß für ein bezüglich einer Kugel sternförmiges Gebiet die Räume $W_p^{(l)}$ und $\tilde{W}_p^{(l)}$

aus ein und derselben Menge von Funktionen bestehen und die Normierungen dieser Räume auf Grund von (192) äquivalent sind.¹⁾

Die Integraldarstellung (188) der Funktionen aus $W_p^{(l)}(D)$ wurde in etwas anderer Gestalt von SOBOLEW angegeben.²⁾

Wir beweisen nun in [117] für Gebiete, die bezüglich einer Kugel sternförmig sind, die in [114] formulierten Einbettungssätze und zeigen dann, wie sich diese Sätze auf eine allgemeine Klasse von Gebieten übertragen lassen. Die Aussage bezüglich der Äquivalenz von $W_p^{(l)}(D)$ und $\tilde{W}_p^{(l)}(D)$ wird damit ebenfalls auf diese Klasse übertragen.

117. Einbettungssätze. Wir wenden uns jetzt wieder der Untersuchung der Integraldarstellung (188) zu. Die integralfreien Glieder auf der rechten Seite von (188) sind vollstetige Operatoren von $W_p^{(l)}(D)$ in $C(D)$. Jede Funktion $u(x) \in W_p^{(l)}(D)$ wird nämlich durch einen solchen Operator in ein und dieselbe, in D stetige und mit der Zahl $U_{t_1, \dots, t_k}$ multiplizierte Funktion $b_{t_1, \dots, t_k}(x)$ übergeführt, wobei $U_{t_1, \dots, t_k}$ ein in $W_p^{(l)}(D)$ stetiges Funktional darstellt. Ist die Menge der Funktionen $u(x)$ in $W_p^{(l)}(D)$ beschränkt, so läßt sich aus den Zahlen $U_{t_1, \dots, t_k}$ eine konvergente Folge auswählen. Die entsprechende Folge der Funktionen $U_{t_1, \dots, t_k} b_{t_1, \dots, t_k}(x)$ konvergiert in D gleichmäßig. Auf die Glieder von (188), welche Integrale enthalten, sind die Sätze 1 und 2 aus [115] anwendbar. Somit gelangen wir zu den beiden folgenden Einbettungssätzen [115].

Satz 1. Ist $pl > n$, so ist jede Funktion $u(x) \in W_p^{(l)}(D)$ in D stetig ($W_p^{(l)}(D) \subset C(D)$); der Einbettungsoperator ist beschränkt,

$$\max_D |u(x)| = \|u\|_{C(D)} \leq K \|u\|_{W_p^{(l)}(D)} \quad (K > 0, \text{konstant}),$$

und vollstetig, d. h., er bildet jede in $W_p^{(l)}(D)$ beschränkte Menge von Funktionen in eine in $C(D)$ kompakte Menge ab.

Satz 2. Es sei $pl \leq n$ und $s > n - pl$. Dann gehören die Funktionen $u(x) \in W_p^{(l)}(D)$ auf jedem s -dimensionalen Durchschnitt D_s von D mit einer Ebene zu $L_q(D_s)$ für jedes

$$q < q^* = \frac{ps}{n - pl}. \quad (194)$$

Der Einbettungsoperator von $W_p^{(l)}(D)$ in $L_q(D_s)$ ist beschränkt und vollstetig. Als Element von $L_q(D_s)$ ist die Funktion $u(x)$ in der Metrik von $L_q(D_s)$ bezüglich einer Parallelverschiebung des Durchschnitts D_s stetig, falls diese Parallelverschiebung zulässig ist.

Bemerkungen. 1. Die Funktionen $u(x) \in W_p^{(l)}(D)$ sind bis auf Äquivalenz bestimmt, und die Aussage des Satzes über das Verhalten der Funktionen $u(x)$ auf den Durchschnittsnissen bezieht sich auf eine gewisse Auswahl aus der Klasse der äquivalenten Funktionen (vgl. [118]). Wir weisen noch darauf hin, daß die Formel (188), wie aus dem Vorhergehenden folgt, gerade eine solche Funktion definiert.

2. Wird in Satz 2 der Raum $L_q(D_s)$ durch $L_{q^*}(D_s)$ ersetzt, so bleibt, wie sich zeigen läßt, der Satz einschließlich des Wortes „beschränkt“ gültig.

3. Wird in D eine s -dimensionale Mannigfaltigkeit T_s gewählt (sie darf auch auf der Berandung von D liegen), die mit Hilfe einer l -mal (eventuell stückweise) stetig differenzierbaren und eindeutig umkehrbaren Variablensubstitution $y_i = y_i(x_1, \dots, x_n)$, $i = 1, \dots, n$, in eine s -dimensionale Ebene abgebildet werden kann, so behält Satz 2 seine Gültigkeit, wenn D_s durch T_s ersetzt wird. Dabei wird gefordert, daß die Variablensubstitution in einer n -dimensionalen Umgebung von T_s definiert ist.

¹⁾ Im folgenden brauchen wir $\tilde{W}_p^{(l)}(D)$ und $W_p^{(l)}(D)$ nicht zu unterscheiden.

²⁾ S. L. SOBOLEW, Einige Anwendungen der Funktionalanalysis in der mathematischen Physik (vgl. das Literaturverzeichnis).

Wir bemerkten schon, daß sich für Funktionen $u(x) \in W_p^{(l)}(D)$ die verallgemeinerten Ableitungen der Ordnung $m < l$ mit Hilfe von Formeln, die zu (188) analog sind, durch Ableitungen l -ter Ordnung mit Hilfe von Integraloperatoren mit einem polaren Kern $[n - (l - m)]$ -ter Ordnung ausdrücken lassen. Benutzen wir die Sätze aus [115], so erhalten wir wie oben die folgenden Sätze:

Satz 3. Ist $pl > n$ und $0 < m < l - \frac{n}{p}$, so sind die verallgemeinerten m -ten Ableitungen der Funktionen $u(x) \in W_p^{(l)}(D)$ in D stetig, und der Einbettungsoperator von $W_p^{(l)}(D)$ in $C^{(m)}(D)$ ist beschränkt und vollstetig.

Satz 4. Ist $m \geq l - \frac{n}{p}$ und $s > n - (l - m)p$, so gehören die verallgemeinerten m -ten Ableitungen der Funktionen $u(x) \in W_p^{(l)}(D)$ auf s -dimensionalen Durchschnitten D_s des Gebietes D mit einer Ebene für alle

$$q < q^* = \frac{ps}{n - (l - m)p} \quad (195)$$

zu $L_q(D_s)$; dabei ist

$$a) \quad \|D^m u\|_{L_q(D_s)} \leq K \|u\|_{W_p^{(l)}(D)}; \quad (196)$$

b) für eine in $W_p^{(l)}(D)$ beschränkte Menge von Funktionen $u(x)$ ist die Menge der $D^m u(x)$ in $L_q(D_s)$ kompakt;

c) die Funktionen $D^m u(x)$ sind in der Metrik von $L_q(D_s)$ bezüglich einer zulässigen Parallelverschiebung von D_s stetig.

Zu Satz 4 können wir analoge Bemerkungen wie zu Satz 2 machen.

Wir beweisen jetzt Satz 3 aus [114] über äquivalente Normen in $W_p^{(l)}(D)$. Die entsprechende Formulierung dieses Satzes lautete:

Gegeben seien lineare und in $W_p^{(l)}(D)$ beschränkte Funktionale $l_k(u)$, $k = 1, 2, \dots, N$, derart, daß sie nicht alle gleichzeitig gleich 0 sind auf jedem nicht identisch verschwindenden Polynom höchstens $(l - 1)$ -ten Grades. Dann ist die durch

$$\|u\|^p = \int_D \sum_{i_1, \dots, i_l} \left| \frac{\partial^l u(x)}{\partial x_{i_1} \dots \partial x_{i_l}} \right|^p dx + \sum_{k=1}^N |l_k(u)|^p \quad (197)$$

definierte Norm in $W_p^{(l)}(D)$ mit der Norm (191) äquivalent.

Beweis. Die Norm (197) läßt sich offenbar mit Hilfe der Norm (191) abschätzen, da in der Norm (197) die Funktionale $l_k(u)$, $k = 1, 2, \dots, N$, beschränkt sind. Wir wollen nun zeigen, daß sich auch die Norm (191) mit Hilfe der Norm (197) abschätzen läßt. Zeitweilig werden wir die Norm (197) einfach mit $\|u\|$ und die zu (154) äquivalente Norm (191) mit $\|u\|_{W_p^{(l)}(D)}$ bezeichnen. Wir müssen zeigen, daß für alle Funktionen aus $W_p^{(l)}(D)$ die Abschätzung

$$\|u\|_{W_p^{(l)}(D)} \leq A \|u\| \quad (198)$$

gilt. Dazu nehmen wir das Gegenteil an, d. h., es existiere eine unendliche Folge von positiven Zahlen A_m ($m = 1, 2, \dots$) und von Elementen $u_m(x) \in W_p^{(l)}(D)$ derart, daß $A_m \rightarrow \infty$ für $m \rightarrow \infty$ und

$$\|u_m\|_{W_p^{(l)}(D)} \geq A_m \|u_m\| \quad (199)$$

gilt. Führen wir in $u_m(x)$ konstante Faktoren ein, so können wir

$$\|u_m\|_{W_p^{(l)}(D)} = 1 \quad (200)$$

annehmen. Aus (199) und (200) ergibt sich, daß $\|u_m\|$ für $m \rightarrow \infty$ gegen 0 strebt und daß folglich alle verallgemeinerten l -ten Ableitungen der Funktionen $u_m(x)$ in $L_p(D)$ gegen 0 konvergieren. Auf Grund von (200) und Satz 2 ist die Folge der $u_m(x)$ in $L_p(D)$ kompakt. Aus ihr wählen wir eine konvergente Teilfolge aus, deren Elemente wir ebenfalls mit $u_m(x)$ bezeichnen wollen, und es sei $u_m(x) \rightarrow u_0(x)$ in $L_p(D)$ für $m \rightarrow \infty$. Aus Satz 2 von [109] folgt dann, daß alle verallgemeinerten l -ten Ableitungen von $u_0(x)$ existieren und gleich 0 sind. Wir weisen noch darauf hin, daß die $u_m(x)$ gegen $u_0(x)$ im Sinne der Norm (191) streben.

Wir zeigen nun, daß $u_0(x) = 0$ ist. Dazu betrachten wir ein echt in D enthaltenes Teilgebiet D' . In D' stimmen die l -ten Ableitungen der Mittelfunktionen $u_{0h}(x)$ für hinreichend kleines h mit den Mittelfunktionen der Ableitungen $D^l u_0(x)$ überein [109]. Daraus folgt unmittelbar, daß alle l -ten Ableitungen der Mittelfunktionen $u_{0h}(x)$ in D' verschwinden. Also sind die $u_{0h}(x)$ Polynome in $x_1, \dots, x_n$ von höchstens $(l-1)$ -tem Grad in dem entsprechenden Teilgebiet. Da die Menge dieser Polynome einen Unterraum von $L_p(D')$ bildet und $u_{0h}(x) \rightarrow u_0(x)$ in $L_p(D)$ gilt, ist $u_0(x)$ ein gewisses Polynom von höchstens $(l-1)$ -ten Grad in einem beliebigen Teilgebiet, also auch überall in D . Auf Grund der Stetigkeit der Funktionale $l_k(u)$ in $W_p^{(l)}(D)$ gilt $l_k(u_m) \rightarrow l_k(u_0)$ für $m \rightarrow \infty$ ($k=1, 2, \dots, N$). Gleichzeitig folgt aus (199) und (200) $l_k(u_m) \rightarrow 0$ für $m \rightarrow \infty$ ($k=1, 2, \dots, N$). Wir erhalten somit $l_k(u_0) = 0$ ($k=1, 2, \dots, N$) und auf Grund der Voraussetzung des zu beweisenden Satzes $u_0(x) = 0$. Dies widerspricht jedoch der Beziehung (200) und der Konvergenz der $u_m(x)$ gegen $u_0(x)$ in $W_p^{(l)}(D)$. Damit ist der Satz bewiesen.

Wir geben nun einige einfache Beispiele zur Benutzung der bewiesenen Sätze an.

1. Es sei $u(x) \in W_2^{(2)}(D)$, d. h., die Funktion $u(x)$ sei nebst ihren verallgemeinerten Ableitungen einschließlich zweiter Ordnung in D quadratisch summierbar. Für $n \leq 3$ folgt aus Satz 1, daß $u(x)$ in D stetig ist. Für $n \geq 4$ braucht diese Behauptung nicht zu gelten.

2. Die Norm (162) in $W_p^{(1)}(D)$, die wir mit Hilfe von Satz 3 aus [114] erhalten haben, führt für $p=2$ auf die bekannte *Poincarésche Ungleichung*

$$\int_D u^2(y) \, dy \leq c \left[\int_D \sum_{k=1}^n \left(\frac{\partial u(y)}{\partial y_k} \right)^2 dy + \left(\int_D u(y) \, dy \right)^2 \right]. \quad (201)$$

3. Analog ergibt die Norm (164) für $p=2$ und $q=2$

$$\int_D u^2(y) \, dy \leq c_1 \left[\int_D \sum_{k=1}^n \left(\frac{\partial u(y)}{\partial y_k} \right)^2 dy + \int_S u^2(y) \, ds \right]. \quad (202)$$

Hier ist S eine hinreichend glatte $(n-1)$ -dimensionale Mannigfaltigkeit in D . Ist insbesondere die Berandung von D stückweise glatt, so kann sie als Mannigfaltigkeit S in (202) benutzt werden. In diesem Fall ist (202) unter dem Namen *Friedrichssche Ungleichung* bekannt.

118. Gebiete allgemeinerer Art. Wir wollen uns nun damit beschäftigen, die Einbettungssätze auf umfassendere Klassen von Gebieten zu übertragen.

Ein beschränktes Gebiet D werde mit Hilfe von stückweise glatten $(n-1)$ -dimensionalen Mannigfaltigkeiten in endlich viele Gebiete zerlegt, wobei in jedem dieser Teilgebiete das in [117] Bewiesene gelte. Dann gilt es auch in D .

Offenbar genügt es, den Fall zu untersuchen, daß D durch eine Fläche S_0 in zwei punktfremde Teilgebiete D_1 und D_2 zerlegt wird. Es sei $u(x) \in \bar{W}_p^{(l)}(D)$. Zunächst zeigen wir, daß $u(x)$ in D alle verallgemeinerten Ableitungen niedrigerer Ordnung aus $L_p(D)$ hat. Offenbar ist $u(x) \in \bar{W}_p^{(l)}(K)$, wobei K eine beliebige Kugel aus D ist. Wenden wir auf diese Kugel das in [116] Gesagte an, so können wir uns davon überzeugen, daß $u(x)$ in K alle verallgemeinerte Ableitungen der Form $\chi(x) = D^k u(x)$, $0 \leq k < l$, hat, die zu $L_p(K)$ gehören. Die Funktionen $\chi(x)$ sind auch überall in D definiert und gehören zu

$L_p(D)$; in D_1 bzw. D_2 hat $u(x)$ nämlich die Ableitungen $D^k u(x)$ aus $L_p(D_1)$ bzw. $L_p(D_2)$, und die Eindeutigkeit der verallgemeinerten Ableitung besagt, daß $\chi(x)$ mit $D^k u(x)$ in D_1 und D_2 übereinstimmt und folglich zu $L_p(D)$ gehört.

Wir zeigen nun, daß $\chi(x)$ mit $D^k u(x)$ in D übereinstimmt. Es sei D' ein beliebiges, echt in D enthaltenes Teilgebiet, $\delta > 0$ der Abstand von D' zur Berandung von D und x ein beliebiger Punkt von D' , und es gelte $0 < h < \delta$. In einer Kugel mit dem Radius δ und dem Punkt x als Mittelpunkt hat $u(y)$ die verallgemeinerte Ableitung $D^k u(y) = \chi(y)$. Bilden wir die Mittelfunktionen mit dem Mittelungsradius h , so können wir behaupten [109], daß im Mittelpunkt der Kugel die Beziehung $D^k u_h(x) = \chi_h(x)$ gilt. Da $u_h(x) \rightarrow u(x)$ und $D^k u_h(x) \rightarrow \chi(x)$ für $h \rightarrow 0$ in $L_p(D')$ ist, stimmt $\chi(x)$ auf Grund von Definition 2 aus [109] mit der verallgemeinerten Ableitung $D^k u(x)$ der Funktion $u(x)$ in D überein. Wir brauchen uns nur zu erinnern, daß $\chi(x)$ zu $L_p(D)$ gehört.

Bemerkung. Diese Überlegungen haben gezeigt, daß die Funktionen aus $W_p^{(l)}(D)$ in einem beliebigen Gebiet D verallgemeinerte Ableitungen aller niedrigeren Ordnungen aus $L_p(D')$ besitzen.

Wir untersuchen nun die Frage, wie Satz 1 aus [117] auf das Gebiet D übertragen werden kann. Wir nehmen an, es sei $n < pl$, also $u(x)$ sei in D_1 bzw. D_2 stetig, und zeigen, daß $u(x)$ auch in D stetig ist. Dazu brauchen wir nur die Stetigkeit von $u(x)$ in den Punkten der Fläche S_0 nachzuweisen. Die Stetigkeit von $u(x)$ in einem inneren Punkt von D folgt aus dem auf eine hinreichend kleine Kugel angewendeten Satz über die Einbettung von $W_p^{(l)}$ in C . Für die Punkte von S_0 auf der Berandung von D stimmt der Grenzwert von $u(x)$ auf einem beliebigen, in D_1 liegenden Weg (auf Grund der Stetigkeit von $u(x)$ in D_1) mit dem Grenzwert auf einem in S_0 liegenden Weg überein. Dasselbe kann man auch von den Grenzwerten bei Annäherung aus D_2 sagen. Daraus folgt, daß die Grenzwerte der Funktion $u(x)$ auf allen Wegen übereinstimmen und $u(x)$ in D stetig ist. Die Vollstetigkeit (und damit die Beschränktheit) des Einbettungsoperators beruht darauf, daß für eine in $W_p^{(l)}(D)$ beschränkte Menge zuerst eine in $C(D_1)$ konvergente Folge und dann aus dieser Folge eine in $C(D_2)$ konvergente Teilfolge ausgewählt werden kann. Diese Teilfolge konvergiert offenbar in D gleichmäßig.

Analog läßt sich zeigen, wie in diesem Fall Satz 3 aus [117] übertragen werden kann. Auch für Satz 2 aus [117] ergeben sich keine Schwierigkeiten. Man braucht im allgemeinen den s -dimensionalen Durchschnitt D_s nur in zwei Teile D'_s und D''_s zu zerlegen, wobei $D'_s = D_s \cdot D_1$ und $D''_s = D_s \cdot D_2$ ist. Wenden wir Satz 2 aus [117] in D_1 und D_2 an, so erhalten wir

$$\|u\|_{L_q(D)_s} \leq \|u\|_{L_q(D'_s)} + \|u\|_{L_q(D''_s)} \leq c [\|u\|_{W_p^{(l)}(D_1)} + \|u\|_{W_p^{(l)}(D_2)}] \leq 2c \|u\|_{W_p^{(l)}(D)}. \quad (203)$$

Damit ist die Beschränktheit des Operators der Einbettung von $W_p^{(l)}(D)$ in $L_q(D_s)$ nachgewiesen.

Auf die gleiche Art läßt sich auch die starke Stetigkeit der Funktionen aus $W_p^{(l)}(D)$ in $L_q(D_s)$ bezüglich einer Parallelverschiebung der Durchschnitts D_s übertragen. Die Vollstetigkeit des Einbettungsoperators von $W_p^{(l)}(D)$ in $L_q(D_s)$ folgt unmittelbar aus (203) und der starken Stetigkeit bezüglich einer Parallelverschiebung.

Die Übertragung von Satz 4 aus [117] erfordert keine neuen Überlegungen.

Auch Satz 3 aus [114] über äquivalente Normen in $W_p^{(l)}(D)$ bleibt gültig, da sich sein in [117] angegebener Beweis nur auf die Einbettungssätze stützte.

Wir können also behaupten, daß für das Gebiet D alle in [114] und [117] betrachteten Einbettungssätze ihre Gültigkeit behalten, wenn jedes der Teilgebiete, aus denen D besteht, bezüglich einer in D liegenden Kugel sternförmig ist.

119. Der Raum $\dot{C}^{(l)}(D)$. Es sei D ein endliches oder unendliches Gebiet im R_n , und $\dot{C}^{(l)}(D)$ sei die Menge aller in D finiten und stetigen Funktionen $\varphi(x)$, die in D stetige Ableitungen bis zur l -ten Ordnung besitzen. Offenbar ist $\dot{C}^{(l)}(D)$ ein linearer

Raum. Analog zum Fall von $C^{(l)}(D)$ [114] führen wir in $\dot{C}^{(l)}(D)$ die Norm

$$\|\varphi\|_{\dot{C}^{(l)}(D)} = \max_{\substack{x \in D \\ 0 \leq k < l}} |D^k \varphi(x)| \quad (204)$$

ein. Die abgeschlossene Hülle von $\dot{C}^{(l)}(D)$ bezüglich dieser Norm führt uns auf einen Banachraum, den wir mit $\check{C}^{(l)}(D)$ bezeichnen. Die Elemente dieses Raumes sind beschränkte Funktionen, die l -mal stetig differenzierbar in $\bar{D}$ sind, wobei auf der Berandung von D die Funktion selbst und alle ihre Ableitungen verschwinden.

Für verschiedene Werte $l=0, 1, \dots$ können die Räume $\check{C}^{(l)}(D)$ ineinander eingebettet werden, d. h., es ist $\check{C}^{(l_1)}(D) \subset \check{C}^{(l_2)}(D)$ für $l_1 > l_2$, wobei die Menge der Elemente von $\check{C}^{(l_1)}(D)$ in $\check{C}^{(l_2)}(D)$ dicht ist. Dies läßt sich leicht zeigen, wenn der Mittelungsprozeß benutzt wird.

Wir untersuchen nun den zu $\check{C}^{(l)}(D)$ adjungierten Raum $U^{(l)}(D)$ (den Raum der linearen Funktionalen für $\check{C}^{(l)}(D)$). Es ist leicht zu sehen, daß $U^{(l_1)}(D)$ für $l_1 > l_2$ in $U^{(l_2)}(D)$ enthalten ist.

Elemente von $U^{(l)}(D)$ sind z. B. die Funktionalen, die mit Hilfe eines in D summierbaren Kerns $\psi(x)$ durch die Beziehung

$$(m, \varphi) = \int_D \psi(x) \varphi(x) dx \quad (205)$$

definiert werden. Ihre Norm genügt der Ungleichung $\|m\| \leq \int_D |\psi(x)| dx$.

Diese Funktionalen werden nicht selten *Funktionalen vom Typ einer Funktion* genannt und mit den Kernen identifiziert, mit deren Hilfe sie definiert sind. Die übrigen Funktionalen werden *verallgemeinerte Funktionen* genannt. Die Funktionalen vom Typ einer Funktion schöpfen nicht den ganzen Raum $U^{(l)}(D)$ aus. So erlaubt z. B. das Funktional $\delta(x-x_0)$, das durch

$$(\delta(x-x_0), \varphi) = \varphi(x_0) \quad (206)$$

definiert wird, wobei x_0 ein fester Punkt von D ist, nicht die Darstellung (205) mit einem in D summierbaren Kern. Man sagt, der diesem Funktional entsprechende Kern sei die *im Punkt x_0 konzentrierte Delta-Funktion* $\delta(x-x_0)$, und schreibt

$$\int_D \delta(x-x_0) \varphi(x) dx = \varphi(x_0). \quad (207)$$

Die Funktion $\delta(x-x_0)$ ist keine Funktion im gewöhnlichen Sinne. Wie wir zeigen werden, sind die in der Form (205) darstellbaren Elemente von $U^{(l)}(D)$ mit einem stückweise stetigen Kern in $U^{(l)}(D)$ dicht. Genauer, es gilt der folgende

Satz. *Die Funktionalen vom Typ einer Funktion mit stückweise stetigem Kern sind in $\check{C}^{(l)}(D)$ im Sinne der schwachen Konvergenz von Funktionalen dicht.*

Zum Beweis dieses Satzes benötigen wir den folgenden

Hilfssatz. *Für ein beliebiges, aber festes $\varphi(x) \in \check{C}^{(l)}(D)$ und ein vorgegebenes $\varepsilon > 0$ existiert ein echt in D enthaltenes Gebiet D_ε derart, daß*

$$\max_{\substack{x \in D-D_\varepsilon \\ 0 \leq k \leq l}} |D^k \varphi(x)| \leq \varepsilon \quad (208)$$

gilt.

Nach Definition von $\dot{C}^{(l)}(D)$ existiert eine Folge von Elementen $\varphi_m(x) \in \dot{C}^{(l)}(D)$, die der Norm (204) nach gegen $\varphi(x)$ konvergiert; daher gibt es eine Zahl m_ε derart, daß für jedes $p > 0$

$$\|\varphi_{m_\varepsilon} - \varphi_{m_\varepsilon+p}\|_{\dot{C}^{(l)}(D)} \leq \varepsilon$$

gilt. Die Funktionen $\varphi_{m_\varepsilon}(x)$ sind nur in einem Gebiet D_ε der im Hilfssatz genannten Art von 0 verschieden, so daß sich für $x \in D - D_\varepsilon$ aus der letzten Ungleichung

$$\max_{\substack{x \in D - D_\varepsilon \\ 0 \leq k \leq l}} |D^k \varphi_{m_\varepsilon+p}(x)| \leq \varepsilon$$

und für $p \rightarrow \infty$ dann (208) ergibt.

Wir kommen nun zum Beweis des oben formulierten Satzes. Es sei m ein Element aus $U^{(l)}(D)$, und wir betrachten den Mittelungskern $\omega_\varrho(|x-y|)$. Offenbar gehört $\omega_\varrho(|x-y|)$ für ein x aus einem beschränkten Gebiet D_ϱ , das in D liegt und von der Berandung von D mindestens den Abstand 2ϱ hat, als Funktion von y zu $\dot{C}^{(l)}(D)$, so daß die Funktion

$$m_\varrho(x) = (m, \omega_\varrho(|x-y|)) \quad (209)$$

für $x \in D_\varrho$ definiert ist. Da $\omega_\varrho(|x-y|)$ unendlich oft differenzierbar und das Funktional m stetig ist, hat (209) stetige Ableitungen beliebiger Ordnung nach x . Für alle $x \in D$ definieren wir nun die stückweise stetige Funktion

$$\tilde{m}_\varrho(x) = \begin{cases} m_\varrho(x) & \text{für } x \in D_\varrho, \\ 0 & \text{für } x \in D - D_\varrho \end{cases}$$

und wählen eine gegen 0 konvergierende Folge $\varrho = \varrho_1, \varrho_2, \dots$, wobei wir D_{ϱ_k} so annehmen, daß D_{ϱ_k} in $D_{\varrho_{k+1}}$ enthalten ist und für $k \rightarrow \infty$ gegen D strebt. Wir zeigen, daß die durch

$$(\tilde{m}_{\varrho_k}, \varphi) = \int_D \tilde{m}_{\varrho_k}(x) \varphi(x) dx = \int_{D_{\varrho_k}} m_{\varrho_k}(x) \varphi(x) dx$$

definierten Funktionale $\tilde{m}_{\varrho_k}$ gegen das obige Funktional m streben. Dazu nehmen wir eine beliebige Funktion $\varphi(x) \in \dot{C}^{(l)}(D)$ und beweisen, daß die Differenz

$$r_k = (m, \varphi) - (\tilde{m}_{\varrho_k}, \varphi) = (m, \varphi) - \int_{D_{\varrho_k}} (m, \omega_{\varrho_k}(|x-y|)) \varphi(x) dx \quad (210)$$

für $k \rightarrow \infty$ verschwindet, woraus sich dann der Satz ergibt. Auf Grund des Hilfssatzes gibt es zu einem vorgegebenen $\varepsilon > 0$ ein echt in D enthaltenes Gebiet D_ε derart, daß für $\varphi(x)$ die Ungleichung (208) gilt. Wir nehmen an, k sei so groß, daß D_ε in D_{ϱ_k} liegt. Wir setzen

$$\varphi_k(y) = \int_{D_{\varrho_k}} \varphi(x) \omega_{\varrho_k}(|x-y|) dx.$$

Die Riemannschen Summen für dieses Integral lauten

$$I(y) = \sum_s \varphi(\xi_s) \omega_{\varrho_k}(|\xi_s - y|) \Delta_s x;$$

dabei ist $\Delta_s x$ das Maß des Teilgebiets, zu dem ξ_s gehört. Die Ableitungen von $I(y)$ nach y bis zur l -ten Ordnung konvergieren gleichmäßig bezüglich y gegen die Funktion $\varphi_k(y)$ und gegen die entsprechenden Ableitungen, wenn $\Delta_s x$ unbeschränkt

verkleinert wird; dies beruht auf der Stetigkeit von $\varphi(x)$ und der Tatsache, daß $\omega_{\varrho_k}(|x-y|)$ unendlich oft nach y differenzierbar ist. Nun gilt auf Grund der Distributivität des Funktional und seiner Stetigkeit bezüglich der Norm (204)

$$\sum_s (m, \varphi(\xi_s) \omega_{\varrho_k}(|\xi_s - y|) \Delta_s x) = (m, \sum_s \varphi(\xi_s) \omega_{\varrho_k}(|\xi_s - y|) \Delta_s x)$$

und nach Grenzübergang

$$\int_{D_{\varrho_k}} (m, \varphi(x) \omega_{\varrho_k}(|x-y|)) dx = (m, \int_{D_{\varrho_k}} \varphi(x) \omega_{\varrho_k}(|x-y|) dx),$$

so daß sich der Ausdruck für r_k in der Form

$$r_k = (m, \varphi(y) - \int_{D_{\varrho_k}} \varphi(x) \omega_{\varrho_k}(|x-y|) dx) \quad (211)$$

schreiben läßt. Die Funktion $\varphi_k(x)$ ist nicht die gewöhnliche Mittelung von $\varphi(x)$ mit dem Kern $\omega_{\varrho_k}(|x-y|)$, denn das Integrationsgebiet D_{ϱ_k} kann nicht die ganze Kugel $|x-y| \leq \varrho_k$ für $y \in D$ enthalten. Gehört jedoch y zu einem beliebigen, echt in D enthaltenen Teilgebiet, so liegt die Kugel $|x-y| \leq \varrho_k$ in D_{ϱ_k} für alle hinreichend großen k , und für diese y ist $\varphi_k(y)$ die Mittelung von $\varphi(x)$ mit dem Kern $\omega_{\varrho_k}(|x-y|)$. Offenbar ist $\varphi_k(y) \in \dot{C}^{(l)}(D)$, und wir wollen zeigen, daß $\varphi_k(y)$ der Norm (204) nach gegen $\varphi(y)$ konvergiert. Dazu nehmen wir ein $\delta > 0$, das kleiner ist als der Abstand von D_ε bis zur Berandung von D , und bezeichnen mit D_δ das Gebiet, das wir erhalten, wenn wir zu D_ε alle Kugeln mit dem Radius δ und dem Mittelpunkt in D_ε hinzufügen. Für $y \in D_\delta$ und alle hinreichend großen k gilt $D_\delta \subset D_{\varrho_k}$ und

$$\varphi_k(y) = \int_{|x-y| \leq \varrho_k} \varphi(x) \omega_{\varrho_k}(|x-y|) dx.$$

Daher streben die $\varphi_k(y)$ nebst ihren Ableitungen nach y bis zur l -ten Ordnung für $k \rightarrow \infty$ gleichmäßig bezüglich $y \in D_\delta$ gegen $\varphi(y)$ und deren entsprechende Ableitungen [71]. Ist $y \in D - D_\delta$, so sind die beiden Funktionen $\varphi(y)$ und $\varphi_k(y)$ gleichmäßig klein und nebst ihren Ableitungen einschließlich l -ter Ordnung, denn für $\varphi(y)$, $y \in \bar{D} - D_\varepsilon$, gilt (208). Somit konvergieren die $\varphi_k(y)$ der Norm (204) nach gegen $\varphi(y)$, und aus (211) folgt $r_k \rightarrow 0$ für $k \rightarrow \infty$. Damit ist der Satz bewiesen.

Man könnte auch zeigen, daß die Funktionale vom Typ einer Funktion mit stetig differenzierbaren Kernen ebenfalls in $U^{(l)}(D)$ dicht sind.

Wir definieren nun für die Elemente von $U^{(l)}(D)$ die Multiplikation mit einer Funktion $a(x)$ und die Differentiation $\frac{\partial}{\partial x_k}$. Es sei $a(x) \in C^{(r)}(D)$. Ist D ein unendliches Gebiet, so nehmen wir an, die Funktion $a(x)$ und ihre Ableitungen seien beschränkt. Ferner führen wir einen linearen Operator ein,

$$A\varphi = a(x) \varphi(x),$$

der $\varphi(x) \in C^{(l)}(D)$ in $a(x) \varphi(x) \in \dot{C}^{(s)}(D)$ überführt, wobei $s = \min(l, r)$ ist. Dann definieren wir die Operation der Multiplikation eines Elements $m \in U^{(l)}(D)$ mit $a(x)$ als den zu A adjungierten Operator A^* , d. h. durch die Beziehung

$$(m, A\varphi) = (A^*m, \varphi), \quad (212)$$

die für alle $\varphi(x) \in \dot{C}^{(l)}(D)$ erfüllt sein muß. Dabei ist der Operator A^* auf die Elemente aus $U^{(s)}(D)$ anwendbar, und es ist $A^*m \in U^{(l)}(D)$. Hat das Funktional m die

Form (205) mit einem in D summierbaren Kern, so ist

$$(m, A\varphi) = \int_D \psi(x) [a(x) \varphi(x)] dx = \int_D [\psi(x) a(x)] \varphi(x) dx,$$

d. h., das Funktional A^*m ist ebenfalls ein Funktional vom Typ einer Funktion mit dem Kern $a(x) \psi(x)$. Wir betrachten nun den Operator der Differentiation:

$$B\varphi = -\frac{\partial \varphi(x)}{\partial x_k};$$

B ist ein beschränkter Operator von $\dot{C}^{(l)}(D)$, $l \geq 1$, in $\dot{C}^{(l-1)}(D)$. Der zu ihm adjungierte Operator B^* läßt sich formal durch eine Gleichung der Form (212) definieren,

$$(m, B\varphi) = (B^*m, \varphi), \quad (213)$$

und führt die Elemente m aus $U^{(l-1)}(D)$ in $U^{(l)}(D)$ über. Für Funktionale m , die sich in der Gestalt (205) mit einem stetig differenzierbaren Kern $\psi(x)$ darstellen lassen, besagt diese Gleichung, daß dem Funktional B^*m der Kern $\frac{\partial \psi(x)}{\partial x_k}$ entspricht, denn es ist

$$(m, B\varphi) = - \int_D \psi(x) \frac{\partial \varphi(x)}{\partial x_k} dx = \int_D \frac{\partial \psi(x)}{\partial x_k} \varphi(x) dx = (B^*m, \varphi).$$

Wir wollen nun sehen, wie sich die Operatoren A^* und B^* von einem Funktional berechnen lassen, das durch die Delta-Funktion, d. h. durch (206) gegeben ist. Es sei $\varphi(x) \in \dot{C}^{(l)}(D)$, $l \geq 1$. Dann gilt

$$\begin{aligned} (\delta(x-x_0), A\varphi) &= (A^*\delta(x-x_0), \varphi) = a(x_0) \varphi(x_0), \\ (\delta(x-x_0), B\varphi) &= (B^*\delta(x-x_0), \varphi) = -\left. \frac{\partial \varphi(x)}{\partial x_k} \right|_{x=x_0}. \end{aligned} \quad (214)$$

Wir können auch die höheren Differentialquotienten einführen, die, wie leicht zu sehen ist, nicht von der Reihenfolge der Differentiation abhängen.

Somit lassen sich für die Elemente aus $U^{(l)}(D)$ Ableitungen bis zur l -ten Ordnung und verschiedene Differentialoperatoren definieren. Für diese Operatoren können wir die gleichen Probleme stellen wie für die gewöhnlichen Funktionen, und zwar das Cauchysche Problem und verschiedene Randwertprobleme. Als erster führte SOBOLEW im Jahre 1936 verallgemeinerte Funktionen bei der Lösung des Cauchyschen Problems für lineare hyperbolische partielle Differentialgleichungen ein. Eine solche Erweiterung der Klasse der Objekte erwies sich von zwei Gesichtspunkten als nützlich: Erstens braucht ein Problem in der Klasse der gewöhnlichen Funktionen keine Lösung zu haben, aber in der Klasse der verallgemeinerten Funktionen (Funktionale) existiert trotzdem eine Lösung; zweitens kann man manchmal leicht die Existenz einer „schlechten“ Lösung, die eine verallgemeinerte Funktion ist, nachweisen und danach untersuchen, wann diese Funktion mit einer gewöhnlichen Funktion übereinstimmt.

Diese beiden Umstände sind deutlich sichtbar beim Cauchyschen Problem für verschiedene Systeme von partiellen Differentialgleichungen mit konstanten Koeffizienten:

$$\begin{aligned} \frac{\partial u_i(x, t)}{\partial t} &= \sum_{j=1}^N \sum_{k=1}^n a_{ij}^k \frac{\partial u_j(x, t)}{\partial x_k} + \sum_{j=1}^N a_{ij} u_j(x, t) + f_i(x, t) \\ u_i(x, 0) &= \varphi_i(x) \quad (i=1, 2, \dots, N). \end{aligned} \quad (215)$$

Wenn wir hier eine Fouriertransformation bezüglich x vornehmen, so können wir dieses Problem auf das Cauchysche Problem für ein System gewöhnlicher Differentialgleichungen mit konstanten Koeffizienten zurückführen, die von einem Zahlenparameter α_k abhängen. Die Anwendung der inversen Fouriertransformation erlaubt dann, von der Lösung dieser gewöhnlichen Differentialgleichungen zu den Lösungen des Ausgangsproblems überzugehen. Bleibt man im Rahmen der klassischen Fouriertransformationen, so muß man sich auf die Untersuchung der in einem bestimmten Sinne abnehmenden Anfangsfunktionen und der Funktionen f_i beschränken. Unter dieser Voraussetzung wurde das Problem von I. G. PETROWSKI untersucht. In seinen Arbeiten ist aus den zusätzlichen Überlegungen ersichtlich, daß es eine Klasse sogenannter hyperbolischer Systeme gibt, für welche die Lösung in einem beliebigen Punkt (x_0, t_0) des Raumes durch die Werte der Anfangsfunktionen $\varphi_i(x)$ bestimmt sind, die nur in einem beschränkten und vom Punkt (x_0, t_0) abhängigen Teil des Raumes definiert sind. Damit ist das Problem (215) für hyperbolische Systeme eindeutig lösbar, unabhängig davon, wie sich die $\varphi_i(x)$ für $|x| \rightarrow \infty$ verhalten.

Ferner wurde in Arbeiten von A. N. TYCHONOFF, O. A. LADYSHENSKAJA und S. D. EIDELMAN bewiesen, daß die Anfangsfunktionen auch bei sogenannten parabolischen Systemen für $|x| \rightarrow \infty$ nicht nur als nicht abnehmend, sondern sogar als unbeschränkt (exponentiell) wachsend angenommen werden können. Jedoch müssen für die Erhaltung der Eindeutigkeit in diesem Fall bestimmte Einschränkungen bezüglich der Ordnung des Wachstums gemacht werden. Bei der Umkehrung des Cauchyschen Problems für die Wärmeleitungsgleichung (d. h. des bezüglich abnehmender t lösaren Cauchyschen Problems für die Gleichung $\frac{\partial u}{\partial t} = \Delta u$) gibt es bekanntlich nur für spezielle Anfangswerte eine gewöhnliche Lösung.

Alle diese Tatsachen erfordern ein eingehenderes Studium des Problems (215) und, im Zusammenhang damit, der Fouriertransformierten von Funktionen, die auf beliebige Weise unbeschränkt wachsen. Diese Untersuchung wurde in Arbeiten von L. SCHWARTZ begonnen und in Arbeiten von I. M. GELFAND und G. E. SCHILOW weitergeführt. Die Fouriertransformierte von unbeschränkt wachsenden Funktionen ist im allgemeinen keine Funktion, sondern ein Funktional in einem gewissen Raum $\hat{C}^{(0)}(R_n)$. In der Klasse dieser Funktionale muß man das Cauchysche Problem für Systeme gewöhnlicher Differentialgleichungen betrachten und dann zu den Lösungen von (215) übergehen, die entweder gewöhnliche oder verallgemeinerte Funktionen sind. Wir wollen hier nicht die Resultate dieser von GELFAND, SCHILOW und ihren Schülern durchgeführten Untersuchungen angeben, sondern wir verweisen auf die Literatur über verallgemeinerte Funktionen und ihre Anwendungen.¹⁾

Wir wollen nun einige Tatsachen angeben, die die Lösungen in Form von Funktionalen betreffen, und zwar die Lösungen für lineare Differentialgleichungen zweiter Ord-

¹⁾ I. M. GELFAND und G. E. SCHILOW, Verallgemeinerte Funktionen und ihre Anwendungen, Band I: Verallgemeinerte Funktionen und das Rechnen mit ihnen, 2. Aufl., VEB Deutscher Verlag der Wissenschaften, Berlin 1967; Band II: Räume von Grundfunktionen und verallgemeinerten Funktionen, VEB Deutscher Verlag der Wissenschaften, Berlin 1962; Band III: Einige Fragen zur Theorie der Differentialgleichungen, VEB Deutscher Verlag der Wissenschaften, Berlin 1964; Band IV (von I. M. GELFAND und N. J. WILENKIN): Einige Anwendungen der harmonischen Analyse, VEB Deutscher Verlag der Wissenschaften, Berlin 1964 (Übersetzungen aus dem Russischen).

nung vom elliptischen, parabolischen und hyperbolischen Typ mit variablen, aber stetig differenzierbaren Koeffizienten. Wir betrachten die Gleichung

$$L(u) = f(x), \quad (216)$$

in der $f(x)$ eine Funktion mit einer Singularität im Punkt $x = x_0$ ist. Für Gleichungen vom elliptischen und vom parabolischen Typ sind alle Lösungen der Gleichung (216) gewöhnliche Funktionen, die überall (eventuell mit Ausnahme des Punktes x_0 , in dem sie eine Singularität haben können) stetig differenzierbar sind. Später werden wir am Beispiel des Laplaceschen Operators zeigen, daß sogar dann, wenn $f(x)$ die im Punkt x_0 konzentrierte Delta-Funktion ist, die Lösungen der Gleichung (216) gewöhnliche Funktionen sind, die im Punkt x_0 einen Pol besitzen. Die zu (216) homogene Gleichung besitzt nur gewöhnliche Funktionen als Lösungen.

Dies gilt jedoch nicht für hyperbolische Gleichungen. Für sie überträgt sich die Singularität von $f(x)$ auf ganze Gebiete, und die Lösung braucht nicht eine gewöhnliche Funktion zu sein, sondern kann eine verallgemeinerte Funktion sein. Dafür geben wir ein Beispiel an, und zwar betrachten wir die Wellengleichung

$$u_{tt} = u_{x_1 x_1} + u_{x_2 x_2} + u_{x_3 x_3}.$$

Eine ihrer Lösungen wird durch die Poissonsche Formel [II, 184] angegeben:

$$u(x, t) = \frac{t}{4\pi} \int_0^{2\pi} \int_0^\pi \varphi(x + t\vec{n}) \sin \theta d\theta d\varphi. \quad (217)$$

Diese Formel ergibt bekanntlich für eine dreimal stetig differenzierbare Funktion $\varphi(x)$ eine zweimal stetig differenzierbare Lösung der Wellengleichung, die den Anfangsbedingungen $u(\dot{x}, 0) = 0$, $u_t(x, 0) = \varphi(x)$ genügt. Es seien nun $\varphi_m(x)$ dreimal stetig differenzierbare nichtnegative Funktionen, die für $m \rightarrow \infty$ gegen die Funktion $\varphi(x) = \frac{1}{x_1^2 + x_2^2}$ streben. Aus (217) ist ersichtlich, daß die ihnen entsprechenden Lösungen $u_m(x, t)$ in jedem Punkt (x, t) , der in dem Gebiet $\sqrt{x_1^2 + x_2^2} \leq t$ des Halbraumes $t \geq 0$ liegt, gegen ∞ streben. Dies legt den Gedanken nahe, daß die den Anfangsbedingungen $u(x, 0) = 0$,

$u_t(x, 0) = \frac{1}{x_1^2 + x_2^2}$ entsprechende Lösung der Wellengleichung als gewöhnliche Funktion nicht existiert. In der Klasse der Funktionale aber existiert die eindeutig bestimmte Lösung. Analog kann man sich mit Hilfe der Kirchhoffschen Formel¹⁾ davon überzeugen, daß die inhomogene Wellengleichung mit der Delta-Funktion $\delta(x - x_0)$ als Störfunktion ebenfalls keine Lösungen in der Klasse der gewöhnlichen Funktionen, sondern nur in der Klasse der Funktionale Lösungen besitzt.

Wie wir uns erinnern, stammen die ersten Arbeiten, in welchen Probleme aus der Klasse der verallgemeinerten Funktionen gestellt und gelöst wurden, von SOBOLEW, und zwar für das Cauchysche Problem für Differentialgleichungen vom hyperbolischen Typ.

Abschließend wollen wir die obige Behauptung über die Lösungen der Laplaceschen Gleichung beweisen.

Es sei D ein beschränktes dreidimensionales Gebiet und $|x - x_0| = r$ der Abstand zwischen einem variablen Punkt x und einem festen Punkt x_0 . Ist die Berandung von D hinreichend glatt und benutzen wir die Greensche Formel, so erhalten wir für $\varphi(x) \in \dot{C}^{(2)}(D)$

$$\varphi(x_0) = -\frac{1}{4\pi} \int_D \frac{\Delta \varphi(x)}{r} dx \quad \text{oder} \quad \int_D \left(-\frac{1}{4\pi r} \right) \Delta \varphi(x) dx = \delta(x - x_0) \varphi(x),$$

¹⁾ Vgl. [II, 212] (Anm. d. Red.).

woraus auf Grund der Definition der Ableitung eines Funktional ersichtlich ist, daß ein Funktional m_0 mit dem Kern $-\frac{1}{4\pi r}$ der inhomogenen Laplaceschen Differentialgleichung

$$\frac{\partial^2 m_0}{\partial x_1^2} + \frac{\partial^2 m_0}{\partial x_2^2} + \frac{\partial^2 m_0}{\partial x_3^2} = \delta(x - x_0) \quad (218)$$

genügt. Damit haben wir gezeigt, daß eine der Lösungen von (218) ein Funktional vom Typ einer Funktion ist. Zur Bestimmung aller Lösungen von (218) genügt es, alle Lösungen der homogenen Laplaceschen Gleichung

$$\Delta m = \frac{\partial^2 m}{\partial x_1^2} + \frac{\partial^2 m}{\partial x_2^2} + \frac{\partial^2 m}{\partial x_3^2} = 0 \quad (219)$$

zu suchen, die die Form von Funktionalen haben.

Wir zeigen nun, daß alle Funktionale m , die der Gleichung (219) genügen, durch Kerne darstellbar sind, die sich aus in D harmonischen Funktionen zusammensetzen. Die Gleichung (219) ist äquivalent mit

$$(m, \Delta \varphi) = 0, \quad (220)$$

wobei $\varphi(x)$ eine beliebige Funktion aus $\hat{C}^{(2)}(D)$ ist. Für $\varphi(x)$ wählen wir die Funktion

$$\varphi(x) = \frac{1}{4\pi |x-y|} \left[\psi\left(\frac{|x-y|}{\varrho_1}\right) - \psi\left(\frac{|x-y|}{\varrho_2}\right) \right]; \quad (221)$$

dabei ist $\psi(\xi)$ eine nichtnegative, unendlich oft differenzierbare Funktion, die für $\xi \in \left[0, \frac{1}{2}\right]$ gleich 1 und für $\xi \geq 1$ gleich 0 ist. Liegt der Punkt y im Innern von D , so gehört $\varphi(x)$ für hinreichend kleine ϱ_1 und ϱ_2 zu $\hat{C}^{(2)}(D)$. Die Funktion

$$\omega_\varrho(|x-y|) = \begin{cases} 0 & \text{für } x=y, \\ \Delta_x \left[\frac{1}{4\pi |x-y|} \psi\left(\frac{|x-y|}{\varrho}\right) \right] & \text{für } |x-y| > 0, \end{cases}$$

die für $|x-y| \leq \frac{1}{2}\varrho$ und $|x-y| \geq \varrho$ verschwindet, läßt sich als Mittelungskern benutzen, denn es ist

$$\begin{aligned} \int \omega_\varrho(|x-y|) dx &= \int_{|x-y| < \varrho} \Delta_x \left[\frac{1}{4\pi |x-y|} \psi\left(\frac{|x-y|}{\varrho}\right) \right] dx \\ &= \int_{r=1} \frac{\partial}{\partial r} \left[\frac{1}{4\pi r} \psi(r) \right] dS - \int_{r=\frac{1}{2}} \frac{\partial}{\partial r} \left[\frac{1}{4\pi r} \psi(r) \right] dr = 1 \end{aligned}$$

und

$$\omega_\varrho(|x-y|) = \frac{1}{\varrho^3} \chi\left(\frac{|x-y|}{\varrho}\right) \quad \text{mit} \quad \chi(\xi) = \left(\frac{\partial^2}{\partial \xi^2} + \frac{2}{\xi} \frac{\partial}{\partial \xi} \right) \left[\frac{1}{4\pi \xi} \psi(\xi) \right].$$

Setzen wir in (220) statt $\varphi(x)$ die Funktion (221) ein, so erhalten wir

$$0 = (m, \omega_{\varrho_1}(|x-y|) - \omega_{\varrho_2}(|x-y|)). \quad (222)$$

Wegen (209) können wir diese Gleichung in der Gestalt $m_{\varrho_1}(y) = m_{\varrho_2}(y)$ für diejenigen y schreiben, deren Abstand von der Berandung von D größer als $\max(\varrho_1, \varrho_2)$ ist. Wir betrachten die Gesamtheit V_δ aller $\varphi(x) \in \hat{C}^{(2)}(D)$, die außerhalb eines gewissen Gebietes $D_\delta \subset D$, das vom Rand von D den Abstand $\delta > 2 \max(\varrho_1, \varrho_2)$ hat, gleich 0 sind. Für diese $\varphi(x)$ gilt

$$\int_{D_\delta} \varphi(y) m_{\varrho_1}(y) dy = (\tilde{m}_{\varrho_1}, \varphi) = \int_{D_\delta} \varphi(y) m_{\varrho_2}(y) dy = (\tilde{m}_{\varrho_2}, \varphi).$$

Andererseits ist, wie bereits bewiesen, $(\tilde{m}_{\varrho_1}, \varphi) \rightarrow (m, \varphi)$ für $\varrho_1 \rightarrow 0$. Folglich gilt für alle $\varrho_k < \frac{\delta}{2}$ und diese $\varphi(x)$

$$(m, \varphi) = (\tilde{m}_{\varrho_k}, \varphi) = \int_{D_\delta} \varphi(y) m_{\varrho_k}(y) dy \quad (k=1, 2),$$

d. h., das Funktional m wird durch den Kern

$$\tilde{m}_{\varrho_k}(x) = \begin{cases} 0 & \text{für } x \in D - D_{\varrho_k}, \\ m_{\varrho_k}(x) & \text{für } x \in D_{\varrho_k} \end{cases}$$

bestimmt. Die Funktion $m_{\varrho_k}(x)$ ist in D_δ harmonisch; denn für $\varphi(x) \in V_\delta$ folgt aus (220)

$$0 = (m, \Delta_x \varphi) = \int_{D_\delta} m_{\varrho_1}(x) \Delta_x \varphi(x) dx = \int_{D_\delta} \varphi(x) \Delta_x m_{\varrho_1}(x) dx,$$

und da V_δ in $L_2(D_\delta)$ dicht ist, gilt $\Delta_x m_{\varrho_1}(x) = 0$ für $x \in D_\delta$.

Da die Zahl $\delta > 0$ beliebig und $m_{\delta'}(x) = m_\delta(x)$ für $\delta' < \delta$ und $x \in D_\delta$ ist, können wir behaupten, daß die Schar der harmonischen Funktionen $m_\delta(x)$ die in D harmonische Funktion $m(x)$ bestimmt, die auf D_δ mit $m_\delta(x)$ übereinstimmt. Diese harmonische Funktion erzeugt das oben untersuchte Funktional m ; denn nehmen wir ein beliebiges $\varphi(x) \in \dot{C}^{(l)}(D)$, so ist dieses außerhalb eines gewissen Gebietes D_δ gleich 0, und für $\varphi(x)$ gilt auf Grund der Überlegungen

$$(m, \varphi) = \int_{D_\delta} m_\delta(x) \varphi(x) dx = \int_D m(x) \varphi(x) dx. \quad (223)$$

Das Verhalten von $m(x)$ bei Annäherung von x an die Berandung von D wird dadurch charakterisiert, daß das Integral (223) für jede Funktion $\varphi(x) \in \dot{C}^{(l)}(D)$ konvergieren muß.

Wir erwähnen noch eine Folgerung dieses Resultats. Wir nehmen an, $\psi(x)$ sei eine in einem endlichen Gebiet D summierbare Funktion, und es gelte

$$\int_D \psi(x) \Delta_x \varphi(x) dx = 0$$

für jede Funktion $\varphi(x) \in \dot{C}^{(2)}(D)$. Das Funktional

$$(m, \varphi) = \int_D \psi(x) \varphi(x) dx$$

genügt der Gleichung (219), und auf Grund des oben Erwähnten sehen wir, daß $\psi(x)$ mit einer in D harmonischen Funktion äquivalent ist.

Analog zum Vorhergehenden lassen sich auch lineare Funktionale auf verschiedenen Funktionenfamilien untersuchen. Dazu betrachten wir das folgende Beispiel. Es sei K die Familie der reellen Funktionen $\varphi(x)$, die im ganzen R_n definiert und finit sind und stetige Ableitungen beliebig hoher Ordnung besitzen. Die Familie K ist ein linearer Raum. Er ist nicht im gewöhnlichen Sinne normierbar, und wir führen für ihn nur die folgende Definition ein:

Definition. Eine Folge von Funktionen $\varphi_k(x)$, $k=1, 2, \dots$, aus K strebt gegen 0, wenn ein beschränktes Gebiet existiert, in dessen Äußerem alle Funktionen $\varphi_k(x)$ gleich 0 sind, und wenn die $\varphi_k(x)$ und jede Ableitung dieser Funktionen für $k \rightarrow \infty$ gleichmäßig gegen 0 streben.

Ein Funktional (m, φ) auf K wird dadurch definiert, daß jedem $\varphi(x) \in K$ eine reelle Zahl (m, φ) zugeordnet wird. Ein solches Funktional heißt linear (oder linear und stetig), wenn es distributiv ist, d. h., wenn

$$(m, c_1 \varphi_1 + c_2 \varphi_2) = c_1 (m, \varphi_1) + c_2 (m, \varphi_2)$$

gilt und es die Eigenschaft besitzt, daß bei Konvergenz der Folge der $\varphi_k(x)$ gegen 0 auch $(m, \varphi_k) \rightarrow 0$ gilt. Funktionale vom Typ einer Funktion werden durch die Formel (205) definiert, wobei D mit R_n übereinstimmt und $\varphi(x)$ eine in einem beliebigen beschränkten Gebiet summierbare Funktion ist. Die Multiplikation eines Funktional mit einer beliebig oft differenzierbaren Funktion $\omega(x)$ ist durch $(\omega m, \varphi) = (m, \omega \varphi)$ und die Differentiation eines Funktional durch

$$(D^k m, \varphi) = (-1)^k (m, D^k \varphi)$$

definiert. Ein Funktional besitzt Ableitungen beliebig hoher Ordnung.

Die Theorie der Funktionale auf dem Raum K wird in Band I der in der Fußnote auf S. 312 zitierten Bücher behandelt.

V. Der Hilbertsche Raum

§1. Die Theorie der beschränkten Operatoren

120. Die Axiome des Raumes. Bei der Behandlung des Funktionenraumes L_2 und des Folgenraumes l_2 haben wir gezeigt, daß die beiden Räume in ihrer geometrischen Struktur übereinstimmen. Sie stellen verschiedene Realisierungen ein und desselben abstrakten Raumes dar, den wir in diesem Kapitel behandeln wollen. Er wurde zum erstenmal in der Gestalt l_2 von HILBERT eingeführt und wird gewöhnlich der Raum H oder Hilbertraum genannt.

Wie wir im folgenden sehen werden, ist H ein Spezialfall des Banachraumes; alles über Räume dieses Typs Gesagte läßt sich auch auf den Raum H übertragen. Jedoch besitzt H darüber hinaus noch besondere spezifische Eigenschaften.

Wir wollen nun die Axiome, durch die der Hilbertraum H definiert wird, aufzählen. H ist ein linearer Raum, dessen Elemente dem Axiom A aus [95] genügen. Dabei berücksichtigen wir, daß wir die Elemente von H mit komplexen Zahlen multiplizieren können (komplexer linearer Raum).

Ferner nehmen wir, falls nichts Gegenteiliges vorausgesetzt wird, an, daß n linear unabhängige Elemente für jedes natürliche n existieren (Axiom B aus [95]). Jetzt führen wir ein neues Axiom über das Skalarprodukt ein.

Axiom C. *Jedes Paar von Elementen x, y aus H definiert eine komplexe Zahl, die wir das Skalarprodukt aus x und y nennen und mit (x, y) bezeichnen. Dieses Skalarprodukt besitzt die Eigenschaften*

$$\left. \begin{aligned} (y, x) &= \overline{(x, y)}, \\ (x' + x'', y) &= (x', y) + (x'', y), \\ (x, x) &> 0 \quad \text{für} \quad x \neq 0, \\ (ax, y) &= a(x, y). \end{aligned} \right\} \quad (1)$$

Dabei bedeutet $x \neq 0$, daß x nicht das Nullelement ist. Aus diesen Eigenschaften läßt sich unmittelbar schließen:

$$\left. \begin{aligned} (x, y' + y'') &= (x, y') + (x, y''), \\ (x, ay) &= \bar{a}(x, y), \\ (x, x) &= 0 \quad \text{für} \quad x = 0, \\ (x, y) &= 0 \quad \text{für} \quad x = 0 \quad \text{oder} \quad y = 0. \end{aligned} \right\} \quad (2)$$

Der Ausdruck $\sqrt{(x, x)}$, der positiv oder gleich 0 sein kann, heißt die *Norm* des Elements x und wird wie in [95] mit $\|x\|$ bezeichnet. Wir weisen jetzt nach, daß die so definierte Norm den drei Bedingungen des Axioms C aus [95] genügt. Aus ihrer Definition folgt sofort $\|x\| \geq 0$, wobei das Gleichheitszeichen nur für das Null-

element gilt. Ferner ist $||ax||^2 = (ax, ax) = |a|^2 (x, x) = |a|^2 ||x||^2$, also

$$||ax|| = |a| \cdot ||x||. \quad (3)$$

Daraus folgt $||-x|| = ||x||$ [95].

Nun ist noch die Gültigkeit der Ungleichung

$$||x+y|| \leq ||x|| + ||y|| \quad (4)$$

nachzuweisen. Dazu beweisen wir vorbereitend die Ungleichung

$$|(x, y)| \leq ||x|| \cdot ||y||, \quad (5)$$

die wir im folgenden stets *Schwarzsche Ungleichung*¹⁾ nennen werden.

Für beliebige Elemente x und y aus H und beliebige komplexe Zahlen a und b gilt

$$\begin{aligned} ||ax+by||^2 &= (ax+by, ax+by) \\ &= a\bar{a}(x, x) + a\bar{b}(x, y) + \bar{a}b(y, x) + b\bar{b}(y, y) \geq 0. \end{aligned}$$

Die Diskriminante dieser (bezüglich der Veränderlichen a und b) positiven hermiteschen Form muß nichtnegativ sein, d. h. $(x, x)(y, y) - (x, y)(y, x) \geq 0$ oder

$$||x||^2 ||y||^2 - |(x, y)|^2 \geq 0,$$

woraus (5) folgt. Wir gehen nun zum Beweis von (4) über. Beachten wir, daß

$$(x, y) + (y, x) = 2 \operatorname{Re} (x, y) \quad (6)$$

ist, so erhalten wir

$$||x+y||^2 = (x+y, x+y) = ||x||^2 + ||y||^2 + 2 \operatorname{Re} (x, y)$$

oder wegen $|\operatorname{Re} (x, y)| \leq |(x, y)|$ und (5)

$$||x+y||^2 \leq ||x||^2 + ||y||^2 + 2||x|| \cdot ||y|| = (||x|| + ||y||)^2;$$

hieraus folgt (4).

Wie in [95] ergeben sich aus (4) die Ungleichungen

$$||x-y|| \geq ||x|| - ||y||$$

und

$$||x-y|| \leq ||x-z|| + ||z-y||. \quad (7)$$

Aus dem Normbegriff folgt so wie in [95] der Begriff des Abstandes zweier Elemente x und y voneinander, $\varrho(x, y) = ||x-y||$, und der Begriff des Limes einer Folge $\{x_n\}$ (starke Konvergenz),

$$x_n \Rightarrow x_0 \text{ im Fall } ||x_0 - x_n|| \rightarrow 0. \quad (8)$$

Das früher über den Limes Gesagte behält seine Gültigkeit; wir sahen [95], daß $a_n x_n \Rightarrow a_0 x_0$ und $x_n + y_n \Rightarrow x_0 + y_0$ für $a_n \rightarrow a_0$, $x_n \Rightarrow x_0$ und $y_n \Rightarrow y_0$ ist.

Satz. Für $x_n \Rightarrow x_0$ und $y_n \Rightarrow y_0$ ist $(x_n, y_n) \rightarrow (x_0, y_0)$.

Beweis. Wir setzen $u_n = x_n - x_0$, $v_n = y_n - y_0$. Nach Voraussetzung streben dann $||u_n||$ und $||v_n||$ gegen 0. Nun gilt

$$\begin{aligned} (x_0, y_0) - (x_n, y_n) &= (x_0, y_0) - (x_0 + u_n, y_0 + v_n) \\ &= -(x_0, v_n) - (u_n, y_0) - (u_n, v_n) \end{aligned}$$

¹⁾ Im Original Bunjakowskische Ungleichung genannt (Anm. d. Red.).

oder, wenn wir (5) berücksichtigen,

$$|(x_0, y_0) - (x_n, y_n)| \leq \|x_0\| \cdot \|v_n\| + \|u_n\| \cdot \|y_0\| + \|u_n\| \cdot \|v_n\|,$$

so daß wir wegen $\|u\|_n \rightarrow 0$, $\|v_n\| \rightarrow 0$

$$(x_n, y_n) \rightarrow (x_0, y_0)$$

erhalten.

Für $y_n = x_n$ folgt hieraus $\|x_n\| \rightarrow \|x_0\|$, sobald $x_n \Rightarrow x_0$ ist (vgl. [95]).

Hat die Folge x_n ein Grenzelement, so ist sie in sich konvergent, d. h., es ist $\|x_n - x_m\| \rightarrow 0$ für $n, m \rightarrow \infty$ [95].

Wir nehmen nun an, H sei ein vollständiger Raum:

Axiom D. *Ist die Folge x_n in sich konvergent, so existiert ein Element $x_0 \in H$ derart, daß $x_n \Rightarrow x_0$ ist.*

Ferner verwenden wir das folgende Axiom:

Axiom E. *Der Raum H ist separabel.*

Mit anderen Worten, es gibt eine abzählbare Menge von Elementen aus H , die in H dicht ist.

Aus dem Vorhergehenden folgt unmittelbar, daß H ein Banachraum ist.

121. Orthogonalität und Orthogonalsysteme von Elementen. Aus $(x, y) = 0$ folgt $(y, x) = 0$ wegen (1). In diesem Fall heißen die Elemente x und y *orthogonal*, und man schreibt $x \perp y$. Wegen (2) ist das Nullelement zu jedem Element orthogonal.

Es seien $x_1, x_2, \dots, x_m$ paarweise orthogonale Elemente, d. h., es sei $(x_p, x_q) = 0$ für $p \neq q$. Das Quadrat der Norm aus der Summe dieser Elemente ist

$$\|x_1 + x_2 + \dots + x_m\|^2 = (x_1 + x_2 + \dots + x_m, x_1 + x_2 + \dots + x_m).$$

Hieraus erhalten wir, wenn wir das Skalarprodukt mit Hilfe von (1) und (2) umformen, für paarweise orthogonale Elemente den Lehrsatz des PYTHAGORAS

$$\|x_1 + x_2 + \dots + x_m\|^2 = \|x_1\|^2 + \|x_2\|^2 + \dots + \|x_m\|^2. \quad (9)$$

Mit Hilfe des Limesbegriffs können wir wie in [95] den Begriff der Konvergenz einer unendlichen Reihe

$$u_1 + u_2 + u_3 + \dots, \quad (10)$$

deren Glieder Elemente aus H sind, einführen. Eine solche Reihe heißt *konvergent*, wenn die n -te Partialsumme $s_n = u_1 + u_2 + \dots + u_n$ für $n \rightarrow \infty$ stark gegen ein Element u konvergiert, das wir die *Summe* der Reihe (10) nennen. Aus dem Vollständigkeitsaxiom folgt sofort die notwendige und hinreichende Bedingung für die Konvergenz von (10): *Für jedes gegebene $\varepsilon > 0$ existiert ein N derart, daß*

$$\|u_{n+1} + u_{n+2} + \dots + u_{n+p}\| \leq \varepsilon \quad \text{für } n \geq N, p \geq 1 \quad (10_1)$$

ist. Diese Bedingung hat dann eine besonders einfache Form, wenn die Glieder der Reihe (10) paarweise orthogonal sind:

Satz. *Sind die Glieder der Reihe (10) paarweise orthogonal, so konvergiert (10) genau dann, wenn die aus nichtnegativen Zahlen bestehende Reihe*

$$\sum_{k=1}^{\infty} \|u_k\|^2 \quad (11)$$

konvergent ist.

In diesem Fall kann nämlich (10_1) auf Grund des Satzes des PYTHAGORAS in der Form

$$||u_{n+1}||^2 + ||u_{n+2}||^2 + \dots + ||u_{n+p}||^2 \leq \varepsilon \quad \text{für } n \geq N, p \geq 1$$

geschrieben werden, und diese Bedingung ist notwendig und hinreichend für die Konvergenz von (11).

Bei einer beliebigen Umordnung der Glieder von (10) erfahren die Glieder von (11) die gleiche Umordnung. Dadurch wird jedoch die Konvergenz von (11) nicht gefährdet. Folglich hat auch die Umordnung der Glieder von (10) keinen Einfluß auf die Konvergenz; falls die Reihe (10) konvergent ist, bleibt sie es auch nach Umordnen der Glieder. Das Analoge gilt, wenn (10) divergent ist. Es läßt sich mit Hilfe des Satzes des PYTHAGORAS und der Konvergenz von (11) leicht zeigen, daß *im betrachteten Fall die Summe der Reihe nicht von der Reihenfolge der Summanden abhängt*.

Wir sagen, die Folge der Elemente

$$x_1, x_2, x_3, \dots \quad (12)$$

bilde ein *normiertes Orthogonalsystem* (oder *orthonormiertes System* oder *Orthonormalsystem*), wenn

$$(x_p, x_q) = \begin{cases} 0 & \text{für } p \neq q, \\ 1 & \text{für } p = q \end{cases} \quad (13)$$

ist. Unter Berücksichtigung des obigen Satzes können wir behaupten, daß die Reihe

$$\sum_{k=1}^{\infty} a_k x_k \quad (14)$$

genau dann konvergiert, wenn die aus nichtnegativen Zahlen bestehende Reihe

$$\sum_{k=1}^{\infty} |a_k|^2 \quad (15)$$

konvergent ist.

Wir setzen voraus, daß diese Bedingung erfüllt ist, und bezeichnen die Summe von (15) mit x . Das Skalarprodukt $\left(\sum_{k=1}^n a_k x_k, x_p\right)$ ist für $n \geq p$ wegen (13) gleich a_p und wir erhalten folglich für $n \rightarrow \infty$

$$a_p = (x, x_p). \quad (16)$$

Die durch diese Formel definierte Zahl a_p nennen wir den *Fourierkoeffizienten* des Elements x bezüglich des Systems (12) und die Reihe (14) die *Fourierreihe* von x . Offenbar ist

$$\left\| x - \sum_{k=1}^n a_k x_k \right\|^2 = ||x||^2 - \sum_{k=1}^n |a_k|^2, \quad (17)$$

und für $n \rightarrow \infty$ ergibt sich hieraus die Parsevalsche Gleichung (Vollständigkeitsrelation)

$$||x||^2 = \sum_{k=1}^{\infty} |a_k|^2, \quad (18)$$

Aus diesen Überlegungen folgt: *Wenn die Reihe (14) konvergiert, so ist sie die Fourierreihe für ihre Summe x , und es gilt die Gleichung (18).*

Wir nehmen nun an, daß ein Element x aus H gegeben ist, bestimmen seine Fourierkoeffizienten (16) und stellen die Gleichung (17) auf. Aus ihr folgt die Besselsche Ungleichung

$$\sum_{k=1}^{\infty} |a_k|^2 \leq \|x\|^2. \quad (19)$$

Die links stehende Reihe, d. h. die Fourierreihe eines beliebigen Elements x , ist unbedingt konvergent. Gilt in (19) das Gleichheitszeichen, so bedeutet dies wegen (17), daß die Summe der Fourierreihe von x gerade gleich x ist. Das System (12) heißt *vollständig*, wenn in (19) für jedes Element x aus H das Gleichheitszeichen gilt. Ferner wollen wir das System (12) *abgeschlossen* nennen, wenn es außer dem Nullelement kein anderes Element aus H gibt, das zu allen x_k orthogonal ist. Wie früher [58] können wir zeigen, daß Abgeschlossenheit und Vollständigkeit äquivalente Begriffe sind. Ist das System (12) vollständig, so ist jedes Element x aus H als konvergente Reihe (14), nämlich als seine Fourierreihe, eindeutig darstellbar. Sind a_k die Fourierkoeffizienten von x und b_k die eines Elements y und ist das System (12) vollständig, so erhalten wir wie in [58] die *verallgemeinerte Parsevalsche Gleichung oder verallgemeinerte Vollständigkeitsrelation*

$$(x, y) = \sum_{k=1}^{\infty} a_k b_k. \quad (18_1)$$

Wir bemerken noch, daß

$$\left\| x - \sum_{k=1}^n c_k x_k \right\|^2 = \|x\|^2 - \sum_{k=1}^n |a_k|^2 + \sum_{k=1}^n |c_k - a_k|^2$$

(vgl. [58]) gilt, wobei die c_k beliebige komplexe Zahlen und die a_k die Fourierkoeffizienten von x sind. Ein Vergleich mit (17) zeigt, daß die linke Seite dieser Formel ihren kleinsten Wert annimmt, wenn die c_k die Fourierkoeffizienten von x sind.

Fassen wir den Raum H als den Funktionenraum L_2 auf, so entspricht die Konvergenz in H der Konvergenz im Mittel in L_2 , mit der wir uns in [56] beschäftigten. Die Konvergenz der Fourierreihe führt dabei auf die Parsevalsche Gleichung aus [58].

Wir vergegenwärtigen uns nun wieder das Orthogonalisierungsverfahren, das wir schon für den n -dimensionalen komplexen Raum benutzt haben [III/1, 31]. Gegeben sei eine unendliche Folge von Elementen

$$z_1, z_2, z_3, \dots \quad (20)$$

aus H , die nicht das Nullelement enthält. Wir bilden das normierte Element $x_1 = z_1 / \|z_1\|$. Ist z_s das erste der Elemente (20) nach z_1 , das nicht in der Form $a_1 x_1$ dargestellt werden kann, so bilden wir das sicher von 0 verschiedene Element

$$y_2 = z_s - (z_s, x_1) x_1$$

und normieren es, d. h., wir bestimmen $x_2 = y_2 / \|y_2\|$. Ist nun z_t das erste der Elemente (20) nach z_2 , das sich nicht in der Form $a_1 x_1 + a_2 x_2$ darstellen läßt, so konstruieren wir das von 0 verschiedene Element

$$y_3 = z_t - (z_t, x_1) x_1 - (z_t, x_2) x_2$$

und normieren es: $x_3 = y_3 / \|y_3\|$. Setzen wir diesen Prozeß fort, so erhalten wir das normierte Orthogonalsystem (12). Dieses System hat die Eigenschaft, daß sich jedes Element x_k als Linearkombination der Elemente $z_1, z_2, \dots, z_k$ darstellen läßt

und daß jedes Element z_k allein durch die Elemente $x_1, x_2, \dots, x_k$ ausgedrückt werden kann. Wir bemerken noch, daß die orthogonalen und von 0 verschiedenen y_s ($s=1, 2, \dots, m$) linear unabhängig sind. Um dies zu zeigen, nehmen wir $c_1y_1 + c_2y_2 + \dots + c_my_m = 0$ an. Multiplizieren wir beide Seiten dieser Gleichung mit y_k , so erhalten wir wegen der Orthogonalität $c_k||y_k||^2 = 0$, also $c_k = 0$ ($k=1, 2, \dots, m$), woraus die lineare Unabhängigkeit der y_s ($s=1, 2, \dots, m$) folgt.

Da H separabel ist, existiert eine in H dichte abzählbare Menge M von Elementen

$$u_1, u_2, u_3, \dots \quad (21)$$

Orthogonalisieren wir die Folge u_k , so erhalten wir ein abgeschlossenes (vollständiges) normiertes Orthogonalsystem x_k ($k=1, 2, \dots$) aus abzählbar unendlich vielen Elementen. Die Vollständigkeit folgt unmittelbar daraus, daß die Menge (21) überall in H dicht ist. Blieben nach der Orthogonalisierung nur endlich viele Elemente übrig, so wäre H endlichdimensional.

Existiert umgekehrt in H ein abgeschlossenes normiertes Orthogonalsystem x_k ($k=1, 2, \dots$) aus abzählbar unendlich vielen Elementen, so läßt sich leicht zeigen, daß die Summen

$$c_1x_1 + c_2x_2 + \dots + c_kx_k$$

mit komplexen rationalen Koeffizienten $c_s = a_s + ib_s$ (a_s, b_s reell und rational) eine in H dichte abzählbare Menge bilden, d. h., *die Separabilität ist der Existenz eines abgeschlossenen normierten Orthogonalsystems in H aus abzählbar unendlich vielen Elementen gleichwertig.*

Aus der Separabilität von H folgt ferner, daß *jedes normierte Orthogonalsystem endlich oder abzählbar unendlich viele Elemente enthält.*

Es seien x und y zwei zueinander orthogonale und normierte Elemente, es sei also $(x, y) = 0$ und $||x|| = ||y|| = 1$. Dann gilt $||x - y||^2 = (x - y, x - y) = 2$ oder $||x - y|| = \sqrt{2}$; d. h., der Abstand zwischen zwei orthogonalen und normierten Elementen ist gleich $\sqrt{2}$. Wir nehmen nun an, daß eine Menge $\mathcal{S}(v)$ von orthogonalen und normierten Elementen existiert, und wählen ein ε mit $0 < \varepsilon < \sqrt{2}/2$. Zu jedem $v \in \mathcal{S}(v)$ gibt es in der H dichten Menge (21) ein Element u_k mit der Eigenschaft $||u_k - v|| \leq \varepsilon$. Andererseits kann für ein festes k nur ein Element aus $\mathcal{S}(v)$ der Ungleichung $||u_k - v|| \leq \varepsilon$ genügen. Würden nämlich zwei Elemente v_1 und v_2 diese Ungleichung erfüllen, so erhielten wir nach der Dreiecksungleichung $||v_2 - v_1|| \leq 2\varepsilon < \sqrt{2}$; es muß jedoch $||v_2 - v_1|| = \sqrt{2}$ sein. Es folgt also unmittelbar, daß $\mathcal{S}(v)$ eine endliche oder abzählbare Menge ist.

122. Die Projektion. Im folgenden spielen lineare Mannigfaltigkeit und Unterraum [95] eine wesentliche Rolle.

Die Menge der Elemente, die einem beliebig festgelegten Unterraum L angehören, genügt allen oben angegebenen Axiomen, außer möglicherweise dem Axiom B, da der Unterraum L auch endlichdimensional sein kann. *Folglich kann jeder unendlichdimensionale Unterraum L als selbständiger komplexer Hilbertscher Raum aufgefaßt werden.* Diese Behauptung gilt offenbar in bezug auf alle Axiome mit Ausnahme des Separabilitätsaxioms. Bezüglich dieses Axioms müssen wir beweisen: Ist H separabel, so stellt auch jeder Unterraum L von H einen separablen Hilbertschen Raum dar. Der Beweis dieser Behauptung bereitet keine Schwierigkeiten [94].

Zwei Unterräume L und M heißen *orthogonal*, wenn jedes Element aus L zu jedem Element aus M orthogonal ist. Man schreibt dafür $L \perp M$. Ein Element x heißt *orthogonal zum Unterraum L* , wenn es zu jedem Element aus L orthogonal

ist. Man schreibt dafür $x \perp L$. Wir beweisen nun einen Satz, der für das Folgende von grundlegender Bedeutung ist.

Satz. Ist L ein Unterraum von H , so kann jedes Element x aus H in der Gestalt

$$x = y + z \quad (22)$$

dargestellt werden, wobei $y \in L$ und $z \perp L$ ist. Die Darstellung (22) ist eindeutig.

Für $x \in L$ erhalten wir die Darstellung (22), wenn wir $x = x + 0$ setzen. Wir nehmen nun an, x liege nicht in L , und

$$d = \inf_{y \in L} \|x - y\|^2 \quad (23)$$

sei die untere Grenze der Menge der positiven Zahlen $\|x - y\|^2$, die entsteht, wenn y die Elemente des Unterraumes L durchläuft. Es existiert eine Folge $\{y_n\}$ von Elementen aus L , für die

$$(x - y_n, x - y_n) = \|x - y_n\|^2 = d_n \rightarrow d \quad (24)$$

gilt. Es sei u ein beliebiges Element aus L . Da L ein Unterraum ist, liegen die Elemente $y_n + \varepsilon u$ für ein beliebiges reelles (und sogar komplexes) ε in L , und mit (23) können wir $(x - y_n - \varepsilon u, x - y_n - \varepsilon u) \geq d$ schreiben. Durch Ausrechnen des Skalarprodukts erhalten wir daraus die Ungleichung

$$(u, u) \varepsilon^2 - 2 \operatorname{Re} (x - y_n, u) \cdot \varepsilon + (d_n - d) \geq 0.$$

Dieser Ausdruck ist für alle reellen ε nichtnegativ; folglich gilt

$$|\operatorname{Re} (u, x - y_n)| \leq \sqrt{d_n - d} \|u\|. \quad (25)$$

Wir wollen diese Abschätzung noch verbessern. Es sei φ das Argument der komplexen Zahl $(u, x - y_n)$, d. h.

$$(u, x - y_n) = |(u, x - y_n)| e^{i\varphi}.$$

Ersetzen wir in der Ungleichung (25) das Element u durch das ebenfalls in L liegende Element $ue^{-i\varphi}$, so erhalten wir wegen $\|ue^{-i\varphi}\| = \|u\|$ und $(ue^{-i\varphi}, x - y_n) = e^{-i\varphi} (u, x - y_n) = |(u, x - y_n)|$ aus (25) die schärfere Abschätzung

$$|(u, x - y_n)| \leq \sqrt{d_n - d} \|u\|. \quad (26)$$

In dieser Ungleichung ist x ein gegebenes Element aus H , $\{y_n\}$ eine der Bedingung (24) genügende Folge von Elementen aus L und u ein beliebiges Element aus L . Wir schätzen nun den Absolutbetrag des Skalarprodukts $(u, y_n - y_m)$ ab. Wenn wir die Differenz $y_n - y_m$ in der Gestalt $y_n - y_m = (y_n - x) + (x - y_m)$ darstellen und die Ungleichung (26) verwenden, erhalten wir

$$|(u, y_n - y_m)| \leq |(u, x - y_n)| + |(u, x - y_m)| \leq (\sqrt{d_n - d} + \sqrt{d_m - d}) \|u\|.$$

Setzen wir hier $u = y_n - y_m$ und dividieren beide Seiten durch $\|y_n - y_m\|$, so kommen wir zu der Ungleichung $\|y_n - y_m\| \leq \sqrt{d_n - d} + \sqrt{d_m - d}$. Offenbar gilt sie auch für $\|y_n - y_m\| = 0$. Ihre rechte Seite strebt für $m, n \rightarrow \infty$ wegen (23) gegen 0; also ist die Folge $\{y_n\}$ in sich konvergent. Nach dem Vollständigkeitsaxiom existiert ein Element y , für das $y_n \rightarrow y$ gilt, und weil L ein Unterraum ist, gehört y zu L . Andererseits erhalten wir durch Grenzübergang in (26) für jedes Element u aus L die Gleichung $(u, x - y) = 0$; die Differenz $x - y$ ist also orthogonal zu L . Wenn wir diese Differenz gleich z setzen, erhalten wir genau die Gleichung (22), und es gilt $y \in L$ und $z \perp L$.

Es bleibt nun nur noch die Eindeutigkeit der erhaltenen Darstellung (22) zu beweisen. Es seien zwei Darstellungen $x = y + z = y_1 + z_1$ gegeben, und es sei $y \in L$, $y_1 \in L$ und $z \perp L$, $z_1 \perp L$. Offenbar gilt $y - y_1 = z_1 - z$. Die linke Seite dieser Gleichung stellt ein Element aus L dar, die rechte Seite ein zu L orthogonales Element. Daraus folgt $(y - y_1, y - y_1) = 0$, also $\|y - y_1\| = 0$, $y_1 = y$ und folglich $z_1 = z$. Damit ist der Satz vollständig bewiesen.

Das in (22) auftretende Element y aus L heißt *Projektion* des Elements x in den Unterraum L . Die Menge der zu dem Unterraum L orthogonalen Elemente bildet offenbar einen gewissen Unterraum, den wir mit M bezeichnen. Nach dem bewiesenen Satz läßt sich jedes Element x aus H eindeutig als Summe zweier Elemente darstellen, von denen das eine in L , das andere in M liegt. Die Menge der zu M orthogonalen Elemente bildet den Unterraum L . Dieser Zusammenhang zwischen den Elementen L und M besteht wechselseitig, und zwei solche Unterräume heißen *komplementäre Unterräume*. Im Fall des reellen dreidimensionalen Raumes sind z. B. die x, y -Ebene und die z -Achse zueinander komplementäre Räume.

Im allgemeinen schreibt man im obigen Fall

$$H = L \oplus M \quad (27)$$

oder

$$L = H \ominus M, \quad M = H \ominus L, \quad (28)$$

so daß $H \ominus M$ der Unterraum der zu M orthogonalen Elemente aus H ist.

123. Lineare Funktionale. Wir hatten früher ein lineares Funktional $l(x)$ über einem Banachschen Raum oder über dem Hilbertschen Raum H definiert und nehmen nun an, daß es über ganz H definiert ist. Seine Norm n_l wird durch

$$n_l = \sup_{\|x\|=1} |l(x)| \quad (29)$$

gegeben, und es ist

$$|l(x)| \leq n_l \|x\|. \quad (30)$$

Beispielsweise ist

$$l(x) = (x, y) \quad (31)$$

mit einem festen Element y aus H ein lineares Funktional. Seine Distributivität folgt aus (1), seine Beschränktheit aus (5):

$$|l(x)| \leq \|y\| \cdot \|x\|. \quad (32)$$

Hier gilt für $x = y$ das Gleichheitszeichen, d. h., den Faktor $\|y\|$ kann man nicht durch einen kleineren ersetzen; $\|y\|$ ist also die *Norm* des Funktionals (31). Ist y das Nullelement, so ist $(x, y) = 0$ für jedes x (*Nullfunktional*).

Die Formel (31) gibt alle möglichen Funktionale über H an; es gilt also der folgende wichtige Satz.

Satz. *Jedes lineare Funktional $l(x)$ läßt sich eindeutig durch die Gleichung (19) darstellen, in der y ein festes Element aus H ist.*

Aus der Distributivität des Funktionals folgt, wenn θ das Nullelement von H ist, $l(\theta) = 0$. Ist L die Menge aller Elemente x , für die $l(x) = 0$ gilt, so ist, da $l(x)$

distributiv und stetig ist, L ein Unterraum. Es kann vorkommen, daß L der gesamte Raum H ist, daß also $l(x)=0$ für jedes Element x gilt. Ein solches Funktional können wir offenbar in der Form $l(x)=(x, \theta)$ darstellen. Wir betrachten nun den allgemeinen Fall, daß der Unterraum L echt in H enthalten ist, und es sei z ein festes Element von H , das nicht in L liegt. Wir können es mit $u \in L$ und $v \perp L$ in der Gestalt $z = u + v$ schreiben; dabei ist $v \neq 0$ [122]. Da v nicht in L liegt, gilt $l(v) \neq 0$.

Mit einem beliebigen Element $x \in H$ bilden wir das Element $w = x - \frac{l(x)}{l(v)} v$ und betrachten

$$l(w) = l(x) - \frac{l(x)}{l(v)} l(v) = l(x) - l(x) = 0.$$

Wir sehen also, daß das Element $w = x - \frac{l(x)}{l(v)} v$ zu L gehört. Wegen $v \perp L$ gilt somit

$$\left(x - \frac{l(x)}{l(v)} v, v\right) = 0$$

oder, nach Einsetzen des Skalarprodukts,

$$(x, v) - \frac{l(x)}{l(v)} \|v\|^2 = 0;$$

daraus folgt schließlich die Darstellung von $l(x)$ als Skalarprodukt:

$$l(x) = \left(x, \frac{\overline{l(v)}}{\|v\|^2} v\right) = (x, y) \quad \text{mit} \quad y = \frac{\overline{l(v)}}{\|v\|^2} v.$$

Es bleibt nun nur noch die Eindeutigkeit der Darstellung von $l(x)$ als Skalarprodukt zu beweisen. Es sei $l(x) = (x, y) = (x, y_1)$. Daraus folgt für jedes $x \in H$ die Gleichung $(x, y - y_1) = 0$. Wenn wir $x = y - y_1$ wählen, erhalten wir $\|y - y_1\| = 0$ und damit $y = y_1$, womit der Satz vollständig bewiesen ist.

Das oben definierte Funktional heißt zuweilen *lineares Funktional erster Art*. Unter einem *linearen Funktional zweiter Art* verstehen wir dann ein beschränktes Funktional, das die Eigenschaft

$$l_1(c_1 x_1 + c_2 x_2 + \dots + c_m x_m) = \bar{c}_1 l_1(x_1) + \bar{c}_2 l_1(x_2) + \dots + \bar{c}_m l_1(x_m)$$

besitzt, bei dem also die konstanten Faktoren, wenn man sie vor das Funktionalzeichen schreibt, in ihre konjugiert komplexen Werte übergehen. Ein Beispiel für ein lineares Funktional zweiter Art stellt ein Skalarprodukt dar, in dem das variable Element x an zweiter und das feste Element y an erster Stelle steht:

$$l_1(x) = (y, x). \quad (31_1)$$

Ist $l_1(x)$ ein lineares Funktional zweiter Art, so ist $l(x) = \overline{l_1(x)}$ ein lineares Funktional erster Art. Aus dieser Bemerkung und aus dem obigen Satz folgt, daß (31₁) die allgemeine Form eines linearen Funktionals zweiter Art ist.

Aus dem obigen Satz folgt, daß jedes lineare Funktional $l(x)$ durch das Element y aus H vollständig bestimmt ist, d. h., der zu H konjugierte Raum H^* ist H selbst. Wir erinnern noch daran, daß ein auf einer überall in H dichten linearen Mannigfaltigkeit L_1 distributives beschränktes Funktional $l(x)$ eindeutig auf ganz H fortgesetzt werden kann derart, daß es über H linear (beschränkt) ist und dieselbe Norm wie in L_1 hat [98].

124. Lineare Operatoren. Wir betrachten nun auf ganz H definierte lineare (beschränkte) Operatoren, deren Werte ebenfalls zu H gehören. Im folgenden wollen wir, wenn nichts anderes gesagt wird, unter linearen Funktionalen und linearen Operatoren stets distributive beschränkte Funktionale bzw. Operatoren, die auf ganz H gegeben sind, verstehen [97, 98]. Die Norm eines Operators A bezeichnen wir mit $\|A\|$ oder n_A . Es ist also

$$\|A\| = n_A = \sup_{\|x\|=1} \|Ax\|. \quad (33)$$

Mit E bezeichnen wir den Einsoperator, es sei also $Ex = x$ für jedes $x \in H$, $\|E\| = 1$. Ist $\|A\| = 0$, so ist A der Nulloperator, d. h., es ist $Ax = 0$ für jedes $x \in H$. Es sei nun L ein Unterraum und $y \in L$, $z \perp L$; dann läßt sich jedes $x \in H$ auf Grund des Satzes aus [122] eindeutig in der Gestalt $x = y + z$ darstellen. Der Operator P_L , der x in y überführt, heißt *Projektionsoperator* oder *Projektor* in L , und man schreibt

$$y = P_L x. \quad (34)$$

Stimmt L mit H überein, so ist $P_L = E$; besteht L aus dem Nullelement, so ist P_L der Nulloperator. Im allgemeinen Fall haben wir $\|P_L x\| \leq \|x\|$, wobei das Gleichheitszeichen genau für $x \in L$ gilt. Hieraus ist ersichtlich, daß $\|P_L\| = 1$ sein muß, wenn P_L nicht der Nulloperator ist. Die Distributivität von P_L folgt daraus, daß sich aus den beiden Zerlegungen $x_1 = y_1 + z_1$ und $x_2 = y_2 + z_2$ ($y_1, y_2 \in L$; $z_1, z_2 \perp L$) die Beziehung

$$x_1 + x_2 = (y_1 + y_2) + (z_1 + z_2)$$

ergibt, wobei $y_1 + y_2 \in L$ und $z_1 + z_2 \perp L$ ist, d. h.

$$P_L(x_1 + x_2) = P_L x_1 + P_L x_2.$$

Analog erhalten wir

$$P_L(ax) = aP_L(x).$$

Wir führen nun einige neue Begriffe ein und erwähnen die elementaren Eigenschaften der linearen Operatoren (vgl. [97]). Das Wort „linear“ werden wir der Kürze wegen oft fortlassen.

Gilt für zwei Operatoren A und B und jedes Element x die Gleichung $Ax = Bx$, so sagen wir, die beiden Operatoren A und B stimmen überein, und schreiben $A = B$. Ist auf einer in H überall dichten linearen Mannigfaltigkeit L ein distributiver und beschränkter Operator A gegeben, so kann man ihn wie im Fall des Funktionalen in eindeutiger Weise auf den gesamten Raum H fortsetzen; Distributivität und Beschränktheit bleiben dabei bestehen, und die Norm ist in H nicht größer als in L .

Sind A und B zwei Operatoren und a und b zwei komplexe Zahlen, so ist $aA + bB$ ein linearer Operator, der durch die Gleichung

$$(aA + bB)x = aAx + bBx \quad (35)$$

definiert wird. Berücksichtigen wir die Ungleichungen

$$\|aAx + bBx\| \leq |a| \cdot \|Ax\| + |b| \cdot \|Bx\| \leq (|a| n_A + |b| n_B) \|x\|,$$

so erhalten wir für die Norm des Operators $aA + bB$

$$\|aA + bB\| \leq |a| n_A + |b| n_B.$$

Operatoren können also mit komplexen Zahlen multipliziert und auch addiert werden. Diese Operationen genügen den gewöhnlichen Gesetzen der Algebra. Die nacheinander erfolgende Anwendung der Operatoren A und B stellt ebenfalls einen linearen Operator dar, den wir mit BA bezeichnen. Die Anwendung derselben Operatoren in umgekehrter Reihenfolge führt zu dem linearen Operator AB , der im allgemeinen von BA verschieden ist. Der Operator BA bzw. AB heißt *Produkt* der Operatoren A und B . Diese Definition läßt sich unmittelbar auf eine beliebige endliche Anzahl von Faktoren erweitern. Im allgemeinen ändert sich das Produkt von Operatoren bei Vertauschung der Faktoren. Ist $AB = BA$, so heißen die Operatoren *vertauschbar*. Es ist

$$||BAx|| \leq n_B ||Ax|| \leq n_B n_A ||x||,$$

die Norm der Produkte BA und AB ist folglich höchstens gleich $n_B n_A$. Ist a eine komplexe Zahl und A ein Operator, so ist die Norm von aA gleich $|a| n_A$. Das Produkt von Operatoren genügt offenbar dem Assoziativgesetz,

$$C(BA) = (CB) A,$$

und dem Distributivgesetz,

$$(A + B) C = AC + BC \quad \text{und} \quad C(A + B) = CA + CB.$$

Wir führen nun den Begriff des *adjungierten Operators* ein. Es sei A ein linearer Operator. Wir betrachten das Skalarprodukt (Ax, y) . Dies ist für jedes feste Element y ein Funktional von x , dessen Distributivität aus der Distributivität des Skalarprodukts und dessen Beschränktheit aus der offenbar geltenden Abschätzung

$$|(Ax, y)| \leq n_A ||y|| \cdot ||x||$$

folgt. Wir können aber jedes Funktional in eindeutiger Weise als Skalarprodukt darstellen. Folglich existiert zu jedem festen Element y ein bestimmtes Element y^* , so daß für jedes x aus H

$$(Ax, y) = (x, y^*) \tag{36}$$

gilt. Folglich stellt diese Gleichung eine bestimmte Vorschrift dar, durch die jedem Element y ein bestimmtes Element y^* zugeordnet wird. Wir schreiben diese Vorschrift in der Form $y^* = A^*y$, wobei A^* ein Operator ist. Die Distributivität von A^* folgt aus der Distributivität der Skalarprodukte (Ax, y) und (x, y^*) bezüglich des zweiten Arguments. Die Beschränktheit von A^* wird sogleich bewiesen. Der Operator A^* heißt *adjungiert* zu A . Die Gleichung (36) läßt sich jetzt in der Form

$$(Ax, y) = (x, A^*y) \tag{37}$$

schreiben.

Aus dieser Definition des adjungierten Operators ergibt sich für die Bildung des adjungierten Operators einer Summe oder eines Produkts von Operatoren

$$(aA)^* = \bar{a}A^*, \quad (A + B)^* = A^* + B^*, \quad (AB)^* = B^*A^*, \quad (A^*)^* = A. \tag{38}$$

Wir beweisen z. B. die dritte Formel. Durch zweimalige Anwendung der Definitionsgleichung (37) erhalten wir $(ABx, y) = (Bx, A^*y) = (x, B^*A^*y)$, woraus $(AB)^* = B^*A^*$ folgt. Wir beweisen noch die letzte Formel (38). Wenn wir die Definitions-

gleichung (37) und die Eigenschaft $(u, v) = \overline{(v, u)}$ verwenden, so erhalten wir $(A^*x, y) = \overline{(y, A^*x)} = \overline{(Ay, x)} = (x, Ay)$; es gilt also $(A^*)^* = A$.

Wir beweisen nun die Beschränktheit von A^* .

Satz 1. *Die Norm des adjungierten Operators ist gleich der Norm des ursprünglichen Operators, $n_{A^*} = n_A$.*

Wenn wir in (37) $x = A^*y$ annehmen und die Ungleichungen (5) und (33) benutzen, erhalten wir

$$\|A^*y\|^2 = |(A(A^*y), y)| \leq \|A(A^*y)\| \cdot \|y\| \leq n_A \|A^*y\| \cdot \|y\|,$$

woraus sich $\|A^*y\| \leq n_A \|y\|$ und damit $n_{A^*} \leq n_A$ ergibt. Wegen $(A^*)^* = A$ können wir in derselben Weise $n_A \leq n_{A^*}$ beweisen. Daraus folgt schließlich $n_{A^*} = n_A$.

Ein Operator A heißt *selbstadjungiert*, wenn $A^* = A$ gilt. Für einen selbstadjungierten Operator ist also die Gleichung

$$(Ax, y) = (x, Ay) \quad (39)$$

charakteristisch. Setzen wir hier $y = x$ und berücksichtigen $(Ax, x) = \overline{(x, Ax)}$, so erhalten wir, daß (Ax, x) im Fall eines selbstadjungierten Operators für jedes Element x reell ist. Auch die umgekehrte Behauptung ist richtig.

Satz 2. *Für die Selbstadjungiertheit von A ist notwendig und hinreichend, daß (Ax, x) für jedes Element x reell ist.*

Die Notwendigkeit haben wir soeben bewiesen. Wir nehmen nun an, daß (Ax, x) für jedes x reell ist, und zeigen, daß A ein selbstadjungierter Operator ist. Nach Voraussetzung gilt

$$(A(x+y), x+y) = (x+y, A(x+y))$$

und

$$(A(x+iy), x+iy) = (x+iy, A(x+iy)).$$

Wenn wir die Skalarprodukte auflösen und berücksichtigen, daß $(Ax, x) = (x, Ax)$ und $(Ay, y) = (y, Ay)$ gilt, so erhalten wir

$$(Ay, x) + (Ax, y) = (y, Ax) + (x, Ay),$$

$$(Ay, x) - (Ax, y) = (y, Ax) - (x, Ay).$$

Durch Subtraktion folgt (39); A ist also ein selbstadjungierter Operator. *

Unter Berücksichtigung von (38) erkennen wir, daß jede Linearkombination $a_1A_1 + a_2A_2 + \dots + a_mA_m$ selbstadjungierter Operatoren A_k mit reellen Koeffizienten a_k ebenfalls ein selbstadjungierter Operator ist und daß das Produkt AB zweier selbstadjungierter Operatoren dann und nur dann selbstadjungiert ist, wenn A und B vertauschbar sind.

Es sei L ein Unterraum und M der zu ihm komplementäre Raum. Auf Grund des Satzes aus [122] ist

$$E = P_L + P_M. \quad (40)$$

Man sieht leicht, daß jeder Projektionsoperator P_L selbstadjungiert ist. Da L und M orthogonal sind und (40) gilt, folgt nämlich

$$(P_Lx, y) = (P_Lx, P_Ly + P_My) = (P_Lx, P_Ly) = (P_Lx + P_Mx, P_Ly) = (x, P_Ly).$$

Ist A ein beliebiger linearer Operator, so lassen sich die beiden Operatoren

$$A_1 = \frac{1}{2} (A + A^*), \quad A_2 = \frac{1}{2i} (A - A^*) \quad (41)$$

bilden. Mit Hilfe von (38) können wir erkennen, daß A_1 und A_2 selbstadjungiert sind. Wir erhalten so die Darstellung $A = A_1 + iA_2$ eines beliebigen linearen Operators mit Hilfe zweier selbstadjungierter Operatoren.

125. Bilineare und quadratische Funktionale. Wir weisen nun auf die Möglichkeit hin, jeden linearen Operator mit Hilfe einer besonderen Art von Funktionalen zu definieren. Ein *bilineares Funktional* ist eine bestimmte Vorschrift, nach der jedem Paar x, y von Elementen aus H eine bestimmte komplexe Zahl $l(x, y)$ zugeordnet wird, wobei $l(x, y)$ bezüglich des ersten Arguments ein distributives Funktional erster Art und bezüglich des zweiten Arguments ein distributives Funktional zweiter Art ist:

$$\begin{aligned} l(ax_1 + bx_2, y) &= al(x_1, y) + bl(x_2, y), \\ l(x, ay_1 + by_2) &= \bar{a}l(x, y_1) + \bar{b}l(x, y_2). \end{aligned} \quad (42)$$

Wir nehmen an, daß die betrachteten bilinearen Funktionale auch beschränkt sind, d. h., daß eine positive Zahl N existiert derart, daß für beliebige Elemente x und y aus H die Ungleichung

$$|l(x, y)| \leq N \|x\| \cdot \|y\| \quad (43)$$

gilt. Der kleinste Wert von N in dieser Ungleichung, die Norm n_l des bilinearen Funktionals, wird durch

$$n_l = \sup_{\|x\|=1, \|y\|=1} |l(x, y)| \quad (44)$$

definiert.

Ist A ein beliebiger linearer Operator, so wird durch

$$l(x, y) = (Ax, y), \quad (45)$$

wie leicht zu verifizieren ist, ein bilineares Funktional gegeben. Dabei gilt $|l(x, y)| \leq n_A \|x\| \cdot \|y\|$, so daß für das bilineare Funktional (45) $n_l \leq n_A$ ist.

Wir zeigen nun, daß jedes bilineare Funktional in der Form (45) dargestellt werden kann.

Satz 1. *Jedes bilineare Funktional ist in eindeutiger Weise in der Form (45) darstellbar, wobei A ein gewisser linearer Operator und die Norm n_l des bilinearen Funktionals gleich der Norm n_A des Operators ist.*

Wird das Element x festgehalten, so entsteht aus dem bilinearen Funktional $l(x, y)$ ein Funktional zweiter Art bezüglich y ; wir können also $l(x, y) = (z, y)$ schreiben [123], wobei z für ein festes x eindeutig bestimmt ist, d. h., mit einem gewissen Operator A aus H ist $z = Ax$. Die Distributivität von A folgt aus der zweiten Gleichung (42) und der Distributivität von (z, y) bezüglich z . Wir beweisen zunächst die Beschränktheit von A . Unter Berücksichtigung von (43) können wir für $N = n_l$

$$|(Ax, y)| \leq n_l \|x\| \cdot \|y\|$$

schreiben. Setzen wir in dieser Ungleichung $y = Ax$ und dividieren beide Seiten durch $\|Ax\|$, so erhalten wir $\|Ax\| \leq n_l \|x\|$ (für $\|Ax\| \neq 0$ gilt dies sicher). Hieraus folgen aber die Beschränktheit von A und die Ungleichung $n_A \leq n_l$. Da wir früher $n_l \leq n_A$ bewiesen hatten, muß folglich $n_A = n_l$ sein.

Wir beweisen nun noch die Eindeutigkeit der Darstellung (45). Es sei $l(x, y) = (Ax, y) = (A_1x, y)$. Daraus folgt, daß für beliebige x und y die Gleichung $(Ax - A_1x, y) = 0$ gilt. Setzen wir $y = Ax - A_1x$, so erhalten wir $\|Ax - A_1x\| = 0$, d. h., für jedes x gilt $A_1x = Ax$; die Operatoren A und A_1 stimmen also überein. Damit ist der Satz vollständig bewiesen.

Aus Satz 1 folgt, daß die Vorgabe eines linearen Operators gleichbedeutend ist mit der Vorgabe eines bilinearen Funktionals. Analog dazu ist in der Algebra die Vorgabe der Elemente a_{ik} einer Matrix gleichbedeutend mit der Vorgabe der Bilinearform

$$\sum_{i,k=1}^n a_{ik} x_k \bar{y}_i.$$

Jedes bilineare Funktional $l(x, y)$ erzeugt ein ihm zugeordnetes quadratisches Funktional (eine quadratische oder hermitesche Form in der Algebra), wenn wir $y = x$ setzen:

$$l(x, x) = (Ax, x).$$

Ein bilineares Funktional läßt sich leicht durch das ihm zugeordnete quadratische Funktional ausdrücken, und zwar gilt mit

$$x_1 = \frac{1}{2}(x+y), \quad x_2 = \frac{1}{2}(x-y), \quad x_3 = \frac{1}{2}(x+iy), \quad x_4 = \frac{1}{2}(x-iy) \quad (46)$$

die leicht zu verifizierende Gleichung

$$(Ax, y) = [(Ax_1, x_1) - (Ax_2, x_2)] + i[(Ax_3, x_3) - (Ax_4, x_4)]. \quad (47)$$

Auf der rechten Seite dieser Gleichung stehen vier quadratische Funktionale. Die Tatsache, daß das quadratische Funktional (Ax, x) für alle Elemente x reell ist, ist, wie wir gesehen haben, für selbstadjungierte Operatoren charakteristisch.

Der Operator A habe nun die Eigenschaft, daß für alle Elemente x die Beziehung $(Ax, x) = 0$ gilt. Aus (47) folgt dann $(Ax, y) = 0$ für alle x und y . Offenbar besitzt das bilineare Funktional diese Eigenschaft, wenn A der Nulloperator ist. Unter Verwendung der in Satz 1 festgestellten Eindeutigkeit kann man also sagen: *Gilt für alle x die Gleichung $(Ax, x) = 0$, so ist A der Nulloperator.* Daraus ergibt sich unmittelbar: *Sind die Operatoren A und B so beschaffen, daß für alle x die Gleichung $(Ax, x) = (Bx, x)$ folgt, so ist $A = B$.*

126. Die Grenzen eines selbstadjungierten Operators. Es sei A ein selbstadjungierter Operator. Unter Berücksichtigung von (5) und (33) können wir $|(Ax, x)| \leq \leq n_A \|x\|^2$ oder, wenn wir $\|x\| = 1$ annehmen, $|(Ax, x)| \leq n_A$ schreiben. Folglich ist für alle normierten Elemente x , d. h. für alle Elemente x mit $\|x\| = 1$, die Menge der reellen Zahlen (Ax, x) nach unten und nach oben beschränkt. Wir bezeichnen mit m die untere und mit M die obere Grenze dieser Menge:

$$m = \inf_{\|x\|=1} (Ax, x), \quad M = \sup_{\|x\|=1} (Ax, x). \quad (48)$$

Die Zahlen m und M heißen im allgemeinen die *untere* bzw. die *obere Grenze* des selbstadjungierten Operators A . Nach Definition dieser Grenzen gelten die Unglei-

chungen

$$m \leq (Ax, x) \leq M \quad \text{für} \quad \|x\| = 1. \quad (49)$$

Für die Elemente x mit beliebiger Norm gelten statt (49) die Ungleichungen

$$m\|x\|^2 \leq (Ax, x) \leq M\|x\|^2. \quad (50)$$

Im Fall $m = M$ haben wir $(Ax, x) = m\|x\|^2$ für $x \in H$, und der Operator A multipliziert jedes Element x mit der Zahl m :

$$Ax = mx \quad \text{für} \quad x \in H.$$

Es zeigt sich, daß die Norm n_A des Operators A in einfacher Weise durch seine Grenzen m und M ausgedrückt werden kann, und zwar gilt der folgende Satz.

Satz 1. Die Norm n_A ist gleich der größeren der beiden Zahlen $|m|$ und $|M|$.

Der Beweis dieses Satzes ist eine wörtliche Wiederholung des Beweises von Satz 2 aus [IV/1, 34], worin wir zeigten, daß $n_A = \sup_{\|x\|=1} |(Ax, x)|$ ist, und dies ist genau die

Behauptung des Satzes. Wir weisen noch darauf hin, daß sich Satz 2 aus [IV/1, 34] mit der Behauptung $n_l = n_A$ deckt, die wir in [125] bewiesen haben.

Wir führen nun einige weitere Begriffe ein.

Definition. Ein selbstadjungierter Operator A heißt *positiv*, wenn das ihm zugeordnete quadratische Funktional (Ax, x) nichtnegativ ist.

Charakteristisch für einen positiven Operator ist die Ungleichung $m \geq 0$, d. h. die Tatsache, daß seine untere Grenze nichtnegativ ist. Weiterhin heißt der selbstadjungierte Operator A *größer* als der selbstadjungierte Operator B , in Zeichen $A > B$, wenn A und B nicht übereinstimmen und die Differenz $A - B$ ein positiver Operator ist. Analog dazu wird ein negativer Operator definiert. Wir erinnern daran, daß im n -dimensionalen Raum eine selbstadjungierte Matrix positiv heißt, wenn die zugeordnete hermitesche Form

$$\sum_{i,k=1}^n a_{ik} x_i \bar{x}_k \quad (a_{ki} = \bar{a}_{ik})$$

nur nichtnegative Werte annimmt. Eine Matrix ist dann und nur dann positiv, wenn sie keine negativen Eigenwerte besitzt. Wenn (Ax, x) bei verschiedener Wahl von x das Vorzeichen wechselt, so ist der selbstadjungierte Operator weder positiv noch negativ. Im endlichdimensionalen Raum entsprechen diesen Operatoren selbstadjungierte Matrizen, welche Eigenwerte verschiedenen Vorzeichens besitzen.

Satz 2. Ist A ein selbstadjungierter Operator, so ist A^2 ein positiver Operator. Ist A ein beliebiger linearer Operator, so sind die Operatoren AA^* und A^*A selbstadjungiert und positiv.

Die erste Behauptung folgt unmittelbar aus

$$(A^2x, x) = (Ax, Ax) = \|Ax\|^2 \geq 0,$$

die zweite aus den Formeln

$$(AA^*x, x) = (A^*x, A^*x) = \|A^*x\|^2 \geq 0, \quad (51)$$

$$(A^*Ax, x) = (Ax, Ax) = \|Ax\|^2 \geq 0. \quad (52)$$

127. Der inverse Operator. In der Theorie der Operatoren ist der Begriff des inversen Operators von Wichtigkeit (vgl. den Begriff der inversen Matrix in [III/1, 25]. Dieser Begriff kann auf verschiedene Arten definiert werden. Das Ziel dieses Abschnitts ist, verschiedene Definitionen des inversen Operators zu geben.

Definition. Ein linearer Operator A besitzt einen *beschränkten inversen Operator* (oder ein *beschränktes Inverses*) B , wenn B ein beschränkter Operator auf ganz H ist und

$$AB = BA = E \quad (53)$$

gilt. E ist dabei der *identische Operator* (*Einsoperator*), d. h., es ist $Ex = x$ [124]. Die Beschränktheit von B wird wie oben durch $\|Bx\| \leq N\|x\|$ definiert.

Man erkennt leicht, daß es zu einem linearen Operator A nur einen einzigen beschränkten inversen Operator B geben kann. Gilt nämlich $AC = E$, so erhält man durch linksseitige Multiplikation mit B und durch Anwendung von (53) wegen $BE = B$ und $EC = C$ die Gleichung $B = C$. Der inverse Operator B wird gewöhnlich mit A^{-1} bezeichnet; es gilt also

$$AA^{-1} = A^{-1}A = E. \quad (54)$$

Multiplizieren wir in der Gleichung

$$y = Ax \quad (x \in H) \quad (55)$$

beide Seiten mit A^{-1} , was möglich ist, da A^{-1} auf ganz H definiert ist, so erhalten wir

$$x = A^{-1}y. \quad (56)$$

Hat A einen beschränkten inversen Operator A^{-1} , so transformiert folglich A den Raum H eineindeutig auf sich, d. h., jedem Element x aus H ist nach (55) eindeutig ein Element y zugeordnet, und umgekehrt ist jedem Element y aus H eindeutig ein durch (56) definiertes Element x zugeordnet. Auch A^{-1} bildet H eineindeutig auf sich ab. Aus der Distributivität von A folgt, daß auch A^{-1} ein distributiver und damit ein linearer Operator ist. Aus (54) ergibt sich unmittelbar

$$(A^{-1})^{-1} = A. \quad (57)$$

Man kann den inversen Operator auch auf allgemeinere Weise definieren. Wir bemerken zunächst, daß die Menge der durch (55) definierten Elemente y wegen der Distributivität des linearen Operators eine lineare Mannigfaltigkeit bildet, die wir mit $R(A)$ bezeichnen wollen. Wir leiten nun die Eigenschaften her, die der Operator A haben muß, damit die Zuordnung zwischen den Elementen x aus H und den Elementen y aus $R(A)$ eineindeutig ist. Wegen (55) ist jedem Element x aus H ein bestimmtes Element y aus $R(A)$ zugeordnet. Umgekehrt muß auch jedem Element y aus $R(A)$ ein bestimmtes Element x aus H entsprechen. Es seien x_1 und x_2 zwei verschiedene Elemente aus H , y_1 und y_2 die zugeordneten Elemente aus $R(A)$:

$$y_1 = Ax_1, \quad y_2 = Ax_2.$$

Durch Subtraktion folgt

$$y_2 - y_1 = A(x_2 - x_1).$$

Für $y_2 = y_1$, d. h., wenn den verschiedenen Elementen x_1 und x_2 aus H ein und dasselbe Element aus $R(A)$ zugeordnet ist, erhalten wir $A(x_2 - x_1) = 0$, d. h., die Gleichung

$$Ax = 0 \quad (58)$$

hat von 0 verschiedene Lösungen. Hat umgekehrt (58) eine von 0 verschiedene Lösung x_0 , so ist den voneinander verschiedenen Elementen $x = x_0$ und $x = 0$ ein und dasselbe Element $y = 0$ zugeordnet. *Folglich existiert eine eindeutige Zuordnung der Form (55) zwischen den Elementen x aus H und den Elementen y aus $R(A)$ dann und nur dann, wenn die Gleichung (58) nur die Lösung $x = 0$ hat.* Dann ist in der linearen Mannigfaltigkeit $R(A)$ ein zu A inverser Operator B definiert. Er transformiert die Elemente y aus $R(A)$ in die Elemente x aus H derart, daß ein Element y vermöge (55) durch ein Element x ausgedrückt wird. Dieser Operator heißt einfach der zu A *inverse Operator* zum Unterschied von dem oben definierten beschränkten inversen Operator. Der Operator B ist nur auf der linearen Mannigfaltigkeit $R(A)$ definiert, die nicht mit H übereinzustimmen braucht, und wir können nicht behaupten, daß B beschränkt ist. Wegen der Distributivität von A können wir aber behaupten, daß B ein auf der linearen Mannigfaltigkeit $R(A)$ *distributiver Operator* ist. Bezeichnen wir B wie oben mit A^{-1} , so können wir

$$A^{-1}(Ax) = x \quad \text{für } x \in H, \quad A(A^{-1}x) = x \quad \text{für } x \in R(A)$$

schreiben.

Wir werden im folgenden zeigen, daß $B = A^{-1}$ beschränkt ist, d. h., daß A ein beschränktes Inverses besitzt, wenn die Gleichung (58) nur die Nulllösung besitzt und $R(A)$ mit dem Raum H übereinstimmt (vgl. [97]).

Wir nehmen nun an, daß der Operator A einen beschränkten inversen Operator hat, und gehen in (54) zu den adjungierten Operatoren über:

$$(A^{-1})^* A^* = A^* (A^{-1})^* = E^* = E.$$

Daraus ergibt sich

$$(A^{-1})^* = (A^*)^{-1}. \quad (59)$$

Nach Gleichung (53) ist der beschränkte Operator B zu A linksseitig und auch rechtsseitig invers; in diesem Fall nennen wir B einfach den beschränkten inversen Operator.

Wir untersuchen nun beschränkte Operatoren, die nur linksseitig oder nur rechtsseitig invers sind.

Wir sagen, der Operator A habe ein *beschränktes Linksinverses* (oder einfach Linksinverses) B , wenn ein linearer Operator B existiert, für den $BA = E$ gilt. Ist entsprechend $AC = E$, so heißt C ein *beschränktes Rechtseinverses* (oder einfach ein Rechtseinverses).

Satz 1. *Hat A mindestens ein Linksinverses B und mindestens ein Rechtseinverses C , so existiert nur ein Links- und nur ein Rechtseinverses, und sie stimmen beide überein, d. h., es existiert der beschränkte inverse Operator A^{-1} .*

Beweis. Nach Voraussetzung ist $BA = E$ und $AC = E$, woraus $(BA)C = C$ und $B(AC) = B$ folgt. Die linken Seiten der letzten beiden Gleichungen sind gleich, also ist $B = C$, d. h., jedes Linksinverse stimmt mit jedem Rechtseinversen überein; daher existiert nur ein Linksinverses und nur ein Rechtseinverses.

Satz 2. *Gibt es nur ein einziges Linksinverses, so existiert auch ein Rechtsinverses und umgekehrt. In beiden Fällen sind beide Inversen die einzigen, und sie stimmen (nach Satz 1) überein.*

Beweis. Um die erste Behauptung des Satzes zu beweisen, nehmen wir an, es möge nur einen einzigen linksinversen Operator B geben. Für ihn gilt $BA = E$. Multiplizieren wir von links mit A , so erhalten wir $ABA = A$ oder $(AB - E)A = 0$, wobei hier 0 der Nulloperator ist. Durch Addition der Gleichung $BA = E$ ergibt sich $(AB - E + B)A = E$. Nun ist B nach Voraussetzung das einzige Linksinverse. Daher folgt $AB - E + B = B$ oder $AB = E$, d. h., B ist auch ein Rechtsinverses.

Bemerkung. Gibt es zu A zwei verschiedene linksinverse (oder rechtsinverse) Operatoren B und C , so existieren unendlich viele linksinverse (rechtsinverse) Operatoren. Ist nämlich $BA = E$ und $CA = E$, so läßt sich leicht nachprüfen, daß der Operator $B + a(C - B)$ bei beliebiger Wahl der Zahl a linksinvers ist:

$$(B + aC - aB)A = BA + aCA - aBA = E + aE - aE = E.$$

Aus diesen Resultaten ergibt sich, daß die folgenden vier Fälle möglich sind:

- I. Es existiert ein einziges Links- und ein einziges Rechtsinverses (A besitzt einen beschränkten inversen Operator).
- II. Es existiert weder ein Links- noch ein Rechtsinverses.
- III. Es existieren unendlich viele Linksinverse, jedoch kein Rechtsinverses.
- IV. Es existieren unendlich viele Rechtsinverse, jedoch kein Linksinverses.

Wir werden sehen, daß alle diese Fälle auftreten können. Wir geben nun ein einfaches theoretisches Kriterium an, mit dem man diese Fälle unterscheiden kann. Wir betrachten die selbstadjungierten positiven (nichtnegativen) Operatoren A^*A und AA^* . Die unteren Grenzen $m(A^*A)$ bzw. $m(AA^*)$ dieser Operatoren sind größer als 0 oder gleich 0 [126]. Nehmen wir an, daß mindestens ein Linksinverses existiert, $BA = E$, und daß k die Norm von B ist, so gilt $\|BAx\| = \|x\|$ und andererseits $\|BAx\| \leq k\|Ax\|$, also $k\|Ax\| \geq \|x\|$ oder $\|Ax\| \geq \frac{1}{k}\|x\|$. Dann gilt die Unglei-

chung $(A^*Ax, x) \geq \frac{1}{k^2}\|x\|^2$ und schließlich $m(A^*A) \geq \frac{1}{k^2}$, also $m(A^*A) > 0$. Wir beweisen nun, daß auch die Umkehrung gilt: Ist $m(A^*A) > 0$, so besitzt A ein beschränktes Linksinverses. Später werden wir zeigen, daß ein selbstadjungierter Operator F ein beschränktes Inverses hat, wenn die untere Grenze von F positiv ist [129]. Wenden wir dieses Ergebnis auf $F = A^*A$ an, so sehen wir, daß ein beschränkter Operator D existiert, für den $DA^*A = E$ oder $(DA^*)A = E$ gilt. Daraus folgt schließlich, daß DA^* ein beschränktes Linksinverses für A ist. Ebenso existiert dann und nur dann mindestens ein beschränktes Rechtsinverses, wenn die Ungleichung $m(AA^*) > 0$ gilt.

Aus diesen Überlegungen ergeben sich die folgenden für die Realisierung der obigen vier Fälle notwendigen und hinreichenden Bedingungen:

- I. $m(A^*A) > 0$ und $m(AA^*) > 0$;
- II. $m(A^*A) = 0$ und $m(AA^*) = 0$;
- III. $m(A^*A) > 0$ und $m(AA^*) = 0$;
- IV. $m(A^*A) = 0$ und $m(AA^*) > 0$.

Die Fälle III und IV können nicht auftreten, wenn die Operatoren A und A^* vertauschbar sind (z. B. wenn $A = A^*$ gilt, d. h., wenn A ein selbstadjungierter

Operator ist). Unter Berücksichtigung von (51) und (52) können wir das Ergebnis im ersten Fall folgendermaßen formulieren: *Ein links- und rechtsinverser Operator existiert dann und nur dann, wenn es eine positive Zahl l gibt derart, daß für jedes Element x die Ungleichungen $\|Ax\| \geq l\|x\|$ und $\|A^*x\| \geq l\|x\|$ gelten.* In den obigen Überlegungen haben wir nirgends die Distributivität der Operatoren B und C benutzt. Wichtig ist nur, daß sie in ganz H definiert und beschränkt sind. Im Fall I ist der einzige links- und rechtsinverse Operator sicher linear. Im Fall III gibt es einen linearen linksinversen Operator $B = DA^*$; Analoges gilt im Fall IV. Im folgenden werden wir es mit einem inversen distributiven Operator A^{-1} zu tun haben, der auf $R(A)$ definiert ist. Wir erwähnen noch, daß aus $BA = E$ die Gleichung $A^*B^* = E$ folgt; wenn für A der Fall III gilt, so gilt für A^* also der Fall IV.

Die inversen Operatoren spielen eine wichtige Rolle bei der Auflösung der Gleichung $Ax = y$, in der y das gegebene und x das gesuchte Element ist. Existiert ein linksinverser Operator B , so erhalten wir durch Multiplikation dieser Gleichung mit B die Gleichung $x = By$, d. h., bei Vorhandensein eines linksinversen Operators ist die Lösung, wenn sie existiert, in der Form $x = By$ darstellbar, und daher kann die Gleichung $Ax = y$ nur eine einzige Lösung haben. Existiert dagegen ein rechtsinverser Operator C , so genügt $x = Cy$ offenbar, wie man durch Einsetzen erkennt, der Gleichung $Ax = y$; die Existenz eines rechtsinversen Operators garantiert also die Existenz einer Lösung $x = Cy$.

128. Das Spektrum eines Operators. Die Hauptaufgabe bei der Anwendung der Theorie der Operatoren in der Analysis besteht in der Auflösung der homogenen Gleichung

$$Ax = \lambda x, \quad (60)$$

d. h.

$$(A - \lambda E)x = 0, \quad (60_1)$$

und der inhomogenen Gleichung

$$Ax = \lambda x + y, \quad (61)$$

d. h.

$$(A - \lambda E)x = y, \quad (61_1)$$

wobei x das gesuchte Element, y das gegebene Element und λ ein Parameter ist. Ein Wert von λ , für den die homogene Gleichung (60) vom Nullelement verschiedene Lösungen hat, heißt *Eigenwert* des Operators A ; die vom Nullelement verschiedenen Lösungen heißen die zu dem Eigenwert gehörigen *Eigenelemente* des Operators A .

Ist λ ein Eigenwert von A und fügen wir zu den zugehörigen Eigenelementen das Nullelement hinzu, so können wir, da die Gleichung (61) linear und homogen und der Operator A stetig ist, zeigen, daß *die Menge dieser Elemente einen Unterraum bildet*. Wir nennen ihn den *Unterraum der zu dem Eigenwert λ gehörigen Eigenelemente*.

Wenn dieser Unterraum der Eigenelemente eine endliche Dimension r hat, wenn also die Maximalzahl linear unabhängiger Elemente dieses Unterraums gleich r ist, sagt man, daß der zugehörige Eigenwert λ den *Rang* oder die *Vielfachheit* r besitzt. Ist der Unterraum der Eigenelemente unendlichdimensional, so sagt man, daß der Rang des zugehörigen Eigenwertes *unendlich* ist. Für einen selbstadjungierten Operator gilt der folgende Satz.

Satz 1. Die Eigenwerte eines selbstadjungierten Operators sind reell, und die zu verschiedenen Eigenwerten gehörigen Eigenelemente sind zueinander orthogonal.

Beweis. Es sei λ ein Eigenwert des selbstadjungierten Operators A und x ein zugehöriges (vom Nullelement verschiedenes) Eigenelement. Multiplizieren wir beide Seiten von (60₁) von rechts mit x , so erhalten wir $(Ax, x) = \lambda \|x\|^2$. Da A selbstadjungiert ist, ist die linke Seite dieser Gleichung und damit auch die Zahl λ reell. Es seien nun λ' und λ'' zwei verschiedene Eigenwerte und x' und x'' zugehörige Eigenelemente: $Ax' = \lambda'x'$, $Ax'' = \lambda''x''$. Wir multiplizieren die erste Gleichung skalar mit x'' von rechts, die zweite skalar mit x' von links und subtrahieren die erhaltenen Gleichungen voneinander:

$$(Ax', x'') - (x', Ax'') = (\lambda' - \lambda'')(x', x'').$$

Wegen der Selbstadjungiertheit von A ist die linke Seite gleich 0. Da $\lambda' - \lambda'' \neq 0$ ist, gilt also $(x', x'') = 0$, womit der Satz bewiesen ist.

Die Auflösung der inhomogenen Gleichung (61₁) führt auf das Aufsuchen des zum Operator $A - \lambda E$ inversen Operators $(A - \lambda E)^{-1}$. Ist λ ein Eigenwert des Operators A , so hat die homogene Gleichung (60₁) vom Nullelement verschiedene Lösungen, und aus dem in [127] Dargestellten folgt, daß der inverse Operator $(A - \lambda E)^{-1}$ nicht existiert. Ist λ jedoch kein Eigenwert von A , so existiert entweder der beschränkte inverse Operator oder einfach der inverse Operator $(A - \lambda E)^{-1}$, der auf $R(A - \lambda E)$ definiert ist.

Wir geben nun folgende Definition.

Definition. Ein Wert λ oder ein Punkt λ (der komplexen Ebene) heißt *regulärer Punkt* des Operators A , wenn der Operator $A - \lambda E$ einen beschränkten inversen Operator

$$R_\lambda = (A - \lambda E)^{-1} \quad (62)$$

besitzt. Dieser für alle regulären Punkte λ definierte lineare Operator R_λ heißt die *Resolvente* von A . Die Menge der Punkte λ , die keine regulären Punkte von A sind, heißt das *Spektrum* von A .

Nach dieser Definition gehört jeder Eigenwert des Operators A zu dessen Spektrum. Wir werden sehen, daß in dem Spektrum auch solche λ liegen können, die keine Eigenwerte sind.

Ist λ ein regulärer Punkt, so hat die inhomogene Gleichung (61₁) für jedes Element y eine einzige, durch

$$x = (A - \lambda E)^{-1} y \quad (63)$$

definierte Lösung. Ist dagegen λ kein regulärer Punkt und stimmt λ nicht mit einem Eigenwert von A überein, so hat (61₁) ebenfalls eine einzige Lösung, wenn y in der linearen Mannigfaltigkeit $R(A - \lambda E)$ liegt. Diese lineare Mannigfaltigkeit besteht aus den Elementen y , die durch

$$y = (A - \lambda E) x \quad (x \in H) \quad (64)$$

definiert sind, wenn x ganz H durchläuft.

Ist also λ kein Eigenwert, so ist auf der linearen Mannigfaltigkeit $R(A - \lambda E)$ der inverse Operator $(A - \lambda E)^{-1}$ definiert. Ist dabei λ kein regulärer Punkt von A , so nennt man den Operator $(A - \lambda E)^{-1}$ auch in diesem Fall die *Resolvente* von A .

Satz 2. Die Elemente der linearen Mannigfaltigkeit $R(A - \lambda E)$ sind zu allen Lösungen der Gleichung $(A^* - \bar{\lambda}E)z = 0$ orthogonal.

Diese Behauptung folgt unmittelbar aus der evidenten Beziehung

$$((A - \lambda E)x, z) = (x, (A^* - \bar{\lambda}E)z).$$

Bemerkung. Ist A ein selbstadjungierter Operator und λ ein (reeller) Eigenwert von A , so folgt aus Satz 2, daß die Elemente von $R(A - \lambda E)$ zu den Eigenelementen von A , die dem Eigenwert λ entsprechen, orthogonal sind. In [129] werden wir einige Sätze über das Spektrum eines selbstadjungierten Operators beweisen.

Wir möchten nun noch auf die Eigenelemente eines selbstadjungierten Operators A hinweisen. Wie wir sahen, bilden die einem Eigenwert $\lambda = \lambda'$ entsprechenden Eigenelemente einen Unterraum. In ihm können wir ein vollständiges normiertes Orthogonalsystem einführen. Hat der Eigenwert $\lambda = \lambda'$ den endlichen Rang r , so besteht dieses System aus r Elementen. Da die den verschiedenen Eigenwerten entsprechenden Eigenelemente zueinander orthogonal sind, erhalten wir im Raum H ein Orthogonalsystem K , wenn wir wie oben in jedem der Unterräume, die zu einem festen Eigenwert gehören, ein normiertes Orthogonalsystem einführen. Wir sagen dann, daß K durch den selbstadjungierten Operator erzeugt wird. Bei der Konstruktion von K können wir die Orthogonalsysteme in den oben erwähnten Unterräumen völlig willkürlich wählen. Hat A überhaupt keine Eigenwerte, so existiert das System K nicht. Da ein normiertes Orthogonalsystem nur endlich oder abzählbar unendlich viele Elemente enthält, können wir unmittelbar schließen, daß die Menge der Eigenwerte von K abzählbar ist, wenn K unendlich viele verschiedene Eigenwerte besitzt.

Das normierte Orthogonalsystem K kann sowohl vollständig als auch nicht vollständig in H sein. Diese Eigenschaft hängt, wie man leicht sieht, nicht von der Wahl des vollständigen normierten Systems im Unterraum der einem festen Eigenwert entsprechenden Eigenelemente ab. Ist K vollständig, so sagt man im allgemeinen, der Operator A habe ein reines Punktspektrum.

129. Das Spektrum eines selbstadjungierten Operators. In diesem Abschnitt betrachten wir ausschließlich selbstadjungierte Operatoren.

Satz 1. Ist λ kein Eigenwert des selbstadjungierten Operators A , so wird durch (64) eine in H dichte lineare Mannigfaltigkeit $R(A - \lambda E)$ definiert.

Beweis. Wir führen den Beweis indirekt und nehmen an, daß die lineare Mannigfaltigkeit $R(A - \lambda E)$ in H nicht dicht ist, daß also das Abschließen von $R(A - \lambda E)$ zu einem von H verschiedenen Unterraum führt. Auf Grund des Satzes aus [122] existiert dann ein vom Nullvektor verschiedener Vektor x_0 , der zu diesem Unterraum und damit auch zu der linearen Mannigfaltigkeit $R(A - \lambda E)$ orthogonal ist; für jedes x aus H gilt also $((A - \lambda E)x, x_0) = 0$ oder, wenn man beachtet, daß A selbstadjungiert ist, $(x, (A - \bar{\lambda}E)x_0) = 0$. Wenn wir $x = (A - \bar{\lambda}E)x_0$ annehmen, erhalten wir $|(A - \bar{\lambda}E)x_0| = 0$, also $Ax_0 = \bar{\lambda}x_0$. Ist λ reell, so gilt $\bar{\lambda} = \lambda$, und λ erweist sich als Eigenwert des Operators A , was der Voraussetzung des Satzes widerspricht. Ist λ komplex, so besagt die Gleichung $Ax_0 = \bar{\lambda}x_0$, daß der selbstadjungierte Operator A einen nicht reellen Eigenwert $\bar{\lambda}$ besitzt, was ebenfalls nicht möglich ist. Damit ist der Satz bewiesen.

Die lineare Mannigfaltigkeit $R(A - \lambda E)$ stimmt mit H überein, wenn λ ein regulärer Punkt ist. Das folgt unmittelbar aus der Definition eines regulären Punktes. Ist λ Eigenwert, so sind alle Elemente von $R(A - \lambda E)$ zu den entsprechenden Eigenelementen von A orthogonal, und $R(A - \lambda E)$ kann in H nicht dicht sein. Später werden wir sehen, daß $R(A - \lambda E)$ nicht mit H übereinstimmt (sie ist in H dicht), wenn λ weder ein regulärer Punkt noch ein Eigenwert ist.

Wir stellen nun eine notwendige und hinreichende Bedingung für die Regularität von λ auf.

Satz 2. *Der Punkt λ eines selbstadjungierten Operators A ist genau dann regulär, wenn es zu jedem x aus H eine positive Zahl p gibt derart, daß*

$$|(A - \lambda E)x| \geq p|x| \quad (65)$$

gilt. Jeder nichtreelle Wert und die außerhalb des Intervalls $[m, M]$ liegenden reellen Werte λ sind regulär; dabei ist m die untere und M die obere Grenze von A .

Beweis. Wir zeigen zunächst die Notwendigkeit der Bedingung (65). Es sei λ ein regulärer Punkt. Dann existiert der beschränkte inverse Operator (62). Bezeichnen wir seine Norm mit q , so können wir

$$|(A - \lambda E)^{-1}y| \leq q|y|$$

schreiben. Setzen wir hier $y = (A - \lambda E)x$, so gelangen wir mit $p = 1/q$ zur Ungleichung (65). Wir beweisen nun, daß die Bedingung (65) hinreichend ist. Aus (65) folgt zunächst, daß λ kein Eigenwert ist; also ist die in Satz 1 definierte lineare Mannigfaltigkeit $R(A - \lambda E)$ in H dicht. Wir zeigen, daß sie abgeschlossen ist und demnach mit H übereinstimmt. Dazu nehmen wir an, daß die $y_n = (A - \lambda E)x_n$ zu $R(A - \lambda E)$ gehören und $y_n \rightarrow y$ gilt. Dann haben wir zu beweisen, daß auch y in $R(A - \lambda E)$ liegt. Wegen (65) gilt $|y_n - y_m| \geq p|x_n - x_m|$. Die Folge der Elemente y_n ist stark konvergent, und auf Grund der letzten Ungleichung hat auch die Folge der x_n diese Eigenschaft, d. h., es gibt ein Element x , für das $x_n \rightarrow x$ gilt. Aus $y_n = (A - \lambda E)x_n$ folgt $y_n \rightarrow y = (A - \lambda E)x$ und damit $y \in R(A - \lambda E)$. Die lineare Mannigfaltigkeit $R(A - \lambda E)$ stimmt also mit H überein, und der zu $A - \lambda E$ inverse Operator $(A - \lambda E)^{-1}$ ist in ganz H definiert. Für den Beweis der Regularität des Punktes λ haben wir nur noch zu zeigen, daß der Operator $(A - \lambda E)^{-1}$ beschränkt ist. Setzen wir in (65) $x = (A - \lambda E)^{-1}y$, so erhalten wir

$$|(A - \lambda E)^{-1}y| \leq \frac{1}{p}|y|,$$

woraus die Beschränktheit von $(A - \lambda E)^{-1}$ folgt. Damit ist der erste Teil des Satzes bewiesen.

Wir nehmen nun an, daß $\lambda = \sigma + \tau i$ und $\tau \neq 0$ ist. Mit $(A - \lambda E)x = y$ können wir

$$((A - \lambda E)x, x) = (y, x)$$

und

$$((A - \bar{\lambda}E)x, x) = (x, (A - \lambda E)x) = (x, y)$$

schreiben. Subtrahieren wir die zweite Gleichung von der ersten, so finden wir

$$(\lambda - \bar{\lambda})(x, x) = (y, x) - (x, y)$$

oder

$$2|\tau| \cdot |x|^2 \leq |(y, x)| + |(x, y)|,$$

und durch Anwendung von (6) gelangen wir zu der Ungleichung

$$2|\tau| \cdot \|x\| \leq 2\|y\|$$

oder

$$\|(A - \lambda E)x\| \leq |\tau| \cdot \|x\| \quad (\lambda = \sigma + \tau i).$$

Für $p = |\tau|$ kommen wir zu der Ungleichung (65); folglich ist jeder nichtreelle Wert von λ regulär. Wir nehmen nun an, daß λ reell ist, aber außerhalb des Intervalls $[m, M]$ liegt. Wir nehmen beispielsweise $\lambda > M$ an und beweisen, daß auch dann die Bedingung (65) erfüllt ist. Wir können

$$((A - \lambda E)x, x) = (Ax, x) - \lambda\|x\|^2$$

oder

$$((A - \lambda E)x, x) = [(Ax, x) - M\|x\|^2] - (\lambda - M)\|x\|^2$$

schreiben. Die Definition der oberen Grenze M von A besagt, daß die Differenz in den eckigen Klammern nicht positiv ist. Außerdem gilt nach Voraussetzung $\lambda > M$, so daß die letzte Gleichung zu

$$|((A - \lambda E)x, x)| \leq (\lambda - M)\|x\|^2$$

führt. Andererseits gilt

$$|((A - \lambda E)x, x)| \leq \|(A - \lambda E)x\| \cdot \|x\|.$$

Aus diesen beiden Ungleichungen ergibt sich

$$\|(A - \lambda E)x\| \leq (\lambda - M)\|x\|.$$

Die Bedingung (65) ist also für $\lambda > M$ erfüllt. Damit ist Satz 2 vollständig bewiesen.

Aus dem eben bewiesenen Satz ergeben sich unmittelbar zwei Folgerungen.

Folgerung 1. *Der Punkt λ gehört genau dann zum Spektrum, wenn eine Folge normierter Elemente x_n existiert, für die*

$$\|(A - \lambda E)x_n\| \rightarrow 0$$

gilt.

Existiert eine Folge mit dieser Eigenschaft, so kann (65) für $p > 0$ nicht erfüllt sein, d. h., λ gehört zum Spektrum. Ist umgekehrt λ ein Punkt des Spektrums, so ist die Bedingung (65) für kein $p > 0$ erfüllt, d. h., es existiert eine Folge von normierten Elementen x_n , für die $\|(A - \lambda E)x_n\|$ gegen 0 strebt. Wir bemerken noch, daß wir für alle x_n ein und dasselbe Element, und zwar das normierte Eigenelement x_0 nehmen können, wenn λ ein Eigenwert ist. Dann gilt $(A - \lambda E)x_n = 0$ für jedes n .

Folgerung 2. *Ist die untere Grenze $m(A)$ des selbstadjungierten Operators A positiv, so liegt $\lambda = 0$ außerhalb des Intervalls $[m, M]$, und A hat einen beschränkten inversen Operator.*

Diese Tatsache benutzten wir schon in [127].

Folgerung 3. *Die Gesamtheit aller regulären Punkte λ der reellen Achse ist eine offene Menge.*

Es sei λ ein regulärer Punkt. Wir müssen zeigen, daß auch alle Punkte $\lambda \pm \varepsilon$

für alle hinreichend kleinen positiven ε regulär sind. Nach Voraussetzung existiert ein positives p , für das (65) gilt, woraus

$$|(A - (\lambda \pm \varepsilon) E) x| \geq |(A - \lambda E) x| - \varepsilon \|x\| \geq (p - \varepsilon) \|x\|$$

folgt; aus dieser Formel ergibt sich, daß für $\varepsilon < p$ alle Punkte $\lambda \pm \varepsilon$ regulär sind.

Folgerung 4. *Die Punkte des Spektrums eines selbstadjungierten Operators bilden eine abgeschlossene Menge.*

Dies ergibt sich unmittelbar aus Folgerung 3 (vgl. [32]).

Satz 3. *Die Werte $\lambda = m$ und $\lambda = M$ gehören zum Spektrum.*

Beweis. Wir nehmen $M > m$ an und beweisen den Satz für $\lambda = M$. Dazu führen wir den selbstadjungierten Operator $B = A - mE$ ein, der die Werte 0 und $M_1 = M - m > 0$ als Grenzen hat. Seine Norm ist gleich M_1 [126]. Die Definition der oberen Grenze besagt, daß es eine Folge von normierten Elementen x_n gibt, für die $(Bx_n, x_n) = M_1 - \delta_n$ mit $\delta_n \geq 0$ und $\delta_n \rightarrow 0$ gilt. Dann ist

$$\begin{aligned} \|Bx_n - M_1 x_n\|^2 &= \|Bx_n\|^2 - 2M_1(Bx_n, x_n) + M_1^2 \\ &\leq M_1^2 - 2M_1(M_1 - \delta_n) + M_1^2 = 2M_1\delta_n. \end{aligned}$$

Hieraus ergibt sich $\|(B - M_1) x_n\| \rightarrow 0$, und wegen Folgerung 1 besitzt der Operator $B - M_1 E = A - ME$ kein beschränktes Inverses.

Nun beweisen wir, daß auch $\lambda = m$ zum Spektrum gehört. Für den selbstadjungierten Operator $-A$ sind $-M$ und $-m$ mit $-M < -m$ die Grenzen, so daß nach dem eben Bewiesenen der Operator $-A + mE$ kein beschränktes Inverses hat, d. h. nicht der Bedingung (65) genügt. Dann wird (65) auch von $A - mE$ nicht erfüllt.

Bemerkung. Im Fall $m = M$ multipliziert der Operator A jedes Element x mit der Zahl m , und es gilt $Ax = mx$ für $x \in H$. Die Zahl m ist der Eigenwert von A , und der Unterraum der zu diesem Eigenwert gehörigen Eigenelemente stimmt mit H überein.

Gehört λ zum Spektrum, ohne Eigenwert zu sein, so ist die lineare Mannigfaltigkeit $R(A - \lambda E)$ überall in H dicht. Wir werden beweisen, daß $R(A - \lambda E)$ in diesem Fall nicht gleich H ist. Dies ergibt sich sofort aus dem folgenden Satz (vgl. [97]).

Satz 4. *Ist $R(A)$ gleich dem ganzen Raum H , so ist der inverse Operator A^{-1} beschränkt.*

Beweis. Zunächst folgt aus dem oben Gesagten, daß $\lambda = 0$ kein Eigenwert ist, wenn $R(A)$ gleich dem Raum H ist; also existiert der in ganz H definierte inverse Operator A^{-1} .

Es seien nun x, y Elemente aus H , und wir setzen $x' = Ax, y' = Ay$. Der Operator A^{-1} besitzt, wie wir leicht sehen, die Symmetrieeigenschaft

$$(A^{-1}x', y') = (x', A^{-1}y'), \quad (66)$$

wobei x' und y' aus H sind. In der Tat ist (66) der Gleichung $(x, Ay) = A x, y)$ äquivalent, die wegen der Selbstadjungiertheit von A gilt. Der Ausdruck (66) ist also für ein beliebig gewähltes normiertes Element y' ein von x' abhängendes lineares Funktional $l_{y'}(x')$. Für diese Schar von Funktionalen sind die Zahlen $|l_{y'}(x')|$ endlich (x' ist beliebig, aber fest); denn es ist

$$|l_{y'}(x')| = |(A^{-1}x', y')| \leq \|A^{-1}x'\|.$$

Hieraus folgt, daß die Normen der Funktionale (66) für $\|y'\| = 1$ beschränkt sind [100]. Diese Normen sind aber gleich $\|A^{-1}y'\|$ [123], so daß also eine Zahl q existiert derart, daß $\|A^{-1}y'\| \leq q$ für $\|y'\| = 1$ ist. Damit ist der Satz bewiesen.

Der Satz 4 gilt offenbar auch für den Operator $A - \lambda E$ (λ reell). Den Fall, daß λ nicht reell ist, haben wir schon früher untersucht.

Im folgenden Abschnitt, der den nichtbeschränkten Operatoren gewidmet ist, werden wir das Spektrum selbstadjungierter nichtbeschränkter Operatoren genauer untersuchen.

130. Die Resolvente. Ist λ kein Eigenwert des Operators A , so hat A die Resolvente

$$R_\lambda = (A - \lambda E)^{-1}.$$

Sie ist auf der linearen Mannigfaltigkeit $R(A - \lambda E)$ definiert und bildet sie eindeutig auf H ab. Aus der Definition des inversen Operators folgt $x=0$, wenn x zu $R(A - \lambda E)$ gehört und $R_\lambda x = 0$ ist.

Jetzt werden wir den Fall betrachten, daß λ ein regulärer Punkt von A ist. Dabei stimmt $R(A - \lambda E)$ mit H überein, und R_λ ist ein in ganz H definierter beschränkter Operator.

Wir beweisen nun (für reguläre λ und μ) die beiden Formeln

$$R_\lambda^* = R_{\bar{\lambda}}, \quad R_\mu - R_\lambda = (\mu - \lambda) R_\mu R_\lambda. \quad (67)$$

Ist λ komplex, so ist auch $\bar{\lambda}$ komplex und somit auch regulär. Für reelle und komplexe λ kann man also behaupten, daß sich beliebige Elemente $x', y' \in H$ in der Form

$$x' = (A - \lambda E)x, \quad y' = (A - \bar{\lambda} E)y$$

darstellen lassen. Daraus folgt

$$(R_\lambda x', y') = (x, (A - \bar{\lambda} E)y) = ((A - \lambda E)x, y) = (x', R_{\bar{\lambda}} y').$$

Aus dieser Gleichung ergibt sich sofort die erste Gleichung (67). Zur zweiten Gleichung (67) gelangt man, wenn man mit Hilfe der Definition der Resolvente die Formeln

$$R_\lambda x = R_\mu (A - \mu E) R_\lambda x \quad \text{und} \quad R_\mu x = R_\mu (A - \lambda E) R_\lambda x$$

aufstellt und sie dann voneinander subtrahiert.

131. Folgen von Operatoren. Das früher [104] über Folgen linearer Operatoren in Banachräumen Gesagte gilt auch für den Raum H . Zu den Haupttatsachen machen wir einige Ergänzungen. Eine Folge von linearen Operatoren A_n konvergiert der Norm nach gegen A , wenn $\|A - A_n\|$ für $n \rightarrow \infty$ gegen 0 strebt. Das ist genau dann der Fall, wenn $\|A_n - A_m\|$ für $n, m \rightarrow \infty$ gegen 0 strebt.

Die starke Konvergenz wird durch die Beziehung $A_n x \rightarrow Ax$ für jedes $x \in H$ definiert. Dabei sind die Normen $\|A_n\|$ beschränkt. Notwendig und hinreichend für die starke Konvergenz ist die Bedingung $\|A_n x - A_m x\| \rightarrow 0$ für $n, m \rightarrow \infty$ und jedes $x \in H$. Aus der Konvergenz der Norm nach folgt die starke Konvergenz. Ist $A_n \rightarrow A$, $B_n \rightarrow B$ (im Sinne einer der beiden Konvergenzen) und streben die Zahlen a_n gegen a , so gilt $a_n A_n \rightarrow aA$, $A_n + B_n \rightarrow A + B$ und $B_n A_n \rightarrow BA$.

Sind die A_n selbstadjungierte Operatoren und ist $A_n \rightarrow A$, so ist A ebenfalls selbst-

adjungiert. Für jedes n und jedes $x \in H$ ist nämlich $(A_n x, x)$ reell; also muß auch $(Ax, x) = \lim_{n \rightarrow \infty} (A_n x, x)$ für jedes $x \in H$ reell und folglich A selbstadjungiert sein.

Mit Hilfe des Grenzwertbegriffs können wir die unendliche Reihe $B_1 + B_2 + B_3 + \dots$ untersuchen, deren Glieder lineare Operatoren aus H sind, und ihre Konvergenz (in diesem oder jenem Sinne) feststellen. Zunächst betrachten wir noch ein für das Folgende wichtiges Beispiel. Es sei A ein linearer Operator mit $\|A\| = q < 1$. Wir konstruieren dann die Reihe

$$S = E + \alpha A + \alpha^2 A^2 + \dots, \quad (68)$$

wobei α eine komplexe Zahl ist. Bezeichnen wir mit S_n den Operator, der gleich der n -ten Partialsumme von (68) ist, so erhalten wir

$$\|S_{n+p} - S_n\| = \|\alpha^{n+1} A^{n+1} + \alpha^{n+2} A^{n+2} + \dots + \alpha^{n+p-1} A^{n+p-1}\|,$$

woraus sich, wenn wir $|\alpha| \leq 1$ annehmen,

$$\|S_{n+p} - S_n\| \leq q^{n+1} + q^{n+2} + \dots = \frac{q^{n+1}}{1-q}$$

ergibt. Folglich strebt $\|S_{n+p} - S_n\|$ für $n \rightarrow \infty$ und jedes $p > 0$ gegen 0, d. h., für $|\alpha| \leq 1$ konvergiert (68) der Norm nach. Da die Abschätzung für $\|S_{n+p} - S_n\|$ nicht α enthält, konvergiert die Reihe (68) gleichmäßig bezüglich α für $|\alpha| \leq 1$. Mit Berücksichtigung von $\|A\| = q < 1$ können wir schließlich behaupten, daß (68) der Norm nach gleichmäßig bezüglich α konvergiert, wenn $|\alpha| \leq 1 + \varepsilon$ ist, wobei das positive ε so klein gewählt werden muß, daß $(1 + \varepsilon)q < 1$ ist.

Multiplizieren wir die Reihe (68) mit $E - \alpha A$, so erhalten wir unter Berücksichtigung des Grenzübergangs bei Folgen von Operatoren die Beziehung $(E - \alpha A)S = S(E - \alpha A) = E$, d. h., die Reihe (68) ist für $|\alpha| \leq 1 + \varepsilon$ der zu $E - \alpha A$ gehörige beschränkte inverse Operator, also $S = (E - \alpha A)^{-1}$.

Genauso wie früher läßt sich die folgende Tatsache beweisen: Sind die Normen der Operatoren A_k nicht größer als gewisse positive Zahlen δ_k , die eine konvergente Reihe bilden, so konvergiert die Reihe

$$A = A_1 + A_2 + \dots$$

der Norm nach, und die Norm von A ist nicht größer als die Summe der aus den δ_k gebildeten Reihe. Wäre nämlich die Norm von A größer als diese Summe, so müßte für hinreichend großes n auch die Norm des Operators

$$S_n = A_1 + A_2 + \dots + A_n$$

größer als diese Summe sein; dies widerspricht aber der sicher gültigen Ungleichung

$$\|S_n\| \leq \|A_1\| + \|A_2\| + \dots + \|A_n\| \leq \delta_1 + \delta_2 + \dots + \delta_n.$$

Die obige Behauptung gilt offenbar auch für normierte Räume.

132. Schwache Konvergenz. Auf Grund der allgemeinen Form eines linearen Funktionalen in H [123] ist die schwache Konvergenz $x_n \rightarrow x_0$ gleichbedeutend mit der Beziehung

$$(x_n, y) \rightarrow (x_0, y) \quad \text{für jedes } y \in H.$$

Wir erinnern daran, daß aus $x_n \rightarrow x_0$ die Existenz einer solchen Zahl $m > 0$ folgt, daß $\|x_n\| \leq m$ für alle n gilt. Ferner ist, da der konjugierte Raum von H mit H über-

einstimmt, jede in H beschränkte Menge schwach kompakt, und H selbst ist schwach vollständig, d. h., aus $(x_n - x_m, y) \rightarrow 0$ für $n, m \rightarrow \infty$ ergibt sich, daß die Folge der x_n schwach konvergiert. Wir wissen auch, daß Ax_n schwach gegen Ax_0 strebt, wenn $x_n \rightarrow x_0$ gilt und A ein linearer Operator ist. Wir zeigen jetzt, daß es zum Beweis von $x_n \rightarrow x_0$ genügt, die Beziehung $(x_n, z_k) \rightarrow (x_0, z_k)$, $k=1, 2, \dots$, nachzuweisen, wobei z_k ($k=1, 2, \dots$) ein in H vollständiges normiertes Orthogonalsystem und $\|x_n\| \leq m$ ist. Zum Beweis nehmen wir an, es gelte $(x_n, z_k) \rightarrow (x_0, z_k)$, $k=1, 2, \dots$, und es sei ein $\varepsilon > 0$ gegeben. Da sich jedes Element y aus H in der Form

$$y = \sum_{k=1}^{\infty} b_k z_k \quad \text{mit} \quad \sum_{k=1}^{\infty} |b_k|^2 = \|y\|^2$$

darstellen läßt, können wir

$$(x_n - x_0, y) = \left(x_n - x_0, \sum_{k=1}^N b_k z_k \right) + \left(x_n - x_0, \sum_{k=N+1}^{\infty} b_k z_k \right)$$

schreiben. Wegen $\|x_n\| \leq m$ ist

$$\left| \left(x_n - x_0, \sum_{k=N+1}^{\infty} b_k z_k \right) \right| \leq \|x_n - x_0\| \cdot \left\| \sum_{k=N+1}^{\infty} b_k z_k \right\| \leq (m + \|x_0\|) \sqrt{\sum_{k=N+1}^{\infty} |b_k|^2}$$

[121], und wir können ein N wählen, für das die rechte Seite dieser Ungleichung höchstens gleich $\varepsilon/2$ ist. Also gilt

$$|(x_n - x_0, y)| \leq \left| \sum_{k=1}^N (x_n - x_0, b_k z_k) \right| + \frac{\varepsilon}{2}.$$

Wegen $(x_n, z_k) \rightarrow (x_0, z_k)$ für alle hinreichend großen n ist der erste Summand der rechten Seite höchstens gleich $\varepsilon/2$ und demzufolge $|(x_n - x_0, y)|$ höchstens gleich ε , woraus sich $(x_n, y) \rightarrow (x_0, y)$ ergibt.

Wir beweisen nun den folgenden Satz.

Satz 1. Ist $x_n \rightarrow x_0$ und $y_n \rightarrow y_0$, so gilt $(x_n, y_n) \rightarrow (x_0, y_0)$ und $(y_n, x_n) \rightarrow (y_0, x_0)$.

Beweis. Es genügt zu zeigen, daß (x_n, y_n) gegen (x_0, y_0) strebt. Die zweite Behauptung ergibt sich dann durch Vertauschen der Elemente. Wir können

$$|(x_0, y_0) - (x_n, y_n)| \leq |(x_0, y_0) - (x_n, y_0)| + |(x_n, y_0) - (x_n, y_n)|$$

oder

$$|(x_0, y_0) - (x_n, y_n)| \leq |(x_0, y_0) - (x_n, y_0)| + |(x_n, y_0 - y_n)|$$

schreiben. Da wegen $x_n \rightarrow x_0$ eine solche positive Zahl m existiert, daß $\|x_n\| \leq m$ ist, erhalten wir mit Hilfe der Schwarzschen Ungleichung

$$|(x_0, y_0) - (x_n, y_n)| \leq m \|y_0 - y_n\| + |(x_0, y_0) - (x_n, y_0)|.$$

Auf der rechten Seite strebt der erste Summand wegen $y_n \rightarrow y_0$ und der zweite Summand wegen $x_n \rightarrow x_0$ gegen 0, d. h., es ist $(x_n, y_n) \rightarrow (x_0, y_0)$. Damit ist der Satz bewiesen.

Satz 2. Aus $x_n \rightarrow x_0$ und $\|x_n\| \rightarrow \|x_0\|$ folgt $x_n \rightarrow x_0$.

Beweis. Es gilt

$$\|x_0 - x_n\|^2 = \|x_0\|^2 + \|x_n\|^2 - (x_n, x_0) - (x_0, x_n).$$

Hieraus folgt $\|x_0 - x_n\|^2 \rightarrow 0$, denn nach Voraussetzung ist

$$(x_n, x_0) \rightarrow \|x_0\|^2, \quad (x_0, x_n) \rightarrow \|x_0\|^2 \quad \text{und} \quad \|x_n\|^2 \rightarrow \|x_0\|^2.$$

Dieser Satz gilt, wie in [127] gezeigt wurde, auch für Banachsche Räume.

133. Vollstetige Operatoren. Wir kennen schon die Definition eines vollstetigen Operators für einen Banachschen Raum und deshalb auch für den Raum H . Ein *vollstetiger Operator* in H ist ein linearer Operator, der jede in H beschränkte Menge in eine kompakte Menge abbildet.

Wir wissen, daß ein linearer Operator eine kompakte Menge wieder in eine kompakte Menge transformiert. Der Einsoperator ist nicht vollstetig; er bildet die Kugel $\|x\| \leq 1$ (eine beschränkte Menge) eineindeutig auf sich ab, aber diese Kugel ist nicht kompakt. Um sich davon zu überzeugen, genügt es, eine unendliche ortho-normierte Folge von Elementen z_k ($k=1, 2, \dots$) zu wählen; sie ist beschränkt wegen $\|z_k\| = 1$, aber nicht kompakt wegen $\|z_p - z_q\| = \sqrt{2}$ für $p \neq q$.

Wir geben im folgenden noch zwei Definitionen eines vollstetigen Operators und werden zeigen, daß sie der oben angegebenen Hauptdefinition gleichwertig sind. Dazu machen wir die folgende einfache Bemerkung.

Konvergiert die eine der beiden Folgen $\{x_n\}$ und $\{y_n\}$ schwach und die andere stark gegen x_0 bzw. y_0 und ist A ein beschränkter linearer Operator, so gilt

$$\lim_{n \rightarrow \infty} (Ax_n, y_n) = (Ax_0, y_0). \quad (69)$$

Dies folgt aus [132] und der Tatsache, daß $Ax_n \rightarrow Ax_0$ für $x_n \rightarrow x_0$ sowie $Ax_n \rightharpoonup Ax_0$ für $x_n \rightharpoonup x_0$ ist.

Definition 1. Der lineare Operator A heißt *vollstetig*, wenn (69) für alle Folgen von Elementen x_n und y_n gilt, die schwach gegen x_0 bzw. y_0 streben.

Definition 2. Der lineare Operator A heißt *vollstetig*, wenn $Ax_n \rightharpoonup Ax_0$ aus $x_n \rightarrow x_0$ folgt.

Wir zeigen nun, daß diese Definitionen äquivalent sind. Der Operator A genüge der Bedingung (69) für die schwach konvergenten Folgen $\{x_n\}$ und $\{y_n\}$, und wir können

$$\|Ax_n - Ax_0\|^2 = (Ax_n, Ax_n - Ax_0) - (Ax_0, Ax_n - Ax_0)$$

schreiben. Wegen $x_n \rightarrow x_0$ strebt $Ax_n - Ax_0$ schwach gegen 0, und beide Glieder der rechten Seite streben wegen (69) gegen 0; es ist also $\|Ax_n - Ax_0\| \rightarrow 0$ oder $Ax_n \rightarrow Ax_0$. Aus der ersten Definition folgt somit die zweite. Wir nehmen nun an, aus $x_n \rightarrow x_0$ folge $Ax_n \rightharpoonup Ax_0$. Dabei ergibt sich (69) unmittelbar aus Satz 1 [132]. Haben wir die Äquivalenz beider Definitionen bewiesen, so brauchen wir für den Beweis der Äquivalenz dieser Definitionen mit der Hauptdefinition nur zu zeigen, daß beispielsweise die Hauptdefinition und die Definition 2 gleichwertig sind. Dazu nehmen wir an, A genüge der Definition 2, und $\{x_n\}$ sei eine beschränkte Folge von Elementen. Aus dieser Folge läßt sich eine Teilfolge $\{x_{n_k}\}$ mit $x_{n_k} \rightarrow x$ auswählen, und damit ist $Ax_{n_k} \rightarrow Ax$ wegen Definition 2, d. h., die Menge Ax_n ist kompakt, und der Operator A genügt der Hauptdefinition. Wir nehmen nun umgekehrt an, daß A der Hauptdefinition genügt und $x_n \rightarrow x_0$ ist. Wir müssen also beweisen, daß $Ax_n \rightharpoonup Ax_0$ gilt. Den Beweis führen wir indirekt und setzen voraus, Ax_n konvergiere nicht stark gegen Ax_0 , d. h., es existiere eine Teilfolge, für die $\|Ax_{n_k} - Ax_0\| \geq a > 0$

gilt. Aus der Hauptdefinition folgt die Kompaktheit der Menge Ax_{n_k} , und wir können annehmen, daß Ax_{n_k} stark gegen ein Element x' konvergiert, welches wegen $\|Ax_{n_k} - Ax_0\| \geq a > 0$ von Ax_0 verschieden sein muß. Nun ist $x_{n_k} \rightarrow x_0$ und folglich $Ax_{n_k} \rightarrow Ax_0$; nach dem Vorhergehenden ist jedoch $Ax_{n_k} \rightarrow x' \neq Ax_0$ und erst recht $Ax_{n_k} \rightarrow x' \neq Ax_0$. Damit sind wir auf einen Widerspruch gestoßen.

Wir haben somit bewiesen, daß die neuen Definitionen eines vollstetigen Operators der Hauptdefinition äquivalent sind.

Im folgenden werden wir den Begriff der schwachen Konvergenz in den Räumen und L_2 erklären.

Satz. *Ist A ein linearer Operator und A^*A vollstetig, so ist A ebenfalls vollstetig.*

Beweis. Es sei x_n ($n = 1, 2, \dots$) eine beschränkte Folge von Elementen ($\|x_n\| \leq a$). Nach Voraussetzung ist A^*Ax_n kompakt, d. h., A^*Ax_n besitzt eine konvergente Teilfolge $A^*Ax_{n_l}$. Wir zeigen, daß auch Ax_{n_k} eine konvergente Teilfolge ist, womit dann der Satz bewiesen sein wird. Es gilt

$$\begin{aligned} \|Ax_{n_k} - Ax_{n_l}\|^2 &= (A^*A(x_{n_k} - x_{n_l}), x_{n_k} - x_{n_l}) \\ &\leq \|A^*Ax_{n_k} - A^*Ax_{n_l}\| \cdot \|x_{n_k} - x_{n_l}\| \leq 2a \|A^*Ax_{n_k} - A^*Ax_{n_l}\|. \end{aligned}$$

Da $A^*Ax_{n_k}$ konvergiert, strebt die rechte Seite und folglich auch $\|Ax_{n_k} - Ax_{n_l}\|$ für $n_k, n_l \rightarrow \infty$ gegen 0, d. h., Ax_{n_k} ist eine konvergente Teilfolge.

Folgerung. *Ist A vollstetig, so ist auch A^* vollstetig.*

Ist A vollstetig, so hat der Operator $AA^* = (A^*)^*A^*$ dieselbe Eigenschaft. Hieraus folgt mit Hilfe des eben bewiesenen Satzes, auf A^* statt A angewendet, daß A^* vollstetig ist.

Wir erinnern hier an die folgende Eigenschaft von Operatorenfolgen: *Konvergiert die aus vollstetigen Operatoren gebildete Folge A_n der Norm nach gegen einen linearen Operator A , so ist A ebenfalls vollstetig* [106].

Wir erwähnen noch eine spezielle Klasse vollstetiger Operatoren.

Definition. Ein linearer Operator D heißt *endlichdimensional*, wenn er sich in der Form

$$Dx = \sum_{k=1}^m (x, v_k) u_k \quad (70)$$

darstellen läßt, wobei u_k und v_k ($k = 1, 2, \dots, m$) feste Elemente aus H sind.

Ein endlichdimensionaler Operator ist vollstetig, denn aus $x_n \rightarrow x_0$ folgt

$$(x_n, v_k) \rightarrow (x_0, v_k),$$

und wegen (70) ist $Dx_n \rightarrow Dx_0$.

Aus dem oben Gesagten ergibt sich: Ist A_n eine Folge von endlichdimensionalen Operatoren, die der Norm nach gegen einen linearen Operator A strebt, so ist A vollstetig.

Im nächsten Abschnitt werden wir zeigen, daß jeder vollstetige Operator als Grenzwert (der Norm nach) einer Folge endlichdimensionaler Operatoren dargestellt werden kann.

134. Die Räume H und l_2 . Es sei

$$z_1, z_2, z_3, \dots \quad (71)$$

ein vollständiges normiertes Orthogonalsystem in H . Mit Hilfe dieses Systems können wir H eineindeutig in den Raum l_2 abbilden, dessen Elemente unendliche Folgen $(\xi_1, \xi_2, \dots)$ komplexer Zahlen sind, die der Bedingung

$$\sum_{k=1}^{\infty} |\xi_k|^2 < \infty \quad (72)$$

unterworfen sind [121]. Jedes Element $x \in H$ ist durch seine Fourierkoeffizienten $\xi_k = (x, z_k)$ charakterisiert, und es gilt

$$x = \sum_{k=1}^{\infty} \xi_k z_k. \quad (73)$$

Ist umgekehrt ein Element $(\xi_1, \xi_2, \dots)$ aus l_2 gegeben, so konvergiert die Reihe (73) in H und ergibt das ihr entsprechende Element aus H . Diese Zuordnung ist eineindeutig, wobei das Skalarprodukt in H gleich dem der entsprechenden Elemente von l_2 ist [60, 121]:

$$(x, y) = \sum_{k=1}^{\infty} \xi_k \bar{\eta}_k;$$

x und y entsprechen dabei $(\xi_1, \xi_2, \dots)$ bzw. $(\eta_1, \eta_2, \dots)$. Somit ist $\|x\|$ gleich der Norm der zugeordneten Elemente aus l_2 , und die Konvergenz in H und die in l_2 sind äquivalent. Wir haben also einen Isomorphismus von H in l_2 . Den Elementen z_k des normierten Orthogonalsystems (71) sind die Elemente

$$(1, 0, 0, 0, \dots), \quad (0, 1, 0, 0, \dots), \quad (0, 0, 1, 0, \dots), \dots$$

aus l_2 zugeordnet.

Es sei $\xi(\xi_1, \xi_2, \dots)$ ein Element aus l_2 und $\xi^{(n)}(\xi_1, \xi_2, \dots, \xi_n, 0, \dots)$ das Abschnittelelement, bei dem die ersten n Komponenten gleich den entsprechenden Komponenten von ξ und die übrigen gleich 0 sind. Wegen

$$\|\xi - \xi^{(n)}\|^2 = \sum_{k=n+1}^{\infty} |\xi_k|^2$$

und wegen (72) ist $\xi^{(n)} \Rightarrow \xi$ in l_2 .

Es sei A ein linearer Operator aus H . Berücksichtigen wir seine Stetigkeit und die Formel (73), so können wir

$$y = Ax = \sum_{k=1}^{\infty} \xi_k A z_k \quad (74)$$

schreiben. Die Komponenten $\eta_1, \eta_2, \dots$ desjenigen Elements aus l_2 , das dem Element y entspricht, werden durch die Formel

$$\eta_l = (y, z_l) = \sum_{k=1}^{\infty} \xi_k (A z_k, z_l) \quad (75)$$

definiert. Wir benutzen dabei die Stetigkeit des Skalarprodukts. Führen wir nun die Zahlen

$$a_{lk} = (A z_k, z_l) \quad (76)$$

ein, so sehen wir, daß dem linearen Operator A aus H in l_2 der Operator

$$\eta_l = \sum_{k=1}^{\infty} a_{lk} \xi_k \quad (l=1, 2, \dots) \quad (77)$$

zugeordnet ist, dem die unendliche Matrix mit den Elementen a_{lk} entspricht. Dem adjungierten Operator A^* entspricht dann die Matrix mit den Elementen

$$a_{lk}^* = (A^* z_k, z_l) = (z_k, A z_l) = \overline{(A z_l, z_k)} ;$$

also ist

$$a_{lk}^* = \overline{a_{kl}} . \quad (78)$$

Der selbstadjungierte Operator wird dann durch

$$a_{kl} = \overline{a_{kl}} \quad (79)$$

charakterisiert.

Wir bilden nun die Menge L der Elemente aus H , die in der Form $x = \sum_{k=1}^m \xi_k z_k$ darstellbar sind, wobei die ξ_k beliebige komplexe Zahlen sind und m eine feste natürliche Zahl ist. Es gilt $\xi_k = (x, z_k)$, und es läßt sich leicht zeigen, daß L ein Unterraum ist. Der zu L orthogonale Unterraum M ist offenbar die Menge der in der Form $x = \sum_{k=m+1}^{\infty} \xi_k z_k$ darstellbaren Elemente; dabei sind die ξ_k diejenigen komplexen Zahlen, für die die Reihe

$$\sum_{k=m+1}^{\infty} |\xi_k|^2 \quad (80)$$

konvergiert. Wir können also

$$H = L \oplus M \quad (81)$$

schreiben [122]. Sind P_L und P_M Projektionsoperatoren in L bzw. M , so ist

$$E = P_L + P_M . \quad (82)$$

Es sei A wieder ein linearer Operator, und wir führen die beiden Operatoren

$$A_1 = P_L A, \quad A_2 = P_M A \quad (83)$$

ein. Wegen (82) ist $A = A_1 + A_2$. Da $P_L A x$ für jedes $x \in H$ zu L gehört, gilt

$$P_L A x = \sum_{k=1}^m a_k z_k$$

mit

$$a_k = (P_L A x, z_k) = (A x, P_L z_k) = (A x, z_k) = (x, A^* z_k) ,$$

d. h.

$$P_L A x = A_1 x = \sum_{k=1}^m (x, A^* z_k) z_k . \quad (84)$$

Daraus ergibt sich, daß $P_L A = A_1$ ein endlichdimensionaler Operator ist.

Analog dazu gilt

$$A_2 x = \sum_{k=m+1}^{\infty} (A x, z_k) z_k = \sum_{k=m+1}^{\infty} (x, A^* z_k) z_k . \quad (85)$$

Dem Element A_2x entspricht also ein Element $(\xi_1, \xi_2, \dots)$, dessen Komponenten durch

$$\xi_k = \begin{cases} 0 & \text{für } k \leq m, \\ (Ax, z_k) = (x, A^*z_k) & \text{für } k > m \end{cases} \quad (86)$$

definiert sind.

Wir setzen nun voraus, daß A vollstetig und U eine Menge normierter Elemente ist ($\|x\| = 1$). Dabei ist Ax für $x \in U$ kompakt; dieselbe Eigenschaft hat dann auch die entsprechende Menge in l_2 . Die Komponenten der Elemente dieser Menge werden durch $\xi_k = (Ax, z_k)$ bestimmt, und auf Grund der Kompaktheit können wir behaupten: Zu jedem normierten x gibt es eine positive Zahl C , für die

$$\sum_{k=1}^{\infty} |(Ax, z_k)|^2 \leq C$$

ist, und zu jedem $\varepsilon > 0$ existiert eine solche natürliche Zahl m_ε , daß

$$\sum_{k=m_\varepsilon+1}^{\infty} |(Ax, z_k)|^2 \leq \varepsilon^2$$

gilt [92].

Aus (85) folgt nun

$$\|A_2x\|^2 = \sum_{k=m+1}^{\infty} |(Ax, z_k)|^2.$$

Daher läßt sich zu jedem $\varepsilon > 0$ eine Zahl $m = m_\varepsilon$ finden, für die $\|A_2x\| \leq \varepsilon$ mit $\|x\| = 1$, also $\|A_2\| \leq \varepsilon$ gilt. Somit haben wir den folgenden Satz bewiesen.

Satz. *Ist A ein vollstetiger Operator und ε eine beliebig gegebene positive Zahl, so existiert eine natürliche Zahl m derart, daß für den oben definierten Operator A_2 die Ungleichung $\|A_2\| \leq \varepsilon$ gilt.*

Wählen wir eine gegen 0 strebende Folge ε_n von positiven Zahlen, so erhalten wir die Folge der endlichdimensionalen Operatoren $A_1^{(\varepsilon_n)}$, für die $\|A - A_1^{(\varepsilon_n)}\| \rightarrow 0$ gilt, d. h., *jeder vollstetige Operator ist Limes der Norm nach von endlichdimensionalen Operatoren.*

Berücksichtigen wir die in [133] erhaltenen Ergebnisse, so können wir behaupten, daß die folgende Definition eines vollstetigen Operators der ursprünglichen Definition (linearer Operator, der jede beschränkte Menge in eine kompakte Menge abbildet) äquivalent ist.

Definition. Ein linearer Operator heißt *vollstetig*, wenn er der Limes der Norm nach einer Folge endlichdimensionaler Operatoren ist.

135. Lineare Gleichungen mit vollstetigen Operatoren. Wir untersuchen nun im Raum H die Gleichungen der Gestalt

$$x - Ax = y, \quad (87)$$

$$x - A^*x = y. \quad (88)$$

Dabei sei A ein vollstetiger Operator, A^* der zu A adjungierte Operator, $y \in H$ ein gegebenes und $x \in H$ das gesuchte Element. Wie F. RIESZ zeigte, gelten die Fredholmschen Sätze aus der Theorie der Integralgleichungen für die Gleichungen (87) und (88) nicht nur im Raum H , sondern, wie wir in [133] erwähnten, auch in Banachschen Räumen. Wir betrachten die obigen Gleichungen im Raum H .

Die in [134] bei der Konstruktion der Operatoren A_1 und A_2 auftretende Zahl m wählen wir derart, daß $\|A_2\| < 1$ ist. Damit ist auch $\|A_2^*\| < 1$, also

$$\|A_2\| = \|A_2^*\| < 1. \quad (89)$$

Die Gleichungen (87) und (88) lassen sich in der Gestalt

$$(E - A_2) x - A_1 x = y, \quad (90)$$

$$(E - A_2^*) x - A_1^* x = y \quad (91)$$

schreiben. Wegen (89) haben die Operatoren $E - A_2$ und $E - A_2^*$ beschränkte Inverse.

Wir setzen nun

$$\tilde{x} = (E - A_2)^{-1} x, \quad \tilde{y} = (E - A_2^*)^{-1} y. \quad (92)$$

Die Gleichung (90) geht, wenn wir $\tilde{x}$ an Stelle von x einführen, in

$$\tilde{x} - B\tilde{x} = y \quad (93)$$

über, und aus (91) erhalten wir, wenn wir auf beide Seiten den Operator $(E - A_2^*)^{-1}$ anwenden,

$$x - B^*x = \tilde{y}; \quad (94)$$

dabei ist $\tilde{y}$ das gegebene Element, $B = A_1 (E - A_2)^{-1}$, $B^* = (E - A_2^*)^{-1} A_1^*$, und B^* ist der zu B adjungierte Operator. Die Gleichungen (94) und (91) sind äquivalent, und die Auflösung von (90) führt auf die Auflösung von (93) und auf die Formel $x = (E - A_2)^{-1} \tilde{x}$. Somit führt die Auflösung von (90) und (91) auf die von (93) und (94). Die entsprechenden homogenen Gleichungen lauten

$$\tilde{x} - B\tilde{x} = 0, \quad x - B^*x = 0. \quad (95)$$

Der Operator $B = P_L A (E - A_2)^{-1}$ ist endlichdimensional, und Bx wird durch (84) gegeben, wenn dort A durch $A(E - A_2)^{-1}$ ersetzt wird. Die dem Operator B aus l_2 entsprechende Matrix hat die Elemente

$$a_{kl} = (P_L A (E - A_2)^{-1} z_l, z_k) = (A (E - A_2)^{-1} z_l, P_L z_k). \quad (96)$$

Nun ist $P_L z_k = 0$ für $k > m$ und folglich $a_{kl} = 0$ für $k > m$. Sind ξ_l und η_l die Komponenten derjenigen Elemente aus l_2 , die den Elementen $\tilde{x}$ bzw. y aus H entsprechen, so nimmt die Gleichung (93) in l_2 die Gestalt

$$\xi_k - \sum_{l=1}^{\infty} a_{kl} \xi_l = \eta_k \quad (k = 1, 2, \dots, m), \quad (97)$$

$$\xi_k = \eta_k \quad (k = m+1, m+2, \dots) \quad (98)$$

an, wobei die ξ_k die gesuchten und die η_k die gegebenen Zahlen sind. Somit sind alle ξ_k für $k > m$ bekannt, und die Auflösung von (93) führt in l_2 auf die Auflösung des Systems

$$\xi_k - \sum_{l=1}^m a_{kl} \xi_l = \eta_k + \sum_{l=m+1}^{\infty} a_{kl} \eta_l \quad (99)$$

von m Gleichungen mit den m Unbekannten ξ_k ($k = 1, 2, \dots, m$). Dem Operator B^* entspricht die Matrix mit den Elementen $a_{kl}^* = \overline{a_{lk}}$ [134], so daß (94) in l_2 die Gestalt

$$\xi_k - \sum_{l=1}^m \overline{a_{lk}} \xi_l = \tilde{\eta}_k \quad (k = 1, 2, \dots) \quad (100)$$

erhält, wobei ξ_l und η_l die Komponenten der den Elementen $x, \tilde{y} \in H$ entsprechenden Elemente aus l_2 sind.

Jeder Lösung $(\xi_1^{(0)}, \xi_2^{(0)}, \dots, \xi_m^{(0)})$ der ersten m Gleichungen des Systems

$$\xi_k - \sum_{l=1}^m \overline{a_{lk}} \xi_l = \tilde{\eta}_k \quad (k = 1, 2, \dots, m) \quad (101)$$

ist eine bestimmte Lösung $(\xi_1^{(0)}, \xi_2^{(0)}, \dots, \xi_m^{(0)}, \xi_{m+1}^{(0)}, \dots)$ des ganzen Systems (100) bei willkürlichen $\tilde{\eta}_k$ ($k = m+1, m+2, \dots$) zugeordnet. In dieser Lösung sind die übrigen ξ_k ($k = m+1, m+2, \dots$) durch

$$\xi_k = \tilde{\eta}_k + \sum_{l=1}^m \overline{a_{lk}} \xi_l \quad (k = m+1, m+2, \dots) \quad (102)$$

definiert.

Die ξ_k und ξ_k , die sich aus (98) und (99) bzw. (101) und (102) errechnen lassen, sind so beschaffen, daß die Reihen mit dem allgemeinen Glied $|\xi_k|^2$ bzw. $|\xi_k|^2$ konvergieren. Für ξ_k folgt dies unmittelbar aus (98), für ξ_k aus (102), wenn man berücksichtigt, daß die Reihen mit dem allgemeinen Glied $|\tilde{\eta}_k|^2$ bzw. $|\overline{a_{lk}}|^2$ konvergieren. Das letzte Glied ergibt sich aus $\overline{a_{lk}} = ((E - A_2^*)^{-1} A^* P_{L^2}, z_k)$, und diese Gleichung folgt aus (96).

Im homogenen Fall muß $\eta_k = \tilde{\eta}_k = 0$ gesetzt werden. Dann geht (99) in

$$\xi_k - \sum_{l=1}^m a_{kl} \xi_l = 0 \quad (k = 1, 2, \dots, m) \quad (103)$$

(für $k > m$ ist $\xi_k = 0$) und (101) in

$$\xi_k - \sum_{l=1}^m \overline{a_{lk}} \xi_l = 0 \quad (k = 1, 2, \dots, m) \quad (104)$$

mit

$$\xi_k = \sum_{l=1}^m \overline{a_{lk}} \xi_l \quad \text{für } k > m \quad (105)$$

über.

Wir bemerken, daß die linear unabhängigen Lösungen des endlichen homogenen Systems (104) wegen (105) die linear unabhängigen Lösungen des unendlichen homogenen Systems erzeugen, das der homogenen Gleichung (94) in H entspricht. Linear abhängige Lösungen von (104) erzeugen auch linear abhängige Lösungen des unendlichen Systems. Berücksichtigen wir die Hauptresultate über die Auflösung von Gleichungssystemen und die Tatsache, daß die Matrizen aus den Koeffizienten von (103) und (104) von gleichem Rang sind, so erhalten wir den folgenden Satz.

Satz 1. *Die inhomogenen Gleichungen (87) und (88) sind für alle y genau dann auflösbar, wenn die entsprechenden homogenen Gleichungen nur die Nulllösung besitzen. In diesem Fall haben die Gleichungen (87) und (88) für jedes y eine einzige Lösung. Die homogenen Gleichungen $x - Ax = 0$ und $x - A^*x = 0$ haben die gleiche Anzahl von linear unabhängigen Lösungen.*

Wir untersuchen nun die inhomogenen Gleichungen (87) für den Fall, daß die homogene Gleichung von 0 verschiedene Lösungen hat, und beweisen den folgenden Satz.

Satz 2. *Im eben erwähnten Fall hat die inhomogene Gleichung (87) genau dann*

eine Lösung, wenn y zu allen Lösungen der homogenen Gleichung

$$x - A^*x = 0 \quad (106)$$

orthogonal ist.

Beweis. Daß die Bedingung notwendig ist, können wir zeigen, ohne nach l_2 überzugehen. Die Gleichung (87) habe die Lösung x_0 , also sei $x_0 - Ax_0 = y$, und z sei eine Lösung von (106), also $z - A^*z = 0$. Wir müssen nachweisen, daß $(y, z) = 0$ ist. Dies ist tatsächlich der Fall, denn es ist

$$(y, z) = (x_0 - Ax_0, z) = (x_0, z - A^*z) = (x_0, 0) = 0.$$

Die Bedingung ist auch hinreichend. Es sei also y orthogonal zu allen Lösungen von (106), und wir zeigen, daß (87) eine Lösung besitzt. Gehen wir nach l_2 über, so gilt nach Voraussetzung

$$\sum_{k=1}^{\infty} \eta_k \xi_k = 0, \quad (107)$$

wobei $(\xi_1, \dots, \xi_m)$ eine Lösung von (104) ist und die ξ_k für $k > m$ durch die Formeln (105) definiert werden. Setzen wir diese ξ_k ($k > m$) ein, so läßt sich (107) in der Form

$$\sum_{k=1}^m \left(\eta_k + \sum_{l=m+1}^{\infty} a_{kl} \eta_l \right) \xi_k = 0$$

schreiben. Beachten wir, daß die Summen in den runden Klammern gleich den rechten Seiten der Gleichungen (99) sind und $(\xi_1, \dots, \xi_m)$ eine Lösung des Systems (104) ist, so können wir behaupten, daß das System (99) Lösungen besitzt [III/1, 15] und die Gleichung (87) ebenfalls. Damit ist der Satz bewiesen.

Wir betrachten nun die Gleichung

$$x - \mu Ax = y, \quad (108)$$

wobei A ein vollstetiger Operator und μ ein komplexer Parameter ist. Da der Operator μA ebenfalls vollstetig ist, können wir auf die Gleichung (108) den oben bewiesenen Satz anwenden. Insbesondere ist (108) für jedes y (sogar eindeutig) auflösbar, wenn die homogene Gleichung

$$x - \mu Ax = 0 \quad \text{oder} \quad Ax = \lambda x \quad \left(\lambda = \frac{1}{\mu} \right) \quad (109)$$

nur die Nulllösung hat (für $\mu = 0$ ist es klar). Hat (109) von 0 verschiedene Lösungen, so ist der entsprechende Wert von λ ein Eigenwert des Operators A .

Wir beweisen jetzt den folgenden Satz.

Satz 3. *Es können nur endlich viele Eigenwerte existieren, die der Bedingung $|\lambda| \leq r$ genügen, wobei r eine beliebig vorgegebene positive Zahl ist.*

Beweis. Wir müssen beweisen, daß es nur endlich viele μ gibt, die der Bedingung $|\mu| \leq 1/r$ genügen und für die die Gleichung (109) von 0 verschiedene Lösungen hat. Der Beweis läßt sich mit derselben Konstruktion durchführen, die wir schon für den Beweis von Satz 1 benutzt haben.

Wir setzen wie dort $A = A_1 + A_2 = P_L A + P_M A$ und wählen ein so großes m , daß die Ungleichung $\|A_2\|/r = q < 1$ erfüllt ist. Der Operator $E - \mu A_2$ hat dabei ein be-

schränktes Inverses für $|\mu| \leq 1/r$, das sich in die Reihe

$$(E - \mu A_2)^{-1} = E + \mu A_2 + \mu^2 A_2^2 + \dots \quad (110)$$

entwickeln läßt, die der Norm nach gleichmäßig bezüglich μ für $|\mu| \leq \frac{1}{r} + \varepsilon$ konvergiert [131], wobei ε eine hinreichend kleine positive Zahl ist. Die Werte von μ , für welche (109) von 0 verschiedene Lösungen hat, erhalten wir, wenn wir die Determinante des Systems (103) gleich 0 setzen, d. h. die Determinante Δ mit den Elementen $\delta_{kl} - a_{kl}$, wobei

$$\delta_{kl} = \begin{cases} 0 & \text{für } k \neq l, \\ 1 & \text{für } k = l \end{cases} \quad \text{und} \quad a_{kl} = (P_L \mu A (E - \mu A_2)^{-1} z_l, z_k)$$

ist. Berücksichtigen wir die Konvergenz von (110), das oben Gesagte über den Limes für Folgen von Operatoren und die Stetigkeit des Skalarprodukts, so können wir behaupten, daß die a_{kl} reguläre Funktionen im Kreis $|\mu| \leq 1/r$ sind. Dieselbe Eigenschaft hat offenbar auch die Determinante Δ ; daher kann die Gleichung $\Delta = 0$ nur endlich viele Wurzeln haben, die der Bedingung $|\mu| \leq 1/r$ genügen, was zu beweisen war.

Der Satz 3 läßt sich auch folgendermaßen formulieren: *Die Eigenwerte λ eines vollstetigen Operators können als Häufungspunkt nur den Punkt $\lambda = 0$ besitzen.*

Der Rang jedes Eigenwertes λ , der der Bedingung $|\lambda| \leq 1/r$ genügt, ist nicht größer als die Zahl m aus (103) bei der Bedingung $\|A_2\|/r = q < 1$. Ist A kein selbstadjungierter Operator, so braucht er überhaupt keine Eigenwerte zu besitzen [IV/1, 13].

136. Vollstetige selbstadjungierte Operatoren. Wir untersuchten die Eigenschaften des Spektrums und die Entwicklung nach Eigenfunktionen eines vollstetigen selbstadjungierten Operators in [IV/1, 31]. Alle Beweise lassen sich ohne Änderungen auch für den Raum H übertragen. In H setzten wir aber Vollstetigkeit voraus, die wir bei den Beweisen in Teil IV/1 noch nicht benutzten. Daher erhalten wir für den Raum H neue Ergebnisse. Zuerst formulieren wir einen Satz, der sich aus den Resultaten von Teil IV/1 ergibt; wir erinnern daran, daß alle Eigenwerte eines selbstadjungierten Operators reell sind.

Satz 1. *Jeder vollstetige selbstadjungierte Operator A , der vom annullierenden Operator verschieden ist, hat wenigstens einen von 0 verschiedenen Eigenwert. Alle Eigenwerte von A sind von endlichem Rang, und außerhalb jedes Intervalls $[-\varepsilon, \varepsilon]$ mit $\varepsilon > 0$ können sich nur endlich viele Eigenwerte befinden. Jedes Element der Gestalt Ax ($x \in H$) läßt sich in eine Fourierreihe nach dem normierten Orthogonalsystem der Eigenelemente x_k entwickeln, die den von 0 verschiedenen Eigenwerten entsprechen:*

$$Ax = \sum_k (Ax, x_k) x_k = \sum_k (x, x_k) \lambda_k x_k. \quad (111)$$

Die Reihe (111) kann entweder endlich oder unendlich viele Summanden enthalten. Wir erinnern und noch daran, daß sich die Eigenwerte λ_k und die Eigenelemente x_k , die ein normiertes Orthogonalsystem bilden, bei der Lösung der Folgen-Extremalprobleme für die quadratische Form (Ax, x) ergeben. Darauf stützt sich der Beweis des Hauptsatzes in Teil IV/1.

Wir nehmen an, (111) enthalte unendlich viele Summanden, und x sei ein belie-

biges Element aus H . Wir setzen

$$z = x - \sum_{k=1}^{\infty} (x, x_k) x_k. \quad (112)$$

Diese Reihe konvergiert [121]. Auf Grund von (111) ist

$$A \left[x - \sum_{k=1}^{\infty} (x, x_k) x_k \right] = 0;$$

z genügt also der Gleichung

$$Az = 0, \quad (113)$$

d. h., z ist entweder ein Nullelement oder ein Eigenelement von A , das zum Eigenwert $\lambda=0$ gehört. Es sei $z_1, z_2, \dots$ ein vollständiges orthonormiertes System von Eigenelementen, die dem Eigenwert $\lambda=0$ entsprechen. Ist $\lambda=0$ kein Eigenwert, so gibt es solche Elemente nicht; ist $\lambda=0$ Eigenwert, so kann er entweder von endlichem oder von unendlichem Rang sein. Ist z Lösung von (113), so können wir behaupten, daß

$$x - \sum_{k=1}^{\infty} (x, x_k) x_k = \sum c_l z_l \quad (114)$$

gilt, wobei

$$c_k = \left(x - \sum_{k=1}^{\infty} (x, x_k) x_k, z_l \right)$$

ist, oder wir erhalten $c_k = (x, z_l)$, wenn wir $(x_k, z_l) = 0$ berücksichtigen [128]. Dann folgt aus (114), daß sich jedes Element x in eine Fourierreihe nach Eigenelementen von A entwickeln läßt, wenn wir die Eigenelemente beachten, die dem Eigenwert $\lambda=0$ entsprechen. Damit haben wir den folgenden Satz bewiesen.

Satz 2. *Ein normiertes Orthogonalsystem von Eigenelementen eines vollstetigen selbstadjungierten Operators ist vollständig.*

Mit der Terminologie aus [128] können wir sagen, daß ein vollstetiger selbstadjungierter Operator ein reines Punktspektrum hat.

Alle oben durchgeführten Untersuchungen lassen sich auch dann anwenden, wenn die Summe (111) aus endlich vielen Summanden besteht. Ist $\lambda=0$ kein Eigenwert, so enthält (111) unendlich viele Summanden (Axiom B aus [95]), und für jedes Element aus H gilt

$$x = \sum_{k=1}^{\infty} (x, x_k) x_k. \quad (115)$$

Bemerkung. Wie auch bei Integralgleichungen [IV/1, 28] erhalten wir für einen vollstetigen selbstadjungierten Operator A das folgende Resultat. Ist λ kein Eigenwert und von 0 verschieden, so hat die inhomogene Gleichung

$$Ax = \lambda x + y \quad (116)$$

für jedes gegebene y nur eine einzige Lösung, und diese Lösung hat die Gestalt

$$x = -\frac{1}{\lambda} y + \sum_k \frac{\lambda_k(y, x_k)}{\lambda(\lambda_k - \lambda)} x_k. \quad (117)$$

Ist λ ein Eigenwert und ist die Bedingung für die Auslösbarkeit einer Gleichung erfüllt, d. h., ist y orthogonal zu allen Eigenelementen, so hat die allgemeine Lösung von (116) die Form (117), in der alle Faktoren von x_k , bei denen der Nenner verschwindet, durch willkürliche Konstanten ersetzt werden müssen.

Wir nehmen an, daß es sowohl positive als auch negative Eigenwerte gibt, und numerieren sie in der Reihenfolge nichtwachsender Absolutbeträge. Die positiven Eigenwerte bezeichnen wir mit λ_k^+ , die negativen mit λ_k^- und die ihnen entsprechenden Eigenfunktionen mit x_k^+ bzw. x_k^- . Unter Berücksichtigung der Entwicklung (111) erhalten wir

$$(Ax, x) = \sum_k \lambda_k^+ |(x, x_k^+)|^2 + \sum_k \lambda_k^- |(x, x_k^-)|^2, \quad (118)$$

woraus sich unmittelbar die neue Formulierung der Extremaleigenschaften von λ_k und x_k ergibt, von denen wir schon sprachen (vgl. [IV/1, 36]).

Satz 3. *Der Eigenwert λ_1^+ ist der größte Wert von (Ax, x) für $\|x\| = 1$ und wird für $x = x_1^+$ angenommen; der Eigenwert λ_n^+ ($n > 1$) ist der größte Wert von (Ax, x) für $\|x\| = 1$ und $(x, x_1^+) = (x, x_2^+) = \dots = (x, x_{n-1}^+) = 0$; er wird für $x = x_n^+$ erreicht.*

Analog ist λ_1^- der kleinste Wert von (Ax, x) für $\|x\| = 1$ und wird für $x = x_1^-$ angenommen; λ_n^- ($n > 1$) ist der kleinste Wert von (Ax, x) für $\|x\| = 1$ und $(x, x_1^-) = (x, x_2^-) = \dots = (x, x_{n-1}^-) = 0$; er wird für $x = x_n^-$ erreicht.

Wir beweisen nun den Satz von COURANT [IV/2, 89] für den Raum H .

Satz 4. *Es seien $z_1, \dots, z_{n-1}$ beliebige feste Elemente aus H , und $m(z_1, \dots, z_{n-1})$ sei die obere Grenze der Werte von (Ax, x) bei den Bedingungen*

$$\|x\| = 1 \quad \text{und} \quad (x, z_1) = (x, z_2) = \dots = (x, z_{n-1}) = 0. \quad (119)$$

Dann ist μ_n^+ die kleinste der Zahlen $m(z_1, \dots, z_{n-1})$ bei jeder beliebigen Wahl der Elemente $z_1, \dots, z_{n-1}$.

Beweis. Dieser Beweis verläuft analog zu dem aus [IV/2, 89]. Wir haben $m(x_1^+, x_2^+, \dots, x_{n-1}^+) = \mu_n^+$ und brauchen nur zu zeigen, daß für jede Wahl der z_k ($k = 1, 2, \dots, n-1$)

$$m(z_1, z_2, \dots, z_{n-1}) \geq \mu_n^+ \quad (120)$$

gilt. Ein Element x , das den Bedingungen (119) genügen soll, habe die Gestalt

$$x = \sum_{k=1}^n c_k x_k^+. \quad (121)$$

Dann kann man statt (119) die folgenden Gleichungen für die c_k schreiben:

$$\sum_{k=1}^n c_k (x_k^+, z_s) = 0 \quad (s = 1, 2, \dots, n-1), \quad (122)$$

$$\sum_{k=1}^n |c_k|^2 = 1. \quad (123)$$

Das homogene System (122) von $n-1$ Gleichungen mit n Unbekannten c_k hat von 0 verschiedene Lösungen. Fügen wir einer dieser Lösungen einen geeigneten konstanten Faktor hinzu, so ist damit (123) erfüllt. Das Element x der Form (121)

genügt also tatsächlich den Bedingungen (119). Für dieses Element gilt

$$(Ax, x) = \left(\sum_{k=1}^n c_k \lambda_k^+ x_k^+, \sum_{k=1}^n c_k x_k^+ \right) = \sum_{k=1}^n \lambda_k^+ |c_k|^2.$$

Wegen $\mu_1^+ \cong \mu_2^+ \cong \dots \cong \mu_n^+$ und der Gültigkeit von (123) erhalten wir $(Ax, x) \cong \mu_n^+$, wobei x den Bedingungen (119) genügt. Also ist erst recht $m(z_1, z_2, \dots, z_{n-1})$, als obere Grenze von (Ax, x) unter der Bedingung (119), nicht kleiner als μ_n^+ . Damit ist der Satz bewiesen.

Analog läßt sich der entsprechende Satz für λ_n^- formulieren.

Bemerkung. Man kann leicht zeigen, daß die obere Grenze $m(z_1, z_2, \dots, z_{n-1})$ der Werte von (Ax, x) für ein gewisses x_0 erreicht wird, das den Bedingungen (119) genügt.

Nach Voraussetzung gibt es nämlich eine Folge von Elementen y_n , die den Bedingungen (119) genügen, derart, daß (Ay_n, y_n) gegen $m(z_1, z_2, \dots, z_{n-1})$ strebt. Wegen $\|y_n\| = 1$ können wir annehmen, daß y_n schwach gegen ein Element x_0 konvergiert; dabei folgt aus der schwachen Konvergenz [132]

$$\|x_0\| \leq 1 \quad \text{und} \quad (x_0, z_1) = (x_0, z_2) = \dots = (x_0, z_{n-1}) = 0.$$

Wegen der Vollstetigkeit von A ist

$$(Ax_0, x_0) = m(z_1, z_2, \dots, z_{n-1}).$$

Wir brauchen nur noch zu zeigen, daß $\|x_0\| = 1$ ist. Aus $m(z_1, z_2, \dots, z_{n-1}) \cong \mu_n^+$ folgt $m(z_1, z_2, \dots, z_{n-1}) > 0$ und $\|x_0\| > 0$. Ist $\|x_0\| < 1$, so würden wir, wenn wir ein der Bedingung (119) genügendes normiertes Element $y_0 = \frac{x_0}{\|x_0\|}$ einführen,

$$(Ay_0, y_0) = \frac{1}{\|x_0\|^2} A(x_0, x_0) = \frac{m(z_1, z_2, \dots, z_{n-1})}{\|x_0\|^2} > m(z_1, z_2, \dots, z_{n-1})$$

erhalten. Nach Definition von $m(z_1, z_2, \dots, z_{n-1})$ ist jedoch

$$(Ay_0, y_0) \leq m(z_1, z_2, \dots, z_{n-1}).$$

Der Widerspruch zeigt, daß also $\|x_0\| = 1$ ist. Damit ist bewiesen, daß (Ax, x) seine obere Grenze $m(z_1, z_2, \dots, z_{n-1})$ erreicht. Mit Hilfe von Satz 2 aus [132] kann man sehen, daß $y_n \rightarrow x_0$ ist.

Dieser Satz läßt sich beim Vergleich der Eigenwerte verschiedener Operatoren anwenden (vgl. [IV/2, 90]).

Wir wollen noch eine unmittelbare Folgerung aus Formel (118) erwähnen: Ein vollstetiger selbstadjungierter Operator A ist genau dann positiv ($(Ax, x) \geq 0$ für $x \in H$), wenn er keine negativen Eigenwerte hat.

Wir zeigen jetzt, daß ein vollstetiger selbstadjungierter Operator durch den in den Sätzen 1 und 2 beschriebenen Charakter des Spektrums vollständig bestimmt ist.

Satz 5. Ein linearer selbstadjungierter Operator A besitze die folgenden Eigenschaften:

a) Das normierte Orthogonalsystem seiner Eigenelemente x_k ($k=1, 2, \dots$) ist vollständig;

- b) alle von 0 verschiedenen Eigenwerte λ_k sind von endlichem Rang;
 c) außerhalb jedes Intervalls $[-\varepsilon, +\varepsilon]$ mit $\varepsilon > 0$ befinden sich nur endlich viele Eigenwerte.

Dann ist A vollstetig.

Beweis. Wir ordnen die Eigenwerte λ_k nach nichtwachsenden Absolutbeträgen:

$$|\lambda_1| \geq |\lambda_2| \geq |\lambda_3| \geq \dots \quad (124)$$

mit $\lambda_n \rightarrow 0$ für $n \rightarrow \infty$. Hat ein Eigenwert den Rang r , so tritt er in (124) r -mal auf (der Eigenwert $\lambda = 0$ kann von unendlichem Rang sein).

Jedes Element $x \in H$ kann man wegen der Vollständigkeit des Systems der x_k in eine Fourierreihe entwickeln:

$$x = \sum_{k=1}^{\infty} a_k x_k \quad (125)$$

und

$$Ax = \sum_{k=1}^{\infty} a_k \lambda_k x_k. \quad (126)$$

Wir nehmen an, U sei eine beschränkte Menge von Elementen x , d. h., es existiere eine positive Zahl l derart, daß

$$\sum_{k=1}^{\infty} |a_k|^2 \leq l^2 \quad (127)$$

für $x \in U$ gilt. Wir müssen zeigen, daß die Menge der Ax kompakt ist. Da die Beschränktheit dieser Menge schon aus der Ungleichung $\|Ax\| \leq n_A l$ folgt, brauchen wir also nur noch zu beweisen, daß für jedes gegebene $\varepsilon > 0$ eine solche natürliche Zahl m_ε existiert, daß $\sum_{k=m_\varepsilon}^{\infty} |a_k|^2 \lambda_k^2 \leq \varepsilon^2$ ist. Wegen $\lambda_n \rightarrow 0$ für $n \rightarrow \infty$ gibt es ein n_ε derart, daß $|\lambda_n| < \varepsilon/l$ für $n \geq n_\varepsilon$ gilt. Also ist

$$\sum_{k=n_\varepsilon}^{\infty} |a_k|^2 \lambda_k^2 < \frac{\varepsilon^2}{l^2} \sum_{k=m_\varepsilon}^{\infty} |a_k|^2 \leq \frac{\varepsilon^2}{l^2} l^2 = \varepsilon^2.$$

Damit ist der Satz bewiesen.

137. Unitäre Operatoren. Neben den selbstadjungierten Operatoren betrachten wir noch eine Klasse linearer Operatoren, die unitären Operatoren.

Definition. Ein linearer Operator U heißt *unitär*, wenn er die Norm invariant läßt, d. h., wenn $\|Ux\| = \|x\|$ ist, und H auf sich abbildet, d. h., zu jedem $y \in H$ existiert ein Urbild x , so daß

$$y = Ux \quad (128)$$

gilt.

Aus dieser Definition folgt, daß die Norm eines unitären Operators gleich 1 ist. Im folgenden Satz werden die Haupteigenschaften unitärer Operatoren angegeben.

Satz 1. Ein unitärer Operator U bildet den Raum H eineindeutig in sich ab, besitzt einen beschränkten inversen Operator, der durch die Formel

$$U^{-1} = U^*, \quad (129)$$

d. h. durch

$$UU^* = U^*U = E \quad (130)$$

definiert ist, wobei U^{-1} ebenfalls unitär ist, und läßt das Skalarprodukt invariant. Die Bedingung (130) ist hinreichend dafür, daß U ein unitärer Operator ist.

Beweis. Sind x_1 und x_2 zwei Elemente aus H , so ist nach Definition des unitären Operators $\|Ux_1 - Ux_2\| = \|U(x_1 - x_2)\| = \|x_1 - x_2\|$. Ist also $Ux_1 = Ux_2$, so gilt auch $x_1 = x_2$, d. h., zu verschiedenen Elementen x liefert (128) auch verschiedene Elemente y . Der Operator U definiert also tatsächlich eine eindeutige Zuordnung von H in sich. Es existiert ein in ganz H definierter beschränkter Operator U^{-1} , wobei, da U die Norm invariant läßt, $\|U^{-1}y\| = \|y\|$ ist, d. h., U^{-1} ist ebenfalls unitär. Wir können also $(Ux, Ux) = (x, x)$ schreiben, woraus wir sofort $(U^*Ux, x) = (x, x)$ erhalten. Eine Gleichung der quadratischen Funktionale ist aber der Gleichung der in ihnen auftretenden Operatoren äquivalent, d. h., es ist $U^*U = E$. Daraus folgt, daß U^* zu U linksinvers ist, und da zu U ein beschränkter inverser Operator existiert, ist auch $UU^* = E$. Damit ist (129) bewiesen. Daß U das Skalarprodukt invariant läßt, folgt unmittelbar aus den Beziehungen

$$(Ux, Uy) = (U^*Ux, y) = (x, y). \quad (131)$$

Zum Schluß zeigen wir, daß U auf Grund von (130) unitär ist: Wegen (130) hat U einen durch (129) definierten inversen Operator; U läßt die Norm invariant, wie wegen (130) aus (131) folgt, wenn man $y = x$ setzt, und ist demzufolge unitär.

Wir weisen noch darauf hin, daß das Produkt U_1U_2 zweier unitärer Operatoren U_1 und U_2 ebenfalls ein unitärer Operator ist. Dies folgt unmittelbar daraus, daß auch das Produkt den Raum H eindeutig in sich abbildet und die Norm invariant läßt, wenn seine Faktoren diese Eigenschaft besitzen. Somit ist ein zu einem unitären Operator inverser Operator sowie das Produkt zweier unitärer Operatoren ebenfalls unitär, d. h., die unitären Operatoren bilden eine Gruppe.

Es sei

$$x_1, x_2, x_3, \dots \quad (132)$$

ein vollständiges normiertes Orthogonalsystem. Wenden wir auf (132) eine unitäre Transformation U an, so erhalten wir wegen der eben bewiesenen Eigenschaften von U das normierte Orthogonalsystem

$$y_1 = Ux_1, \quad y_2 = Ux_2, \quad y_3 = Ux_3, \dots \quad (133)$$

Jeder Vektor x läßt sich nach den Elementen des Systems (132) entwickeln:

$$x = a_1x_1 + a_2x_2 + a_3x_3 + \dots \quad (134)$$

Daher gilt für das transformierte Element Ux eine Entwicklung nach den Elementen des Systems (133) mit denselben Koeffizienten:

$$Ux = a_1y_1 + a_2y_2 + a_3y_3 + \dots \quad (135)$$

Das Element Ux kann jedes Element aus H sein; daher ist auch das System (133) vollständig. Haben wir umgekehrt zwei vollständige normierte Orthogonalsysteme x_k und y_k ($k = 1, 2, \dots$) und definieren wir durch die Formel (135) einen Operator U für ein beliebiges Element x mit der Darstellung (134), so bildet dieser Operator

den Raum H eineindeutig in sich ab und läßt die Norm invariant:

$$||Ux||^2 = ||x||^2 = \sum_{k=1}^{\infty} |a_k|^2,$$

d. h., U ist unitär. Somit läßt sich jeder unitäre Operator mit Hilfe der Transformation der Elemente eines vollständigen normierten Orthogonalsystems in die Elemente eines anderen solchen Systems definieren.

Es sei nun A ein linearer Operator und $y = Ax$. Wir wählen einen beliebigen unitären Operator U und setzen $y' = Uy$, $x' = Ux$. Wegen $y = Ax$ sind y' und x' vermöge der Formel

$$y' = (UAU^{-1}) x' \quad (136)$$

verknüpft. Man sagt, der Operator $B = UAU^{-1}$ sei *unitär äquivalent* zu A . Aus $B = UAU^{-1}$ folgt $A = U^{-1}BU$; wenn B also unitär äquivalent zu A ist, so ist A unitär äquivalent zu B . Ist P ein Projektionsoperator in den Unterraum L_P , so ist UPU^{-1} ebenfalls ein Projektionsoperator in den Unterraum, den man erhält, wenn man U auf L_P anwendet. Ist x_0 das dem Eigenwert λ_0 entsprechende Eigenelement von A , d. h. ist $Ax_0 = \lambda_0 x_0$, so erhalten wir $(UAU^{-1}) x'_0 = \lambda_0 x'_0$, wenn wir $x'_0 = Ux_0$ setzen. *Unitär äquivalente Operatoren besitzen also die gleichen Eigenwerte, und die Eigenelemente solcher Operatoren sind durch die entsprechende unitäre Transformation verknüpft.* Mit Hilfe von (129) können wir leicht nachprüfen, daß B selbstadjungiert ist, wenn A selbstadjungiert ist.

Satz 2. *Die Eigenwerte eines unitären Operators haben den Absolutbetrag 1, und die den verschiedenen Eigenwerten entsprechenden Eigenelemente sind zueinander orthogonal.*

Beweis. Es sei U ein unitärer Operator und x_0 dasjenige Eigenelement von U , das dem Eigenwert λ_0 entspricht, $Ux_0 = \lambda_0 x_0$. Da U die Norm invariant läßt, können wir

$$(x_0, x_0) = (Ux_0, Ux_0) = (\lambda_0 x_0, \lambda_0 x_0) = |\lambda_0|^2 (x_0, x_0)$$

schreiben. Daraus folgt $||x_0|| = |\lambda_0| \cdot ||x_0||$, also $|\lambda_0| = 1$ wegen $||x_0|| \neq 0$. Um die zweite Behauptung nachzuweisen, nehmen wir an, daß x_0 und x_1 die den voneinander verschiedenen Eigenwerten λ_0 und λ_1 entsprechenden Eigenelemente sind, d. h., es sei $Ux_0 = \lambda_0 x_0$ und $Ux_1 = \lambda_1 x_1$. Berücksichtigen wir wieder die Eigenschaft von U , die Norm invariant zu lassen, so gilt

$$(x_0, x_1) = (Ux_0, Ux_1) = (\lambda_0 x_0, \lambda_1 x_1) = \lambda_0 \bar{\lambda}_1 (x_0, x_1).$$

Wäre $(x_0, x_1) \neq 0$, so würde aus diesen Gleichungen $\lambda_0 \bar{\lambda}_1 = 1$ folgen. Nun ist auf Grund des oben Bewiesenen $|\lambda_0| = 1$ und folglich $\bar{\lambda}_1 = 1/\lambda_0 = \bar{\lambda}_0$, also $\lambda_1 = \lambda_0$. Das ist aber nicht möglich, da λ_0 und λ_1 verschieden sein sollten. Also ist $(x_0, x_1) = 0$, was zu beweisen war.

Wir führen nun noch eine Klasse von Operatoren ein.

Definition. Ein linearer Operator V heißt *isometrisch*, wenn er die Norm eines Elements invariant läßt, d. h., wenn $||Vx|| = ||x||$ für $x \in H$ ist.

Ein isometrischer Operator braucht nicht unitär zu sein. Dies wollen wir an Hand eines Beispiels erklären. Es sei wie oben ein vollständiges normiertes Orthogonalsystem von Elementen x_k ($k = 1, 2, \dots$) aus H gegeben, so daß jedes Element x

durch seine Fourierreihe (134) dargestellt werden kann. Wir definieren V durch

$$Vx = \sum_{k=1}^{\infty} a_k x_{k+1} . \quad (137)$$

Wir sehen, daß V ein linearer Operator ist und

$$||Vx||^2 = ||x||^2 = \sum_{k=1}^{\infty} |a_k|^2$$

gilt. Aus (137) folgt, daß V den Raum H eineindeutig in den Raum der zu x_1 orthogonalen Elemente abbildet. Nach Definition bildet ein unitärer Operator den Raum H eineindeutig auf sich ab.

138. Die absolute Norm eines Operators. Wir führen jetzt einen neuen Normbegriff für lineare Operatoren ein. Es sei A ein linearer Operator, und x_k, y_k ($k = 1, 2, \dots$) seien zwei beliebige vollständige normierte Orthogonalsysteme in H . Dann bilden wir die aus nicht-negativen Summanden bestehende Summe

$$\sum_{p,q=1}^{\infty} (Ax_p, y_q) (y_q, Ax_p) = \sum_{p,q=1}^{\infty} |(Ax_p, y_q)|^2 \quad (138)$$

und bezeichnen die positive Quadratwurzel aus (138) mit $N(A; x_p, y_q)$. Diese Größe kann auch unendlich groß sein. Wir wollen zeigen, daß sie nicht von der Wahl der Systeme x_p, y_q abhängt. Da die (Ax_p, y_q) die Fourierkoeffizienten des Elements Ax_p bezüglich des Systems y_q sind, können wir wegen der Parsevalschen Gleichung statt (138)

$$N^2(A; x_p, y_q) = \sum_{p=1}^{\infty} ||Ax_p||^2 \quad (139)$$

schreiben. Andererseits erhalten wir wegen $(Ax_p, y_q) = (x_p, A^*y_q)$

$$N^2(A; x_p, y_q) = N^2(A^*; y_q, x_p) = \sum_{q=1}^{\infty} ||(A^*y_q)^2|| . \quad (140)$$

Aus (139) und (140) folgt, daß $N^2(A; x_p, y_q)$ weder von der Wahl der y_q noch der x_p abhängt, so daß wir also einfach $N^2(A)$ setzen können. Die positive Zahl $N(A)$ heißt die *absolute Norm* des Operators A . Sie kann auch unendlich sein. Aus (139) und (140) und der Unabhängigkeit der Norm $N(A)$ von der Wahl der Systeme x_p und y_q ergibt sich

$$N(A) = N(A^*) . \quad (141)$$

Ferner liefern die Formel

$$N^2(A + B) = \sum_{p=1}^{\infty} ||Ax_p + Bx_p||^2$$

und die Ungleichung (107) aus [59] unmittelbar

$$N(A + B) \leq N(A) + N(B) . \quad (142)$$

Es sei U ein unitärer Operator. Die Elemente $U^{-1}x_p$ bilden ein vollständiges normiertes Orthogonalsystem, und es ist $||UAz|| = ||Az||$. Daraus folgt auf Grund von (139)

$$N(UAU^{-1}) = N(A) , \quad (143)$$

d. h., *unitär äquivalente Operatoren haben die gleiche absolute Norm*. Setzen wir voraus, daß $N(A)$ endlich und x ein beliebiges normiertes Element ist, so können wir x als erstes Element eines normierten Orthogonalsystems auffassen. Aus (139) erhalten wir also

$N^2(A) \geq \|Ax\|^2$, d. h.

$$\|Ax\| \leq N(A) \quad \text{für} \quad \|x\| = 1;$$

die gewöhnliche Norm eines Operators ist demnach höchstens gleich seiner absoluten Norm.

Satz. Ist die absolute Norm des Operators A endlich, so ist A vollstetig. Ist außerdem A selbstadjungiert, so gilt

$$N^2(A) = \sum_{k=1}^{\infty} \lambda_k^2, \quad (144)$$

wobei die λ_k die Eigenwerte von A sind, die je nach ihrer Vielfachheit mehrfach auftreten.

Beweis. Es sei U eine beliebige beschränkte Menge. Wir müssen zeigen, daß die Menge der Ax ($x \in U$) für $N(A) < \infty$ kompakt ist. Nach Voraussetzung existiert eine positive Zahl l derart, daß $\|x\| \leq l$ für $x \in U$ ist. Die Beschränktheit der Menge AU folgt unmittelbar aus $\|Ax\| \leq N(A)l$. Wir führen ein beliebiges vollständiges normiertes Orthogonalsystem y_k ($k=1, 2, \dots$) ein und bilden H in l_2 ab. Dann bleibt nur noch zu zeigen, daß für jedes gegebene positive ε eine natürliche Zahl n_ε existiert, für die

$$\sum_{k=n_\varepsilon}^{\infty} |(Ax, y_k)|^2 \leq \varepsilon^2 \quad (145)$$

ist. Es gilt

$$\sum_{k=n_\varepsilon}^{\infty} |(Ax, y_k)|^2 = \sum_{k=n_\varepsilon}^{\infty} |(x, A^*y_k)|^2 \leq l^2 \sum_{k=n_\varepsilon}^{\infty} \|A^*y_k\|^2.$$

Die Reihe (140) konvergiert, da $N(A)$ endlich sein sollte, und es existiert ein nicht von $x \in U$ abhängiges n_ε , für das

$$\sum_{k=n_\varepsilon}^{\infty} \|A^*y_k\|^2 \leq \frac{\varepsilon^2}{l^2}$$

gilt. Daraus folgt (145); der Operator A ist also vollstetig. Ist A außerdem selbstadjungiert, so nehmen wir für y_k das vollständige normierte Orthogonalsystem der Eigen-elemente von A , so daß $Ay_k = \lambda_k y_k$ und $\|Ay_k\|^2 = \|A^*y_k\|^2 = \lambda_k^2$ ist. Damit erhalten wir auf Grund von (140) die Formel (144), und die Endlichkeit von $N(A)$ ist gleichbedeutend mit der Konvergenz der in (144) rechts stehenden Reihe.

Im folgenden werden wir beweisen, daß zu einem selbstadjungierten positiven Operator A , für den also $(Ax, x) \geq 0$ für $x \in H$ gilt, ein linearer positiver Operator B mit $B^2 = A$ existiert. Man schreibt im allgemeinen $B = \sqrt{A}$. Mit Hilfe dieses Operators können wir den Begriff der Spur eines linearen positiven vollstetigen selbstadjungierten Operators einführen. Aus

$$N^2(B) = \sum_{p=1}^{\infty} \|Bx_p\|^2 = \sum_{p=1}^{\infty} (Bx_p, Bx_p) = \sum_{p=1}^{\infty} (B^2x_p, x_p) = \sum_{p=1}^{\infty} (Ax_p, x_p)$$

folgt, daß die Summe $\sum_{p=1}^{\infty} (Ax_p, x_p)$ nicht von der Wahl der x_p abhängt. Diese Summe nennen wir die *Spur* des Operators A und bezeichnen sie mit $\text{Sp}(A)$. Aus dem Vorhergehenden ergibt sich

$$\text{Sp}(A) = N^2(\sqrt{A}) = \sum_{p=1}^{\infty} (Ax_p, x_p).$$

Besitzt A ein reines Punktspektrum und nehmen wir für x_p das vollständige normierte Orthogonalsystem der Eigen-elemente von A , so erhalten wir wegen $Ax_p = \mu_p x_p$

$$\text{Sp}(A) = \sum_{p=1}^{\infty} \mu_p.$$

139. Operatoren über Unterräumen. In diesem und dem folgenden Abschnitt werden die Operatoren über Unterräumen und die Eigenschaften der Projektion untersucht. Im folgenden verwenden wir dann die Ergebnisse in der Theorie der selbstadjungierten Operatoren.

Es seien L_k ($k=1, 2, \dots, m$) paarweise orthogonale Unterräume. Ihre Summe (vgl. [122]) definieren wir durch

$$L = L_1 \oplus L_2 \oplus \dots \oplus L_m. \quad (146)$$

Mit L werde die Menge der Elemente x bezeichnet, die die Gestalt

$$x = x_1 + x_2 + \dots + x_m, \quad x_k \in L_k, \quad (147)$$

haben. Aus der Orthogonalität der L_k folgt $x_k = P_{L_k}x$ ($k=1, 2, \dots, m$), und es gilt die Gleichung

$$\|x\|^2 = \|x_1\|^2 + \|x_2\|^2 + \dots + \|x_m\|^2.$$

Es läßt sich leicht beweisen, daß L ein Unterraum ist. Er heißt die *orthogonale Summe* der Unterräume L_k .

Wir betrachten nun die unendliche Summe

$$L = L_1 \oplus L_2 \oplus L_3 \oplus \dots \quad (148)$$

von paarweise orthogonalem Unterräumen. Mit L bezeichnen wir jetzt die Menge der Elemente x , die sich als Summe einer konvergenten Reihe darstellen lassen:

$$x = x_1 + x_2 + x_3 + \dots, \quad x_k \in L_k. \quad (149)$$

Diese Gleichung ist der Beziehung

$$\|x\|^2 = \|x_1\|^2 + \|x_2\|^2 + \|x_3\|^2 + \dots \quad (150)$$

äquivalent [122], und ist (150) erfüllt, so gilt $x_k = P_{L_k}x$. Ist x ein beliebiges Element von H , $x_k = P_{L_k}x$ und $s_m(x) = x_1 + x_2 + \dots + x_m$, so gilt

$$\|x - s_m(x)\|^2 = \|x\|^2 - \sum_{k=1}^m \|x_k\|^2. \quad (151)$$

Die Gleichung (150) und die Beziehung $\|x - s_m(x)\| \rightarrow 0$ für $m \rightarrow \infty$ sind gleichwertig. Wir können leicht sehen, daß L eine lineare Mannigfaltigkeit ist. Wollen wir zeigen, daß L auch Unterraum ist, so nehmen wir $x^{(n)} \in L$ und $x^{(n)} \rightarrow x$ für $n \rightarrow \infty$ an und beweisen, daß auch x zu L gehört. Offenbar gilt

$$\|x - s_m(x)\| \leq \|x - x^{(n)}\| + \|x^{(n)} - s_m(x^{(n)})\| + \|s_m(x^{(n)} - x)\|.$$

Da $s_m(x^{(n)} - x)$ die Projektion von $x^{(n)} - x$ in den Unterraum $L_1 \oplus L_2 \oplus \dots \oplus L_m$ ist, gilt $\|s_m(x^{(n)} - x)\| \leq \|x^{(n)} - x\|$, und wir schreiben

$$\|x - s_m(x)\| \leq 2 \|x - x^{(n)}\| + \|x^{(n)} - s_m(x^{(n)})\|. \quad (152)$$

Es sei ein positives ε gegeben, und wir wählen ein n derart, daß $\|x - x^{(n)}\| \leq \varepsilon/3$ ist. Nun ist $\|x^{(n)} - s_m(x^{(n)})\| \leq \varepsilon/3$ für hinreichend große m , denn $x^{(n)}$ gehört zu L , und aus (152) folgt $\|x - s_m(x)\| \leq \varepsilon$, d. h., x liegt in L , und L ist ein Unterraum. Das Element $P_L y$ ist für jedes $y \in H$ in der Form

$$P_L y = z_1 + z_2 + \dots \quad (153)$$

mit $z_k = P_{L_k}(P_L y)$ darstellbar. Da der Unterraum L_k in L vorkommt, erhalten wir $P_L(P_{L_k} y) = P_{L_k} y$, also $P_L P_{L_k} = P_{L_k}$; darauf folgt $z_k = P_{L_k} y$, wenn wir zu den adjungierten Operatoren $P_{L_k} P_L = P_{L_k}$ übergehen. Somit erhält (153) die Gestalt

$$P_L y = P_{L_1} y + P_{L_2} y + P_{L_3} y + \dots, \quad (154)$$

d. h.

$$P_L = P_{L_1} + P_{L_2} + P_{L_3} + \dots, \quad (155)$$

wobei die Konvergenz der Reihe im Sinne der starken Konvergenz von Operatorfolgen zu verstehen ist.

Sind $x_1, x_2, \dots$ paarweise orthogonale und normierte Elemente, so erhalten wir, wenn wir annehmen, daß jedes x_k einen eindimensionalen Unterraum L_k der Elemente ax_k (a beliebige komplexe Zahl) erzeugt, die orthogonale Summe dieser von Elementen der Form $\sum_k c_k x_k$ gebildeten Unterräume. Die aus den Zahlen $|c_k|^2$ gebildete Reihe konvergiert, und der Projektionsoperator in den Unterraum L hat die Gestalt

$$P_L y = \sum_k a_k x_k \quad \text{mit} \quad a_k x_k = P_{L_k} y, \quad a_k = (y, x_k).$$

Man sagt, daß ein Unterraum M ein *Teil* des Unterraumes L sei ($M \subset L$), wenn alle Elemente von M in L liegen. Dabei ist die Differenz $L \ominus M$ der Unterräume die Menge der zu M orthogonalen Elemente von L [122]. Setzen wir $L \ominus M = M_1$, so ist $L = M \oplus M_1$, d. h., die Unterräume M und M_1 sind bezüglich L komplementär [122].

Das Produkt $L_1 L_2$ zweier Unterräume ist die Menge der Elemente, die L_1 und L_2 angehören. Es läßt sich leicht zeigen, daß diese Menge ein Unterraum ist. Diese Produktdefinition ist auch auf endlich oder abzählbar unendlich viele Unterräume anwendbar.

140. Projektionsoperatoren. Wir sahen, daß der Projektionsoperator P_L in den Unterraum L ein selbstadjungierter Operator mit der Norm 1 ist (ausgeschlossen ist der Fall, daß P_L der Nulloperator ist [124]). Aus der Definition von P_L folgt sofort

$$P_L^2 = P_L \quad (156)$$

und daher

$$(P_L x, x) = (P_L^2 x, x) = (P_L x, P_L x) = \|P_L x\|^2 \geq 0,$$

d. h., P_L ist ein positiver Operator.

Wir beweisen nun einige Sätze über Projektionsoperatoren.

Satz 1. Ist A ein selbstadjungierter Operator, der der Beziehung

$$A^2 = A \quad (157)$$

genügt, so ist A Projektionsoperator in den von den Elementen $y = Ax$ gebildeten Raum L , wobei x den ganzen Raum H durchläuft.

Beweis. Die Menge L der Elemente $y = Ax$, wobei x den Raum H durchläuft, ist infolge der Distributivität des Operators A linear. Wir wollen zeigen, daß L ein Unterraum ist. Es sei y_n ($n = 1, 2, \dots$) eine Folge von Elementen aus L , und es gelte

$y_n \Rightarrow y$. Wir müssen also nachweisen, daß auch y zu L gehört. Wegen $y_n \in L$ können wir behaupten, daß es Elemente x_n gibt, für die $y_n = Ax_n$ oder wegen (157) $y_n = A(Ax_n)$, also $y_n = Ay_n$ ist. Daraus erhalten wir, wenn wir zur Grenze übergehen und die Stetigkeit von A berücksichtigen, $y = Ay$ und folglich $y \in L$. Wir brauchen jetzt nur noch zu zeigen, daß das Element $x - Ax$ zu jedem Element aus L , d. h. zu jedem Element Az orthogonal ist, wobei z ein Element aus H ist. Es gilt

$$(x - Ax, Az) = (x, Az) - (Ax, Az)$$

oder, da A selbstadjungiert ist,

$$(x - Ax, Az) = (x, Az) - (x, A^2z).$$

Wegen (157) verschwindet die rechte Seite dieser Gleichung. Also ist $(x - Ax, Az) = 0$. Damit ist der Satz bewiesen.

Zwei Projektionsoperatoren P_L und P_M sind zueinander orthogonal, wenn die Bedingung

$$P_L P_M = 0 \tag{158}$$

erfüllt ist; 0 bezeichnet hier den Nulloperator. Gehen wir in (158) zu den adjungierten Operatoren über und berücksichtigen wir, daß ein Projektionsoperator selbstadjungiert ist, so erhalten wir neben (158) noch die Bedingung

$$P_M P_L = 0. \tag{159}$$

Satz 2. *Die Projektionsoperatoren P_L und P_M sind dann und nur dann zueinander orthogonal, wenn die Unterräume L und M zueinander orthogonal sind.*

Beweis. Die Bedingung ist notwendig, Sind nämlich L und M nicht zueinander orthogonal, so existiert ein Element $x_0 \in M$, das zu L nicht orthogonal ist. Für dieses Element erhalten wir $P_M x = x$ und folglich $P_L(P_M x) = P_L x \neq 0$. Dies widerspricht jedoch der Beziehung (158). Die Bedingung ist auch hinreichend; denn ist $L \perp M$, so ist $P_M x$ orthogonal zu L für jedes Element x , und es gilt $P_L(P_M x) = 0$, d. h., die Formel (158) ist richtig.

Satz 3. *Die Summe $P_L + P_M$ ist dann und nur dann ein Projektionsoperator, wenn die Unterräume L und M zueinander orthogonal sind. Ist diese Bedingung erfüllt, so ist $P_L + P_M$ Projektionsoperator in $L \oplus M$.*

Beweis. Um die Notwendigkeit der Bedingung zu zeigen, nehmen wir an, $P_L + P_M$ sei ein Projektionsoperator, für den auf Grund von (156)

$$(P_L + P_M)(P_L + P_M) = P_L + P_M \tag{160}$$

gilt. Nach Umformung ergibt sich

$$P_L P_M + P_M P_L = 0, \tag{161}$$

denn es ist $P_L^2 = P_L$, $P_M^2 = P_M$. Multiplizieren wir (161) von links mit P_L , so finden wir

$$P_L P_M + P_L P_M P_L = 0. \tag{162}$$

Multiplizieren wir diese Gleichung von rechts mit P_L , so ist $P_L P_M P_L = 0$. Wir gelangen also mit Hilfe von (162) zu der Formel $P_L P_M = 0$, aus der wegen Satz 1 die Beziehung $L \perp M$ folgt. Damit ist die Notwendigkeit bewiesen. Sind nun L und M

zueinander orthogonal, so ist wegen (158) und (159) die Gleichung (161) und somit auch die Gleichung $(P_L + P_M)^2 = P_L + P_M$ erfüllt. Auf Grund von Satz 1 ist $P_L + P_M$ ein Projektionsoperator. Der diesem Operator entsprechende Unterraum wird von den Elementen

$$y = (P_L + P_M)x = P_Lx + P_Mx \quad (163)$$

aufgespannt (x durchläuft ganz H). Dabei gehört P_Lx zu L und P_Mx zu M . Somit liegt jedes durch (163) definierte Element y in $L \oplus M$. Ist dagegen $u + v$ ein beliebiges Element aus $L \oplus M$, wobei $u \in L$ und $v \in M$ ist, so erhalten wir, wenn wir $x = u + v$ in (163) setzen, $y = u + v$. Die Formel (163) definiert also tatsächlich den Unterraum $L \oplus M$. Damit ist der Satz bewiesen.

Man sagt, der Operator P_M sei ein Teil des Operators P_L , wenn

$$P_L P_M = P_M \quad (164)$$

gilt. Gehen wir zu den adjungierten Operatoren über, so erhalten wir

$$P_M P_L = P_M. \quad (165)$$

Satz 4. *Der Operator P_M ist dann und nur dann ein Teil von P_L , wenn der Unterraum M ein Teil von L ist. Diese Bedingung ist der Ungleichung*

$$\|P_M x\| \leq \|P_L x\| \quad (166)$$

für jedes x oder der Ungleichung

$$P_M \leq P_L \quad (167)$$

äquivalent.

Beweis. Ist (164) erfüllt und ist $x_0 \in M$, so gilt $P_M x_0 = x_0$. Aus (164) folgt ferner $P_L x_0 = x_0$, d. h., x_0 gehört zu L , und M ist ein Teil von L . Ist dagegen M ein Teil von L , so gehört $P_M x$ für jedes x zu M und infolgedessen auch zu L . Daher ist

$$P_L(P_M x) = P_M x,$$

d. h., die Bedingung (164) ist erfüllt. Wir können also auf Grund von (165) für jedes Element x die Beziehung $P_M x = P_M(P_L x)$ schreiben, woraus sich (166) ergibt. Es läßt sich auch zeigen, daß M auf Grund von (166) ein Teil von L ist. Wir führen diesen Beweis indirekt. Ist M nicht Teil von L , so gibt es in M ein Element x_0 , das nicht zu L gehört. Für dieses Element gilt $\|P_M x_0\| = \|x_0\|$ und $\|P_L x_0\| < \|x_0\|$. Das ist aber ein Widerspruch zu (166). Die beiden Ungleichungen (166) und (167) sind äquivalent, denn auf Grund von (157) können wir (166) in der Form

$$(P_L x, x) \geq (P_M x, x) \quad \text{oder} \quad ((P_L - P_M)x, x) \geq 0$$

schreiben. Damit ist der Satz vollständig bewiesen.

Satz 5. *Die Differenz $P_L - P_M$ ist genau dann ein Projektionsoperator, wenn M ein Teil von L ist. Ist diese Bedingung erfüllt, so ist $P_L - P_M$ Projektionsoperator in $L \oplus M$.*

Beweis. Ist $P_L - P_M$ ein Projektionsoperator, so muß

$$(P_L - P_M)(P_L - P_M) = P_L - P_M \quad (168)$$

oder, wenn wir die Klammern ausrechnen,

$$P_L P_M + P_M P_L = 2P_M \quad (169)$$

gelten. Multiplizieren wir diese Gleichung erst von links und dann von rechts mit P_L , so erhalten wir die Beziehungen

$$P_L P_M + P_L P_M P_L = 2P_L P_M \quad \text{und} \quad P_L P_M P_L + P_M P_L = 2P_M P_L.$$

Aus ihnen folgt $P_L P_M = P_M P_L$, und wegen (169) ist $P_L P_M = P_M P_L = P_M$; also gilt (164), und M ist Teil von L . Nehmen wir nun an, M ist Teil von L , d. h., es gelten die beiden Formeln (164) und (165), so folgt aus ihnen (169) und somit auch (168). Also ist $P_L - P_M$ auf Grund von Satz 1 ein Projektionsoperator. Der ihm entsprechende Unterraum wird durch

$$y = (P_L - P_M)x = P_L x - P_M x \quad (170)$$

definiert, wobei x ganz H durchläuft. Die Elemente $P_L x$ und $P_M x$ gehören zu L , denn nach Voraussetzung ist M ein Teil von L . Die Beziehung (170) gibt also die zu L gehörenden Elemente an. Diese Elemente sind außerdem orthogonal zu M . Ist nämlich z ein Element vom M , d. h., gilt $P_M z = z$, so können wir

$$(P_L x - P_M x, z) = (P_L x - P_M x, P_M z)$$

schreiben. Nehmen wir auf der rechten Seite dieser Gleichung den Operator P_M nach links und benutzen wir (165), so finden wir

$$(P_L x - P_M x, z) = (P_M x - P_M x, z) = 0,$$

d. h., $P_L x - P_M x$ ist orthogonal zu M . Also definiert (170) die in $L \ominus M$ liegenden Elemente y . Ist u ein beliebiges Element aus $L \ominus M$, ist also $u \in L$ und $u \perp M$, so gilt

$$y = P_L u - P_M u = P_L u = u.$$

Daher definiert (170) den Unterraum $L \ominus M$.

Satz 6. *Das Produkt $P_L P_M$ ist dann und nur dann ein Projektionsoperator, wenn P_L und P_M vertauschbar sind,*

$$P_L P_M = P_M P_L. \quad (171)$$

Ist diese Bedingung erfüllt, so ist $P_L P_M$ ein Projektionsoperator in den Raum LM .

Beweis. Die Bedingung (171) ist notwendig, da sie für die Selbstadjungiertheit von $P_L P_M$ notwendig und hinreichend ist. Der Operator $P_L P_M$ genügt, wenn (171) erfüllt ist, auch der Bedingung (157), denn es ist

$$(P_L P_M)(P_L P_M) = P_L^2 P_M^2 = P_L P_M.$$

Damit ist der erste Teil des Satzes bewiesen. Es sei nun x ein beliebiges Element von H . Dann gehört das Element

$$y = (P_L P_M)x = P_L(P_M x) = P_M(P_L x) \quad (172)$$

offenbar sowohl zu L als auch zu M , d. h., es gehört zu LM . Ist dagegen x_0 ein beliebiges Element aus LM , so ergibt (172) für $x = x_0$ den Wert $y = x_0$. Durch (172) ist also der Unterraum LM definiert.

Satz 7. *Der Grenzoperator einer konvergenten Folge von Projektionsoperatoren ist ebenfalls ein Projektionsoperator.*

Beweis. Es gilt $P_n \rightarrow P$, wobei die P_n Projektionsoperatoren sind und P ein selbstadjungierter Operator ist [131]. Gehen wir in der Gleichung $P_n^2 = P_n$ zur Grenze über, so ergibt sich $P^2 = P$, woraus auf Grund von Satz 1 folgt, daß P ein Projektionsoperator ist.

Satz 8. *Jede monotone Folge von Projektionsoperatoren hat einen Grenzoperator.*

Beweis. Wir betrachten zunächst die nichtfallende Folge von Projektionsoperatoren

$$P_1 \leq P_2 \leq P_3 \leq \dots \quad (173)$$

Diese Folge hat einen Grenzoperator, wenn $P_n x$ für jedes x ein Grenzelement hat; für ein beliebig gegebenes positives ε muß also ein solches N existieren, daß

$$\|P_n x - P_m x\| \leq \varepsilon \quad \text{für } n > m > N \quad (174)$$

gilt. Auf Grund von (173) und Satz 4 ist

$$\|P_1 x\| \leq \|P_2 x\| \leq \|P_3 x\| \leq \dots ;$$

dabei ist $\|P_n x\| \leq \|x\|$ für jedes n . Die nichtfallende Folge der nichtnegativen Zahlen $\|P_n x\|$ hat also einen Grenzwert, und zu einem beliebig gegebenen positiven ε existiert eine Zahl N , für die

$$\|P_n x\|^2 - \|P_m x\|^2 \leq \varepsilon^2 \quad \text{für } n > m > N$$

gilt. Dieser Ungleichung können wir auf Grund von (157) die Form

$$((P_n - P_m)x, x) \leq \varepsilon^2 \quad \text{für } n > m > N$$

geben. Da P_m ein Teil von P_n ist, stellt $P_n - P_m$ einen Projektionsoperator dar. Die letzte Ungleichung führt wegen (157) auf (174). Damit ist der Satz bewiesen.

Wir weisen noch darauf hin, daß auf Grund von Satz 7 der Grenzoperator P für die Folge (173) ein Projektionsoperator ist. Lassen wir n in der Ungleichung $((P_n - P_m)x, x) \geq 0$ gegen Unendlich streben, so erhalten wir $((P - P_m)x, x) \geq 0$, also $P \geq P_m$. Analog läßt sich zeigen, daß auch eine fallende Folge von Projektionsoperatoren einen Grenzoperator hat und dieser ebenfalls ein Projektionsoperator ist.

Satz 9. *Es seien abzählbar unendlich viele paarweise orthogonale Teilräume L_k ($k = 1, 2, \dots$) gegeben. Dann ist die Summe*

$$\sum_k P_{L_k} \quad (175)$$

ein Projektionsoperator in den Unterraum

$$L = L_1 \oplus L_2 \oplus \dots \quad (176)$$

Diese Behauptung folgt sofort aus [139].

141. Zerlegung der Einheit. Das Stieltjesche Integral. Die weitere Entwicklung der Theorie der selbstadjungierten Operatoren stützt sich auf die allgemeine Formel

zur Darstellung eines selbstadjungierten Operators. Zur Herleitung dieser Formel müssen wir noch einen neuen wichtigen Begriff einführen.

Definition. Die *Zerlegung der Einheit* ist eine Schar von Projektionsoperatoren $\mathfrak{E}_\lambda$, die von einem reellen Parameter λ abhängt und den folgenden Bedingungen genügt:

1. Der Projektionsoperator $\mathfrak{E}_\lambda$ wird mit wachsendem λ nicht kleiner, d. h., es ist $\mathfrak{E}_\mu \supseteq \mathfrak{E}_\lambda$ für $\mu > \lambda$;
2. es existieren endliche Zahlen $\lambda = a$ und $\lambda = b$ derart, daß $\mathfrak{E}_a = O$ und $\mathfrak{E}_b = E$ ist;
3. der Projektionsoperator $\mathfrak{E}_\lambda$ ist bezüglich λ von rechts stetig, d. h., es gilt

$$\lim_{\lambda \rightarrow \lambda' + 0} \mathfrak{E}_\lambda = \mathfrak{E}_{\lambda'} \quad (177)$$

Auf Grund von Satz 7 aus [140] existiert zu jedem λ' sowohl für $\lambda \rightarrow \lambda' - 0$ als auch für $\lambda \rightarrow \lambda' + 0$ ein Grenzooperator $\mathfrak{E}_\lambda$. Diese Grenzooperatoren sind Projektionsoperatoren, die wir mit $\mathfrak{E}_{\lambda'-0}$ bzw. $\mathfrak{E}_{\lambda'+0}$ bezeichnen wollen. Wegen (177) ist $\mathfrak{E}_{\lambda'+0} = \mathfrak{E}_{\lambda'}$, und wir wollen sagen, $\mathfrak{E}_\lambda$ ist im Punkt λ stetig, wenn $\mathfrak{E}_\lambda = \mathfrak{E}_{\lambda-0}$ ist. Die Bedingung (177) fordert, daß der Projektionsoperator $\mathfrak{E}_\lambda$ in jedem Punkt von rechts stetig ist; sie wird nur deshalb hinzugenommen, damit der Wert von $\mathfrak{E}_\lambda$ in jedem seiner Unstetigkeitspunkte bezüglich λ bestimmt werden kann.

Wir erwähnen noch einige Eigenschaften der Zerlegung der Einheit $\mathfrak{E}_\lambda$. Es sei m die obere Grenze derjenigen λ , für welche $\mathfrak{E}_\lambda = O$ ist, d. h.

$$\mathfrak{E}_\lambda = O \quad \text{für } \lambda < m \quad \text{und} \quad \mathfrak{E}_\lambda > O \quad \text{für } \lambda > m. \quad (178)$$

Im Punkt $\lambda = m$ selbst ist der Projektionsoperator $\mathfrak{E}_\lambda$ vom Nulloperator verschieden, wenn $\mathfrak{E}_\lambda$ in diesem Punkt eine Sprungstelle besitzt. Mit M bezeichnen wir die untere Grenze derjenigen λ , für die $\mathfrak{E}_\lambda = E$ ist. Da $\mathfrak{E}_\lambda$ von rechts stetig ist, muß $\mathfrak{E}_M = E$ sein; somit ist M durch die Bedingungen

$$\mathfrak{E}_\lambda < E \quad \text{für } \lambda < M \quad \text{und} \quad \mathfrak{E}_\lambda = E \quad \text{für } \lambda \geq M \quad (179)$$

definiert.

Ist ε_0 eine beliebige feste positive Zahl, so können wir sagen, daß der Projektionsoperator $\mathfrak{E}_\lambda$ zwischen O und E variiert, wenn λ das Intervall $[m - \varepsilon_0, M]$ durchläuft. Ferner können wir auf Grund von Satz 5 aus [140] behaupten, daß für $\mu > \lambda$ die Differenz $\mathfrak{E}_\mu - \mathfrak{E}_\lambda$ ein Projektionsoperator ist und

$$\mathfrak{E}_\lambda \mathfrak{E}_\mu = \mathfrak{E}_\mu \mathfrak{E}_\lambda = \mathfrak{E}_\lambda \quad (\mu > \lambda) \quad (180)$$

gilt. Lassen wir in der Differenz $\mathfrak{E}_\mu - \mathfrak{E}_\lambda$ die Zahl λ gegen $\mu - 0$ streben, so sehen wir, daß $\mathfrak{E}_\mu - \mathfrak{E}_{\mu-0}$ ein Projektionsoperator ist. Genauso läßt sich beweisen, daß $\mathfrak{E}_\nu - \mathfrak{E}_{\nu-0}$ für $\nu > \mu$ ein Projektionsoperator ist.

Wir führen nun einen neuen Begriff ein, den wir im folgenden oft benutzen werden. Mit Δ bezeichnen wir das Intervall $[\alpha, \beta]$ und setzen

$$\Delta \mathfrak{E}_\lambda = \mathfrak{E}_\beta - \mathfrak{E}_\alpha. \quad (181)$$

Sind Δ' und Δ'' zwei Intervalle, die keine gemeinsamen inneren Punkte besitzen, so gilt wegen (180)

$$\Delta' \mathfrak{E}_\lambda \cdot \Delta'' \mathfrak{E}_\lambda = O. \quad (182)$$

Auf Grund von Satz 2 aus [140] ist (182) der folgenden Aussage äquivalent: Für beliebige Elemente x und y gilt

$$\Delta' \mathfrak{E}_\lambda x \perp \Delta'' \mathfrak{E}_\lambda y. \quad (183)$$

Ist Δ_0 der Durchschnitt der Intervalle Δ' und Δ'' , so finden wir mit Hilfe von (180)

$$\Delta' \mathbb{S}_\lambda \cdot \Delta'' \mathbb{S}_\lambda = \Delta_0 \mathbb{S}_\lambda. \quad (184)$$

Da wir Operatoren addieren und in Folgen von Operatoren zur Grenze übergehen können, haben wir hier die Möglichkeit, mit Hilfe der Zerlegung der Einheit $\mathbb{S}_\lambda$ das „Stieltjessche Integral“ für eine beliebige stetige Funktion zu konstruieren. Es sei auf dem Intervall $[m - \varepsilon_0, M]$, wobei ε_0 eine feste positive Zahl ist, eine stetige Funktion $f(\lambda)$ gegeben, die auch komplex sein kann. Das Intervall zerlegen wir mit Hilfe von Punkten $\lambda_0, \dots, \lambda_n$ in Teilintervalle:

$$m - \varepsilon_0 = \lambda_0 < \lambda_1 < \lambda_2 < \dots < \lambda_{n-1} < \lambda_n = M, \quad (185)$$

und für diese δ -Unterteilung des Intervalls $[m - \varepsilon_0, M]$ stellen wir die ihr entsprechende „Riemann-Stieltjessche Summe“

$$\sigma_\delta = \sum_{k=1}^n f(v_k) \Delta_k \mathbb{S}_\lambda = \sum_{k=1}^n f(v_k) (\mathbb{S}_{\lambda_k} - \mathbb{S}_{\lambda_{k-1}}) \quad (186)$$

auf, wobei v_k ein Wert aus $[\lambda_{k-1}, \lambda_k]$ ist. Die Summe σ_δ ist ein linearer Operator. Wir bezeichnen mit η_δ die größte der Differenzen $\lambda_k - \lambda_{k-1}$ und beweisen den folgenden grundlegenden Satz.

Satz. Für jede Folge von Unterteilungen δ_n mit $\eta_{\delta_n} \rightarrow 0$ besitzt die Folge der Operatoren σ_{δ_n} im Sinne der starken Konvergenz von Operatoren einen Grenzoperator.

Zum Beweis benötigen wir die beiden folgenden Hilfssätze.

Hilfssatz 1. Sind α und α_k ($k=1, 2, \dots, n$) komplexe Zahlen und ist

$$x = x_1 + x_2 + \dots + x_n,$$

wobei die Elemente x_k paarweise orthogonal sind, so gilt die Ungleichung

$$\left\| \alpha x - \sum_{k=1}^n \alpha_k x_k \right\| \leq \delta \|x\|; \quad (187)$$

dabei ist δ die größte der Zahlen $|\alpha - \alpha_k|$.

Es ist

$$\alpha x - \sum_{k=1}^n \alpha_k x_k = \sum_{k=1}^n (\alpha - \alpha_k) x_k,$$

woraus auf Grund des Satzes des PYTHAGORAS

$$\left\| \alpha x - \sum_{k=1}^n \alpha_k x_k \right\|^2 = \sum_{k=1}^n |\alpha - \alpha_k|^2 \|x_k\|^2 \leq \delta^2 \sum_{k=1}^n \|x_k\|^2 \quad (188)$$

folgt. Aus dem gleichen Satz ergibt sich auch

$$\|x\|^2 = \sum_{k=1}^n \|x_k\|^2,$$

so daß wir aus (188) sofort (187) erhalten.

Hilfssatz 2. Ist δ die Zerlegung (185) des Intervalls $[m - \varepsilon_0, M]$ und δ' eine beliebige andere Zerlegung

$$m - \varepsilon_0 = \lambda'_0 < \lambda'_1 < \lambda'_2 < \dots < \lambda'_{n'-1} < \lambda'_{n'} = M$$

desselben Intervalls, so gilt für jedes x die Ungleichung

$$\|\sigma_\delta x - \sigma_{\delta'} x\| \leq 2\omega \|x\|, \quad (189)$$

wobei ω die größte Schwankung der Funktion $f(\lambda)$ in den Intervallen $[\lambda_{k-1}, \lambda_k]$ und $[\lambda'_{k-1}, \lambda'_k]$ ist, d. h., ω ist diejenige Zahl, für die die Ungleichung

$$|f(\alpha) - f(\beta)| \leq \omega \quad (190)$$

gilt, wenn α und β beide entweder dem Intervall $[\lambda_{k-1}, \lambda_k]$ oder dem Intervall $[\lambda'_{k-1}, \lambda'_k]$ angehören.

Wir untersuchen das Produkt $\delta\delta'$ der beiden Unterteilungen. Beim Übergang von δ zu $\delta\delta'$ wird jedes Teilintervall Δ_k von δ in endlich viele Intervalle $\Delta_k^{(s)}$ ($s = 1, 2, \dots, m_k$) zerlegt. Dabei muß jeder Summand $f(v_k) \Delta_k \mathbb{E}_\lambda x$ von

$$\sigma_\delta x = \sum_{k=1}^n f(v_k) \Delta_k \mathbb{E}_\lambda x \quad (191)$$

durch die Summe $\sum_{s=1}^{m_k} f(v_k^{(s)}) \Delta_k^{(s)} \mathbb{E}_\lambda x$ ersetzt werden; $v_k^{(s)}$ ist ein Wert aus dem Intervall $\Delta_k^{(s)}$. Die Werte v_k und $v_k^{(s)}$ liegen also beide im Intervall Δ_k von δ , und wegen (190) gilt

$$|f(v_k) - f(v_k^{(s)})| \leq \omega \quad (s = 1, 2, \dots, m_k). \quad (192)$$

Wir betrachten nun die Differenz

$$\sigma_\delta x - \sigma_{\delta\delta'} x = \sum_{k=1}^n \left[f(v_k) \Delta_k \mathbb{E}_\lambda x - \sum_{s=1}^{m_k} f(v_k^{(s)}) \Delta_k^{(s)} \mathbb{E}_\lambda x \right].$$

Auf Grund von (183) sind die Elemente $\Delta_k \mathbb{E}_\lambda x$ und $\Delta_k^{(s)} \mathbb{E}_\lambda x$ für verschiedene k zueinander orthogonal. Wir können also, wenn wir den Satz des PYTHAGORAS benutzen,

$$\|\sigma_\delta x - \sigma_{\delta\delta'} x\|^2 = \sum_{k=1}^n \left\| f(v_k) \Delta_k \mathbb{E}_\lambda x - \sum_{s=1}^{m_k} f(v_k^{(s)}) \Delta_k^{(s)} \mathbb{E}_\lambda x \right\|^2 \quad (193)$$

schreiben. Außerdem gilt

$$\Delta_k \mathbb{E}_\lambda x = \sum_{s=1}^{m_k} \Delta_k^{(s)} \mathbb{E}_\lambda x,$$

und die Elemente $\Delta_k^{(s)} \mathbb{E}_\lambda x$ ($s = 1, 2, \dots, m_k$) sind ebenfalls paarweise orthogonal. Wegen Hilfssatz 1 und Ungleichung (192) ergibt sich die Abschätzung

$$\left\| f(v_k) \Delta_k \mathbb{E}_\lambda x - \sum_{s=1}^{m_k} f(v_k^{(s)}) \Delta_k^{(s)} \mathbb{E}_\lambda x \right\| \leq \omega |\Delta_k \mathbb{E}_\lambda x|$$

und folglich wegen (193)

$$\|\sigma_\delta x - \sigma_{\delta\delta'} x\|^2 \leq \omega^2 \sum_{k=1}^n |\Delta_k \mathbb{E}_\lambda x|^2. \quad (194)$$

Da $\mathbb{E}_{m-\varepsilon_0} = O$ und $\mathbb{E}_M = E$ ist, erhält x die Form

$$x = \sum_{k=1}^n \Delta_k \mathbb{E}_\lambda x; \quad (195)$$

dabei sind die Summanden paarweise orthogonal. Der Satz des PYTHAGORAS ergibt

$$\|x\|^2 = \sum_{k=1}^n |\Delta_k \mathbb{E}_\lambda x|^2, \quad (196)$$

so daß (194) die Gestalt

$$\|\sigma_\delta x - \sigma_{\delta\delta'} x\| \leq \omega \|x\|$$

hat.

Analog läßt sich beweisen, daß

$$||\sigma_{\delta'}x - \sigma_{\delta\delta'}x|| \leq \omega ||x||$$

ist. Die Behauptung aus Hilfssatz 2 folgt dann sofort aus der Ungleichung

$$||\sigma_{\delta}x - \sigma_{\delta'}x|| \leq ||\sigma_{\delta}x - \sigma_{\delta\delta'}x|| + ||\sigma_{\delta'}x - \sigma_{\delta\delta'}x||.$$

Wir gehen nun zum Beweis des obigen Satzes über. Zunächst muß gezeigt werden, daß die Folge der Elemente $\sigma_{\delta_n}x$ für jedes x ein Grenzelement y besitzt: $\sigma_{\delta_n}x \Rightarrow y$. Dabei können wir sofort erkennen, daß das Grenzelement y nicht von der Wahl der Folge δ_n abhängt. Sind nämlich δ_n und δ'_n zwei Folgen von Unterteilungen, die den Voraussetzungen des Satzes genügen, wobei $\sigma_{\delta_n}x \Rightarrow y$ und $\sigma_{\delta'_n}x \Rightarrow y'$ ist, so genügt auch die Folge der Unterteilungen $\delta_1, \delta'_1, \delta_2, \delta'_2, \dots$ den Voraussetzungen; also hat die Folge der Elemente $\sigma_{\delta_1}x, \sigma_{\delta'_1}x, \sigma_{\delta_2}x, \sigma_{\delta'_2}x, \dots$ ein Grenzelement. Daraus folgt sofort $y' = y$.

Ehe wir zum weiteren Beweis des Satzes übergehen, wollen wir noch eine Ungleichung herleiten.

Die in der Summe σ_{δ} auftretenden Elemente $\Delta_k \mathcal{E}_{\lambda}x$ sind paarweise orthogonal, und auf Grund des Satzes des PYTHAGORAS gilt

$$||\sigma_{\delta}x||^2 = \sum_{k=1}^n |f(v_k)|^2 ||\Delta_k \mathcal{E}_{\lambda}x||^2. \quad (197)$$

Ferner ist der Absolutwert der stetigen Funktion $f(\lambda)$ beschränkt, $|f(\lambda)| \leq p$, wobei p eine gewisse positive Zahl ist. Damit erhält (197) die Gestalt

$$||\sigma_{\delta}x||^2 \leq p^2 \sum_{k=1}^n ||\Delta_k \mathcal{E}_{\lambda}x||^2. \quad (197_1)$$

Mit Hilfe von (196) ergibt sich hieraus $||\sigma_{\delta}x|| \leq p||x||$, d. h., bei jeder Zerlegung ist die Norm des Operators σ_{δ} höchstens gleich p .

Wir zeigen nun, daß die Folge der Elemente $\sigma_{\delta_n}x$ für jedes x ein Grenzelement hat. Auf Grund der Voraussetzung über die gleichmäßige Stetigkeit von $f(\lambda)$ auf dem Intervall $[m - \varepsilon_0, M]$ existiert zu jedem positiven ε eine Zahl N derart, daß $|f(\lambda') - f(\lambda'')| \leq \varepsilon$ gilt, wenn λ' und λ'' beide in ein und demselben Teilintervall von δ_n ($n > N$) liegen. Wenden wir jetzt den Hilfssatz 2 an, so erhalten wir die Ungleichung

$$||\sigma_{\delta_n}x - \sigma_{\delta_m}x|| \leq 2\varepsilon ||x|| \quad \text{für } n, m > N,$$

d. h., die Folge der $\sigma_{\delta_n}x$ konvergiert stark und strebt somit gegen ein gewisses Grenzelement. Damit ist der Satz vollständig bewiesen.

Zur Bezeichnung des Grenzoperators einer Operatorenfolge bei unbegrenzter Verfeinerung der Teilintervalle (im Sinne der starken Konvergenz) kann man logischerweise die gewöhnliche Bezeichnung des Stieltjesschen Integrals benutzen:

$$\lim_{k=1}^n f(v_k) \Delta_k \mathcal{E}_{\lambda} = \int_{m-\varepsilon_0}^M f(\lambda) d\mathcal{E}_{\lambda}. \quad (198)$$

Das Grenzelement der Folge der Elemente (191) bei unbegrenzter Verkleinerung der Teilintervalle läßt sich mit

$$\lim_{k=1}^n f(v_k) \Delta_k \mathcal{E}_{\lambda}x = \int_{m-\varepsilon_0}^M f(\lambda) d\mathcal{E}_{\lambda}x \quad (199)$$

bezeichnen.

Ganz analog kann man die Existenz der entsprechenden Integrale

$$\int_{\alpha}^{\beta} f(\lambda) d\mathfrak{E}_{\lambda} \quad \text{und} \quad \int_{\alpha}^{\beta} f(\lambda) d\mathfrak{E}_{\lambda} x \quad (200)$$

über ein beliebiges Teilintervall $[\alpha, \beta]$ des Intervalls $[m - \varepsilon_0, M]$ nachweisen. Außerhalb des Intervalls $[m - \varepsilon_0, M]$ hat sowohl der Operator $\mathfrak{E}_{\lambda}$ als auch das Element $\mathfrak{E}_{\lambda} x$ konstante Werte, und infolgedessen läßt sich das Integral über das endliche Intervall $[m - \varepsilon_0, M]$ in Form eines uneigentlichen Integrals darstellen:

$$\int_{m-\varepsilon_0}^M f(\lambda) d\mathfrak{E}_{\lambda} = \int_{-\infty}^{+\infty} f(\lambda) d\mathfrak{E}_{\lambda}, \quad \int_{m-\varepsilon_0}^M f(\lambda) d\mathfrak{E}_{\lambda} x = \int_{-\infty}^{+\infty} f(\lambda) d\mathfrak{E}_{\lambda} x. \quad (201)$$

Trennen wir den Operator $\mathfrak{E}_m$ an der Sprungstelle $\lambda = m$ des Operators $\mathfrak{E}_{\lambda}$ ab, so können wir diese Integrale in Integrale über das Intervall $[m, M]$ überführen:

$$\int_{m-\varepsilon_0}^M f(\lambda) d\mathfrak{E}_{\lambda} = f(m) \mathfrak{E}_m + \int_m^M f(\lambda) d\mathfrak{E}_{\lambda}, \quad \int_{m-\varepsilon_0}^M f(\lambda) d\mathfrak{E}_{\lambda} x = f(m) \mathfrak{E}_m x + \int_m^M f(\lambda) d\mathfrak{E}_{\lambda} x. \quad (202)$$

Wir weisen noch auf eine zu (197₁) analoge Abschätzung hin: Gilt auf dem Intervall $[\alpha, \beta]$ die Beziehung $|f(\lambda)| \leq p_1$, so ist

$$\left\| \int_{\alpha}^{\beta} f(\lambda) d\mathfrak{E}_{\lambda} x \right\| \leq p_1 \|(\mathfrak{E}_{\beta} - \mathfrak{E}_{\alpha}) x\|. \quad (203)$$

Im folgenden werden wir statt der unteren Grenze $m - \varepsilon_0$ einfach m schreiben. Das so bezeichnete Integral ist dann gleich dem Integral über $[m, M]$ plus den Ausdrücken $f(m) \mathfrak{E}_m$ bzw. $f(m) \mathfrak{E}_m x$.

142. Die Spektralschar eines selbstadjungierten Operators. Sind die Werte der Funktion $f(\lambda)$ reell, so ist der Operator σ_f als Linearkombination von Projektionsoperatoren mit reellen Koeffizienten selbstadjungiert, und der Grenzüperator von σ_f bei unbegrenzter Verfeinerung der Teilintervalle hat ebenfalls diese Eigenschaft. Setzen wir $f(\lambda) = \lambda$, so erhalten wir einen gewissen selbstadjungierten Operator

$$A = \int_m^M \lambda d\mathfrak{E}_{\lambda} \quad (204)$$

oder

$$Ax = \int_m^M \lambda d\mathfrak{E}_{\lambda} x. \quad (205)$$

Die Formel (204) ist grundlegend für die ganze Theorie der selbstadjungierten Operatoren. Wir gelangen zu ihr, wenn wir von einer Zerlegung der Einheit $\mathfrak{E}_{\lambda}$ ausgehen. Jeder solchen Zerlegung ist vermöge (204) ein bestimmter selbstadjungierter Operator A zugeordnet. Es läßt sich aber auch die Umkehrung beweisen.

Satz. Zu jedem gegebenen selbstadjungierten Operator A existiert eine Zerlegung der Einheit $\mathfrak{E}_{\lambda}$ derart, daß A durch (204) ausgedrückt wird.

Der Beweis dieses Satzes ist sehr kompliziert. Um die weiteren Ausführungen nicht zu unterbrechen, werden wir ihn am Ende dieses Paragraphen bringen. Im folgenden beweisen wir eine Formel, mit deren Hilfe wir $\mathfrak{E}_{\lambda}$ bezüglich eines gegebenen selbstadjungierten Operators A definieren können. Aus dieser Formel folgt dann, daß verschiedenen Zerlegungen der Einheit auch verschiedene Operatoren A entsprechen. Auf Grund des obigen Satzes stellt (204) die allgemeine Form eines

beschränkten selbstadjungierten Operators dar. Multiplizieren wir die Summe (191) für $f(\lambda) = \lambda$ mit einem beliebigen Element y und gehen wir zur Grenze über, so erhalten wir für das Skalarprodukt (Ax, y) ein Stieltjessches Integral

$$(Ax, y) = \int_m^M \lambda d(\mathfrak{E}_\lambda x, y). \quad (206)$$

Ist $\mathfrak{E}_m$ von 0 verschieden, so können wir auch

$$(Ax, y) = m(\mathfrak{E}_m x, y) + \int_m^M \lambda d(\mathfrak{E}_\lambda x, y) \quad (207)$$

schreiben, wobei das Integral das gewöhnliche Stieltjessche Integral ist. Wir könnten auch (206) statt (204) als grundlegende Formel auffassen, da der Operator A durch die Vorgabe eines bilinearen Funktional vollständig bestimmt ist. Das Skalarprodukt $(\mathfrak{E}_\lambda x, y)$ läßt sich durch vier Skalarprodukte der Form $(\mathfrak{E}_\lambda z, z) = \|\mathfrak{E}_\lambda z\|^2$ darstellen [125].

Wegen $\mathfrak{E}_\mu \geq \mathfrak{E}_\lambda$ für $\mu > \lambda$ fällt $\|\mathfrak{E}_\lambda z\|^2$ für wachsende λ nicht. Somit ist die im Differential stehende (im allgemeinen komplexe) Funktion $(\mathfrak{E}_\lambda x, y)$ eine Funktion beschränkter Variation von λ . Setzen wir $y = x$, so erhalten wir ein quadratisches Funktional in Form eines Stieltjesschen Integrals:

$$(Ax, x) = \int_m^M \lambda d(\mathfrak{E}_\lambda x, x). \quad (208)$$

Hier steht im Differential die nichtfallende Funktion $(\mathfrak{E}_\lambda x, x) = \|\mathfrak{E}_\lambda x\|^2$.

Die Schar der Projektionsoperatoren $\mathfrak{E}_\lambda$ wird im allgemeinen die *Spektralschar* des durch (204) definierten selbstadjungierten Operators A genannt. Wir zeigen, daß die oben definierten Zahlen m und M mit den in [126] definierten Grenzen des Operators A übereinstimmen, und schreiben dazu das quadratische Funktional (Ax, x) in der Gestalt

$$(Ax, x) = m\|\mathfrak{E}_m x\|^2 + \int_m^M \lambda d\|\mathfrak{E}_\lambda x\|^2.$$

Im Differential steht eine nichtfallende Funktion von λ . Ersetzen wir λ zuerst durch m und dann durch M , so ergeben sich die Ungleichungen

$$m\|\mathfrak{E}_M x\|^2 \leq (Ax, x) \leq m\|\mathfrak{E}_m x\|^2 + M[\|\mathfrak{E}_M x\|^2 - \|\mathfrak{E}_m x\|^2]$$

oder wegen $\mathfrak{E}_M x = x$ die Ungleichungen

$$m\|x\|^2 \leq (Ax, x) \leq M\|x\|^2.$$

Es bleibt noch zu zeigen, daß m und M die Grenzen für (Ax, x) mit $\|x\| = 1$ sind. Wir beweisen dies für M . Die Differenz $\mathfrak{E}_M - \mathfrak{E}_{M-\varepsilon} = E - \mathfrak{E}_{M-\varepsilon}$, wobei ε eine gegebene positive Zahl ist, ist ein vom Nulloperator verschiedener Projektionsoperator. Dabei ist $(E - \mathfrak{E}_{M-\varepsilon})x = x$, also $\mathfrak{E}_{M-\varepsilon}x = 0$ und somit $\mathfrak{E}_\lambda x = 0$ für $\lambda \leq M - \varepsilon$. Wir können also

$$(Ax, x) > (M - \varepsilon) \|(E - \mathfrak{E}_{M-\varepsilon})x\|^2 = M - \varepsilon$$

schreiben, wenn wir in (208) den Faktor λ durch $M - \varepsilon$ ersetzen und $\|x\| = 1$ annehmen. Daraus folgt, da ε beliebig ist, daß M die obere Grenze von (Ax, x) für $\|x\| = 1$ ist.

Wir leiten nun noch eine Formel her. Beide Seiten der Gleichung

$$\sigma_\delta = \sum_{k=1}^n v_k \Delta_k \mathfrak{E}_\lambda = \sum_{k=1}^n v_k (\mathfrak{E}_{\lambda_k} - \mathfrak{E}_{\lambda_{k-1}}) \quad (209)$$

multiplizieren wir mit $\mathfrak{E}_\lambda$, wobei wir annehmen, daß λ einer der Teilpunkte λ_s ist. Dabei gelten wegen (180) die Beziehungen $\mathfrak{E}_\lambda \cdot \Delta_k \mathfrak{E}_\lambda = \Delta_k \mathfrak{E}_\lambda \cdot \mathfrak{E}_\lambda = 0$ für $\lambda < \lambda_k$ und $\mathfrak{E}_\lambda \cdot \Delta_k \mathfrak{E}_\lambda = \Delta_k \mathfrak{E}_\lambda \cdot \mathfrak{E}_\lambda = \Delta_k \mathfrak{E}_\lambda$ für $\lambda \geq \lambda_k$. Somit ist

$$\mathfrak{E}_\lambda \sigma_\delta \mathfrak{E}_\lambda = \sum_{\lambda_k \leq \lambda} v_k \Delta_k \mathfrak{E}_\lambda,$$

und gehen wir zur Grenze über, so erhalten wir

$$\mathfrak{E}_\lambda A = A \mathfrak{E}_\lambda = \int_m^\lambda \lambda d\mathfrak{E}_\lambda \quad (\lambda > m) \quad (210)$$

und analog (für ein bilineares Funktional)

$$(\mathfrak{E}_\lambda A x, y) = \int_m^\lambda \lambda d(\mathfrak{E}_\lambda x, y). \quad (211)$$

143. Stetige Funktionen eines selbstadjungierten Operators. Ist A ein durch (204) definierter Operator, so können wir jeder im Intervall $[m, M]$ stetigen Funktion $f(\lambda)$ den durch

$$f(A) = \int_m^M f(\lambda) d\mathfrak{E}_\lambda \quad (212)$$

definierten Operator $f(A)$ zuordnen. Diese Zuordnung zwischen einer stetigen Funktion $f(\lambda)$ und einem Operator $f(A)$ ist distributiv, d. h., der stetigen Funktion $c_1 f_1(\lambda) + c_2 f_2(\lambda)$ entspricht der Operator $c_1 f_1(A) + c_2 f_2(A)$. Dies folgt unmittelbar aus der Distributivität des Integrals (212) bezüglich $f(\lambda)$. Außerdem ist diese Zuordnung auch multiplikativ, d. h., der Funktion $f_1(\lambda) f_2(\lambda)$ entspricht der Operator $f_1(A) f_2(A)$ oder der ihm gleiche Operator $f_2(A) f_1(A)$. Um dies zu beweisen, bilden wir das Produkt der Summen σ_δ für die Funktionen $f_1(\lambda)$ und $f_2(\lambda)$:

$$\sum_{k=1}^n f_1(v_k) \Delta_k \mathfrak{E}_\lambda \cdot \sum_{k=1}^n f_2(v_k) \Delta_k \mathfrak{E}_\lambda.$$

Unter Berücksichtigung von (182) und (184) können wir dieses Produkt in der Gestalt

$$\sum_{k=1}^n f_1(v_k) \Delta_k \mathfrak{E}_\lambda \cdot \sum_{k=1}^n f_2(v_k) \Delta_k \mathfrak{E}_\lambda = \sum_{k=1}^n f_1(v_k) f_2(v_k) \Delta_k \mathfrak{E}_\lambda \quad (213)$$

schreiben. Nach Grenzübergang erhalten wir hieraus

$$\int_m^M f_1(\lambda) d\mathfrak{E}_\lambda \cdot \int_m^M f_2(\lambda) d\mathfrak{E}_\lambda = \int_m^M f_1(\lambda) f_2(\lambda) d\mathfrak{E}_\lambda, \quad (214)$$

was zu beweisen war.

Neben der Formel (212) können wir für ein bilineares und für ein quadratisches Funktional die entsprechenden Formeln

$$(f(A) x, y) = \int_m^M f(\lambda) d(\mathfrak{E}_\lambda x, y), \quad (f(A) x, x) = \int_m^M f(\lambda) d\|\mathfrak{E}_\lambda x\|^2 \quad (215)$$

aufschreiben. Analog zu (211) gilt ferner

$$\mathfrak{E}_\lambda f(A) = f(A) \mathfrak{E}_\lambda = \int_m^\lambda f(\lambda) d\mathfrak{E}_\lambda. \quad (216)$$

Mit Hilfe von (214) erhalten wir für ganze positive Potenzen von A

$$A^n = \int_m^M \lambda^n d\mathfrak{E}_\lambda \quad (n=1, 2, 3, \dots) \quad (217)$$

und für Polynome

$$a_0 A^n + a_1 A^{n-1} + \dots + a_{n-1} A + a_n = \int_m^M (a_0 \lambda^n + a_1 \lambda^{n-1} + \dots + a_{n-1} \lambda + a_n) d\mathfrak{E}_\lambda. \quad (218)$$

Wie wir oben erwähnten, ist der Operator σ_δ selbstadjungiert, wenn $f(\lambda)$ eine reelle Funktion ist; der Grenzübergang für σ_δ , d. h. $f(A)$, ist ebenfalls selbstadjungiert. Ist $f(\lambda) \geq 0$ auf dem Intervall $[m, M]$, so ist der Operator $f(A)$ wegen (215) positiv. Wir nehmen nun an, daß $f(\lambda)$ eine komplexe Funktion ist: $f(\lambda) = \varphi(\lambda) + i\psi(\lambda)$. Dabei ist $f(A) = \varphi(A) + i\psi(A)$; auch $\varphi(A)$ und $\psi(A)$ sind selbstadjungierte Operatoren. Bilden wir den Operator $F(A) = \varphi(A) - i\psi(A)$, so können wir, wenn wir die Selbstadjungiertheit der Operatoren $\varphi(A)$ und $\psi(A)$ berücksichtigen, $(f(A)x, y) = (x, F(A)y)$ schreiben, d. h., der Operator $F(A)$ ist zu $f(A)$ adjungiert.

Wir erwähnen noch einige Vertauschbarkeitseigenschaften. Aus (210) folgt, daß der Operator $\mathfrak{E}_\lambda$ für jedes λ mit A vertauschbar ist. Daher ist auch der Operator $\Delta\mathfrak{E}_\lambda = \mathfrak{E}_\beta - \mathfrak{E}_\alpha$ für alle α und β mit A vertauschbar. Daraus folgt die Vertauschbarkeit der Summe σ_δ mit A , und durch Grenzübergang erhalten wir, daß der Operator $f(A)$ mit A vertauschbar ist.

Wir beweisen nun den folgenden Satz.

Satz 1. *Ist B ein beliebiger, mit A vertauschbarer Operator, so ist der Operator $f(A)$ mit B vertauschbar.*

Beweis. Es sei ε_n eine Folge gegen 0 strebender positiver Zahlen. Nach dem Weierstraßschen Satz [II, 168] existiert eine Folge von Polynomen $P_n(\lambda)$, für die

$$|f(\lambda) - P_n(\lambda)| \leq \varepsilon_n \quad (m \leq \lambda \leq M) \quad (219)$$

gilt. Wir bilden die Differenz

$$f(A) - P_n(A) = \int_m^M [f(\lambda) - P_n(\lambda)] d\mathfrak{E}_\lambda.$$

Wenn wir (203) und (219) berücksichtigen, können wir

$$|[f(A) - P_n(A)]x| \leq \varepsilon_n \|x\| \quad (220)$$

schreiben; daraus folgt $P_n(A) \rightarrow f(A)$. Der mit A vertauschbare Operator B ist auch mit einem beliebigen Polynom $P_n(A)$ vertauschbar, d. h., es ist $BP_n(A) = P_n(A)B$. Durch Grenzübergang erhalten wir $Bf(A) = f(A)B$. Damit ist der Satz bewiesen.

Später werden wir zeigen [161], daß auch die Spektralschar $\mathfrak{E}_\lambda$ für jedes λ mit einem beliebigen, mit A vertauschbaren Operator B vertauschbar ist. Wenn umgekehrt B mit $\mathfrak{E}_\lambda$ vertauschbar ist, so ist B mit einem beliebigen Operator $\Delta\mathfrak{E}_\lambda$, also auch mit der Summe (209) und damit (nach Grenzübergang) mit dem Operator A vertauschbar. Folglich gilt der Satz:

Satz 2. *Für die Vertauschbarkeit eines Operators mit A ist notwendig und hinreichend, daß er für jedes λ mit $\mathfrak{E}_\lambda$ vertauschbar ist.*

Wir geben nun ein Beispiel für die Funktion eines Operators an, welches wir in [138] benutzten. Es sei A ein positiver Operator, also $m \geq 0$, und wir setzen

$f(\lambda) = \sqrt{\lambda}$ ($\lambda \geq 0$), wobei wir nur das positive Vorzeichen der Wurzel berücksichtigen wollen. Dann läßt sich der positive Operator $\sqrt{A}$ durch

$$\sqrt{A} = \int_m^M \sqrt{\lambda} d\mathbb{E}_\lambda \quad \text{oder} \quad (\sqrt{A}x, y) = \int_m^M \sqrt{\lambda} d(\mathbb{E}_\lambda x, y)$$

darstellen. Wegen (214) gilt $\sqrt{A} \sqrt{A} = A$.

144. Eine Formel für die Resolvente. Die Charakterisierung der regulären Werte von λ . Durch Verwendung der Spektralschar können wir eine Formel für die Resolvente bilden [130] und die regulären Werte von λ neu charakterisieren. Im folgenden wollen wir von der Resolvente R_l nur für reguläre l sprechen.

Satz 1. *Ist l eine nicht reelle oder eine außerhalb des Intervalls $[m, M]$ der reellen Achse liegende reelle Zahl; so läßt sich die Resolvente R_l des Operators A durch*

$$R_l = \int_m^M \frac{1}{\lambda - l} d\mathbb{E}_\lambda \quad (221)$$

definieren.

Beweis. Unter den Voraussetzungen des Satzes ist die Funktion $1/(\lambda - l)$ stetig im Intervall $[m - \varepsilon_0, M]$ für hinreichend kleines ε_0 . Mit Hilfe von (214) können wir

$$\int_m^M \frac{1}{\lambda - l} d\mathbb{E}_\lambda \cdot \int_m^M (\lambda - l) d\mathbb{E}_\lambda = \int_m^M (\lambda - l) d\mathbb{E}_\lambda \cdot \int_m^M \frac{1}{\lambda - l} d\mathbb{E}_\lambda = \int_m^M d\mathbb{E}_\lambda = E \quad (222)$$

schreiben. Da aber

$$\int_m^M (\lambda - l) d\mathbb{E}_\lambda = A - lE$$

ist, folgt aus (222) sofort (221) [130].

Satz 2. *Liegt l im Intervall $[m, M]$, aber innerhalb eines Intervalls $[\alpha, \beta]$, in dem $\mathbb{E}_\lambda$ konstant ist ($\mathbb{E}_\beta = \mathbb{E}_\alpha$), so existiert die Resolvente R_l und wird durch (221) definiert.*

Beweis. Wir zerlegen $[m - \varepsilon_0, M]$ in die drei Teilintervalle $[m - \varepsilon_0, \alpha]$, $[\alpha, \beta]$ und $[\beta, M]$. In $[m - \varepsilon_0, \alpha]$ und $[\beta, M]$ ist die Funktion $1/(\lambda - l)$ stetig; in $[\alpha, \beta]$ ist $\mathbb{E}_\lambda$ konstant, so daß für dieses Intervall alle Operatoren $\Delta_k \mathbb{E}_\lambda$ gleich dem Nulloperator sind. Wir setzen die Funktion $1/(\lambda - l)$ aus den äußeren Intervallen in das mittlere Intervall $[\alpha, \beta]$ so fort, daß sie im ganzen Intervall $[m - \varepsilon_0, M]$ stetig ist, und bezeichnen die so konstruierte Funktion mit $\varphi(\lambda)$. Der Wert des Integrals

$$R_l = \int_m^M \varphi(\lambda) d\mathbb{E}_\lambda \quad (223)$$

hängt offenbar nicht davon ab, welche Werte $\varphi(\lambda)$ in $[\alpha, \beta]$ annimmt. Verwenden wir die Formel (214), so ergibt sich

$$\int_m^M \varphi(\lambda) d\mathbb{E}_\lambda \cdot \int_m^M (\lambda - l) d\mathbb{E}_\lambda = \int_m^M (\lambda - l) d\mathbb{E}_\lambda \cdot \int_m^M \varphi(\lambda) d\mathbb{E}_\lambda = \int_m^M (\lambda - l) \varphi(\lambda) d\mathbb{E}_\lambda$$

oder, wenn wir berücksichtigen, daß $\varphi(\lambda) = 1/(\lambda - l)$ für $\lambda \leq \alpha$ und $\lambda \geq \beta$ und daß $\mathbb{E}_\lambda$ in $[\alpha, \beta]$ konstant ist,

$$\int_m^M \varphi(\lambda) (\lambda - l) d\mathbb{E}_\lambda = \int_m^\alpha d\mathbb{E}_\lambda + \int_\beta^M d\mathbb{E}_\lambda = \mathbb{E}_\alpha + (E - \mathbb{E}_\beta) = E.$$

Hieraus folgt, daß (223) die Resolvente ist.

Statt (223) können wir offenbar auch das Integral (221) nehmen, wenn wir von $m - \varepsilon_0$ bis α und dann von β bis M integrieren. Aus dem eben bewiesenen Satz folgt, daß die in $[m, M]$ liegenden Werte $\lambda = l$ regulär sind, wenn sie zu dem Abschnitt gehören, in dem $\mathbb{S}_\lambda$ konstant ist. Im folgenden Satz zeigen wir, daß diese Bedingung für die Regularität nicht nur hinreichend, sondern auch notwendig ist.

Satz 3. *Existiert für ein reelles $\lambda = l$ die Resolvente R_l , so liegt l innerhalb des Intervalls $[\alpha, \beta]$, in dem $\mathbb{S}_\lambda$ konstant ist.*

Beweis. Es sei $[\alpha, \beta]$ ein beliebiges Intervall, in dem l liegt, und $\Delta \mathbb{S}_\mu = \mathbb{S}_\beta - \mathbb{S}_\alpha$. Nach Definition der Resolvente ist $\Delta \mathbb{S}_\mu x = R_l(A - lE) \Delta \mathbb{S}_\mu x$, und wegen (216) gilt

$$(A - lE) \Delta \mathbb{S}_\mu x = \int_{\alpha}^{\beta} (\lambda - l) d\mathbb{S}_\lambda x,$$

also

$$\Delta \mathbb{S}_\mu x = R_l \left[\int_{\alpha}^{\beta} (\lambda - l) d\mathbb{S}_\lambda x \right] \quad (\alpha < l < \beta).$$

Mit N bezeichnen wir die Norm des Operators R_l :

$$\|\Delta \mathbb{S}_\mu x\| \leq N \left\| \int_{\alpha}^{\beta} (\lambda - l) d\mathbb{S}_\lambda x \right\|. \quad (224)$$

Im Intervall $[\alpha, \beta]$ gilt $|\lambda - l| < \beta - \alpha$, und wegen (203) führt (224) auf die Ungleichung

$$\|\Delta \mathbb{S}_\mu x\| \leq N(\beta - \alpha) \|\Delta \mathbb{S}_\mu x\|. \quad (225)$$

Wir wählen nun das Intervall $[\alpha, \beta]$, in dem l liegt, so klein, daß die Ungleichung $N(\beta - \alpha) < 1$ erfüllt ist. Damit folgt aus (225) sofort $\|\Delta \mathbb{S}_\mu\| = 0$, also $\mathbb{S}_\beta = \mathbb{S}_\alpha$. Im Intervall $[\alpha, \beta]$ ist also $\mathbb{S}_\lambda$ konstant.

Fassen wir die Sätze 2 und 3 zusammen, so gelangen wir zu der

Folgerung. *Ein reeller Wert λ ist dann und nur dann regulär, wenn λ innerhalb desjenigen Intervalls liegt, in dem $\mathbb{S}_\lambda$ konstant ist.*

Aus dieser Folgerung können wir sofort den Schluß ziehen, daß alle reellen Werte in der Umgebung des regulären reellen Punktes λ ebenfalls regulär sind, d. h., die regulären Werte von λ bilden auf der reellen Achse eine offene Menge. Wir kommen also zu dem Schluß: *Die Punkte des Spektrums bilden eine abgeschlossene Menge.*

Für ein bilineares Funktional mit dem Operator R_l gilt die Formel

$$(R_l x, y) = \int_{m_i}^M \frac{1}{\lambda - l} d(\mathbb{S}_\lambda x, y). \quad (226)$$

Als Integrationsintervall können wir hier $(-\infty, \infty)$ wählen [141], und $(\mathbb{S}_\lambda x, y)$ ist eine Funktion von beschränkter Variation von λ . Setzen wir $\lambda = \sigma + i\tau$ und wenden wir die Umkehrformel des Cauchy-Stieltjesschen Integrals [29] an, so läßt sich die Spektralschar durch die Resolvente vermöge der Formel

$$\frac{1}{2} [(\mathbb{S}_{\lambda-0} x, y) + (\mathbb{S}_{\lambda+0} x, y)] = \lim_{\tau \rightarrow +0} \frac{1}{2\pi i} \int_{-\infty}^{\lambda} ((R_{\sigma+i\tau} - R_{\sigma-i\tau}) x, y) d\sigma \quad (227)$$

darstellen. Ist $\mathfrak{E}_\lambda$ im Punkt λ stetig, so ist die linke Seite von (227) gleich $(\mathfrak{E}_\lambda x, y)$. In Unstetigkeitsstellen wird der Wert von $(\mathfrak{E}_\lambda x, y)$ durch die rechtsseitige Stetigkeit definiert. Wir weisen darauf hin, daß die Resolvente R_l durch den Operator A bestimmt ist; damit folgt aus (227), daß es zu jedem gegebenen selbstadjungierten Operator A nur eine Spektralschar gibt, mit deren Hilfe er durch (204) ausgedrückt wird.

Wir erwähnen noch, daß auf Grund von Satz 1 aus [143] die Operatoren R_l für verschiedene l vertauschbar sind.

145. Eigenwerte und Eigenelemente. Mit Hilfe der Spektralschar lassen sich die Eigenwerte und die Eigenelemente eines selbstadjungierten Operators sehr leicht definieren.

Satz. Der Wert $\lambda = \lambda_0$ ist dann und nur dann ein Eigenwert eines selbstadjungierten Operators A mit der Spektralschar $\mathfrak{E}_\lambda$, wenn $\mathfrak{E}_\lambda$ den Punkt λ_0 als Unstetigkeitsstelle besitzt, d. h., wenn $\mathfrak{E}_{\lambda_0} - \mathfrak{E}_{\lambda_0-0} > 0$ ist. Dabei ist $\mathfrak{E}_{\lambda_0} - \mathfrak{E}_{\lambda_0-0}$ der Projektionsoperator in den Unterraum M_0 der den Eigenwerten λ_0 entsprechenden Eigenelemente.

Beweis. Es sei M_0 der dem Projektionsoperator $\mathfrak{E}_{\lambda_0} - \mathfrak{E}_{\lambda_0-0}$ zugeordnete Unterraum. Ist $\mathfrak{E}_{\lambda_0}$ im Punkt λ_0 stetig, so besteht M_0 nur aus dem Nullelement. Der Beweis des Satzes reduziert sich auf den Beweis der folgenden Behauptungen: Liegt x_0 in M_0 , so ist $(A - \lambda_0 E)x_0 = 0$, und gilt $(A - \lambda_0 E)x_0 = 0$, so gehört x_0 zu M_0 . Wir nehmen zunächst $x_0 \in M_0$ an. Dann gilt $(\mathfrak{E}_{\lambda_0} - \mathfrak{E}_{\lambda_0-0})x_0 = x_0$ und erst recht $\mathfrak{E}_{\lambda_0}x_0 = x_0$, so daß also $\mathfrak{E}_{\lambda_0-0}x_0 = 0$ ist. Da $\mathfrak{E}_\lambda$ für wachsendes λ nicht fällt, können wir behaupten, daß $\mathfrak{E}_\lambda x_0 = x_0$ für $\lambda \geq \lambda_0$ und $\mathfrak{E}_\lambda x_0 = 0$ für $\lambda < \lambda_0$ ist. Wenden wir auf das Element x_0 die Formel (205) an, so ergibt sich

$$Ax_0 = \int_m^M \lambda d\mathfrak{E}_\lambda x_0. \quad (228)$$

Bei der Bildung der Summen σ_δ setzen wir voraus, daß λ_0 ein Teilpunkt ist. Auf Grund des Vorhergehenden verschwinden alle Differenzen $\Delta_k \mathfrak{E}_\lambda x_0$ außer derjenigen, die dem Teilintervall mit dem rechten Endpunkt λ_0 entspricht, d. h.

$$Ax_0 = \lim_{\varepsilon \rightarrow 0} v_0 (\mathfrak{E}_{\lambda_0} - \mathfrak{E}_{\lambda_0-\varepsilon}) x_0 \quad \text{mit} \quad \lambda_0 - \varepsilon \leq v_0 \leq \lambda_0.$$

Gehen wir zur Grenze über, so erhalten wir

$$Ax_0 = \lambda_0 (\mathfrak{E}_{\lambda_0} - \mathfrak{E}_{\lambda_0-0}) x_0.$$

Da nach Voraussetzung x_0 in M_0 liegt, ist die rechte Seite dieser Beziehung gleich $\lambda_0 x_0$, d. h., x_0 genügt der Gleichung $(A - \lambda_0 E)x_0 = 0$. Um nun die Umkehrung beweisen zu können, nehmen wir an, daß x_0 dieser Gleichung genügt. Wir müssen also zeigen, daß $x_0 \in M_0$ ist. Aus $(A - \lambda_0 E)x_0 = 0$ folgt

$$((A - \lambda_0 E)^2 x_0, x_0) = 0 \quad (229)$$

oder, wenn wir das quadratische Funktional durch das Stieltjesche Integral ersetzen,

$$\int_m^M (\lambda - \lambda_0)^2 d \|\mathfrak{E}_\lambda x_0\|^2 = 0. \quad (230)$$

Die integrierbare Funktion $(\lambda - \lambda_0)^2$ ist nichtnegativ, und $\|\mathfrak{E}_\lambda x_0\|^2$ ist eine nichtfallende Funktion von λ . Daraus folgt, daß alle Elemente des Integrals (230) nichtnegativ sind; der Wert dieses Integrals über jeden Teil des Integrationsintervalls

muß also ebenfalls gleich 0 sein. Wählen wir eine positive Zahl ε , so können wir

$$\int_{\lambda_0+\varepsilon}^M (\lambda - \lambda_0)^2 d \|\mathbb{S}_\lambda x_0\|^2 = 0 \quad (231)$$

schreiben. Die integrierbare Funktion $(\lambda - \lambda_0)^2$ ist auf dem Integrationsintervall mindestens gleich ε^2 , so daß aus (231)

$$\varepsilon^2 \int_{\lambda_0+\varepsilon}^M d \|\mathbb{S}_\lambda x_0\|^2 = 0,$$

d. h.

$$\varepsilon^2 [\|x_0\|^2 - \|\mathbb{S}_{\lambda_0+\varepsilon} x_0\|^2] = 0$$

folgt. Da ε beliebig ist, erhalten wir hieraus $\mathbb{S}_\lambda x_0 = x_0$ für $\lambda > \lambda_0$.

Ganz analog läßt sich zeigen, daß $\mathbb{S}_\lambda x_0 = x_0$ für $\lambda < \lambda_0$ gilt. Also ergibt sich sofort

$$x_0 = \lim_{\varepsilon \rightarrow 0} (\mathbb{S}_{\lambda_0+\varepsilon} - \mathbb{S}_{\lambda_0-\varepsilon}) x_0 = (\mathbb{S}_{\lambda_0} - \mathbb{S}_{\lambda_0-0}) x_0.$$

Damit ist der Satz bewiesen.

Besitzt der Operator A Eigenwerte, so gelangen wir, wenn wir in jedem Unterraum der einem festen Eigenwert entsprechenden Eigenelemente ein vollständiges normiertes Orthogonalsystem der Eigenelemente von A einführen, zu dem normierten Orthogonalsystem

$$x_1, x_2, x_3, \dots \quad (232)$$

und zur Folge der entsprechenden Eigenwerte

$$\mu_1, \mu_2, \mu_3, \dots \quad (233)$$

des Operators A [128]. Ist der Rang eines Eigenwertes gleich r , so tritt dieser Eigenwert in der Folge (233) r -mal auf. Die Zahl r kann auch unendlich sein.

Bezeichnen wir mit λ_k ($k=1, 2, \dots$) die Unstetigkeitsstellen von $\mathbb{S}_\lambda$ und mit L_k die zugehörigen Unterräume der Eigenelemente, so können wir

$$P_{L_k} = \mathbb{S}_{\lambda_k} - \mathbb{S}_{\lambda_{k-0}} \quad (234)$$

schreiben. Wir bilden die orthogonale Summe der Unterräume L_k :

$$H' = L_1 \oplus L_2 \oplus L_3 \oplus \dots \quad (235)$$

Der Projektionsoperator in den Unterraum H' hat, wie wir wissen, die Form

$$P_{H'} = P_{L_1} + P_{L_2} + P_{L_3} + \dots \quad (236)$$

Der Unterraum H' besteht aus den Elementen x , die durch die Elemente des normierten Orthogonalsystems (232) mit Hilfe einer konvergenten Reihe

$$x = a_1 x_1 + a_2 x_2 + a_3 x_3 + \dots \quad (237)$$

dargestellt werden können.

146. Das reine Punktspektrum. Man sagt, der selbstadjungierte Operator A hat ein *reines Punktspektrum*, wenn das normierte Orthogonalsystem (232) im (separablen) Raum H vollständig ist (vgl. [128]). Das ist gleichbedeutend damit, daß der durch (235) definierte Unterraum H' mit H übereinstimmt, oder damit, daß der durch (236) definierte Projektionsoperator $P_{H'}$ die identische Transformation darstellt:

$$E = \sum_k P_{L_k}. \quad (238)$$

Wenn wir beide Seiten dieser Gleichung mit $\mathfrak{E}_\lambda$ multiplizieren und

$$(\mathfrak{E}_{\lambda_k} - \mathfrak{E}_{\lambda_k-0}) \mathfrak{E}_\lambda = \begin{cases} 0 & \text{für } \lambda < \lambda_k, \\ \mathfrak{E}_{\lambda_k} - \mathfrak{E}_{\lambda_k-0} & \text{für } \lambda \geq \lambda_k \end{cases}$$

berücksichtigen, erhalten wir die Darstellung von $\mathfrak{E}_\lambda$ mit Hilfe der Sprungstellen von $\mathfrak{E}_\lambda$:

$$\mathfrak{E}_\lambda = \sum_{\lambda_k \leq \lambda} P_{L_k} = \sum_{\lambda_k \leq \lambda} (\mathfrak{E}_{\lambda_k} + \mathfrak{E}_{\lambda_k-0}). \quad (239)$$

In dem betrachteten Fall läßt sich ein beliebiges Element x in Form der Reihe (237) darstellen; dabei sind die a_k die Fourierkoeffizienten von x bezüglich des Systems (232). Wenn wir auf beide Seiten der Gleichung (237) den Operator A anwenden und beachten, daß $Ax_k = \mu_k x_k$ gilt, so erhalten wir

$$Ax = \sum_s \mu_s a_s x_s. \quad (240)$$

Werden mit b_s die Fourierkoeffizienten von y bezeichnet, $b_s = (y, x_s)$, $\bar{b}_s = (x_s, y)$, so folgt durch skalare Multiplikation mit y die Darstellung

$$(Ax, y) = \sum_s \mu_s a_s \bar{b}_s \quad (241)$$

für bilineare Funktionale.

Für $y = x$ erhalten wir die Darstellung

$$(Ax, x) = \sum_s \mu_s |a_s|^2 \quad (242)$$

für quadratische Funktionale, die der Darstellung einer quadratischen (hermiteschen) Form in Gestalt einer Quadratsumme vollkommen analog ist. Wir haben folglich im Fall eines reinen Punktspektrums sehr einfache Darstellungen für den Operator A selbst und für das bilineare sowie das quadratische Funktional mit Hilfe des normierten Orthogonalsystems (232).

Wir betrachten nun das sogenannte rein kontinuierliche Spektrum.

147. Das einfache kontinuierliche Spektrum. Man sagt, der selbstadjungierte Operator A hat ein *rein kontinuierliches Spektrum*, wenn die Spektralschar $\mathfrak{E}_\lambda$ für alle Werte von λ stetig ist. Unsere Aufgabe ist es, zu den Gleichungen aus [146] analoge Gleichungen aufzustellen. Wir müssen jedoch zuerst noch einen neuen Begriff einführen.

Es sei e eine Menge von Elementen aus H , und $x_1, x_2, \dots, x_n$ seien beliebige dieser Elemente. Wir bilden ihre Linearkombinationen $c_1 x_1 + c_2 x_2 + \dots + c_n x_n$ mit beliebigen Koeffizienten c_k . Die Menge der Elemente von H , die auf diese Weise in Gestalt endlicher Linearkombinationen von Elementen aus e dargestellt werden können, bilden offenbar eine lineare Mannigfaltigkeit L .

Definition. Die durch Abschließen der betrachteten linearen Mannigfaltigkeit L erhaltene Menge wird *lineare abgeschlossene Hülle* der Menge e genannt.

Eine lineare abgeschlossene Hülle ist ein Unterraum, und die charakteristische Eigenschaft der in ihm liegenden Elemente x besteht in folgendem: Für ein beliebig vorgegebenes positives ε existieren eine endliche Menge von Elementen $x_1, x_2, \dots, x_n$ aus e und Zahlen c_k , so daß

$$||x - (c_1 x_1 + c_2 x_2 + \dots + c_n x_n)|| \leq \varepsilon$$

gilt.

Insbesondere liegen in diesem Unterraum offenbar alle endlichen Linearkombinationen von Elementen aus ϵ .

Es sei $\mathbb{S}_\lambda$ die Spektralschar des Operators A mit rein kontinuierlichem Spektrum. In diesem Fall ist $\mathbb{S}_m = 0$. Wir nehmen ein vom Nullelement verschiedenes Element x , bilden die Menge der Elemente

$$\mathbb{S}_\lambda x, \quad (243)$$

wobei λ alle Werte von m bis M durchläuft, und bezeichnen mit C_x die lineare abgeschlossene Hülle der Elemente (243). Jedem Element y aus H können wir eine stetige, von λ abhängige Funktion

$$\varphi_y(\lambda) = (y, \mathbb{S}_\lambda x) \quad (244)$$

zuordnen. Diese Funktion ist offenbar bezüglich y distributiv:

$$\varphi_{ay+bz}(\lambda) = a\varphi_y(\lambda) + b\varphi_z(\lambda).$$

Wir bilden außerdem die beiden stetigen, von λ abhängigen Funktionen

$$\varrho(\lambda) = (\mathbb{S}_\lambda x, x) = \|\mathbb{S}_\lambda x\|^2, \quad h_y(\lambda) = (\mathbb{S}_\lambda y, y) = \|\mathbb{S}_\lambda y\|^2. \quad (245)$$

Wie wir wissen, nehmen sie mit wachsendem λ nicht ab. Ist $\Delta = [\alpha, \beta]$ ein beliebiges Intervall, so führen wir für eine beliebige Funktion $f(\lambda)$ die übliche Bezeichnung

$$\Delta f(\lambda) = f(\beta) - f(\alpha) \quad (246)$$

ein. Zum Beispiel ist

$$\Delta \varrho(\lambda) = (\Delta \mathbb{S}_\lambda x, x) = ((\mathbb{S}_\beta - \mathbb{S}_\alpha) x, x)$$

oder

$$\Delta \varrho(\lambda) = \|\Delta \mathbb{S}_\lambda x\|^2 \quad (247)$$

und analog

$$\Delta h_y(\lambda) = \|\Delta \mathbb{S}_\lambda y\|^2. \quad (248)$$

Für die Funktion $\varphi_y(\lambda)$ gilt

$$\Delta \varphi_y(\lambda) = (y, \Delta \mathbb{S}_\lambda x) = (y, (\Delta \mathbb{S}_\lambda)^2 x) = (\Delta \mathbb{S}_\lambda y, \Delta \mathbb{S}_\lambda x)$$

und folglich $|\Delta \varphi_y(\lambda)|^2 \leq \|\Delta \mathbb{S}_\lambda x\|^2 \cdot \|\Delta \mathbb{S}_\lambda y\|^2$ oder

$$|\Delta \varphi_y(\lambda)|^2 \leq \Delta \varrho(\lambda) \Delta h_y(\lambda).$$

Daraus ergibt sich die Existenz des Integrals (vgl. [81])

$$\int_m^M \frac{d\varphi_y(\lambda) \overline{d\varphi_y(\lambda)}}{d\varrho(\lambda)} = \int_m^M \frac{|d\varphi_y(\lambda)|^2}{d\varrho(\lambda)}. \quad (249)$$

Wir zeigen weiterhin, daß dieses Integral für $y \in C_x$ gleich $\|y\|^2$ ist. Wir zerlegen dazu das Intervall $[m, M]$ in Teilintervalle Δ_k ($k = 1, 2, \dots, n$) und bilden die Elemente

$$\frac{\Delta_k \mathbb{S}_\lambda x}{\sqrt{\Delta_k \varrho(\lambda)}} \quad (250)$$

des Raumes H . Ist die Funktion $\varrho(\lambda)$ im Intervall Δ_k konstant, so ist auch $\Delta_k \mathbb{S}_\lambda x = 0$, und (250) ist ein unbestimmter Ausdruck. Wir wollen vereinbaren, solche unbestimmten Ausdrücke aus den folgenden Formeln auszuschließen. Wegen (183) und (247) sind die übrigbleibenden Elemente (250) normiert und paarweise orthogonal.

Die Fourierkoeffizienten von y bezüglich des Systems (250) haben die Gestalt

$$\frac{(y, \Delta_k \mathbb{E}_\lambda x)}{\sqrt{\Delta_k \varrho(\lambda)}} = \frac{\Delta_k \varphi_y(\lambda)}{\sqrt{\Delta_k \varrho(\lambda)}}.$$

Das Quadrat der Norm für die Differenz des Elements y und seiner Fourierreihe läßt sich bekanntlich [121] in der Form

$$\left\| y - \sum_{k=1}^n \frac{\Delta_k \varphi_y(\lambda)}{\Delta_k \varrho(\lambda)} \Delta_k \mathbb{E}_\lambda x \right\|^2 = \|y\|^2 - \sum_{k=1}^n \frac{|\Delta_k \varphi_y(\lambda)|^2}{\Delta_k \varrho(\lambda)} \quad (251)$$

ausdrücken, woraus sich die Besselsche Ungleichung

$$\sum_{k=1}^n \frac{|\Delta_k \varphi_y(\lambda)|^2}{\Delta_k \varrho(\lambda)} \leq \|y\|^2 \quad (252)$$

und nach Grenzübergang durch Verfeinern der Teilintervalle

$$\int_m^M \frac{|\mathrm{d}\varphi_y(\lambda)|^2}{\mathrm{d}\varrho(\lambda)} \leq \|y\|^2 \quad (253)$$

ergibt.

Satz 1. Für $y \in C_x$ gilt die Gleichung

$$\|y\|^2 = \int_m^M \frac{|\mathrm{d}\varphi_y(\lambda)|^2}{\mathrm{d}\varrho(\lambda)}. \quad (254)$$

Beweis. Liegt y in C_x , so existiert, da C_x die lineare abgeschlossene Hülle der $\mathbb{E}_\lambda x$ ist, für ein beliebig vorgegebenes ε eine endliche Menge von Elementen $\mathbb{E}_{\lambda_s} x$ ($s = 1, 2, \dots, p$) aus den $\mathbb{E}_\lambda x$ und Zahlen c_s derart, daß

$$y = \sum_{s=1}^p c_s \mathbb{E}_{\lambda_s} x + z \quad \text{und} \quad \|z\| \leq \varepsilon \quad (255)$$

gilt. Wir nehmen die Punkte λ_s ($s = 1, 2, \dots, p$) als Teilpunkte des Intervalls $[m, M]$, fügen noch die Punkte $\lambda_0 = m$ und $\lambda_{p+1} = M$ hinzu, wenn sie nicht schon zu den vorigen λ_s gehören, und führen die so erhaltenen Teilintervalle $[\lambda_{s-1}, \lambda_s]$, $s = 1, 2, \dots, p+1$, ein, die wir mit Δ'_s bezeichnen. Es gilt $\mathbb{E}_{\lambda_0} = 0$ und $\mathbb{E}_{\lambda_{p+1}} = E$, und wir können, wenn wir die übliche Bezeichnung $\Delta'_s \mathbb{E}_\lambda = \mathbb{E}_{\lambda_s} - \mathbb{E}_{\lambda_{s-1}}$ einführen,

$$\begin{aligned} \mathbb{E}_{\lambda_1} &= \Delta'_1 \mathbb{E}_\lambda, \\ \mathbb{E}_{\lambda_2} &= \Delta'_1 \mathbb{E}_\lambda + \Delta'_2 \mathbb{E}_\lambda, \\ \mathbb{E}_{\lambda_3} &= \Delta'_1 \mathbb{E}_\lambda + \Delta'_2 \mathbb{E}_\lambda + \Delta'_3 \mathbb{E}_\lambda, \\ &\dots \end{aligned}$$

schreiben. Damit kann die in (255) auftretende Linearkombination der $\mathbb{E}_{\lambda_s} x$ als Linearkombination der $\Delta'_s \mathbb{E}_\lambda x$ dargestellt werden, und die Gleichung (255) erhält die Gestalt

$$y = \sum_{s=1}^{p+1} b_s \Delta'_s \mathbb{E}_\lambda x + z \quad \text{und} \quad \|z\| \leq \varepsilon;$$

dabei sind die b_s neue Koeffizienten. Anders ausgedrückt gilt

$$\left\| y - \sum_{s=1}^{p+1} b_s \Delta'_s \mathbb{E}_\lambda x \right\| \leq \varepsilon. \quad (255_1)$$

Diese Ungleichung gilt erst recht, wenn man an Stelle der Summe die Fourierreihe von y bezüglich des normierten Orthogonalsystems

$$\frac{\Delta'_s \mathfrak{E}_\lambda x}{\sqrt{\Delta'_s \varrho(\lambda)}} \quad (s = 1, 2, \dots, p+1)$$

nimmt [121]. Dann läßt sich die Ungleichung (255₁) wegen (251) in der Gestalt

$$\|y\|^2 - \sum_{s=1}^{p+1} \frac{|\Delta'_s \varphi_y(\lambda)|^2}{\Delta'_s \varrho(\lambda)} \leq \varepsilon^2 \quad \text{oder} \quad \sum_{s=1}^{p+1} \frac{|\Delta'_s \varphi_y(\lambda)|^2}{\Delta'_s \varrho(\lambda)} \geq \|y\|^2 - \varepsilon^2$$

schreiben. Wenn wir diese Ungleichung mit (253) vergleichen und berücksichtigen, daß ε willkürlich ist, sehen wir, daß die obere Grenze der in (252) auftretenden Summen gleich $\|y\|^2$ ist, d. h., daß (254) gilt. Durch Anwendung dieser Gleichung und der Beziehung

$$(y, z) = \frac{1}{2} \|y + z\|^2 + \frac{i}{2} \|y + iz\|^2 - \frac{1+i}{2} [\|y\|^2 + \|z\|^2]$$

erhalten wir für y und z aus C_x die allgemeinere Formel

$$(y, z) = \int_m^M \frac{d\varphi_y(\lambda) \overline{d\varphi_z(\lambda)}}{d\varrho(\lambda)} \quad (256)$$

mit

$$\varphi_z(\lambda) = (z, \mathfrak{E}_\lambda x). \quad (257)$$

Für die Aufstellung analoger Gleichungen für bilineare Funktionale beweisen wir den folgenden Satz.

Satz 2. *Gehört y zu C_x , so liegen auch $\mathfrak{E}_\lambda y$ und Ay in C_x .*

Beweis. Aus $y \in C_x$ folgt, daß y entweder eine endliche Linearkombination der Elemente $\mathfrak{E}_{\lambda_s} x$ ist,

$$y = \sum_{s=1}^p c_s \mathfrak{E}_{\lambda_s} x, \quad (258)$$

oder daß y Grenzwert solcher Linearkombinationen ist. Im ersten Fall ist

$$\mathfrak{E}_\lambda y = \sum_{s=1}^p c_s \mathfrak{E}_\lambda \mathfrak{E}_{\lambda_s} x.$$

Wegen (180) ist aber $\mathfrak{E}_\lambda \mathfrak{E}_{\lambda_s} = \mathfrak{E}_\lambda$ für $\lambda \leq \lambda_s$ und $\mathfrak{E}_\lambda \mathfrak{E}_{\lambda_s} = \mathfrak{E}_{\lambda_s}$ für $\lambda \geq \lambda_s$, d. h., $\mathfrak{E}_\lambda y$ ist ebenfalls eine endliche Linearkombination von Elementen aus C_x und gehört damit ebenfalls zu C_x . Ist y Grenzwert endlicher Linearkombinationen von Elementen aus C_x ,

$$y = \lim_{n \rightarrow \infty} \sum_{s=1}^{p_n} c_s^{(n)} \mathfrak{E}_{\lambda_s^{(n)}} x,$$

so gilt

$$\mathfrak{E}_\lambda y = \lim_{n \rightarrow \infty} \sum_{s=1}^{p_n} c_s^{(n)} \mathfrak{E}_\lambda \mathfrak{E}_{\lambda_s^{(n)}} x,$$

d. h., $\mathfrak{E}_\lambda y$ ist ebenfalls Grenzwert endlicher Linearkombinationen aus C_x ; also ist auch in diesem Fall $\mathfrak{E}_\lambda y \in C_x$. Das Element Ay ist wegen (205) Grenzwert endlicher Linearkombinationen von Elementen $\mathfrak{E}_{\lambda_s} y$. Nach dem Bewiesenen gehört jedes

$\mathfrak{S}_\lambda y$ auch zu C_x ; damit liegen jede endliche Linearkombination von Elementen $\mathfrak{S}_\lambda y$ und die Grenzwerte solcher Linearkombination ebenfalls in C_x , da C_x ein Unterraum ist. Also ist $Ay \in C_x$, womit der Satz bewiesen ist.

Wir können also in Gleichung (256) y durch Ay ersetzen. An Stelle von $\varphi_y(\lambda)$ steht dann

$$\varphi_{Ay}(\lambda) = (Ay, \mathfrak{S}_\lambda x) = \int_m^M \mu d_\mu(\mathfrak{S}_\mu y, \mathfrak{S}_\lambda x) = \int_m^M \mu d_\mu(y, \mathfrak{S}_\mu \mathfrak{S}_\lambda x),$$

und unter Berücksichtigung von (180) erhalten wir

$$\varphi_{Ay}(\lambda) = \int_m^\lambda \mu d_\mu \varphi_y(\mu).$$

Aus Gleichung (256) ergibt sich somit

$$(Ay, z) = \int_m^M \frac{d \left(\int_m^\lambda \mu d_\mu \varphi_y(\mu) \right) \overline{d\varphi_z(\lambda)}}{d\varrho(\lambda)}$$

oder auf Grund der Eigenschaften der Hellingerschen Integrale [83]

$$(Ay, z) = \int_m^M \lambda \frac{d\varphi_y(\lambda) \overline{d\varphi_z(\lambda)}}{d\varrho(\lambda)}. \quad (259)$$

Ferner strebt wegen (254) bei unbegrenzter Verfeinerung der Teilintervalle Δ_k der Ausdruck auf der rechten Seite von (251) gegen 0; folglich gilt

$$\sum_{k=1}^n \frac{\Delta_k \varphi_y(\lambda)}{\Delta_k \varrho(\lambda)} \Delta_k \mathfrak{S}_\lambda x \Rightarrow y.$$

Hier sind die Summanden Elemente von H , und der Grenzwert dieser Summe läßt sich natürlich entsprechend dem gewöhnlichen Fall in der Gestalt eines Hellingerschen Integrals schreiben:

$$y = \int_m^M \frac{d\varphi_y(\lambda)}{d\varrho(\lambda)} d\mathfrak{S}_\lambda x \quad (y \in C_x). \quad (260)$$

Wenden wir diese Gleichung auf Ay statt auf y an, so erhalten wir

$$Ay = \int_m^M \lambda \frac{d\varphi_y(\lambda)}{d\varrho(\lambda)} d\mathfrak{S}_\lambda x \quad (261)$$

oder, als Grenzwert der Summe,

$$\sum_{k=1}^n \lambda_k \frac{\Delta_k \varphi_y(\lambda)}{\Delta_k \varrho(\lambda)} \Delta_k \mathfrak{S}_\lambda x \Rightarrow Ay. \quad (262)$$

Wir weisen darauf hin, daß wir durch Benutzung analoger Summen und des analogen Grenzübergangs in H das Hellingersche Integral allgemein für Elemente aus H definieren können. Wir wenden nun die Gleichungen (256) und (260) nicht auf das Element Ay , sondern auf das ebenfalls in C_x liegende Element $\mathfrak{S}_\mu y$ an; dabei ist μ eine feste Zahl aus dem Intervall $[m, M]$. Es gilt

$$\varphi_{\mathfrak{S}_\mu y}(\lambda) = (\mathfrak{S}_\mu y, \mathfrak{S}_\lambda x) = (y, \mathfrak{S}_\mu \mathfrak{S}_\lambda x) = \begin{cases} (y, \mathfrak{S}_\lambda x) & \text{für } \lambda \leq \mu, \\ (y, \mathfrak{S}_\mu x) & \text{für } \lambda \geq \mu, \end{cases}$$

also

$$\varphi_{\mu} y(\lambda) = \begin{cases} \varphi_y(\lambda) & \text{für } \lambda \leq \mu, \\ \varphi_y(\mu) & \text{für } \lambda \geq \mu. \end{cases}$$

Die oben erwähnten Gleichungen ergeben also unmittelbar

$$(\mathfrak{E}_{\mu} y, z) = \int_m^{\mu} \frac{d\varphi_y(\lambda) \overline{d\varphi_z(\lambda)}}{d\varrho(\lambda)} \quad (263)$$

und

$$\mathfrak{E}_{\mu} y = \int_m^{\mu} \frac{d\varphi_y(\lambda)}{d\varrho(\lambda)} d\mathfrak{E}_{\lambda} x. \quad (264)$$

Die Gleichung (256) entspricht der verallgemeinerten Parsevalschen Gleichung, die Beziehungen (259) und (261) entsprechen den Gleichungen (241) und (240). Die Formeln aus [147] wurden unter der Voraussetzung hergeleitet, daß y und z in C_x liegen. Man sagt, der selbstadjungierte Operator hat ein *einfaches kontinuierliches Spektrum*, wenn ein Element x aus H existiert, für das C_x mit H übereinstimmt. Wenn das der Fall ist und wenn in den erwähnten Gleichungen für x eben dieses Element genommen wird, so gelten sie für beliebige Elemente y und z aus H .

148. Invariante Unterräume. Wir untersuchen nun das nicht einfache kontinuierliche Spektrum und das gemischte Spektrum, d. h. den Fall, daß Eigenwerte existieren, aber kein vollständiges System bilden. Dazu führen wir erst einen neuen Begriff ein und beweisen einige Tatsachen.

Definition. Der Unterraum L heißt *invarianter Unterraum* für den Operator A , wenn folgende Bedingung erfüllt ist: Mit $x \in L$ ist auch $Ax \in L$. Man sagt in diesem Fall auch, A wird durch L *reduziert*.

Die Bedeutung dieser Definition liegt in folgendem. Wenn der Operator A durch L reduziert wird, kann er gesondert betrachtet werden als ein Operator, der in L definiert ist, wobei L ein endlichdimensionaler Raum oder ein Hilbertscher Raum sein kann. Anders ausgedrückt, der in ganz H definierte Operator A induziert einen in L definierten Operator, der für die Elemente von L mit A übereinstimmt. Die Untersuchung des Operators A in verschiedenen invarianten Unterräumen erleichtert dessen Studium. Wir bemerken, daß der Operator A , wenn er in H selbstadjungiert ist, offenbar in jedem für ihn invarianten Unterraum selbstadjungiert bleibt. Im folgenden untersuchen wir invariante Unterräume nur für selbstadjungierte Operatoren.

Satz 1. *Reduziert der Unterraum L einen selbstadjungierten Operator A , so reduziert auch der komplementäre Unterraum $H \ominus L$ den Operator A . Der Unterraum L reduziert dann und nur dann einen selbstadjungierten Operator A , wenn der Projektionsoperator P_L mit A vertauschbar ist, wenn also*

$$P_L A = A P_L \quad (265)$$

gilt.

Beweis. Wenn L den Operator A reduziert, ist mit $z \in L$ auch $Az \in L$. Es ist zu zeigen, daß mit $x \perp L$ auch $Ax \perp L$ gilt. Es sei z ein beliebiges Element aus L . Dann ist $Az \in L$, und es gilt $(Ax, z) = (x, Az) = 0$, womit der erste Teil des Satzes bewiesen ist. Wir kommen nun zum Beweis der Bedingung (265). Dazu gehen wir

von der offenbar gültigen Gleichung $Ax = AP_Lx + A(E - P_L)x$ aus. Wenn A durch L reduziert wird, ist $A(P_Lx) \in L$ und nach dem soeben Bewiesenen $A[(E - P_L)x] \in H \ominus L$; folglich ist der erste Summand auf der rechten Seite die Projektion von Ax in L , d. h., für jedes x gilt $P_LAx = AP_Lx$, womit die Notwendigkeit von (265) bewiesen ist. Nun sei (265) erfüllt, und x liege in L . Dann ist $Ax = A(P_Lx) = P_L(Ax)$, also $Ax \in L$. Damit ist der Satz vollständig bewiesen.

Durch Anwendung von Satz 2 aus [143] erhalten wir aus Satz 1 die

Folgerung. Der Operator A wird dann und nur dann von dem Unterraum L reduziert, wenn $\mathcal{E}_\lambda$ für jedes λ von L reduziert wird.

Satz 2. *Wenn der Operator A durch die paarweise orthogonalen Unterräume L_k reduziert wird, so wird er auch durch ihre orthogonale Summe*

$$L = L_1 \oplus L_2 \oplus L_3 \oplus \dots$$

reduziert.

Beweis. Nach Voraussetzung des Satzes ist A mit allen P_{L_k} vertauschbar; demzufolge kann man A auch mit der Summe $P_L = P_{L_1} + P_{L_2} + P_{L_3} + \dots$ vertauschen, womit der Satz bewiesen ist. Dieser Satz gilt auch, wenn A nicht selbstadjungiert ist.

Wir erwähnen nun einige einfache Tatsachen, die mit dem Begriff des invarianten Unterraumes eines selbstadjungierten Operators zusammenhängen. Ist der Projektionsoperator P_M mit dem Projektionsoperator P_L vertauschbar, so wird P_M durch L reduziert, und der Operator P_M induziert in L den Projektionsoperator P_{LM} in den Unterraum LM . Ferner möge L den Operator A und damit dessen Spektralschar $\mathcal{E}_\lambda$ reduzieren. Wir bezeichnen mit $A^{(1)}$ und $\mathcal{E}_\lambda^{(1)}$ diejenigen Operatoren, die in L durch die Operatoren A bzw. $\mathcal{E}_\lambda$ induziert werden. Wie man leicht sieht, ist $\mathcal{E}_\lambda^{(1)}$ die Zerlegung der Einheit für $A^{(1)}$. Ist in (205) $x \in L$, so können wir A durch $A^{(1)}$ und $\mathcal{E}_\lambda$ durch $\mathcal{E}_\lambda^{(1)}$ ersetzen; $\mathcal{E}_\lambda^{(1)}$ ist die Spektralschar des in L definierten Operators $A^{(1)}$. Wegen (223) reduziert L die Resolvente R_l des Operators A , wobei R_l in L die Resolvente des Operators $A^{(1)}$ induziert. Sind $A^{(1)}$, $A^{(2)}$, $\mathcal{E}_\lambda^{(1)}$ und $\mathcal{E}_\lambda^{(2)}$ die Operatoren, die durch den selbstadjungierten Operator A bzw. durch $\mathcal{E}_\lambda$ in dem invarianten Unterraum L bzw. in dem komplementären Unterraum $H \ominus L$ induziert wurden, und sind $x = x_1 + x_2$ und $y = y_1 + y_2$ die Zerlegungen von x bzw. y in L bzw. $H \ominus L$, so gelten offenbar die Gleichungen

$$\left. \begin{aligned} Ax &= A^{(1)}x_1 + A^{(2)}x_2, & \mathcal{E}_\lambda x &= \mathcal{E}_\lambda^{(1)}x_1 + \mathcal{E}_\lambda^{(2)}x_2, \\ (Ax, y) &= (A^{(1)}x_1, y) + (A^{(2)}x_2, y). \end{aligned} \right\} \quad (266)$$

Analoge Gleichungen gelten bei der Zerlegung von H in endlich oder abzählbar unendlich viele paarweise orthogonale, den Operator A reduzierende Unterräume. Der gesamte Raum H und der Nullunterraum, d. h. der Unterraum, der nur das Nullelement enthält, bilden die trivialen invarianten Unterräume für einen beliebigen Operator. Wenn ein Operator keine anderen invarianten Unterräume hat, so heißt er *irreduzibel*. Jeder Unterraum von Eigenelementen, die einem Eigenwert λ_0 des Operators A zugeordnet sind, ist ein invarianter Unterraum für A , und in diesem Unterraum bedeutet die Anwendung des Operators A die Multiplikation der Elemente mit der Zahl λ_0 . Ist x_0 ein beliebiges, dem Eigenwert λ_0 zugeordnetes Eigenelement und a eine beliebige komplexe Zahl, so ist die Gesamtheit der Elemente der Gestalt ax_0 ebenfalls ein Unterraum, der A reduziert. Wenn L_k ($k = 1, 2, \dots$) alle Unterräume der Eigenelemente des selbstadjungierten Operators A sind, wird A durch die orthogonale Summe H' der L_k reduziert. Sind A' und $\mathcal{E}'_\lambda$

die in H' durch die Operatoren A bzw. $\mathfrak{E}_\lambda$ induzierten Operatoren, so gilt nach (234) und (236)

$$\begin{aligned} A'x &= \sum_k A(\mathfrak{E}_{\lambda_k} - \mathfrak{E}_{\lambda_{k-0}}) x, \\ \mathfrak{E}'_\lambda x &= \sum_k \mathfrak{E}_\lambda(\mathfrak{E}_{\lambda_k} - \mathfrak{E}_{\lambda_{k-0}}) x = \sum_{\lambda_k \leq \lambda} (\mathfrak{E}_{\lambda_k} - \mathfrak{E}_{\lambda_{k-0}}) x, \end{aligned} \quad (267)$$

d. h., $\mathfrak{E}'_\lambda$ bezeichnet die Summe der Sprungstellen von $\mathfrak{E}_\lambda$ in den Punkten $\lambda_k \leq \lambda$. In dem zu H' komplementären Unterraum H'' erweist sich der Operator $\mathfrak{E}'_\lambda$, der durch $\mathfrak{E}_\lambda$ in diesem Unterraum induziert worden ist, als Projektionsoperator in den Unterraum $H''M_\lambda$, wobei M_λ der dem Projektionsoperator $\mathfrak{E}_\lambda$ zugeordnete Unterraum ist. Jedes Element des Raumes H'' ist orthogonal zu jedem L_k , d. h., es gilt $(\mathfrak{E}_{\lambda_k} - \mathfrak{E}_{\lambda_{k-0}})x = 0$ für $x \in H''$, und $\mathfrak{E}'_\lambda$ läßt sich für jedes x aus H'' in Gestalt der Differenz

$$\mathfrak{E}'_\lambda = \mathfrak{E}_\lambda - \sum_{\lambda_k \leq \lambda} (\mathfrak{E}_{\lambda_k} - \mathfrak{E}_{\lambda_{k-0}})$$

darstellen; daraus folgt, daß $\mathfrak{E}'_\lambda$ für alle λ stetig ist. Ist also das Spektrum von A kein reines Punktspektrum, so enthält der Unterraum H'' vom Nullelement verschiedene Elemente, die Spektralschar ist in ihm stetig, und der Operator hat überhaupt keine Eigenwerte in H'' . In H' bilden die Eigenelemente ein vollständiges System, und der Operator A' hat in H' ein reines Punktspektrum.

149. Der allgemeine Fall eines kontinuierlichen Spektrums. Liegt das Element y in dem in [147] konstruierten Unterraum C_x , so ist, wie wir gesehen haben, auch $Ay \in C_x$, d. h., C_x reduziert A . Hat der Operator A kein Punktspektrum, d. h., ist seine Spektralschar $\mathfrak{E}_\lambda$ für alle Werte von λ stetig, und ist sein kontinuierliches Spektrum nicht einfach, so kann der Raum H , wie wir sogleich zeigen werden, als orthogonale Summe von Unterräumen der Gestalt C_x dargestellt werden. Der in jedem dieser Unterräume durch den Operator A induzierte Operator hat ein einfaches kontinuierliches Spektrum, und für die Elemente eines solchen Unterraumes gelten die in [147] hergeleiteten Formeln. Entsprechende Formeln für beliebige Elemente aus H ergeben sich mit Hilfe der Zerlegung eines solchen Elements nach den obengenannten Unterräumen; man erhält diese Formeln wegen (266) durch Addition der entsprechenden Formeln für die speziellen Unterräume. Wir geben nun die erwähnte Darstellung von H als orthogonale Summe von Unterräumen der Gestalt C_x an, wobei wir den Raum H als separabel voraussetzen. Wir benutzen dazu ein beliebiges vollständiges normiertes Orthogonalsystem $u_1, u_2, u_3, \dots$, nehmen $y_1 = u_1$ an und bilden C_{y_1} . Das Element u_2 können wir in der Gestalt $u_2 = v_2 + y_2$ darstellen, wobei $v_2 \in C_{y_1}$ und $y_2 \perp C_{y_1}$ ist. Für $y_2 \neq 0$ bilden wir C_{y_2} . Wir zeigen, daß $C_{y_2} \perp C_{y_1}$ ist. Für jedes λ gilt $(y_2, \mathfrak{E}_\lambda y_1) = 0$, denn es ist $\mathfrak{E}_\lambda y_1 \in C_{y_1}$. Weiterhin erhalten wir $(\mathfrak{E}_\mu y_2, \mathfrak{E}_\lambda y_1) = (y_2, \mathfrak{E}_\mu \mathfrak{E}_\lambda y_1) = 0$. Daraus folgt, daß jede Linearkombination von Elementen $\mathfrak{E}_\mu y_2$ zu jeder Linearkombination der $\mathfrak{E}_\lambda y_1$ orthogonal ist, und nach Grenzübergang zeigt sich, daß jedes Element aus C_{y_2} zu jedem Element aus C_{y_1} orthogonal ist. Wir betrachten nun das Element u_3 und stellen es in der Form $u_3 = v_3 + y_3$ dar, wobei $v_3 \in C_{y_1} \oplus C_{y_2}$ und $y_3 \perp C_{y_1} \oplus C_{y_2}$ ist. Wir bilden C_{y_3} und beweisen wie oben, daß $C_{y_3} \perp C_{y_1}$ und $C_{y_3} \perp C_{y_2}$ ist, usw. Auf diese Weise erhalten wir endlich oder abzählbar unendlich viele paarweise orthogonale Unterräume C_{y_k} , und da jedes Element x aus H nach den (ein vollständiges System bildenden) Elementen u_k zerlegt wird, ergibt die orthogonale Summe der Unterräume C_{y_k} den gesamten Raum H :

$$H = C_{y_1} \oplus C_{y_2} \oplus C_{y_3} \oplus \dots \quad (268)$$

In jedem der Unterräume C_{y_k} gelten die in [147] aufgeführten Formeln. Wir erhalten daher für beliebige Elemente y und z aus H

$$(y, z) = \sum_k \int_m^M \frac{d(y, \mathbb{S}_\lambda y_k) d(\mathbb{S}_\lambda y_k, z)}{d\rho_k(\lambda)}, \quad (269)$$

$$(\mathbb{S}_\mu y, z) = \sum_k \int_m^\mu \frac{d(y, \mathbb{S}_\lambda y_k) d(\mathbb{S}_\lambda y_k, z)}{d\rho_k(\lambda)}, \quad (270)$$

$$(Ay, z) = \sum_k \int_m^M \lambda \frac{d(y, \mathbb{S}_\lambda y_k) d(\mathbb{S}_\lambda y_k, z)}{d\rho_k(\lambda)}, \quad (271)$$

$$y = \sum_k \int_m^M \frac{d(y, \mathbb{S}_\lambda y_k)}{d\rho_k(\lambda)} d\mathbb{S}_\lambda y_k, \quad Ay = \sum_k \int_m^M \lambda \frac{d(y, \mathbb{S}_\lambda y_k)}{d\rho_k(\lambda)} d\mathbb{S}_\lambda y_k, \quad (272)$$

$$\mathbb{S}_\mu y = \sum_k \int_m^\mu \frac{d(y, \mathbb{S}_\lambda y_k)}{d\rho_k(\lambda)} d\mathbb{S}_\lambda y_k; \quad (273)$$

dabei ist $\rho_k(\lambda) = \|\mathbb{S}_\lambda y_k\|^2$, und die Summen können endlich oder auch unendlich sein. Im letzten Fall hat man die Konvergenz der Reihen in den Gleichungen (272) und (273), welche Elemente von H enthalten, als Konvergenz in H zu verstehen.

Das dargestellte Verfahren zur Bildung der Unterräume C_{y_k} läßt sich in Form von Gleichungen aufschreiben, die wir nun einführen. Ist v ein beliebiges Element von H , so wird seine Projektion in C_{y_k} durch den zugehörigen Summanden der ersten Gleichung (272), d. h. durch

$$v_k = \int_m^M \frac{d(v, \mathbb{S}_\lambda y_k)}{d\rho_k(\lambda)} d\mathbb{S}_\lambda y_k \quad (274)$$

bestimmt, und für $\mathbb{S}_\lambda v_k$ erhalten wir, wenn wir wie immer (180) berücksichtigen,

$$\mathbb{S}_\lambda v_k = \int_m^\lambda \frac{d(v, \mathbb{S}_\mu y_k)}{d\rho_k(\mu)} d\mathbb{S}_\mu y_k.$$

Aus diesen Bemerkungen ergeben sich die folgenden Gleichungen, die dem obigen Bildungsprozeß der Unterräume C_{y_k} entsprechen:

$$\mathbb{S}_\lambda y_1 = \mathbb{S}_\lambda u_1, \quad \mathbb{S}_\lambda y_k = \mathbb{S}_\lambda u_k - \sum_{s=1}^{k-1} \int_m^\lambda \frac{d(u_k, \mathbb{S}_\mu y_s)}{d\rho_s(\mu)} d\mathbb{S}_\mu y_s. \quad (275)$$

Bei verschiedener Wahl des Ausgangssystems der u_k ergeben sich im allgemeinen verschiedene Unterräume C_{y_k} . Es kann vorkommen, daß auch die Anzahl dieser Unterräume verschieden ist. Zweckmäßig ist es, diese Unterräume von vornherein möglichst zu erweitern.

Man kann zeigen, daß eine Bildung der C_{y_k} möglich ist, bei der die folgende Bedingung erfüllt ist: Jede Menge vom Maß 0 bezüglich $\rho_p(\lambda)$ ist auch bezüglich $\rho_{p+1}(\lambda)$, $\rho_{p+2}(\lambda)$, ... eine Menge vom Maß 0. Diese Bedingung ist nach den Ergebnissen von [74] gleichbedeutend mit der folgenden: Jedes $\rho_p(\lambda)$ wird durch die vorher-

gehenden $\varrho_k(\lambda)$ mit Hilfe der Gleichung

$$\varrho_p(\lambda) = \int_m^\lambda \varphi_p^{(k)}(\lambda) d\varrho_k(\lambda) \quad (k=1, 2, \dots, p-1)$$

ausgedrückt; dabei hat man das Integral im Sinne eines Lebesgue-Stieltjesschen Integrals zu verstehen, und die $\varphi_p^{(k)}(\lambda)$ sind nichtnegative, bezüglich $\varrho_k(\lambda)$ meßbare und summierbare Funktionen. Ist diese Bedingung erfüllt, nennen wir die Zerlegung der Spektralschar *normal*. Man kann beweisen, daß in verschiedenen normalen Zerlegungen die Anzahl der Unterräume gleich ist. Auf diese Frage werden wir später noch zurückkommen.

150. Das gemischte Spektrum. Wie wir bereits erwähnt haben, hat der selbst-adjungierte Operator A ein gemischtes Spektrum, wenn es zwar Eigenelemente des Operators A gibt, aber das normierte Orthogonalsystem

$$x_1, x_2, x_3, \dots \quad (276)$$

der Eigenelemente in H nicht vollständig ist. Wie oben seien die L_k die Unterräume der dem Eigenwert λ_k zugeordneten Eigenelemente. Wie wir in [148] gesehen haben, hat der Operator A in dem invarianten Unterraum

$$H' = L_1 \oplus L_2 \oplus L_3 \oplus \dots$$

ein reines Punktspektrum mit dem λ_k als Eigenwerten und den L_k als Unterräumen der Eigenelemente. Die Elemente (276) bilden dann ein aus den Eigenelementen von A bestehendes vollständiges System in H' . In dem komplementären Unterraum H'' hat der Operator A nur ein rein kontinuierliches Spektrum. Unter Berücksichtigung der Gleichungen in [149] und [146] erhalten wir die folgenden allgemeinen Gleichungen, in denen die Summen von dem Punktspektrum in H' und die Integrale von dem kontinuierlichen Spektrum in H'' stammen:

$$(y, z) = \sum_k a_k \bar{b}_k + \sum_k \int_m^M \frac{d(y, \mathcal{E}_\lambda y_k) d(\mathcal{E}_\lambda y_k, z)}{d\varrho_k(\lambda)}, \quad (277)$$

$$(Ay, z) = \sum_k \mu_k a_k \bar{b}_k + \sum_k \int_m^M \lambda \frac{d(y, \mathcal{E}_\lambda y_k) d(\mathcal{E}_\lambda y_k, z)}{d\varrho_k(\lambda)}, \quad (278)$$

$$y = \sum_k a_k x_k + \sum_k \int_m^M \frac{d(y, \mathcal{E}_\lambda y_k)}{d\varrho_k(\lambda)} d\mathcal{E}_\lambda y_k, \quad (279)$$

$$Ay = \sum_k \mu_k a_k x_k + \sum_k \int_m^M \lambda \frac{d(y, \mathcal{E}_\lambda y_k)}{d\varrho_k(\lambda)} d\mathcal{E}_\lambda y_k; \quad (280)$$

a_k und b_k sind dabei Fourierkoeffizienten der Elemente y und z bezüglich des Systems (276), μ_k sind die den Eigenelementen x_k zugeordneten Eigenwerte von A .

Führen wir die obige Zerlegung eines Operators in einen Operator mit reinem Punktspektrum in H' und einen Operator mit kontinuierlichem Spektrum in H'' durch, so können wir eine Klassifikation der Punkte des Spektrums vornehmen.

Definition. Der Punkt λ_0 gehört zum *Punktspektrum*, wenn er ein Eigenwert von A ist. Er gehört zum *Grenzspektrum*, wenn er Häufungspunkt des Punktspektrums ist, d. h., wenn sich in einer beliebigen ε -Umgebung von ihm Eigenwerte befinden, er selbst aber kein Eigenwert ist. Der Punkt λ_0 gehört zum *kontinuierlichen*

Spektrum, wenn er zum Spektrum des durch A in H'' induzierten Operators A'' gehört, d. h., wenn in keinem, λ_0 im Innern enthaltenden Intervall die Spektralschar $\mathfrak{E}_\lambda''$ des Operators A'' konstant ist.

Jeder Punkt des Spektrums von A gehört zu mindestens einer der drei Kategorien; es kann aber z. B. auch vorkommen, daß der Punkt λ_0 gleichzeitig zu allen drei Kategorien gehört. Zuweilen wird noch der Begriff des Verdichtungspunktes eines Spektrums eingeführt, und zwar heißt der Punkt λ_0 *Verdichtungspunkt des Spektrums*, wenn er entweder ein Eigenwert von unendlichem Rang, ein Element des Grenzspektrums oder ein Element des kontinuierlichen Spektrums ist.

151. Differentiallösungen. Wir betrachten einen Operator mit rein kontinuierlichem Spektrum. Die Elemente $\mathfrak{E}_\lambda x$ genügen bei beliebiger Wahl von x einer gewissen Gleichung, die der Gleichung $Ax = \lambda x$ im Fall eines Punktspektrums analog ist. Es sei $A = [\lambda_1, \lambda_2]$ ein beliebiges Intervall. Wenn wir die Eigenschaft (180) verwenden, so erhalten wir analog zu (213) die Beziehung

$$\int_m^M \lambda d\mathfrak{E}_\lambda \cdot A\mathfrak{E}_\lambda = \int_{\lambda_1}^{\lambda_2} \lambda d\mathfrak{E}_\lambda.$$

Wir führen das Element $x(\lambda) = \mathfrak{E}_\lambda x$ ein, das im Intervall $[m, M]$ stetig von dem Parameter λ abhängt, so daß $\|x(\lambda_0) - x(\lambda)\| \rightarrow 0$ für $\lambda \rightarrow \lambda_0$ (λ_0 beliebig aus $[m, M]$) gilt. Aus der letzten Gleichung erkennen wir, daß $x(\lambda)$ der Beziehung

$$A[Ax(\lambda)] = \int_{\lambda} \lambda dx(\lambda) \quad (281)$$

genügt; $x(\lambda)$ nennt man dabei eine *Differentiallösung* der Gleichung $Ax = \lambda x$. Die in [149] gebildeten Vektoren $y_k(\lambda) = \mathfrak{E}_\lambda y_k$ stellen demnach Differentiallösungen dar. Für verschiedene Werte von k befinden sie sich in den orthogonalen Unterräumen C_{y_k} ; daher gilt für beliebige Intervalle Δ_1 und Δ_2

$$(\Delta_1 y_p(\lambda), \Delta_2 y_q(\lambda)) = 0 \quad (p \neq q). \quad (282)$$

Haben Δ_1 und Δ_2 keine gemeinsamen inneren Punkte, so gilt wegen (180)

$$(\Delta_1 y_p(\lambda), \Delta_2 y_p(\lambda)) = 0. \quad (283)$$

Enthalten Δ_1 und Δ_2 ein gemeinsames Intervall $\Delta_{1,2}$, so gilt wegen (180)

$$(\Delta_1 y_p(\lambda), \Delta_2 y_p(\lambda)) = \|\Delta_{1,2} y_p(\lambda)\|^2. \quad (284)$$

Wir führen nun den Begriff des *abgeschlossenen Systems von Differentiallösungen* ein. Ein beliebiges System von im Sinne von (282) orthogonalen Differentiallösungen $y_p(\lambda)$ heißt abgeschlossen, wenn das einzige Element x , das zu allen $y_p(\lambda)$ orthogonal ist, das also für jedes p und jedes λ die Bedingung

$$(y_p(\lambda), x) = 0 \quad (285)$$

erfüllt, das Nullelement ist. Es läßt sich leicht zeigen, daß die soeben gebildeten Lösungen $y_p(\lambda) = \mathfrak{E}_\lambda y_p$ ein abgeschlossenes System bilden. Aus (285) folgt nämlich, daß das Element x orthogonal zu dem Unterraum C_{y_p} ist, der die lineare abgeschlossene Hülle der $y_p(\lambda) = \mathfrak{E}_\lambda y_p$ darstellt, und zwar gilt das für jedes p . Nun ist aber die orthogonale Summe der C_{y_p} der gesamte Raum H ; folglich ist das Element x orthogonal zum gesamten Raum H und damit wirklich das Nullelement.

Bei der Bildung der Lösungen von (281) sind wir von der Spektralschar $\mathfrak{E}_\lambda$ ausgegangen. Jetzt gehen wir von der Gleichung selbst aus. Es sei uns auf irgendeine Weise gelungen, eine Lösung $x(\lambda)$ der Gleichung (281) zu bilden. In dieser Gleichung

ist $x(\lambda)$ nur in der Differenz und im Differential enthalten. Wenn wir also von $x(\lambda)$ ein beliebiges, von λ unabhängiges Element subtrahieren, erhalten wir ebenfalls eine Lösung der Gleichung. Insbesondere ist die Differenz $x(\lambda) - x(m)$ eine Lösung; wir können daher stets $x(m) = 0$ annehmen. Wir zeigen sogleich, daß jede Lösung von (281), die im Intervall $[m, M]$ stetig von λ abhängt und der Bedingung $x(m) = 0$ genügt, notwendigerweise die Gestalt $x(\lambda) = \mathbb{E}_\lambda x$ hat. Angenommen, es sei uns auf irgendeine Weise gelungen, endlich oder unendlich viele, paarweise (im Sinne von (282)) orthogonale Lösungen $y_p(\lambda)$ von (281) zu bilden. Jede von ihnen hat, wie wir eben gesehen haben, die Gestalt $y_p(\lambda) = \mathbb{E}_\lambda y_p$, wobei y_p ein gewisses Element aus H ist. Die lineare abgeschlossene Hülle jedes der $y_p(\lambda)$ ist ein gewisser Unterraum C_{y_p} ; dabei sind diese Unterräume wegen (282) paarweise orthogonal. Die Abgeschlossenheit der Lösungen $y_p(\lambda)$ läuft darauf hinaus, daß die orthogonale Summe der C_{y_p} der gesamte Raum H ist. Wenn diese Abgeschlossenheit besteht, können wir alle Gleichungen aus [149] aufschreiben, wenn wir in ihnen $\mathbb{E}_\lambda y_k$ durch $y_k(\lambda)$ ersetzen. Folglich können wir bei der Bildung dieser Gleichungen von jedem beliebigen abgeschlossenen Orthogonalsystem stetiger Differentiallösungen ausgehen. Die Abgeschlossenheit eines auf beliebige Weise gebildeten Systems paarweise orthogonaler Lösungen läßt sich für ein bilineares Funktional durch (271) oder, falls die Spektralschar $\mathbb{E}_\lambda$ bekannt ist, durch (273) verifizieren.

Wir erwähnen noch, daß sich für $x(\lambda) = \mathbb{E}_\lambda x$ eine von (281) verschiedene Gleichung aufstellen läßt. Unter der Annahme, daß das Intervall Δ nicht den Punkt $\lambda = 0$ enthält, gilt wegen (180)

$$\int_m^M \lambda d\mathbb{E}_\lambda \cdot \int_\Delta \frac{1}{\lambda} d\mathbb{E}_\lambda = \Delta \mathbb{E}_\lambda,$$

woraus für $x(\lambda)$ die Gleichung

$$A \left[\int_\Delta \frac{1}{\lambda} dx(\lambda) \right] = \Delta x(\lambda)$$

oder

$$\int_\Delta \frac{1}{\lambda} d[Ax(\lambda)] = \Delta x(\lambda) \quad (281_1)$$

folgt.

Wir beweisen nun die obige Behauptung.

Satz. Jede Lösung der Gleichung (281), die im Intervall $[m, M]$ stetig und für $\lambda = m$ gleich dem Nullelement ist, hat die Gestalt $x(\lambda) = \mathbb{E}_\lambda x(M)$.

Beweis. Unter der Bedingung $x(m) = 0$ ergibt die Gleichung (281)

$$\int_m^\lambda \mu dx(\mu) = Ax(\lambda),$$

wobei hier und im folgenden die Integrationsvariable mit μ bezeichnet wird. Wenn wir irgendwie zwei Zahlen μ_1 und $\mu_2 > \mu_1$ festlegen, erhalten wir

$$\int_m^\lambda \mu d(x(\mu), x(\mu_2) - x(\mu_1)) = (Ax(\lambda), x(\mu_2) - x(\mu_1)).$$

Wegen der Selbstadjungiertheit von A und wegen (281) können wir

$$(Ax(\lambda), x(\mu_2) - x(\mu_1)) = (x(\lambda), Ax(\mu_2) - Ax(\mu_1)) = \int_{\mu_1}^{\mu_2} \mu d(x(\lambda), x(\mu))$$

schreiben und kommen dadurch zu der Gleichung

$$\int_m^\lambda \mu d(x(\mu), x(\mu_2) - x(\mu_1)) = \int_{\mu_1}^{\mu_2} \mu d(x(\lambda), x(\mu)) .$$

Das rechts stehende Integral formen wir mit Hilfe partieller Integration um und wenden auf das erhaltene Integral den Mittelwertsatz an:

$$\int_m^\lambda \mu d(x(\mu), x(\mu_2) - x(\mu_1)) = \mu_2(x(\lambda), x(\mu_2)) - \mu_1(x(\lambda), x(\mu_1)) - (\mu_2 - \mu_1)(x(\lambda), x(\mu_3))$$

(μ_3 gehört zu $[\mu_1, \mu_2]$) oder

$$\begin{aligned} & \int_m^\lambda \mu d(x(\mu), x(\mu_2) - x(\mu_1)) \\ &= (\mu_2 - \mu_1)[(x(\lambda), x(\mu_2)) + \mu_1(x(\lambda), x(\mu_2) - x(\mu_1)) - (\mu_2 - \mu_1)(x(\lambda), x(\mu_3))]. \end{aligned}$$

Dieser Gleichung können wir die Gestalt

$$\int_m^\lambda (\mu - \mu_1) d \frac{(x(\mu), x(\mu_2) - x(\mu_1))}{\mu_2 - \mu_1} = (x(\lambda), x(\mu_2) - x(\mu_3))$$

geben. Führen wir die stetigen Funktionen

$$\omega(\mu) = \frac{(x(\mu), x(\mu_2) - x(\mu_1))}{\mu_2 - \mu_1}, \quad f(\lambda) = (x(\lambda), x(\mu_2) - x(\mu_3)) \quad (286)$$

ein, so erhalten wir die Gleichung

$$\int_m^\lambda (\mu - \mu_1) d\omega(\mu) = f(\lambda),$$

wobei wir $\lambda < \mu_1 < \mu_2$ annehmen und offenbar $\omega(m) = 0$ gilt. Die letzte Gleichung läßt sich leicht nach $\omega(\mu)$ auflösen. Dafür genügt es, auf der linken Seite eine partielle Integration durchzuführen und $\omega(\lambda) = \frac{f(\lambda)}{\lambda - \mu_1} + u(\lambda)$ zu setzen, wobei $u(\lambda)$ die neue gesuchte Funktion ist, die für $\lambda = m$ gleich 0 ist:

$$\omega(\lambda) = \frac{f(\lambda)}{\lambda - \mu_1} + \int_m^\lambda \frac{f(u)}{(\mu - \mu_1)^2} d\mu$$

oder mit Berücksichtigung der Bezeichnungen (286),

$$\frac{(x(\lambda), x(\mu_2) - x(\mu_1))}{\mu_2 - \mu_1} = \frac{(x(\lambda), x(\mu_2) - x(\mu_3))}{\lambda - \mu_1} + \int_m^\lambda \frac{(x(\mu), x(\mu_2) - x(\mu_3))}{(\mu - \mu_1)^2} d\mu .$$

Gilt $\mu_1 \rightarrow \mu_2$, so ist auch $\mu_3 \rightarrow \mu_2$, und der erste Summand auf der rechten Seite strebt gegen 0. Dasselbe trifft auch für das Integral zu, denn es ist

$$|(x(\mu), x(\mu_2) - x(\mu_3))| \leq c \|x(\mu_2) - x(\mu_3)\| ,$$

wenn c der größte Wert von $\|x(\mu)\|$ im Intervall $[m, M]$ ist. Wir haben den Fall betrachtet, daß μ_1 von kleineren Werten her gegen μ_2 strebt. In derselben Weise könnten wir den Fall $\lambda < \mu_2 < \mu_1$ und $\mu_1 \rightarrow \mu_2$ untersuchen. Aus der letzten Gleichung ergibt sich also

$$\frac{d}{d\mu} (x(\lambda), x(\mu)) = 0 \quad \text{für} \quad \mu > \lambda$$

und folglich

$$(x(\lambda), x(\mu)) = (x(\lambda), x(\lambda)) = \|x(\lambda)\|^2 \quad \text{für} \quad \mu > \lambda .$$

Wenden wir diese Beziehung für den Fall $\mu = M$ auf die Lösung $y(\lambda) = x(\lambda) - \mathfrak{E}_\lambda x(M)$ von (281) an, die für $\lambda = m$ und $\lambda = M$ verschwindet, so erhalten wir $\|y(\lambda)\|^2 = 0$, also $x(\lambda) = \mathfrak{E}_\lambda x(M)$. Damit ist der Satz bewiesen.

152. Die Multiplikation mit der unabhängigen Veränderlichen. Wir kommen auf die Ergebnisse in [147] zurück und betrachten den Funktionenraum $L_2^{(x)}$ der im Intervall $[m, M]$ definierten und bezüglich $\varrho(\lambda) = \|\mathfrak{E}_\lambda x\|^2$ meßbaren Funktionen $f(\lambda)$, für die

$$\int_m^M |f(\lambda)|^2 d\varrho(\lambda) < \infty \quad (287)$$

gilt. Der Raum $L_2^{(x)}$ stellt eine Realisierung des Raumes H dar. In diesem Raum ist der Operator der Multiplikation mit der unabhängigen Veränderlichen,

$$A_0[f(\lambda)] = \lambda f(\lambda), \quad (288)$$

offenbar beschränkt und selbstadjungiert, denn es gilt, wenn mit n die größere der Zahlen $|m|$ und $|M|$ bezeichnet wird,

$$\|\lambda f(\lambda)\| \leq n \|f(\lambda)\|,$$

und es ist, da λ reell ist,

$$(\lambda f(\lambda), g(\lambda)) = (f(\lambda), \lambda g(\lambda)) = \int_m^M \lambda f(\lambda) \overline{g(\lambda)} d\varrho(\lambda).$$

Wir stellen nun eine Verbindung zwischen dem Raum C_x und dem Funktionenraum $L_2^{(x)}$ her. Wegen der Existenz des Hellingerschen Integrals (249) läßt sich jedem Element y aus C_x eine Funktion $y(\lambda)$ aus $L_2^{(x)}$ zuordnen, mit der

$$\varphi_y(\lambda) = (y, \mathfrak{E}_\lambda x) = \int_m^\lambda y(\mu) d\varrho(\mu) \quad (289)$$

gilt [82]. Verschiedenen Elementen y und z aus C_x sind dabei auch verschiedene Elemente $y(\lambda)$ und $z(\lambda)$ aus $L_2^{(x)}$ zugeordnet. In der Tat, wenn die den Elementen y und z zugeordneten Funktionen $y(\lambda)$ und $z(\lambda)$ äquivalent wären, so hätten wir wegen (289) für beliebige λ die Gleichung $(y - z, \mathfrak{E}_\lambda x) = 0$. Die Differenz $y - z$ wäre damit orthogonal zu allen Linearkombinationen der $\mathfrak{E}_\lambda x$, und durch Grenzübergang erhielten wir, daß die Differenz $y - z$ zum gesamten Unterraum C_x orthogonal wäre. Nun ist aber $y - z \in C_x$, also wäre $(y - z, y - z) = \|y - z\|^2 = 0$ und somit $y = z$. Umgekehrt ist es nicht möglich, daß das in (289) auftretende Integral für zwei nicht äquivalente Funktionen aus $L_2^{(x)}$ und für beliebige Werte von λ ein und denselben Wert hat [52]. Folglich wird durch (289) eine eindeutige Zuordnung zwischen den Elementen y aus C_x und den Elementen einer gewissen linearen Mannigfaltigkeit N aus $L_2^{(x)}$ hergestellt. Wir beweisen, daß N mit $L_2^{(x)}$ übereinstimmt. Zuerst zeigen wir, daß die lineare Mannigfaltigkeit N abgeschlossen ist. Wir können der Gleichung (256) die Form

$$(y, z) = \int_m^M y(\lambda) \overline{z(\lambda)} d\varrho(\lambda) \quad (290)$$

geben [81]; dabei ist $z(\lambda)$ dasjenige Element aus $L_2^{(x)}$, das dem Element z aus C_x zugeordnet ist, d. h.

$$(z, \mathfrak{E}_\lambda x) = \int_m^\lambda z(\mu) d\varrho(\mu). \quad (291)$$

Es sei $y^{(n)}(\lambda)$ eine Folge von Elementen aus M , und $y^{(m)}$ seien die zugeordneten Elemente aus C_x . Setzen wir in (290) $y = z = y^{(n)} - y^{(m)}$, so erhalten wir

$$\|y^{(n)} - y^{(m)}\|^2 = \int_m^M |y^{(n)}(\lambda) - y^{(m)}(\lambda)|^2 d\rho(\lambda). \quad (292)$$

Wenn die $y^{(n)}(\lambda)$ im Mittel gegen ein gewisses Element $y_0(\lambda)$ aus $L_2^{(x)}$ streben, konvergiert für $n \rightarrow \infty$ und $m \rightarrow \infty$ die rechte Seite von (292) gegen 0; daher ist die Folge der Elemente $y^{(n)}$ in sich konvergent, und es existiert ein Element u , für das $y^{(n)} \rightarrow u$ gilt. Dabei liegt u in C_x , weil C_x ein Unterraum ist. Es sei $u(\lambda)$ das Element aus N , das dem Element u vermittelt (289) zugeordnet ist. Wir zeigen, daß $u(\lambda)$ dem Element $y_0(\lambda)$ äquivalent ist. Daraus folgt dann, daß $y_0(\lambda) \in N$ ist, d. h., daß N eine abgeschlossene lineare Mannigfaltigkeit ist. Aus (290) ergibt sich für $y = z = u - y^{(n)}$

$$\|u - y^{(n)}\|^2 = \int_m^M |u(\lambda) - y^{(n)}(\lambda)|^2 d\rho(\lambda);$$

darin erkennt man, daß $y^{(n)}(\lambda)$ im Mittel gegen $u(\lambda)$ strebt. Wegen der Eindeutigkeit des Grenzwertes bei Konvergenz im Mittel folgt damit, daß die Funktionen $u(\lambda)$ und $y_0(\lambda)$ äquivalent sind. Wir zeigen nun, daß die abgeschlossene lineare Mannigfaltigkeit N mit $L_2^{(x)}$ übereinstimmt. Wäre das nicht der Fall, so existierte ein nicht dem Nullelement äquivalentes Element $f_0(\lambda)$ aus $L_2^{(x)}$, das zu allen Elementen aus N orthogonal wäre. Für ein Element $y = \mathfrak{E}_\nu x$ nimmt (289) wegen $\mathfrak{E}_\nu \mathfrak{E}_\lambda = \mathfrak{E}_\nu$ im Fall $\nu \leq \lambda$ und $\mathfrak{E}_\nu \mathfrak{E}_\lambda = \mathfrak{E}_\lambda$ im Fall $\nu \geq \lambda$ die Gestalt

$$(\mathfrak{E}_\nu x, \mathfrak{E}_\lambda x) = (x, \mathfrak{E}_\nu \mathfrak{E}_\lambda x) = \|\mathfrak{E}_\nu \mathfrak{E}_\lambda x\|^2 = \int_m^\nu d\rho(\lambda) \quad \text{für } \lambda \geq \nu,$$

$$(\mathfrak{E}_\nu x, \mathfrak{E}_\lambda x) = \int_m^\lambda d\rho(\lambda) \quad \text{für } \lambda \leq \nu$$

an, d. h., die dem Element $\mathfrak{E}_\nu x$ zugeordnete Funktion $f(\lambda)$ aus N ist der Funktion äquivalent, die durch

$$f(\lambda) = \begin{cases} 1 & \text{für } \lambda < \nu, \\ 0 & \text{für } \lambda > \nu \end{cases}$$

definiert wird. Der Wert von $f(\lambda)$ für $\lambda = \nu$ ist wegen der Stetigkeit von $\rho(\lambda)$ unwesentlich. Aus der Orthogonalität von $f_0(\lambda)$ zu der soeben definierten Funktion ergibt sich für jedes ν

$$\int_m^\nu f_0(\lambda) d\rho(\lambda) = 0.$$

Daraus folgt bekanntlich [49], daß $f_0(\lambda)$ bezüglich $\rho(\lambda)$ äquivalent 0 ist und damit der Unterraum N mit $L_2^{(x)}$ übereinstimmt. Die vorhergehenden Überlegungen führen uns zu dem folgenden Satz.

Satz 1. *Durch die Gleichung (289) wird eine eindeutige Zuordnung zwischen den Elementen y aus C_x und den Elementen $y(\lambda)$ aus $L_2^{(x)}$ hergestellt.*

Wegen (290) bleibt bei dieser Zuordnung die Größe des Skalarprodukts invariant. Folglich sind auch die Normen zugeordneter Elemente gleich. Außerdem ist die Zuordnung wegen der Distributivität des Skalarprodukts $(y, \mathfrak{E}_\lambda x)$ bezüglich y und der Distributivität des in (289) auftretenden Integrals offenbar distributiv. Daraus folgt, daß der Funktionenraum $L_2^{(x)}$ bei der angegebenen Zuordnung eine Realisierung

des Hilbertschen Raumes C_x darstellt. In diesem Raum ist der Operator A definiert, für den $\mathcal{E}_\lambda$ die Spektralschar darstellt.

Wir beweisen nun den folgenden Satz.

Satz 2. *Der Ersetzung von y durch Ay in C_x entspricht die Multiplikation von $y(\mu)$ mit μ in $L_2^{(x)}$, d. h., dem Operator A in C_x ist der Operator der Multiplikation mit der unabhängigen Veränderlichen (288) in $L_2^{(x)}$ zugeordnet.*

Beweis. Durch Anwendung der Gleichungen (206) und (289) und der Eigenschaft des Lebesgue-Stieltjesschen Integrals aus [75] können wir, wenn wir $y \in C_x$ annehmen,

$$\begin{aligned}(Ay, \mathcal{E}_\lambda x) &= \int_m^\lambda \mu d_\mu (\mathcal{E}_\mu y, \mathcal{E}_\lambda x) = \int_m^\lambda \mu d_\mu (y, \mathcal{E}_\mu x) \\ &= \int_m^\lambda \mu d_\mu \left[\int_m^\mu y(\nu) d\rho(\nu) \right] = \int_m^\lambda \mu y(\mu) d\rho(\mu),\end{aligned}$$

also

$$(Ay, \mathcal{E}_\lambda x) = \int_m^\lambda \mu y(\mu) d\rho(\mu)$$

schreiben. Daraus folgt durch Vergleich mit (289), daß dem Ersetzen von y durch Ay die Multiplikation von $y(\mu)$ mit μ entspricht.

Wir erwähnen noch, daß die allgemeine Gleichung (259) für ein bilineares Funktional (Ay, z) , wobei y und z in C_x liegen, wegen (290) und wegen Satz 2 mit Hilfe des Lebesgue-Stieltjesschen Integrals folgendermaßen geschrieben werden kann:

$$(Ay, z) = \int_m^M \lambda y(\lambda) \overline{z(\lambda)} d\rho(\lambda). \quad (293)$$

153. Unitär äquivalente selbstadjungierte Operatoren. Es sei $\mathcal{E}_\lambda$ eine Zerlegung der Einheit für einen selbstadjungierten Operator A und U ein unitärer Operator, und es gelte $B = UAU^{-1}$. Der Operator $\mathcal{E}'_\lambda = U\mathcal{E}_\lambda U^{-1}$ ist, wie man leicht sieht, ebenfalls eine Zerlegung der Einheit. Ferner sei B' der dieser Zerlegung zugeordnete selbstadjungierte Operator, so daß $B'x$ als Grenzwert der Summe

$$\sum_{k=1}^n \nu_k A_k (U\mathcal{E}_\lambda U^{-1}) x = U \left(\sum_{k=1}^n \nu_k A_k \mathcal{E}_\lambda \right) U^{-1} x$$

definiert ist. Daraus folgt, daß B' mit B übereinstimmt, d. h., $\mathcal{E}'_\lambda = U\mathcal{E}_\lambda U^{-1}$ ist die Spektralschar des Operators B . Unitär äquivalente Operatoren haben also das gleiche Spektrum.

Bei einem reinen Punktspektrum ist die Übereinstimmung der Eigenwerte nebst ihrem Rang sowohl notwendig als auch hinreichend für die unitäre Äquivalenz. Der zugeordnete unitäre Operator U läßt sich leicht als Operator konstruieren, der den Unterraum der Eigenelemente von B in den derjenigen Eigenelemente von A abbildet, die den gleichen Eigenwerten entsprechen.

Die Frage nach den Bedingungen für die unitäre Äquivalenz ist wesentlich komplizierter, wenn das Spektrum kontinuierlich ist. Wir geben das diesen Fall betreffende Hauptresultat ohne Beweis an (der Raum sei dabei separabel).

Zwei selbstadjungierte Operatoren $A^{(1)}$ und $A^{(2)}$ sind genau dann unitär äquivalent, wenn die folgenden Bedingungen erfüllt sind: a) Die Spektren der beiden Operatoren sind vom gleichen Typ (entweder reines Punktspektrum oder rein

kontinuierliches Spektrum oder gemischtes Spektrum); b) liegt ein Punktspektrum vor, so besteht es für beide Operatoren aus den gleichen Werten mit einem für beide Operatoren gleichen Rang; c) liegt ein kontinuierliches Spektrum vor, so ist die Anzahl der invarianten Unterräume in der normalen Zerlegung des stetigen Teils der Spektralschar [49] für beide Operatoren gleich, und führen wir für normale Zerlegungen des kontinuierlichen Spektrums von $A^{(1)}$ und $A^{(2)}$ die in [149] konstruierten Funktionen

$$||\mathfrak{E}_\lambda^{(1)} y_k^{(1)}||^2 = \varrho_k^{(1)}(\lambda) \quad \text{und} \quad ||\mathfrak{E}_\lambda^{(2)} y_k^{(2)}||^2 = \varrho_k^{(2)}(\lambda)$$

ein, so muß eine Menge vom Maß 0 bezüglich $\varrho_k^{(1)}(\lambda)$ auch bezüglich $\varrho_k^{(2)}(\lambda)$ vom Maß 0 sein und umgekehrt, d. h., für jedes k ist

$$\varrho_k^{(1)}(\lambda) = \int_m^\lambda \varphi_k^{(1)}(\lambda) d\varrho_k^{(2)}(\lambda) \quad \text{und} \quad \varrho_k^{(2)}(\lambda) = \int_m^\lambda \varphi_k^{(2)}(\lambda) d\varrho_k^{(1)}(\lambda),$$

wobei $\varphi_k^{(1)}(\lambda)$ und $\varphi_k^{(2)}(\lambda)$ meßbar bezüglich $\varrho_k^{(2)}(\lambda)$ bzw. $\varrho_k^{(1)}(\lambda)$, nichtnegativ und summierbar sind.

154. Die Spektralzerlegung unitärer Operatoren. Es sei A ein selbstadjungierter Operator und $\mathfrak{E}_\lambda$ seine Spektralschar, für die

$$\mathfrak{E}_\lambda = \begin{cases} 0 & \text{für } \lambda = 0, \\ E & \text{für } \lambda = 1 \end{cases} \quad (294)$$

gilt. Wir bilden einen Operator U nach der Formel

$$U = \int_0^1 e^{2\pi i \lambda} d\mathfrak{E}_\lambda = e^{2\pi i A}. \quad (295)$$

Für die Bildung des adjungierten Operators hat man die Funktion $e^{2\pi i \lambda}$ durch ihre konjugierte komplexe Funktion $e^{-2\pi i \lambda}$ zu ersetzen [143], d. h.

$$U^* = \int_0^1 e^{-2\pi i \lambda} d\mathfrak{E}_\lambda,$$

und mit Hilfe von (214) erhalten wir die Gleichung $UU^* = U^*U = E$, d. h., der durch (295) definierte Operator U ist unter der Bedingung (294) unitär. Ohne Beweis geben wir an, daß auch die umgekehrte Behauptung gilt:

Satz. Geht man von allen möglichen Zerlegungen der Einheit aus, die den Bedingungen (294) genügen, so erhält man durch die Gleichung (295) die allgemeine Form eines unitären Operators; dabei sind verschiedenen Zerlegungen der Einheit $\mathfrak{E}_\lambda$ verschiedene unitäre Operatoren U zugeordnet.

In [143] haben wir Funktionen $f(A)$ eines selbstadjungierten Operators A definiert, die stetigen Funktionen $f(t)$ zugeordnet sind. In [155] werden wir diese Definition auf eine umfassendere Klasse von Funktionen $f(t)$ verallgemeinern.

155. Funktionen eines selbstadjungierten Operators. Es sei A ein selbstadjungierter Operator und $\mathfrak{E}_\lambda$ seine Spektralschar. Ist $f(\lambda)$ im Intervall $[m, M]$ stetig, so haben wir den Operator $f(A)$ durch die Gleichung

$$f(A) = \int_m^M f(\lambda) d\mathfrak{E}_\lambda$$

oder durch die äquivalente Gleichung für ein bilineares Funktional

$$(f(A) x, y) = \int_m^M f(\lambda) d(\mathcal{E}_\lambda x, y) \quad (296)$$

definiert; $(\mathcal{E}_\lambda x, y)$ ist darin eine komplexe Funktion von beschränkter Variation bezüglich λ . Diese Funktion läßt sich bekanntlich [125] als Linearkombination von vier nichtfallenden Funktionen der Gestalt $\|\mathcal{E}_\lambda z\|^2$ darstellen, wobei z ein gewisses Element aus H ist. Wenn $f(\lambda)$ eine beliebige beschränkte Funktion ist, die bezüglich der nichtfallenden Funktion

$$\|\mathcal{E}_\lambda z\|^2 \quad (297)$$

für jedes z meßbar ist, so existiert das Integral (296) für beliebige x und y ; das bilineare Funktional $(f(A) x, y)$ ist also wohlbestimmt. Die Distributivität dieses Funktional ergibt sich offenbar aus der Distributivität von $(\mathcal{E}_\lambda x, y)$. Wir beweisen jetzt die Beschränktheit dieses Funktional und nehmen dabei $f(\lambda)$ als reell an, was übrigens nicht wesentlich ist. Die Funktion $f(\lambda)$ ist nach Voraussetzung beschränkt, es gilt also mit einer gewissen positiven Zahl c die Ungleichung $|f(\lambda)| \leq c$. Nehmen wir $y = x$ an, so erhalten wir für das quadratische Funktional

$$(f(A) x, x) = \int_m^M f(\lambda) d\|\mathcal{E}_\lambda x\|^2. \quad (298)$$

Wir erinnern daran, daß für $\mathcal{E}_m \neq 0$ das Integral auf der rechten Seite gleichbedeutend ist mit der Summe

$$f(m) \|\mathcal{E}_m x\|^2 + \int_m^M f(\lambda) d\|\mathcal{E}_\lambda x\|^2,$$

in der das Integral als gewöhnlicher Summengrenzwert aufzufassen ist. Aus $|f(\lambda)| \leq c$ und aus $\|\mathcal{E}_M x\| = \|x\|$ erhalten wir für das Integral (298) die Abschätzung

$$|(f(A) x, x)| \leq c \|x\|^2. \quad (298_1)$$

Ferner gilt

$$4 \operatorname{Re} \left[\int_m^M f(\lambda) d(\mathcal{E}_\lambda x, y) \right] = \int_m^M f(\lambda) d(\mathcal{E}_\lambda (x+y), x+y) - \int_m^M f(\lambda) d(\mathcal{E}_\lambda (x-y), x-y)$$

oder

$$4 \operatorname{Re} \left[\int_m^M f(\lambda) d(\mathcal{E}_\lambda x, y) \right] = \int_m^M f(\lambda) d\|\mathcal{E}_\lambda (x+y)\|^2 - \int_m^M f(\lambda) d\|\mathcal{E}_\lambda (x-y)\|^2$$

und wegen (298₁)

$$4 \left| \operatorname{Re} \left[\int_m^M f(\lambda) d(\mathcal{E}_\lambda x, y) \right] \right| \leq c [\|x+y\|^2 + \|x-y\|^2] = 2c [\|x\|^2 + \|y\|^2].$$

Folgern wir nun wie in [122], so können wir aus dieser Abschätzung sofort eine Abschätzung ohne die Beschränkung auf den Realteil erhalten:

$$4 \left| \int_m^M f(\lambda) d(\mathcal{E}_\lambda x, y) \right| \leq 2c [\|x\|^2 + \|y\|^2].$$

Für $\|x\| = \|y\| = 1$ folgt die Abschätzung

$$|(f(A) x, y)| = \left| \int_m^M f(\lambda) d(\mathcal{E}_\lambda x, y) \right| \leq c. \quad (299)$$

Wenn x und y eine beliebige positive Norm haben, so können wir

$$(f(A)x, y) = \|x\| \cdot \|y\| \cdot \left(f(A) \frac{x}{\|x\|}, \frac{y}{\|y\|} \right)$$

schreiben, wobei die Norm der Elemente $\frac{x}{\|x\|}$ und $\frac{y}{\|y\|}$ gleich 1 ist. Wegen (299) gilt also

$$|(f(A)x, y)| \leq c \|x\| \cdot \|y\|.$$

Daraus folgt schließlich die Beschränktheit des bilinearen Funktional $(f(A)x, y)$.

Auf diese Weise kommen wir zu dem folgenden allgemeinen Ergebnis: *Ist $f(\lambda)$ eine beschränkte und bezüglich der nichtfallenden Funktionen (297) bei beliebiger Wahl von z meßbare Funktion, so wird durch die Gleichung (296) ein linearer Operator $f(A)$ bestimmt.* Wir erwähnen noch einige Eigenschaften eines solchen Operators $f(A)$. Ist $f(\lambda)$ reell, so folgt aus (296) für $y=x$, daß $f(A)$ ein selbstadjungierter Operator ist. Ist $f(\lambda)$ komplex, so erhalten wir den adjungierten Operator $(f(A))^*$ durch (296), wenn wir $f(\lambda)$ durch die konjugiert komplexe Funktion ersetzen. Im Fall $f(\lambda) \geq 0$ folgt aus (298), daß $f(A)$ ein positiver Operator ist.

Wir betrachten nun eine Funktion $f_\mu(\lambda)$, die wir in folgender Weise definieren:

$$f_\mu(\lambda) = \begin{cases} 1 & \text{für } \lambda \leq \mu, \\ 0 & \text{für } \lambda > \mu. \end{cases} \quad (300)$$

Diese Funktion ist offenbar eine B -Funktion, und wir können das Integral (296) bilden. Zerlegen wir das Integrationsintervall in die Teilintervalle $[m - \varepsilon_0, \mu]$ und $[\mu, M]$, so erhalten wir

$$(f_\mu(A)x, y) = \int_m^\mu d(\mathfrak{E}_\lambda x, y) = (\mathfrak{E}_\mu x, y),$$

woraus

$$\mathfrak{E}_\mu = f_\mu(A) \quad (301)$$

folgt. Wir haben diese Beziehung unter der Annahme erhalten, daß jeder selbstadjungierte Operator A eine Spektralschar $\mathfrak{E}_\lambda$ besitzt, durch die er vermöge (204) dargestellt werden kann, d. h., bei der Herleitung der Gleichung (301) haben wir uns auf den Satz aus [142] gestützt, dessen Beweis wir noch zu führen haben. Dieser Beweis läuft im Grunde genommen darauf hinaus, daß wir für einen beliebigen selbstadjungierten Operator A eine Funktion $f_\mu(A)$ definieren, ohne die Spektralschar $\mathfrak{E}_\lambda$ zu benutzen, und danach unter der Annahme $\mathfrak{E}_\mu = f_\mu(A)$ die grundlegende Gleichung (204) beweisen.

Verwenden wir die bekannten Eigenschaften des Lebesgue-Stieltjesschen Integrals, so können wir leicht die Eigenschaften der Funktion eines selbstadjungierten Operators A bestimmen. Dabei nehmen wir natürlich an, daß alle auftretenden Funktionen der oben definierten Klasse angehören, d. h. beschränkt und in bezug auf die Funktionen (210) bei beliebiger Wahl von z meßbar sind.

Satz 1. *Der Linearkombination $a_1 f_1(\lambda) + a_2 f_2(\lambda) + \dots + a_p f_p(\lambda)$ ist der Operator $a_1 f_1(A) + a_2 f_2(A) + \dots + a_p f_p(A)$ zugeordnet.*

2. *Der Operator $f(A)$ ist mit $\mathfrak{E}_\mu$ und mit A vertauschbar.*

3. *Es gilt*

$$(f_1(A)x, f_2(A)y) = \int_m^M f_1(\lambda) \overline{f_2(\lambda)} d(\mathfrak{E}_\lambda x, y). \quad (302)$$

4. Der Funktion $f_1(\lambda) f_2(\lambda)$ ist der Operator

$$f_1(A) f_2(A) = f_2(A) f_1(A) \quad (303)$$

zugeordnet.

Beweis. Wir zeigen zunächst, daß $f(A)$ mit $\mathcal{E}_\mu$ vertauschbar ist. Es gilt

$$\begin{aligned} (f(A) \mathcal{E}_\mu x, y) &= \int_m^M f(\lambda) d(\mathcal{E}_\lambda \mathcal{E}_\mu x, y) = \int_m^M f(\lambda) d(\mathcal{E}_\lambda x, \mathcal{E}_\mu y) \\ &= (f(A) x, \mathcal{E}_\mu y) = (\mathcal{E}_\mu f(A) x, y); \end{aligned}$$

daraus folgt $f(A) \mathcal{E}_\mu = \mathcal{E}_\mu f(A)$. Wir bemerken noch, daß aus diesen Gleichungen wegen (180)

$$(f(A) \mathcal{E}_\mu x, y) = \int_m^\mu f(\lambda) d(\mathcal{E}_\lambda x, y) \quad (304)$$

folgt.

Die Gleichung (302) ergibt sich aus

$$\begin{aligned} (f_1(A) x, f_2(A) y) &= \int_m^M f_1(\lambda) d(\mathcal{E}_\lambda x, f_2(A) y) = \int_m^M f_1(\lambda) d(f_2(A) \mathcal{E}_\lambda y, x) \\ &= \int_m^M f_1(\lambda) d \left[\int_m^\lambda f_2(\mu) d(\mathcal{E}_\mu y, x) \right] = \int_m^M f_1(\lambda) \overline{f_2(\lambda)} d(\mathcal{E}_\lambda x, y). \end{aligned} \quad (305)$$

Schließlich ist

$$\begin{aligned} (f_1(A) f_2(A) x, y) &= \int_m^M f_1(\lambda) d(\mathcal{E}_\lambda f_2(A) x, y) = \int_m^M f_1(\lambda) d \int_m^\lambda f_2(\mu) d(\mathcal{E}_\mu x, y) \\ &= \int_m^M f_1(\lambda) f_2(\lambda) d(\mathcal{E}_\lambda x, y), \end{aligned}$$

und denselben Ausdruck erhält man für das Produkt $f_2(A) f_1(A)$. Analog wie in [143] können wir zeigen, daß $f(A)$ mit einem beliebigen, mit A vertauschbaren Operator B vertauschbar ist. Auch die Umkehrung ist richtig: Wenn ein beschränkter linearer Operator C mit einem beliebigen, mit A vertauschbaren Operator B vertauschbar ist, existiert eine Funktion $j(\lambda)$, für die $C = f(A)$ gilt. Den Beweis dieses wichtigen Satzes findet man in einer Arbeit von F. RIESZ¹⁾.

Setzen wir in (305) $f_2(\lambda) \equiv f_1(\lambda)$ und $y = x$, so erhalten wir

$$\|f_1(A) x\|^2 = \int_m^M |f_1(\lambda)|^2 d\|\mathcal{E}_\lambda x\|^2. \quad (306)$$

Wir erwähnen noch einige Tatsachen über Funktionen eines selbstadjungierten Operators. Ist $|f(\lambda)| \equiv 1$, so ist $f(A)$ ein unitärer Operator. Nimmt $f(\lambda)$ nur die Werte 0 und 1 an, so ist $f(A)$ ein Projektionsoperator. Wir wollen beweisen: Ist $\lambda = \lambda_0$ ein Eigenwert von A und ist x_0 ein zugeordnetes Eigenelement, so ist $f(\lambda_0)$ Eigenwert von $f(A)$ mit demselben Eigenelement x_0 . Wir wissen, daß im Fall $\lambda < \lambda_0$ die Gleichung $\mathcal{E}_\lambda x_0 = 0$ und im Fall $\lambda \geq \lambda_0$ die Gleichung $\mathcal{E}_\lambda x_0 = x_0$ gilt; (296) liefert uns für jedes y die Beziehung

$$(f(A) x_0, y) = f(\lambda_0) (x_0, y) = (f(\lambda_0) x_0, y),$$

¹⁾ F. RIESZ, Sur les fonctions des transformations hermitiennes dans l'espace de Hilbert, Acta Sci. Math. Szeged 7 (1935), 147–159.

aus der, da y beliebig ist, $f(A)x_0 = f(\lambda_0)x_0$ folgt. Wir erwähnen noch: Besitzt die Funktion $f(\lambda)$ endlich viele Unstetigkeiten, so ist sie bezüglich aller Funktionen (297) meßbar [42], und damit hat $f(A)$ eine wohlbestimmte Bedeutung. Dasselbe gilt auch dann, wenn $f(\lambda)$ eine B -Funktion ist [47], was wir bereits oben benutzten.

156. Vertauschbare Operatoren. Wir untersuchen nun vertauschbare selbstadjungierte Operatoren.

Satz 1. *Zwei selbstadjungierte Operatoren A und B sind dann und nur dann vertauschbar, wenn ihre Spektralscharen $\mathcal{E}_\lambda$ und F_μ für beliebige Werte von λ und μ vertauschbar sind.*

Beweis. Wir wissen, daß die Spektralschar eines beliebigen selbstadjungierten Operators C mit C sowie mit einem beliebigen, mit C vertauschbaren Operator vertauschbar ist [148]. Daraus erkennt man, daß aus der Gleichung $AB = BA$ die Vertauschbarkeit von F_μ mit A und damit die Vertauschbarkeit von $\mathcal{E}_\lambda$ mit F_μ folgt. Ist umgekehrt $\mathcal{E}_\lambda$ mit F_μ vertauschbar, so lassen sich die Riemann-Stieltjes'schen Summen in der Integraldarstellung von A und B und damit auch diese Operatoren selbst vertauschen.

Satz 2. *Sind die selbstadjungierten Operatoren A , B und C mit reinem Punktspektrum paarweise vertauschbar, so gibt es ein vollständiges normiertes Orthogonalsystem von Elementen, die die Eigenelemente jedes der drei Operatoren sind.*

Beweis. Es seien $\mathcal{E}_\lambda$, F_μ und G_ν die Spektralscharen der drei Operatoren. Wegen Satz 1 sind sie paarweise vertauschbar. Ferner seien λ , μ und ν beliebige Eigenwerte von A , B bzw. C , außerdem L_λ , M_μ und N_ν die Unterräume der zugeordneten Eigen-elemente sowie $\Delta_\lambda = \mathcal{E}_\lambda - \mathcal{E}_{\lambda-0}$, $\Delta_\mu = F_\mu - F_{\mu-0}$ und $\Delta_\nu = G_\nu - G_{\nu-0}$ die Projektionsoperatoren in diese Unterräume. Diese drei Projektionsoperatoren sind paarweise vertauschbar; daher ist ihr Produkt

$$\Delta_{\lambda\mu\nu} = \Delta_\lambda \Delta_\mu \Delta_\nu$$

der Projektionsoperator in den Unterraum $R_{\lambda\mu\nu}$, der aus den Elementen besteht, die in L_λ , M_μ und N_ν gemeinsam auftreten [140]. Betrachten wir zwei verschiedene solcher Unterräume $R_{\lambda\mu\nu}$ und $R_{\lambda'\mu'\nu'}$, so besteht mindestens eins der Zahlenpaare (λ, λ') , (μ, μ') und (ν, ν') aus verschiedenen Zahlen. Es sei etwa $\lambda \neq \lambda'$. Dann sind x und x' , falls $x \in R_{\lambda\mu\nu}$ und $x' \in R_{\lambda'\mu'\nu'}$ ist, Eigenwerte von A , die verschiedenen Eigenelementen zugeordnet sind, und daher zueinander orthogonal. Also sind die Unterräume $R_{\lambda\mu\nu}$ paarweise orthogonal. Wir wollen beweisen, daß ihre orthogonale Summe gleich dem gesamten Raum H ist. Dazu brauchen wir nur zu zeigen, daß kein vom Nullelement verschiedenes Element existiert, das zu allen Unterräumen $R_{\lambda\mu\nu}$ orthogonal ist, d. h., ist ein Element x_0 vom Nullelement verschieden, so ist es zu mindestens einem der $R_{\lambda\mu\nu}$ nicht orthogonal. Für die Unterräume L_λ , M_μ und N_ν ist das klar, denn nach Voraussetzung haben A , B und C ein reines Punktspektrum; daher ist die orthogonale Summe z. B. der L_λ der gesamte Raum H . Wir greifen ein Element $x_0 \neq 0$ heraus. Nach dem eben Gesagten existiert ein Eigenwert λ des Operators A , so daß $\Delta_\lambda x_0 \neq 0$ ist. Weiterhin existiert, wiederum auf Grund dieser Überlegungen, ein Eigenwert μ des Operators B , so daß $\Delta_\mu(\Delta_\lambda x_0) \neq 0$ ist, und ein Eigenwert ν des Operators C , so daß $\Delta_\nu(\Delta_\mu \Delta_\lambda x_0) \neq 0$ gilt, woraus folgt, daß x_0 nicht zu $R_{\lambda\mu\nu}$ orthogonal ist. Folglich ist die orthogonale Summe der $R_{\lambda\mu\nu}$ der gesamte Raum H . Wenn wir in jedem $R_{\lambda\mu\nu}$ ein vollständiges normiertes Orthogonalsystem wählen, so erhalten wir ein in H vollständiges normiertes Orthogonalsystem, und jedes seiner Elemente ist, da es in einem gewissen $R_{\lambda\mu\nu}$ liegt, ein Eigenelement aller drei Operatoren A , B und C . Der Satz läßt sich in genau derselben Weise auch für eine beliebige endliche Anzahl paarweise vertauschbarer selbstadjungierter Operatoren beweisen.

Wir haben oben gesehen, daß die verschiedenen Funktionen ein und desselben selbstadjungierten Operators vertauschbare Operatoren sind [155]. Wir beweisen nun die

umgekehrte Behauptung für den Fall, daß die vorgegebenen Operatoren ein reines Punktspektrum haben.

Satz 3. *Wenn die selbstadjungierten Operatoren A , B und C ein reines Punktspektrum haben und paarweise vertauschbar sind, so lassen sie sich als Funktionen ein und desselben selbstadjungierten Operators darstellen.*

Beweis. Der Raum H wird als separabel vorausgesetzt. Nach Satz 2 gibt es ein vollständiges normiertes Orthogonalsystem $x_1, x_2, x_3, \dots$ von Elementen, die die Eigenelemente von A , B und C sind, d. h. für die

$$Ax_n = \lambda_n x_n, \quad Bx_n = \mu_n x_n, \quad Cx_n = \nu_n x_n \quad (n=1, 2, \dots)$$

gilt. Wir setzen $\varrho_m = 1/m$ und entwickeln ein beliebiges Element x nach den Elementen x_k :

$$x = \sum_{k=1}^{\infty} a_k x_k.$$

Dann definieren wir einen selbstadjungierten Operator D durch

$$Dx = \sum_{k=1}^{\infty} a_k \varrho_k x_k.$$

Die Reihe auf der rechten Seite ist offenbar konvergent, denn bilden die Zahlen $|a_k|^2$ eine konvergente Reihe, so gilt dasselbe erst recht für die Zahlen $|a_k \varrho_k|^2$. Aus der Definition von D folgt, daß die x_k die Eigenelemente von D sind und zu den Eigenwerten ϱ_k gehören; D hat also ein reines Punktspektrum. Wir können eine beschränkte Funktion $f_1(\lambda)$ bilden, die in den Punkten $\lambda = \varrho_k$ gleich λ_k und überall stetig ist, außer eventuell im Punkt $\lambda = 0$. Analog lassen sich Funktionen $f_2(\lambda)$ und $f_3(\lambda)$ bilden, für die $f_2(\varrho_k) = \mu_k$ und $f_3(\varrho_k) = \nu_k$ gilt. Nach den Ergebnissen aus [155] können wir für die Funktionen $f_k(\lambda)$ die zugeordneten Operatoren $f_1(D)$, $f_2(D)$ und $f_3(D)$ bilden. Der Operator $f_1(D)$ hat die Eigenelemente x_k , die ein vollständiges System bilden, und die Eigenwerte $f_1(\varrho_k) = \lambda_k$. Dieselben Eigenwerte und Eigenelemente hat auch der Operator A . Wenn aber zwei Operatoren mit reinen Punktspektren die gleichen Eigenwerte und die gleichen, diesen Eigenwerten zugeordneten Eigenelemente haben, so folgt aus ihrer Integraldarstellung durch die Spektralschar, daß sie übereinstimmen müssen, d. h., daß $A = f_1(D)$ und analog $B = f_2(D)$ und $C = f_3(D)$ gilt, womit der Satz bewiesen ist.

Wir bemerken, daß dieser Satz sich in genau derselben Weise auch für den Fall einer beliebigen endlichen Anzahl selbstadjungierter Operatoren beweisen läßt. Er gilt auch dann, wenn die Operatoren keine reinen Punktspektren haben, womit wir uns aber nicht aufhalten wollen.¹⁾

157. Störung des Spektrums eines selbstadjungierten Operators. Wir erinnern daran, daß als *Verdichtungspunkte* des Spektrums eines selbstadjungierten Operators A diejenigen Werte von λ bezeichnet werden, die entweder Häufungspunkte des Punktspektrums oder Eigenwerte von unendlichem Rang oder Punkte des kontinuierlichen Spektrums sind. Unser Ziel ist es, den folgenden Satz zu beweisen:

Satz 1. *Addiert man zu dem selbstadjungierten Operator A einen vollstetigen selbstadjungierten Operator C , so bleibt die Menge der Verdichtungspunkte des Spektrums unverändert.*

Es zeigt sich jedoch, daß die Addition eines vollstetigen Operators den Charakter eines Spektrums wesentlich ändern kann, denn es gilt der folgende Satz:

Satz 2. *Zu jedem vorgegebenen selbstadjungierten Operator A kann man einen solchen selbstadjungierten vollstetigen Operator C mit absoluter Norm, die höchstens gleich einer gegebenen positiven Zahl ϵ ist, addieren, daß $A + C$ ein reines Punktspektrum hat.*

¹⁾ Vgl. J. v. NEUMANN, Über Funktionen von Funktionaloperatoren, Ann. Math. **32** (1931), 191–226.

Mit Hilfe dieses Satzes läßt sich die folgende Behauptung beweisen:

Satz 3. *Haben zwei selbstadjungierte Operatoren A_1 und A_2 in ihrem Spektrum die gleiche Menge von Verdichtungspunkten, so existieren ein unitärer Operator U und ein selbstadjungierter vollstetiger Operator C derart, daß $A_2 = UA_1U^{-1} + C$ ist.*

Zum Beweis von Satz 1 benötigen wir noch zwei Hilfssätze.

Hilfssatz 1. *Ist $\lambda = \mu$ ein Verdichtungspunkt des Spektrums eines selbstadjungierten Operators A , so existiert eine schwach gegen 0 konvergierende Folge normierter Elemente x_n , für die*

$$\|Ax_n - \mu x_n\| \rightarrow 0 \quad (307)$$

gilt.

Ist μ ein Häufungspunkt des Punktspektrums oder ein Eigenwert von unendlichem Rang, so existiert eine unendliche Folge paarweise orthogonaler normierter Elemente x_n , für die die zugeordneten Eigenwerte λ_n gegen μ streben. Ist z ein beliebiges Element, so streben seine Fourierkoeffizienten $c_n = (z, x_n)$ gegen 0. Folglich gilt $x_n \rightarrow 0$, und die Behauptung des Hilfssatzes 1 folgt für den betrachteten Fall aus den Gleichungen

$$\|Ax_n - \mu x_n\| = \|(A - \lambda_n E)x_n + (\lambda_n - \mu)x_n\| = |\lambda_n - \mu| \cdot \|x_n\| = |\lambda_n - \mu|.$$

Wir nehmen nun an, daß μ ein Punkt des kontinuierlichen Spektrums und $\mathcal{E}'_\lambda$ der stetige Teil der Spektralschar ist. Die Differenz $\mathcal{E}'_{\mu+\delta} - \mathcal{E}'_{\mu-\delta}$ ist bei beliebig kleinem positivem δ ein Projektionsoperator in einen gewissen Unterraum L_δ . Wir benutzen eine Folge positiver Zahlen δ_n , für die $\delta_n \rightarrow 0$ gilt, und eine Folge normierter Elemente x_n aus L_{δ_n} und zeigen, daß dann die Behauptung des Hilfssatzes ebenfalls erfüllt ist. Nach Definition der δ_n ist $(\mathcal{E}'_{\mu+\delta_n} - \mathcal{E}'_{\mu-\delta_n})x_n = x_n$ und für ein beliebiges Element z

$$(z, x_n) = (z, (\mathcal{E}'_{\mu+\delta_n} - \mathcal{E}'_{\mu-\delta_n})x_n) = ((\mathcal{E}'_{\mu+\delta_n} - \mathcal{E}'_{\mu-\delta_n})z, x_n),$$

$$|(z, x_n)| \leq \|(\mathcal{E}'_{\mu+\delta_n} - \mathcal{E}'_{\mu-\delta_n})z\|.$$

Es gilt aber $\mathcal{E}'_{\mu+\delta_n} - \mathcal{E}'_{\mu-\delta_n} \rightarrow 0$, womit die schwache Konvergenz $x_n \rightarrow 0$ bewiesen ist. Für den Beweis von (307) hat man die folgenden offenbar gültigen Beziehungen zu benutzen:

$$\begin{aligned} \|Ax_n - \mu x_n\|^2 &= \int_{m-\varepsilon}^M (\lambda - \mu)^2 d(\mathcal{E}'_\lambda x_n, x_n) = \int_{m-\varepsilon}^M (\lambda - \mu)^2 d(\mathcal{E}'_\lambda (\mathcal{E}'_{\mu+\delta_n} - \mathcal{E}'_{\mu-\delta_n})x_n, x_n) \\ &= \int_{\mu-\delta_n}^{\mu+\delta_n} (\lambda - \mu)^2 d\|(\mathcal{E}'_\lambda - \mathcal{E}'_{\mu-\delta_n})x_n\|^2 \leq \delta_n^2 \|(\mathcal{E}'_{\mu+\delta_n} - \mathcal{E}'_{\mu-\delta_n})x_n\|^2 = \delta_n^2 \rightarrow 0. \end{aligned}$$

Hilfssatz 2. *Ist $\lambda = \mu$ kein Verdichtungspunkt des Spektrums, so existiert für jede schwach gegen 0 konvergierende Folge normierter Elemente x_n eine positive Zahl a derart, daß für alle genügend großen n*

$$\|Ax_n - \mu x_n\| \geq a > 0 \quad (308)$$

gilt.

Nach Voraussetzung existiert eine positive Zahl d , für die im Intervall $\mu - d \leq \lambda \leq \mu + d$ die Spektralschar $\mathcal{E}_\lambda$ entweder konstant ist oder ihre Änderung in einem Sprung im Punkt $\lambda = \mu$ besteht, wobei der Unterraum L_μ der diesem Sprung zugeordneten Eigenwerte von endlicher Dimension ist. Es gilt

$$\begin{aligned} \|Ax_n - \mu x_n\|^2 &= \int_{m-\varepsilon}^M (\lambda - \mu)^2 d\|\mathcal{E}_\lambda x_n\|^2 \geq \int_{m-\varepsilon}^{\mu-d} (\lambda - \mu)^2 d\|\mathcal{E}_\lambda x_n\|^2 + \int_{\mu+d}^M (\lambda - \mu)^2 d\|\mathcal{E}_\lambda x_n\|^2 \\ &\geq d^2 [\|\mathcal{E}_{\mu-d} x_n\|^2 + d^2 (\|x_n\|^2 - \|\mathcal{E}_{\mu+d} x_n\|^2)] = d^2 - d^2 [\|\mathcal{E}_{\mu+d} x_n\|^2 - \|\mathcal{E}_{\mu-d} x_n\|^2] \end{aligned}$$

oder

$$\|Ax_n - \mu x_n\|^2 \geq d^2 - d^2 ((\mathcal{E}_{\mu+d} - \mathcal{E}_{\mu-d})x_n, x_n). \quad (309)$$

Für $\mathfrak{E}_{\mu+d} - \mathfrak{E}_{\mu-d} = 0$ erhalten wir (308), wenn wir $a=d$ annehmen. Nun habe $\mathfrak{E}_\lambda$ einen Sprung für $\lambda = \mu$, und es sei $z_1, z_2, \dots, z_m$ ein vollständiges normiertes Orthogonalsystem in L_μ . Dann ist

$$(\mathfrak{E}_{\mu+d} - \mathfrak{E}_{\mu-d}) x_n = \sum_{s=1}^m (x_n, z_s) z_s ;$$

wegen $x_n \rightarrow C$ gilt $(\mathfrak{E}_{\mu+d} - \mathfrak{E}_{\mu-d}) x_n \rightarrow 0$, und aus (309) folgt, daß (308) für genügend große n erfüllt ist, wenn etwa $a=d/2$ angenommen wird.

Aus diesen Hilfssätzen folgt unmittelbar, daß $\lambda = \mu$ dann und nur dann ein Verdichtungspunkt des Spektrums ist, wenn eine Folge normierter Elemente x_n existiert, für die $x_n \rightarrow 0$ ist und die Beziehung (307) gilt.

Nun ist es leicht, den Satz 1 zu beweisen. Wir addieren zu dem Operator A einen vollstetigen selbstadjungierten Operator C , und es sei $A_1 = A + C$. Dann ist $A = A_1 + (-C)$, wobei $-C$ ebenfalls ein vollstetiger selbstadjungierter Operator ist. Ferner sei $\lambda = \mu$ ein Verdichtungspunkt des Spektrums von A und x_n eine der Bedingung (307) genügende, schwach gegen 0 konvergierende Folge. Wegen $x_n \rightarrow 0$ gilt dann $Cx_n \rightarrow 0$, und aus der Ungleichung

$$\|A_1 x_n - \mu x_n\| \leq \|A x_n - \mu x_n\| + \|C x_n\|$$

folgt $\|A_1 x_n - \mu x_n\| \rightarrow 0$, d. h., $\lambda = \mu$ ist ein Verdichtungspunkt des Spektrums auch für A_1 . Ganz analog folgt aus der Gleichung $A = A_1 + (-C)$ umgekehrt, daß jeder Verdichtungspunkt des Spektrums von A_1 auch ein Verdichtungspunkt des Spektrums von A ist. Damit ist der Satz bewiesen.

158. Normale Operatoren. Wir weisen noch auf eine spezielle Art linearer Operatoren hin, auf die sogenannten normalen Operatoren. Ein linearer Operator A heißt *normal*, wenn er mit seinem adjungierten Operator vertauschbar ist [IV/1]:

$$AA^* = A^*A . \quad (310)$$

Einen Spezialfall normaler Operatoren bilden die adjungierten und die unitären Operatoren. Setzen wir

$$A_1 = \frac{1}{2} (A + A^*), \quad A_2 = \frac{1}{2i} (A - A^*) , \quad (311)$$

so können wir A und A^* durch die selbstadjungierten Operatoren A_1 und A_2 ausdrücken:

$$A = A_1 + iA_2, \quad A^* = A_1 - iA_2 . \quad (312)$$

Daraus folgt, daß der lineare Operator A dann und nur dann normal ist, wenn die selbstadjungierten Operatoren A_1 und A_2 vertauschbar sind. Wenn das der Fall ist, lassen sich die Spektralscharen $\mathfrak{E}_\lambda^{(1)}$ und $\mathfrak{E}_\mu^{(2)}$ dieser Operatoren für beliebige λ und μ vertauschen. Wir definieren eine von der komplexen Veränderlichen $\alpha = \lambda + i\mu$ abhängige Schar von Projektionsoperatoren durch

$$\mathfrak{E}_\alpha = \mathfrak{E}_\lambda^{(1)} \mathfrak{E}_\mu^{(2)} \quad (\alpha = \lambda + i\mu) . \quad (313)$$

Der Projektionsoperator $\mathfrak{E}_\alpha$ ist nur in einem gewissen Intervall Δ_0 der Ebene der komplexen Variablen α veränderlich, und für den Operator A gelten Gleichungen, die den Gleichungen für einen selbstadjungierten Operator vollkommen analog sind:

$$A = \iint_{\Delta_0} \alpha d\mathfrak{E}_\alpha, \quad (Ax, y) = \iint_{\Delta_0} \alpha d(\mathfrak{E}_\alpha x, y) . \quad (314)$$

Wir beweisen z. B. die zweite dieser Gleichungen. Dazu nehmen wir an, das Intervall Δ_0 sei durch die Ungleichungen $a \leq \lambda \leq b$, $c \leq \mu \leq d$ definiert. Es gilt

$$\iint_{\Delta_0} \alpha d(\mathfrak{E}_\alpha x, y) = \iint_{\Delta_0} \lambda d_\lambda d_\mu (\mathfrak{E}_\lambda^{(1)} \mathfrak{E}_\mu^{(2)} x, y) + i \iint_{\Delta_0} \mu d_\lambda d_\mu (\mathfrak{E}_\lambda^{(1)} \mathfrak{E}_\mu^{(2)} x, y) .$$

Im ersten dieser Integrale kann man, da der Integrand nicht von μ abhängt, nach Bildung der Riemann-Stieltjesschen Summe die Summation bezüglich μ durchführen und hat dabei $\mathcal{E}_c^{(2)} = 0$ und $\mathcal{E}_a^{(2)} = E$ zu berücksichtigen. Analog kann man im zweiten Integral die Summation bezüglich λ durchführen; man muß dabei nur $\mathcal{E}_a^{(1)} = 0$ und $\mathcal{E}_b^{(1)} = E$ berücksichtigen. Auf diese Weise erhalten wir

$$\int_{A_0} \lambda d(\mathcal{E}_\lambda x, y) = \int_a^b \lambda d(\mathcal{E}_\lambda^{(1)} x, y) + i \int_c^d \mu d(\mathcal{E}_\mu^{(2)} x, y) = (A_1 x, y) + i(A_2 x, y).$$

Andererseits ist

$$(Ax, y) = (A_1 x, y) + i(A_2 x, y);$$

durch Vergleich ergibt sich die zweite Beziehung (314). Die weitere allgemeine Theorie der normalen Operatoren läßt sich analog zur Theorie der selbstadjungierten Operatoren entwickeln.

Im Fall eines normalen Operators A mögen die selbstadjungierten Operatoren A_1 und A_2 ein reines Punktspektrum haben. Wir können ein vollständiges normiertes Orthogonalsystem der Elemente x_k ($k = 1, 2, 3, \dots$) bilden, die die Eigenelemente von A_1 und A_2 sind [156], für die also $A_1 x_k = \mu_k^{(1)} x_k$, $A_2 x_k = \mu_k^{(2)} x_k$ gilt. Dann ist offenbar

$$Ax_k = (A_1 + iA_2) x_k = (\mu_k^{(1)} + \mu_k^{(2)} i) x_k;$$

folglich sind die x_k die den Eigenwerten $\mu_k^{(1)} + \mu_k^{(2)} i$ zugeordneten Eigenelemente von A .

Wir betrachten noch den Fall, daß der normale Operator A vollstetig ist. Wie wir wissen, sind dann auch der Operator A^* und wegen (311) die Operatoren A_1 und A_2 vollstetig. Es ist nicht schwierig, den Satz 1 aus [186] auf normale vollstetige Operatoren auszudehnen. Es seien μ_k die von 0 verschiedenen Eigenwerte von A_1 und x_k die zugeordneten Eigenelemente, d. h. $A_1 x_k = \mu_k x_k$. Berücksichtigen wir, daß A_2 mit A_1 vertauschbar ist, so erhalten wir durch Multiplikation der letzten Gleichung mit A_2

$$A_1(A_2 x_k) = \mu_k A_2 x_k,$$

d. h., $A_2 x_k$ ist entweder das Nullelement oder ein demselben Eigenwert zugeordnetes Eigenelement. Wir nehmen an, daß μ_k ein Eigenwert der Vielfachheit h und $\mu_k = \mu_{k+1} = \dots = \mu_{k+h-1}$ ist. Nach dem oben Gesagten gilt

$$A_2 x_j = \sum_{s=k}^{k+h-1} c_{js} x_s \quad (j = k, k+1, \dots, k+h-1)$$

und

$$c_{js} = (A_2 x_j, x_s) = (x_j, A_2 x_s) = c_{sj};$$

die c_{js} bilden also eine endliche hermitesche Matrix. Wenn wir x_s unitär transformieren (eine solche Änderung ist für die Betrachtung unwesentlich), so können wir diese Matrix auf Diagonalform bringen und unter Beibehaltung der alten Bezeichnung für die Elemente

$$A_1 x_j = \mu_j x_j, \quad A_2 x_j = \nu_j x_j \quad (j = k, k+1, \dots, k+h-1)$$

schreiben. Dabei können einige der Zahlen ν_j oder sogar alle diese Zahlen gleich 0 sein. Diese Operation können wir für alle von 0 verschiedenen Eigenwerte von A_1 durchführen, jedoch nicht alle von 0 verschiedenen Eigenwerte von A_2 erhalten. Wenn wir diese nicht erhaltenen, von 0 verschiedenen Eigenwerte von A_2 betrachten und eine zur vorigen Operation analoge Operation ausüben, wobei wir von A_2 ausgehen und zu A_1 gelangen, finden wir schließlich eine endliche oder abzählbar unendliche Menge paarweise orthogonaler und normierter Elemente y_k ($k = 1, 2, \dots$), für die

$$A_1 y_k = \mu_k^{(1)} y_k, \quad A_2 y_k = \mu_k^{(2)} y_k$$

gilt; dabei ist von den beiden reellen Zahlen $\mu_k^{(1)}$ und $\mu_k^{(2)}$ mindestens eine von 0 verschieden. Außerdem läßt sich jedes Eigenelement von A_1 , das einem von 0 verschiede-

nen Eigenwert zugeordnet ist, durch endlich viele y_k linear ausdrücken; dasselbe gilt auch für A_2 . Weiterhin ist offenbar

$$Ay_k = (\mu_k^{(1)} + \mu_k^{(2)i}) y_k, \quad A^* y_k = (\mu_k^{(1)} - \mu_k^{(2)i}) y_k.$$

Setzen wir $x = Ay = A_1 y + iA_2 y$, so ist [136] $A_1 y = \sum_k a_k y_k$ und $A_2 y = \sum_k b_k y_k$. Folglich kann jedes Element x , das sich in der Gestalt Ay ausdrücken läßt, nach den Elementen y_k entwickelt werden:

$$x = Ay = \sum_k (a_k + b_k i) y_k.$$

Gilt etwa $\mu_k^{(1)} = 0$, so fehlt in der Entwicklung von $A_1 y$ das Glied, das y_k enthält. Wir haben gesehen [155], daß mit dem Operator A auch der Operator A^* eine Funktion eines selbstadjungierten Operators B ist; daher sind A und A^* vertauschbar, d. h., A ist ein normaler Operator. Es gilt auch die Umkehrung: Jeder normale Operator ist eine Funktion eines selbstadjungierten Operators. Um dies zu beweisen, nehmen wir an, es sei ein normaler Operator $A = A_1 + A_2 i$ gegeben. Die selbstadjungierten Operatoren A_1 und A_2 sind vertauschbar und demzufolge, wie wir in [156] erwähnten, Funktionen ein und desselben selbstadjungierten Operators B , also $A_1 = F_1(B)$ und $A_2 = F_2(B)$. Wenn wir die Funktion $F(\lambda) = F_1(\lambda) + iF_2(\lambda)$ bilden, erhalten wir $A = F(B)$, was wir beweisen wollten.

159. Hilfssätze. Aufgabe dieses und des nächsten Abschnitts ist es, den grundlegenden Satz aus [142] und die Tatsache zu beweisen, daß ein Operator, wenn er sich mit einem selbstadjungierten Operator A vertauschen läßt, auch mit dessen Spektralschar $\mathfrak{S}_\lambda$ für jedes λ vertauschbar ist. Bei der Darstellung dieses Beweises können wir die bis [142] erhaltenen Ergebnisse benutzen. Zuerst ist es notwendig, die Gültigkeit einiger Hilfssätze nachzuweisen.

Hilfssatz 1. Sind A und B vertauschbare selbstadjungierte Operatoren, die der Beziehung

$$A^2 = B^2 \tag{315}$$

genügen, und ist P der Projektionsoperator in den Unterraum L , der aus den Elementen x gebildet wird, die die Gleichung

$$(A + B)x = 0 \quad \text{oder} \quad Ax = -Bx \tag{316}$$

erfüllen, so gelten die folgenden Behauptungen:

a) Ist ein gewisser Operator D mit $A + B$ vertauschbar, so läßt er sich auch mit P vertauschen;

b) ist $Ax = 0$, so gilt $x \in L$, also $Px = x$;

c) der Operator A läßt sich durch die Gleichung

$$A = (E - 2P) B \tag{317}$$

ausdrücken.

Beweis. a) Nach Voraussetzung gilt

$$D(A + B) = (A + B) D. \tag{318}$$

Gehört x zu L , so ist wegen (316) $D(A + B)x = 0$ und daher auch $(A + B)Dx = 0$, also $Dx \in L$. Ist z ein beliebiges Element aus H , so liegt Pz und nach dem oben Bewiesenen auch DPz in L . Wir können für ein beliebiges Element $z \in H$ also $PDPz = DPz$ oder

$$PDP = DP \tag{319}$$

schreiben. Gehen wir in (318) zu den adjungierten Operatoren über und beachten die Selbstadjungiertheit von A und B , so erhalten wir [124]

$$(A + B)D^* = D^*(A + B),$$

d. h., auch der Operator D^* ist mit $A + B$ vertauschbar, und wir können für ihn (319), d. h.

$$PD^*P = D^*P$$

schreiben. Gehen wir hier zu den adjungierten Operatoren über, so folgt $PDP = PD$, da P selbstadjungiert ist. Vergleichen wir dieses Ergebnis mit (319), so finden wir $DP = PD$. Somit ist D tatsächlich mit P vertauschbar, was wir beweisen wollten. Insbesondere sind die Operatoren A und B nach Voraussetzung mit $A + B$ vertauschbar; folglich lassen sich A und B mit P vertauschen.

b) Aus den Gleichungen

$$\|Az\|^2 = (Az, Az) = (A^2z, z), \quad \|Bz\|^2 = (Bz, Bz) = (B^2z, z)$$

und der Bedingung (315) folgt $\|Az\| = \|Bz\|$ für jedes Element z . Ist $Ax = 0$, so gilt auch $Bx = 0$, und folglich genügt x der Gleichung (316), d. h., es ist $x \in L$ und $Px = x$, was wir beweisen wollten.

c) Aus der Vertauschbarkeit von A und B und aus (315) folgt $(A + B)(A - B) = 0$, d. h., ist z ein beliebiges Element aus H , so gehört $(A - B)z$ zu L ; also ist $P(A - B)z = (A - B)z$, d. h.

$$P(A - B) = A - B.$$

Für ein beliebiges Element z ist weiterhin $Pz \in L$ und damit $(A + B)Pz = 0$; also ist

$$(A + B)P = 0.$$

Subtrahieren wir hiervon die vorhergehende Gleichung und berücksichtigen, daß A und B mit P vertauschbar sind, so folgt $2PB = -A + B$, woraus sich schließlich (317) ergibt. Damit ist der Hilfssatz vollständig bewiesen.

Hilfssatz 2. *Ist ein selbstadjungierter Operator C nichtnegativ und ein selbstadjungierter Operator F mit C vertauschbar, so gilt $F^2C = CF^2 \geq 0$.*

Beweis. Berücksichtigen wir die Voraussetzung, so können wir, wenn wir $Fx = y$ setzen,

$$(CF^2x, x) = (FCFx, x) = (CFx, Fx) = (Cy, y) \geq 0$$

schreiben, woraus die Behauptung folgt.

Wir erwähnen noch einen Spezialfall von Hilfssatz 2: Ist ein Projektionsoperator P mit C vertauschbar, so gilt $PC \geq 0$ wegen $P^2 = P$.

Ist $P(t) = a_0 + a_1t + \dots + a_nt^n$ ein Polynom und A ein Operator, so läßt sich, wie wir gesehen haben, dem Polynom $P(t)$ der Operator $P(A) = a_0E + a_1A + \dots + a_nA^n$ zuordnen. Ist A selbstadjungiert und sind die Koeffizienten a_k reell, so ist auch $P(A)$ ein selbstadjungierter Operator. Zur Klärung der Eigenschaften der Polynome von Operatoren müssen wir zwei weitere Hilfssätze beweisen.

Hilfssatz 3. *Ist das Polynom $P(t)$ im Intervall $[0, 1]$ positiv, so können wir es für alle hinreichend großen Werte von p in der Gestalt*

$$P(t) = \sum_{s=0}^p c_s t^s (1-t)^{p-s} \quad (320)$$

darstellen; dabei sind sämtliche Koeffizienten c_s positiv.

Beweis. Für Polynome ersten Grades folgt die Behauptung aus der Gleichung

$$P(t) = c_0(1-t) + c_1t \quad \text{mit} \quad c_0 = P(0) \quad \text{und} \quad c_1 = P(1).$$

Wir betrachten nun ein positives Polynom zweiten Grades, das sich nicht in reelle Faktoren ersten Grades zerlegen läßt:

$$P(t) = \alpha + 2\beta t + \gamma t^2 \quad (\alpha > 0, \gamma > 0, \alpha\gamma - \beta^2 > 0).$$

Diesem Polynom geben wir mit Hilfe der Formel

$$[(1-t) + t]^k = \sum_{s=0}^k \binom{k}{s} t^s (1-t)^{k-s} = 1$$

die Gestalt

$$P(t) = \alpha \sum_{s=0}^p \binom{p}{s} t^s (1-t)^{p-s} + 2\beta t \sum_{s=1}^p \binom{p-1}{s-1} t^{s-1} (1-t)^{p-s} + \gamma t^2 \sum_{s=0}^p \binom{p-2}{s-2} t^{s-2} (1-t)^{p-s}$$

oder, wenn wir die Glieder mit gleichen Potenzen zusammenfassen,

$$P(t) = \sum_{s=0}^p \frac{(p-2)!}{s! (p-s)!} t^s (1-t)^{p-s} [p(p-1) \alpha + 2s(p-1) \beta + s(s-1) \gamma]. \quad (321)$$

Der Ausdruck in den eckigen Klammern ist für alle reellen s und für alle hinreichend großen Werte von p positiv, denn die Diskriminante dieses dreigliedrigen Ausdrucks in bezug auf s ,

$$p(p-1) \alpha \gamma - \frac{1}{4} (2p\beta - 2\beta - \gamma)^2 = p^2(\alpha\gamma - \beta^2) + p(2\beta^2 + 2\gamma\beta + \alpha) - \frac{1}{4} (2\beta + \gamma)^2,$$

ist wegen $\alpha\gamma - \beta^2 > 0$ für alle hinreichend großen Werte von p positiv. Daher führt die Gleichung (321) gerade auf die Beziehung (320) mit positiven c_s für alle hinreichend großen Werte von p . Wir betrachten nun ein beliebiges positives Polynom im Intervall $[0, 1]$. Es läßt sich als Produkt positiver Polynome ersten und zweiten Grades mit nicht-reellen Wurzeln darstellen. Für jeden Faktor gilt eine Zerlegung (320). Damit erhalten wir eine ebensolche Zerlegung auch für ihr Produkt, wobei der Grad p gleich der Summe der Gradzahlen der einzelnen Faktoren ist.

Bemerkung. Mit Hilfe der Variablentransformation $t_1 = \frac{t-a}{b-a}$ können wir ein beliebiges endliches Intervall $a \leq t \leq b$ auf das Intervall $0 \leq t_1 \leq 1$ zurückführen, und für Polynome, die im Intervall $[a, b]$ positiv sind, erhalten wir die Formel

$$P(t) = \sum_{s=0}^p c_s (t-a)^s (b-t)^{p-s} \quad (c_s > 0). \quad (322)$$

Hilfssatz 4. Ist m die untere und M die obere Grenze des selbstadjungierten Operators A , d. h. des quadratischen Funktionals (Ax, x) für $\|x\| = 1$, und ist $P(t)$ ein im Intervall $[m, M]$ nichtnegatives Polynom, so ist $P(A)$ ein positiver Operator,

$$(P(A)x, x) \geq 0. \quad (323)$$

Beweis. Es genügt, diesen Hilfssatz für den Fall zu beweisen, daß $P(t)$ im Intervall $[m, M]$ positiv ist. Nehmen wir nämlich an, daß er für diesen Fall bewiesen und $Q(t)$ ein im Intervall $[m, M]$ nichtnegatives Polynom ist, und setzen wir $P(t) = Q(t) + \varepsilon$ mit $\varepsilon > 0$, so haben wir $P(t) > 0$ im Intervall $[m, M]$; folglich ist nach unserer Annahme

$$((Q(A) + \varepsilon)x, x) = (Q(A)x, x) + \varepsilon(x, x) \geq 0.$$

Durch Grenzübergang $\varepsilon \rightarrow 0$ erhalten wir die Ungleichung (323) auch für $Q(A)$. Wir gehen nun zum Beweis des Hilfssatzes für positives $P(t)$ über. Wegen (322) für $a = m$ und $b = M$ brauchen wir nur nachzuweisen, daß der Operator

$$(A - mE)^s (ME - A)^{p-s} \quad (324)$$

positiv ist, wobei wir die Zahl p als ungerade annehmen (die Summe positiver Operatoren ist wieder positiv). Wir nehmen etwa an, daß s gerade ist, $s = 2j$, und stellen den Operator (324) in der Gestalt $A_2^j A_1$ mit

$$A_1 = ME - A, \quad A_2 = (A - mE)^j (ME - A)^{(p-2j-1)/2}$$

dar; dabei ist A_2 mit A_1 vertauschbar, und A_1 ist positiv, denn es ist $(A_1x, x) = M - (Ax, x) \geq 0$ für $\|x\| = 1$. Mit Hilfssatz 2 können wir bestätigen, daß der Operator (324) positiv ist. Für ungerades s haben wir $A_1 = A - mE$ zu nehmen.

Folgerung 1. *Genügen im Intervall $[m, M]$ die Polynome $P_1(t)$ und $P_2(t)$ der Ungleichung $P_2(t) \geq P_1(t)$, so gilt $P_2(A) \geq P_1(A)$.*

Gilt nämlich speziell $|P(t)| \leq \varepsilon$, d. h. $-\varepsilon \leq P(t) \leq \varepsilon$, so ist $-\varepsilon E \leq P(A) \leq \varepsilon E$ oder $-\varepsilon \leq (P(A)x, x) \leq \varepsilon$ für $\|x\| = 1$; folglich ist die Norm von $P(A)$ nicht größer als ε [126].

Folgerung 2. *Aus der Folgerung 1 ergibt sich: Strebt die Folge der Polynome $P_n(t)$ gleichmäßig im Intervall $[m, M]$ gegen ein Polynom $P(t)$, so strebt $P_n(A)$ gegen $P(A)$, und die Norm der Differenz $P(A) - P_n(A)$ strebt sogar gegen 0.*

160. Die Potenzreihe eines Operators. Wir erinnern an den in [181] bewiesenen Hilfssatz, der besagt: Sind die Normen einer Folge von Operatoren A_n ($n = 1, 2, 3, \dots$) nicht größer als die positiven Zahlen δ_n , die eine konvergente Reihe bilden, so ist die Reihe

$A = \sum_{n=0}^{\infty} A_n$ konvergent, und die Norm des Operators A ist nicht größer als die Summe der Zahlen δ_n . Gibt es speziell eine Potenzreihe $\sum_{n=0}^{\infty} a_n t^n$, die im Intervall $|t| \leq k$ absolut konvergent ist, und ist die Norm des Operators A nicht größer als k , so ist die Reihe $\sum_{n=0}^{\infty} a_n A^n$ konvergent. Für das Folgende legen wir nur die binomische Reihe

$$\sqrt{1+t} = \sum_{n=0}^{\infty} \binom{\frac{1}{2}}{n} t^n, \quad |t| \leq 1, \quad (325)$$

mit

$$\binom{\frac{1}{2}}{0} = 1, \quad \binom{\frac{1}{2}}{n} = \frac{\frac{1}{2} \left(\frac{1}{2} - 1\right) \left(\frac{1}{2} - 2\right) \dots \left(\frac{1}{2} - n + 1\right)}{n!} \quad (326)$$

zugrunde. Die Beziehung (325) liefert den positiven Wert der Quadratwurzel und bleibt auch für $t = \pm 1$ richtig [I, 181]. Die Entwicklungskoeffizienten (326) sind positiv für ungerade $n > 0$ und negativ für gerade $n > 0$. Folglich sind, wenn wir in (325) $t = -1$ annehmen, alle Glieder außer dem ersten negativ, woraus

$$0 = 1 - \sum_{n=1}^{\infty} \left| \binom{\frac{1}{2}}{n} \right| \quad \text{oder} \quad \sum_{n=1}^{\infty} \left| \binom{\frac{1}{2}}{n} \right| = 1 \quad (327)$$

folgt, und die Reihe (325) konvergiert absolut und gleichmäßig für $|t| \leq 1$ [I, 146]. Wenn wir in (325) die Größe t durch $t^2 - 1$ ersetzen, erhalten wir auf der linken Seite den absoluten Betrag der Quadratwurzel aus t^2 , d. h. den absoluten Betrag $|t|$; für ihn ergibt sich die folgende Entwicklung in eine im Intervall $|t| \leq 1$ absolut konvergente Reihe:

$$|t| = \sum_{n=0}^{\infty} \binom{\frac{1}{2}}{n} (t^2 - 1)^n. \quad (328)$$

Diese Entwicklung verwenden wir für selbstadjungierte Operatoren. Es sei A ein selbstadjungierter Operator mit der Norm m_A . Wir bilden den selbstadjungierten Operator $C = \frac{1}{m_A^2} A^2 - E$. Es gilt

$$(Cx, x) = \frac{1}{m_A^2} (A^2 x, x) - \|x\|^2 = \frac{1}{m_A^2} \|Ax\|^2 - \|x\|^2,$$

woraus $-1 \leq (Cx, x) \leq 0$ für $\|x\| = 1$ folgt. Die Norm von C ist also höchstens gleich 1. Wir können die Reihe

$$B = m_A \sum_{n=0}^{\infty} \left(\frac{1}{2}\right)^n C^n = m_A \sum_{n=0}^{\infty} \left(\frac{1}{2}\right)^n \left(\frac{1}{m_A^2} A^2 - E\right)^n \quad (329)$$

bilden. Bezeichnen wir mit $S_n(t)$ die Partialsumme der Reihe (328), so gilt $S_n^2(t) \rightarrow t^2$ gleichmäßig im Intervall $[-1, 1]$. Folglich strebt $S_n^2\left(\frac{1}{m_A}\right)$ gegen $\frac{1}{m_A^2} A^2$, und durch Grenzübergang erhält man, daß der durch (329) definierte selbstadjungierte Operator B der Bedingung $B^2 = A^2$ genügt.

Ferner ist der Operator D , wenn er mit A vertauschbar ist, auch mit der Partialsumme der Reihe (329) und folglich (nach Grenzübergang) mit B vertauschbar. Daraus folgt insbesondere, daß A mit B vertauschbar ist, $AB = BA$.

Wir zeigen noch, daß B ein positiver Operator ist. Da die Norm von C nicht größer als 1 ist, erhalten wir $|(C^n x)| \leq \|x\|^2$ und gelangen, wenn wir den Ausdruck (329) in der Gestalt

$$B = m_A \left[E + \sum_{n=1}^{\infty} \left(\frac{1}{2}\right)^n C^n \right]$$

schreiben, zu der Ungleichung

$$(Bx, x) \geq m_A \left[(x, x) - \sum_{n=1}^{\infty} \left| \left(\frac{1}{2}\right)^n (C^n x, x) \right| \right] \geq m_A \left[1 - \sum_{n=1}^{\infty} \left| \left(\frac{1}{2}\right)^n \right| \right] \|x\|^2,$$

aus der wegen (327) $(Bx, x) \geq 0$ folgt.

Wir erkennen also die folgenden Eigenschaften: Der Operator B ist selbstadjungiert und positiv, läßt sich mit A vertauschen und genügt der Beziehung $B^2 = A^2$; jeder mit A vertauschbare Operator ist auch mit B vertauschbar. In [161] konstruieren wir durch Verwendung des Operators B und des Hilfssatzes 1 die Spektralschar $\mathfrak{E}_\lambda$ des Operators A und beweisen die grundlegende Gleichung (204) aus [142].

161. Die Spektralschar. Wir beweisen nun den folgenden Satz:

Satz. *Jedem selbstadjungierten Operator A ist ein Projektionsoperator $\mathfrak{E}_0$ mit folgenden Eigenschaften zugeordnet:*

- Ist ein Operator D mit A vertauschbar, so ist er auch mit $\mathfrak{E}_0$ vertauschbar;*
- ist $Az = 0$, so gilt $\mathfrak{E}_0 z = z$;*
- die selbstadjungierten Operatoren $A\mathfrak{E}_0$ und $A(E - \mathfrak{E}_0)$ genügen den Bedingungen*

$$A\mathfrak{E}_0 \geq 0, \quad A(E - \mathfrak{E}_0) \geq 0. \quad (330)$$

Beweis. Wir verwenden für $\mathfrak{E}_0$ den in Hilfssatz 1 aus [159] auftretenden Projektionsoperator P . Ist der Operator D mit A vertauschbar, so ist er auch mit B und daher auch mit $A + B$ vertauschbar, und die ersten beiden Behauptungen des Satzes ergeben sich aus Hilfssatz 1. Ferner folgen aus $A = (E - 2\mathfrak{E}_0)B$, aus der Vertauschbarkeit von $\mathfrak{E}_0$ mit A und B sowie aus $\mathfrak{E}_0^2 = \mathfrak{E}_0$ die Gleichungen

$$A\mathfrak{E}_0 = -B\mathfrak{E}_0, \quad A(E - \mathfrak{E}_0) = B(E - \mathfrak{E}_0).$$

Nun sind aber die Produkte des positiven Operators B mit den Projektionsoperatoren $\mathfrak{E}_0$ und $E - \mathfrak{E}_0$, mit denen er vertauschbar ist, positive Operatoren; daher folgen aus den letzten Gleichungen unmittelbar die Bedingungen (330). Damit ist der Satz bewiesen.

Es sei λ eine beliebige reelle Zahl. Wir können für den selbstadjungierten Operator $A - \lambda E$ den im letzten Satz betrachteten Projektionsoperator bilden, den wir mit $\mathfrak{E}_\lambda$ bezeichnen. Er besitzt die folgenden Eigenschaften: a) Ist ein Operator D mit $A - \lambda E$ oder, was dasselbe ist, mit A vertauschbar, so ist er auch mit $\mathfrak{E}_\lambda$ vertauschbar; b) ist

$(A - \lambda E)z = 0$, so ist $\mathfrak{E}_\lambda z = z$; c) es gelten die Ungleichungen

$$(A - \lambda E) \mathfrak{E}_\lambda \leq 0, \quad (A - \lambda E) (E - \mathfrak{E}_\lambda) \geq 0. \quad (331)$$

Wir weisen noch darauf hin, daß $\mathfrak{E}_\lambda$ für jedes λ mit $A - \lambda E$, also mit A vertauschbar ist, und wollen zeigen, daß $\mathfrak{E}_\lambda$ eine Zerlegung der Einheit ist. Jedes $\mathfrak{E}_{\lambda_1}$ ist mit A und daher nach dem oben Gesagten auch mit einem beliebigen $\mathfrak{E}_{\lambda_2}$ vertauschbar. Dann gilt $\mathfrak{E}_\lambda = 0$ für $\lambda < m$. Wäre das nicht der Fall, so gäbe es ein Element x , dessen Norm gleich 1 ist und für das $\mathfrak{E}_\lambda x = x$ gilt; also wäre

$$((A - \lambda E) \mathfrak{E}_\lambda x, x) = ((A - \lambda E) x, x) = (Ax, x) - \lambda > 0,$$

denn es ist $\lambda < m$. Das widerspricht jedoch der ersten Ungleichung (331). Somit ist $\mathfrak{E}_\lambda = 0$ für $\lambda < m$. Durch Benutzung der zweiten Ungleichung (331) läßt sich ganz analog $\mathfrak{E}_\lambda = E$ für $\lambda > M$ nachweisen. Es bleibt noch zu zeigen, daß $\mathfrak{E}_\lambda \leq \mathfrak{E}_\mu$, also $\mathfrak{E}_\lambda \mathfrak{E}_\mu = \mathfrak{E}_\lambda$ für $\lambda < \mu$ oder, was auf dasselbe hinausläuft, die Gleichung

$$\mathfrak{E}_\lambda (E - \mathfrak{E}_\mu) = 0 \quad (332)$$

gilt. Die linke Seite dieser Gleichung bezeichnen wir mit R :

$$\mathfrak{E}_\lambda (E - \mathfrak{E}_\mu) = (E - \mathfrak{E}_\mu) \mathfrak{E}_\lambda = R. \quad (333)$$

Für jedes Element x muß also die Beziehung $Rx = 0$ gelten. Wir führen die Bezeichnung $Rx = y$ ein. Aus (333) folgt unmittelbar

$$\left. \begin{aligned} \mathfrak{E}_\lambda R &= \mathfrak{E}_\lambda^2 (E - \mathfrak{E}_\mu) = \mathfrak{E}_\lambda (E - \mathfrak{E}_\mu) = R \\ \text{und analog} \quad (E - \mathfrak{E}_\mu) &= R = R. \end{aligned} \right\} \quad (334)$$

Wegen (331) gilt

$$((A - \lambda E) \mathfrak{E}_\lambda y, y) \leq 0, \quad ((A - \mu E) (E - \mathfrak{E}_\mu) y, y) \geq 0. \quad (335)$$

Andererseits ist wegen (334)

$$\mathfrak{E}_\lambda y = \mathfrak{E}_\lambda R x = R x = y, \quad (E - \mathfrak{E}_\mu) y = (E - \mathfrak{E}_\mu) R x = R x = y.$$

Damit erhält die erste Ungleichung (335) die Form $((A - \lambda E) y, y) \leq 0$, die zweite Ungleichung (335) analog $((A - \mu E) y, y) \geq 0$. Subtraktion der zweiten Ungleichung von der ersten liefert $((\mu - \lambda) y, y) \leq 0$ oder $(\mu - \lambda) \|y\|^2 \leq 0$, woraus wegen $\lambda < \mu$ die Beziehung $y = 0$, also $Rx = 0$ folgt. Damit ist (332) bewiesen. Die rechtsseitige Stetigkeit von $\mathfrak{E}_\lambda$ werden wir später beweisen.

Für den Beweis der Integraldarstellung des Operators A mit Hilfe von $\mathfrak{E}_\lambda$ leiten wir eine Ungleichung her. Wir führen den Projektionsoperator

$$\Delta = \mathfrak{E}_\mu - \mathfrak{E}_\lambda \quad (\mu > \lambda) \quad (336)$$

ein. Für jedes x ist

$$((A - \mu E) \mathfrak{E}_\mu \Delta x, \Delta x) \leq 0, \quad ((A - \lambda E) (E - \mathfrak{E}_\lambda) \Delta x, \Delta x) \geq 0.$$

Wenn wir die offenbar gültigen Gleichungen $\Delta^2 = \Delta$, $\mathfrak{E}_\mu \Delta = (E - \mathfrak{E}_\lambda) \Delta = \Delta$ berücksichtigen, können wir die vorhergehenden Ungleichungen in der Gestalt

$$((A - \mu E) \Delta x, x) \leq 0, \quad ((A - \lambda E) \Delta x, x) \geq 0$$

oder

$$\lambda (\Delta x, x) \leq (A \Delta x, x) \leq \mu (\Delta x, x)$$

schreiben. Mit einer beliebigen, der Bedingung $\lambda \leq \nu \leq \mu$ genügenden Zahl ν erhalten wir daraus $|((A - \nu E) \Delta x, x)| \leq (\mu - \lambda) (\Delta x, x)$ oder, wenn wir $(\Delta x, x) = \|\Delta x\|^2 \leq \|x\|^2$ berücksichtigen, $|((A - \nu E) \Delta x, x)| \leq (\mu - \lambda) \|x\|^2$.

Daraus erkennt man [126], daß die Norm des Operators $(A - \nu E) \Delta$ höchstens gleich $\mu - \lambda$ ist, also

$$|(A - \nu E) \Delta x| \leq (\mu - \lambda) \|x\|.$$

Wenn wir in dieser Ungleichung x durch Δx ersetzen und $\Delta^2 = \Delta$ berücksichtigen, so erhalten wir die für die weiteren Betrachtungen grundlegende Ungleichung

$$\|A \Delta x - \nu \Delta x\| \leq (\mu - \lambda) \|\Delta x\|. \quad (337)$$

Hier ist $\lambda \leq \nu \leq \mu$, und Δ wird durch (336) definiert.

Wir gehen nun zum Beweis der Gleichung (204) über. Wir wählen eine gewisse positive Zahl ε_0 und zerlegen das Intervall $[m - \varepsilon_0, M]$ in Teilintervalle:

$$m - \varepsilon_0 = \lambda_0 < \lambda_1 < \lambda_2 < \dots < \lambda_{n-1} < \lambda_n = M;$$

hierauf führen wir die Projektionsoperatoren $\Delta_k = \mathcal{E}_{\lambda_k} - \mathcal{E}_{\lambda_{k-1}}$ ein, für die

$$E = \sum_{k=1}^n \Delta_k \quad \text{und} \quad \Delta_k \Delta_l = 0 \quad \text{für} \quad k \neq l \quad (338)$$

gilt. Für jedes Element x gilt die Zerlegung $x = \sum_{k=1}^n \Delta_k x = \sum_{k=1}^n x_k$ in paarweise orthogonale Summanden und die Formel $Ax = \sum_{k=1}^n Ax_k$. Man erkennt leicht, daß

$$(Ax_k, Ax_l) = 0 \quad \text{und} \quad (Ax_k, x_l) = 0 \quad \text{für} \quad k \neq l$$

ist. Die erste dieser beiden Gleichungen z. B. läßt sich folgendermaßen beweisen:

$$(Ax_k, Ax_l) = (A \Delta_k x, A \Delta_l x) = (A \Delta_l \Delta_k x, Ax);$$

der letzte Ausdruck ist wegen (338) gleich 0. Wir bilden nun

$$Ax - \sum_{k=1}^n \nu_k \Delta_k x = \sum_{k=1}^n (Ax_k - \nu_k x_k);$$

dabei ist ν_k ein beliebiger Wert aus dem Intervall $[\lambda_{k-1}, \lambda_k]$. Da die Summanden auf der rechten Seite paarweise orthogonal sind, kann man mit Hilfe des Satzes des PYTHAGORAS

$$\left\| Ax - \sum_{k=1}^n \nu_k \Delta_k x \right\|^2 = \sum_{k=1}^n \|Ax_k - \nu_k x_k\|^2 \quad (339)$$

schreiben. Es sei δ die größte der Differenzen $\lambda_k - \lambda_{k-1}$. Wenn wir die Ungleichung (337) verwenden, erhalten wir aus (339)

$$\left\| Ax - \sum_{k=1}^n \nu_k \Delta_k x \right\|^2 \leq \delta^2 \sum_{k=1}^n \|x_k\|^2$$

oder nach dem Satz des PYTHAGORAS

$$\left\| Ax - \sum_{k=1}^n \nu_k \Delta_k x \right\|^2 \leq \delta^2 \|x\|^2.$$

Hieraus folgt, daß für jedes Element x im Fall $\delta \rightarrow 0$

$$Ax = \lim_{\delta \rightarrow 0} \sum_{k=1}^n \nu_k \Delta_k x$$

gilt. Somit erhalten wir die grundlegende Gleichung

$$A = \int_m^M \lambda d\mathcal{E}_\lambda.$$

Es bleibt noch zu beweisen, daß $\mathcal{E}_\lambda$ rechtsseitig stetig ist. Wenn μ gegen λ strebt, nimmt der durch (336) definierte Projektionsoperator Δ nicht zu und strebt gegen einen

Grenzoperator Δ_0 . Wir haben zu zeigen, daß Δ_0 der Nulloperator ist. Wenn wir in (337) den Grenzübergang vornehmen, erhalten wir $(A - \lambda E) \Delta_0 x = 0$. Daraus folgt wegen der zweiten Eigenschaft von $\mathfrak{S}_\lambda$ die Beziehung $\mathfrak{S}_\lambda \Delta_0 x = \Delta_0 x$ oder $(\mathfrak{S} - \mathfrak{S}_\lambda) \Delta_0 x = 0$. Andererseits ist $(E - \mathfrak{S}_\lambda) A = A$, und durch Grenzübergang erhalten wir $(E - \mathfrak{S}_\lambda) \Delta_0 = \Delta_0$. Aus $(E - \mathfrak{S}_\lambda) \Delta_0 x = 0$ folgt $\Delta_0 x = 0$, d. h., Δ_0 annulliert wirklich jedes Element x . Wir erwähnen noch, daß ein beliebiger, mit A vertauschbarer Operator D auch mit $\mathfrak{S}_\lambda$ vertauschbar ist.

§ 2. Die Räume l_2 und L_2

162. Lineare Operatoren im Raum l_2 . Wir wenden jetzt die allgemeine Theorie auf die Räume l_2 und L_2 an. Oben sahen wir, daß wir durch die Wahl eines vollständigen normierten Orthogonalsystems in H zu einer eindeutigen Zuordnung der Elemente des abstrakten Raumes H und des Raumes l_2 gelangen. Wir betrachten l_2 als eine Realisierung des Raumes H , da in l_2 bei der gewöhnlichen Definition der algebraischen Operationen und des Skalarprodukts alle Axiome von H gelten.

Wir führen nun den Begriff des Abschnittselements ein (vgl. [134]). Es sei $x(\xi_1, \xi_2, \dots)$ ein Element aus l_2 , und $x^{(k)}(\xi_1, \xi_2, \dots, \xi_k, 0, \dots)$ habe die ersten k Komponenten wie x , jedoch seien die übrigen Komponenten gleich 0. Dieses Element $x^{(k)}$ nennen wir das *Abschnittselement* von x . Für $x^{(k)}$ gilt

$$\|x - x^{(k)}\|^2 = \sum_{m=k+1}^{\infty} |\xi_m|^2 \rightarrow 0 \quad \text{für } k \rightarrow \infty, \quad (1)$$

also $x^{(k)} \rightarrow x$ für $m \rightarrow \infty$. Mit $\varphi_1, \varphi_2, \dots$ bezeichnen wir die Einheitsvektoren aus l_2 , d. h., φ_k besitzt die Komponente $\xi_k = 1$, die übrigen Komponenten sind gleich 0. Für ein Element x gilt also

$$x = \sum_{m=1}^{\infty} \xi_m \varphi_m. \quad (2)$$

Ist A ein linearer Operator in l_2 und ist $x' = Ax$, so gilt, wenn wir die Bezeichnung $x' = (\xi'_1, \xi'_2, \dots)$ einführen,

$$\xi'_n = (x', \varphi_n) = \sum_{m=1}^{\infty} a_{nm} \xi_m \quad (n=1, 2, \dots), \quad (3)$$

wobei

$$a_{nm} = (A\varphi_m, \varphi_n) \quad (4)$$

ist. Ein linearer Operator in l_2 läßt sich also durch eine Matrix mit den Elementen (4) darstellen. Dem adjungierten Operator A^* entspricht die Matrix mit den Elementen

$$a_{nm}^* = (A^*\varphi_m, \varphi_n) = (\varphi_m, A\varphi_n) = \overline{a_{mn}} \quad (5)$$

(vgl. [134]), und ein selbstadjungierter Operator ist durch die Gleichung

$$a_{nm} = \overline{a_{mn}} \quad (6)$$

charakterisiert. Für ein bilineares Funktional (Ax, y) gilt

$$(Ax, y) = (x, A^*y) = \sum_{n=1}^{\infty} \left(\sum_{m=1}^{\infty} a_{nm} \xi_m \right) \overline{\eta_n} = \sum_{m=1}^{\infty} \xi_m \left(\sum_{n=1}^{\infty} a_{nm} \overline{\eta_n} \right), \quad (7)$$

wenn $\eta_1, \eta_2, \dots$ die Komponenten von y sind.

Bilden wir die Abschnittselemente $x^{(k)}$ und $y^{(l)}$ für x bzw. y , so erhalten wir

$$(Ax^{(k)}, y^{(l)}) = \sum_{n=1}^l \sum_{m=1}^k a_{nm} \xi_m \overline{\eta_n}.$$

Es strebt aber $(Ax^{(k)}, y^{(l)})$ gegen (Ax, y) für $k \rightarrow \infty$ und $l \rightarrow \infty$; folglich gilt

$$\sum_{n=1}^{\infty} \left(\sum_{m=1}^{\infty} a_{nm} \xi_m \right) \overline{\eta_n} = \lim_{\substack{k \rightarrow \infty \\ l \rightarrow \infty}} \sum_{n=1}^l \sum_{m=1}^k a_{nm} \xi_m \overline{\eta_n}. \quad (8)$$

Sind a_{pq} und b_{pq} die Elemente der den Operatoren A bzw. B zugeordneten Matrizen, so ist dem Operator $D = BA$ die Matrix mit den Elementen d_{pq} zugeordnet, die durch

$$d_{pq} = (D\varphi_q, \varphi_p) = (BA\varphi_q, \varphi_p) = (A\varphi_q, B^*\varphi_p)$$

oder auf Grund der Formel für das Skalarprodukt in l_2 durch

$$d_{pq} = \sum_{s=1}^{\infty} (A\varphi_q, \varphi_s) \overline{(B^*\varphi_p, \varphi_s)} = \sum_{s=1}^{\infty} a_{sq} \overline{b_{sp}^*}$$

bestimmt werden. Wenn wir (5) berücksichtigen, so erhalten wir schließlich

$$d_{pq} = \sum_{s=1}^{\infty} b_{ps} a_{sq}. \quad (9)$$

Bezeichnen wir die unendlichen Matrizen mit denselben Buchstaben A bzw. B wie die zugeordneten Operatoren und die Elemente dieser Matrizen mit $\{A\}_{pq}$ bzw. $\{B\}_{pq}$, so können wir (9) in der Gestalt

$$\{BA\}_{pq} = \sum_{s=1}^{\infty} \{B\}_{ps} \{A\}_{sq} \quad (10)$$

schreiben. Wenn drei lineare beschränkte Operatoren A , B und C gegeben sind, so läßt sich mit Hilfe des Assoziativgesetzes $(CB)A = C(BA)$ für die Änderung der Reihenfolge der Summation die Gleichung

$$\sum_{t=1}^{\infty} \left(\sum_{s=1}^{\infty} \{C\}_{ps} \{B\}_{st} \right) \{A\}_{tq} = \sum_{s=1}^{\infty} \{C\}_{ps} \left(\sum_{t=1}^{\infty} \{B\}_{st} \{A\}_{tq} \right) \quad (11)$$

aufstellen.

163. Beschränkte Operatoren. Wie wir wissen, erzeugt jeder beschränkte lineare Operator A eine unendliche Matrix von Elementen a_{pq} . Wir stellen nun die umgekehrte Frage: Wie müssen die Elemente a_{pq} einer unendlichen Matrix beschaffen sein, damit die Gleichung (3) einen beschränkten linearen Operator in l_2 ergibt? Wir verlangen also, daß die Reihen in (3) für ein beliebiges Element $(\xi_1, \xi_2, \dots)$ aus l_2 konvergieren und daß eine Zahl N existiert, mit der für jedes Element $x \in l_2$ die Ungleichung

$$\sum_{n=1}^{\infty} \left| \sum_{k=1}^{\infty} a_{nk} \xi_k \right|^2 \leq N^2 \sum_{k=1}^{\infty} |\xi_k|^2 \quad (12)$$

erfüllt ist. Wir erinnern daran, daß für ein bilineares Funktional wegen der Beschränktheit des Operators A die Abschätzung $|(Ax, y)| \leq N \|x\| \cdot \|y\|$ gilt. Wenn wir diese Abschätzung auf Abschnittselemente anwenden, erhalten wir für die a_{pq} die notwendige Bedingung

$$\left| \sum_{n=1}^l \sum_{m=1}^k a_{nm} \xi_m \overline{\eta_n} \right|^2 \leq N^2 \sum_{m=1}^k |\xi_m|^2 \cdot \sum_{n=1}^l |\eta_n|^2, \quad (13)$$

in der nur endliche Summen auftreten. Diese Bedingung ist aber nicht nur notwendig, sondern auch hinreichend. Es gilt also der folgende Satz:

Satz 1. Die a_{pq} sind dann und nur dann die Elemente der Matrix einer linearen beschränkten Transformation, wenn für beliebige natürliche k und l und für beliebige komplexe Zahlen ξ_m und η_n die Bedingung (13) bei einer bestimmten Wahl der (nicht von ξ_p, η_q, k und l abhängigen) Zahl N erfüllt ist.

Beweis. Die Notwendigkeit dieser Bedingung haben wir vorhin schon erkannt. Wir beweisen nun, daß die Bedingung hinreichend ist. Es sei $(\xi_1, \xi_2, \dots)$ ein beliebiges Element aus l_2 . In (13) setzen wir $l = k$ und

$$\eta_n = \sum_{m=1}^k a_{nm} \xi_m \quad (n = 1, 2, 3, \dots, k).$$

Dann ergibt (13)

$$\left(\sum_{n=1}^k \left| \sum_{m=1}^k a_{nm} \xi_m \right|^2 \right)^2 \leq N^2 \sum_{m=1}^k |\xi_m|^2 \cdot \sum_{n=1}^k \left| \sum_{m=1}^k a_{nm} \xi_m \right|^2$$

oder

$$\sum_{n=1}^k \left| \sum_{m=1}^k a_{nm} \xi_m \right|^2 \leq N^2 \sum_{m=1}^k |\xi_m|^2,$$

also erst recht

$$\sum_{n=1}^k \left| \sum_{m=1}^k a_{nm} \xi_m \right|^2 \leq N^2 \sum_{m=1}^{\infty} |\xi_m|^2. \quad (14)$$

Wir zeigen, daß aus diesen Ungleichungen für jedes Element aus l_2 die Konvergenz der Reihen

$$\sum_{m=1}^{\infty} a_{nm} \xi_m \quad (n = 1, 2, \dots) \quad (15)$$

folgt. Diese Reihe möge bei einer gewissen Wahl des Elements $(\xi_1^{(0)}, \xi_2^{(0)}, \dots)$ und der Zahl n divergieren. Dann divergiert erst recht die Reihe $\sum_{m=1}^{\infty} |a_{nm} \xi_m^{(0)}|$, und die endliche Summe $\sum_{m=1}^k |a_{nm} \xi_m^{(0)}|$ dieser Reihe wächst mit k unbegrenzt an. Wir ändern die Argumente der komplexen Zahlen $\xi_m^{(0)}$ so, daß die Produkte $a_{nm} \xi_m^{(0)}$ positive Zahlen sind. Wenn wir auf das in dieser Weise erhaltene Element von l_2 die Ungleichung (14) anwenden und auf der linken Seite alle Summanden außer denen, die zu dem oben gewählten Wert von n gehören, weglassen, erhalten wir

$$\left(\sum_{m=1}^k a_{nm} \xi_m^{(0)} \right)^2 \leq N^2 \sum_{m=1}^{\infty} |\xi_m^{(0)}|^2.$$

Die linke Seite dieser Ungleichung wächst mit k unbegrenzt an; es ergibt sich also ein Widerspruch. Folglich konvergieren tatsächlich alle Reihen (15) für jedes Element x .

Wir zeigen nun, daß aus (14) die Ungleichung

$$\sum_{n=1}^k \left| \sum_{m=1}^{\infty} a_{nm} \xi_m \right|^2 \leq N^2 \sum_{m=1}^{\infty} |\xi_m|^2 \quad (16)$$

folgt. Wäre für ein gewisses k und ein Element x aus l_2 die entgegengesetzte Ungleichung erfüllt, so würde für eben dieses k und für ein hinreichend großes l (wir könnten dabei offenbar $l \geq k$ annehmen)

$$\sum_{n=1}^k \left| \sum_{m=1}^l a_{nm} \xi_m \right|^2 > N^2 \sum_{m=1}^{\infty} |\xi_m|^2,$$

also erst recht

$$\sum_{n=1}^l \left| \sum_{m=1}^l a_{nm} \xi_m \right|^2 > N^2 \sum_{m=1}^{\infty} |\xi_m|^2$$

gelten, was in Widerspruch zu (14) für $k=l$ steht. Damit ist die Ungleichung (16) bewiesen. Wenn wir in ihr k unbegrenzt wachsen lassen, gelangen wir zu (12). Damit ist der Satz bewiesen.

Bemerkung. Die Bedingung (13) haben wir bei dem Beweis, daß sie hinreichend ist, nur für $l=k$ benutzt. Wir zeigen jetzt, daß man diese Bedingung nur für die quadratische Form zu verifizieren braucht, d. h., daß der Operator sicher beschränkt ist, wenn für jedes k die Ungleichung

$$\left| \sum_{n,m=1}^k a_{nm} \xi_m \overline{\xi_n} \right| \leq N \sum_{m=1}^k |\xi_m|^2 \quad (17)$$

erfüllt ist. Berücksichtigen wir die Gleichung, in der ein bilineares Funktional durch die zugeordnete quadratische Form ausgedrückt wird, und die Ungleichung (17), so können wir

$$\left| \sum_{m,n=1}^k a_{nm} \xi_m \overline{\eta_n} \right| \leq \frac{N}{4} \left[\sum_{m=1}^k |\xi_m + \eta_m|^2 + \sum_{m=1}^k |\xi_m - \eta_m|^2 + \sum_{m=1}^k |\xi_m + i\eta_m|^2 + \sum_{m=1}^k |\xi_m - i\eta_m|^2 \right]$$

schreiben. Nehmen wir an, daß die Normen der Elemente x und y gleich 1 sind, so erhalten wir unter Berücksichtigung von $|\alpha + \beta|^2 \leq 2(|\alpha|^2 + |\beta|^2)$

$$\left| \sum_{m,n=1}^k a_{nm} \xi_m \overline{\eta_n} \right| \leq 4N.$$

Für Elemente von beliebiger Norm erhalten wir dagegen

$$\left| \sum_{m,n=1}^k a_{nm} \xi_m \overline{\eta_n} \right| \leq 4N \left[\sum_{m=1}^k |\xi_m|^2 \right]^{1/2} \left[\sum_{m=1}^k |\eta_m|^2 \right]^{1/2},$$

d. h., aus (17) folgt die Bedingung (13) für $l=k$ und folglich die Beschränktheit des durch die Matrix der a_{nm} definierten Operators.

Wir erwähnen noch einige mit der Bedingung (13) verknüpfte Umstände. Wird die Bedingung (13) von den a_{pq} erfüllt, so genügen die Elemente $\{A^*\}_{pq} = \overline{a_{qp}}$ der dem adjungierten Operator zugeordneten Matrix offenbar gleichfalls dieser Bedingung, was auch aus der allgemeinen Theorie folgt. Wir führen noch die Matrix des transponierten Operators und die Matrix des konjugiert komplexen Operators ein:

$$\{A'\}_{pq} = a_{qp}, \quad \{\bar{A}\}_{pq} = \overline{a_{pq}}. \quad (18)$$

Offenbar gilt

$$A^* = (\bar{A})' = \overline{A'}, \quad (19)$$

und die Elemente der Matrizen $\overline{A'}$ und $\overline{A}$ genügen der Bedingung (13), wenn die Elemente der Ausgangsmatrix A dieser Bedingung genügen. Aus (13) ergibt sich unmittelbar, daß alle a_{pq} dem absoluten Betrag nach durch eine nicht von p und q abhängige Zahl beschränkt sind, und zwar erhalten wir, wenn wir $\xi_p = \eta_q = 1$ und die übrigen ξ_m und η_n gleich 0 annehmen, $|a_{pq}| \leq N$.

Es gibt noch eine notwendige Bedingung, der die Elemente der Matrix einer beschränkten Transformation genügen müssen. Aus (4) folgt, daß die a_{nk} die Komponenten des Elements $A\varphi_k$ sind. Daher muß die Reihe, die aus den Quadraten der absoluten Beträge der Elemente einer beliebigen Spalte besteht, konvergieren:

$$\sum_{n=1}^{\infty} |a_{nk}|^2 < \infty \quad (k=1, 2, \dots). \quad (20)$$

Wenn wir zu A^* übergehen, sehen wir, daß dasselbe auch bezüglich der Zeilen gilt:

$$\sum_{m=1}^{\infty} |a_{km}|^2 < \infty \quad (k=1, 2, \dots). \quad (21)$$

Wir bringen nun eine einfache hinreichende Bedingung für die Beschränktheit der Transformation, die der Matrix der a_{nm} zugeordnet ist.

Satz 2. *Existiert eine (nicht von m und n abhängige) positive Zahl l , für die die Ungleichungen*

$$\sum_{m=1}^{\infty} |a_{nm}| \leq l \quad (n=1, 2, \dots) \quad (22)$$

und

$$\sum_{n=1}^{\infty} |a_{nm}| \leq l \quad (m=1, 2, \dots) \quad (23)$$

erfüllt sind, so liefert die Matrix der a_{nm} eine beschränkte Transformation.

Beweis. Es genügt zu zeigen, daß für $\|x\| \leq 1$ und $\|y\| \leq 1$ die Summe

$$S = \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} |a_{nm}| \cdot |\xi_m| \cdot |\eta_n| \quad (24)$$

beschränkt ist. Dann ist erst recht (13) erfüllt. Wegen $|ab| \leq \frac{1}{2}(|a|^2 + |b|^2)$ erhalten wir

$$S \leq \frac{1}{2} \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} |a_{nm}| (|\xi_m|^2 + |\eta_n|^2) = \frac{1}{2} \sum_{m=1}^{\infty} \left[|\xi_m|^2 \sum_{n=1}^{\infty} |a_{nm}| \right] + \frac{1}{2} \sum_{n=1}^{\infty} \left[|\eta_n|^2 \sum_{m=1}^{\infty} |a_{nm}| \right]$$

oder wegen (22) und (23)

$$S \leq \frac{l}{2} \sum_{m=1}^{\infty} |\xi_m|^2 + \frac{l}{2} \sum_{n=1}^{\infty} |\eta_n|^2 = \frac{l}{2} (\|x\|^2 + \|y\|^2) \leq l,$$

womit der Satz bewiesen ist. Im Fall einer selbstadjungierten Matrix führen die Bedingungen (22) auf die Bedingungen (23). Wir weisen noch darauf hin, daß die Reihe (24) nicht für jede beschränkte Matrix konvergiert.

Als Beispiel betrachten wir die beiden unendlichen Matrizen

$$A = \begin{pmatrix} 0 & 0 & 0 & \dots \\ 1 & 0 & 0 & \dots \\ 0 & 1 & 0 & \dots \\ 0 & 0 & 1 & \dots \\ \dots & \dots & \dots & \dots \end{pmatrix}, \quad B = \begin{pmatrix} a_1 & 1 & 0 & 0 & \dots \\ a_2 & 0 & 1 & 0 & \dots \\ a_3 & 0 & 0 & 1 & \dots \\ a_4 & 0 & 0 & 0 & 1 & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \end{pmatrix},$$

wobei die a_k eine Folge solcher reellen Zahlen bilden, für die die Reihe mit dem allgemeinen Glied $|a_k|$ konvergiert. Auf Grund von Satz 2 können wir leicht nachprüfen, daß den Matrizen A und B lineare Transformationen entsprechen. Ferner läßt sich mit Hilfe von (9) sofort zeigen, daß für jede Wahl der a_k (die der obigen Forderung genügen müssen) die Beziehung $BA = E$ gilt. Daher besitzt der lineare Operator A unendlich viele beschränkte linksinverse Operatoren und infolgedessen keinen beschränkten rechtsinversen Operator. Gehen wir zu den adjungierten, d. h., da die a_k reell sind, zu den transponierten Matrizen A' und B' über, so erhalten wir $A'B' = E$. Die Gleichung $Ax = y$ erhält die Gestalt $\xi_1 = \eta_2, \xi_2 = \eta_3, \dots$, und für jedes $y \in l_2$ hat sie eine einzige Lösung in l_2 . Die Gleichung $A'x = y$ erhält dagegen die Gestalt $\xi_2 = \eta_1, \xi_3 = \eta_2, \dots$, und sie hat unendlich viele Lösungen, da ξ_1 beliebig ist. Das hängt damit zusammen, daß A' keinen beschränkten linksinversen Operator besitzt.

164. Unitäre Matrizen und Projektionsmatrizen. Wir erinnern an die grundlegende charakteristische Eigenschaft einer unitären Transformation U :

$$U^*U = UU^* = E.$$

Sind u_{pq} die Elemente der dem unitären Operator U zugeordneten Matrix, so lassen sich diese Bedingungen in der Form

$$\sum_{s=1}^{\infty} u_{ps}^* u_{sq} = \sum_{s=1}^{\infty} u_{ps} u_{sq}^* = \delta_{pq} \quad (25)$$

schreiben; dabei ist δ_{pq} das Kroneckersymbol ($\delta_{pq} = 0$ für $p \neq q$ und $\delta_{pp} = 1$). Mit $u_{mn}^* = \overline{u_{nm}}$ erhalten diese Beziehungen die Gestalt

$$\sum_{s=1}^{\infty} \overline{u_{sp}} u_{sq} = \delta_{pq}, \quad (26)$$

$$\sum_{s=1}^{\infty} u_{ps} \overline{u_{qs}} = \delta_{pq}. \quad (27)$$

Die Matrix der u_{pq} ist also bezüglich ihrer Zeilen und Spalten orthogonal. Bei endlichen Matrizen sind die Bedingungen (27) eine Folge von (26) [III/1, 28]. Bei unendlichen Matrizen sind diese Bedingungen voneinander unabhängig.

Satz. Die komplexen Zahlen u_{pq} bilden dann und nur dann eine Matrix, die einer unitären Transformation zugeordnet ist, wenn die Bedingungen (26) und (27) erfüllt sind.

Beweis. Die Notwendigkeit der Bedingungen (26) und (27) ergibt sich aus der obigen Betrachtung. Es bleibt zu beweisen, daß die Bedingungen auch hinreichend sind, daß also, wenn sie erfüllt sind, der Matrix der u_{pq} eine lineare (beschränkte) Transformation zugeordnet ist. Daß diese Transformation unitär ist, folgt nachträglich aus der Äquivalenz der Bedingungen (26) und (27) mit den für eine unitäre Transformation charakteristischen Bedingungen (25) [137].

Die Bedingung (27) zeigt, daß $\sum_{s=1}^{\infty} |u_{ps}|^2 = 1$ gilt; folglich konvergieren für jedes Element x die Reihen [59]

$$\xi'_n = \sum_{k=1}^{\infty} u_{nk} \xi_k. \quad (28)$$

Wir bilden den Ausdruck

$$\sum_{p=1}^m \left| \sum_{q=1}^n u_{pq} \xi_q \right|^2 = \sum_{p=1}^m \sum_{s=1}^n \sum_{t=1}^n u_{ps} \overline{u_{pt}} \xi_s \overline{\xi_t} = \sum_{s=1}^n \sum_{t=1}^n \left(\sum_{p=1}^m u_{ps} \overline{u_{pt}} \right) \xi_s \overline{\xi_t}.$$

Bei unbegrenztem Wachsen von m erhalten wir mit Hilfe von (26)

$$\sum_{p=1}^{\infty} \left| \sum_{q=1}^n u_{pq} \xi_q \right|^2 = \sum_{q=1}^n |\xi_q|^2,$$

also erst recht

$$\sum_{p=1}^m \left| \sum_{q=1}^n u_{pq} \xi_q \right|^2 \leq \sum_{q=1}^{\infty} |\xi_q|^2,$$

wobei m eine feste endliche Zahl sei. Daraus folgt ganz analog zu Satz 1 aus [163] durch eine indirekte Beweisführung, daß diese Gleichung für $n = \infty$ gilt, und lassen wir danach m unbegrenzt wachsen, so gelangen wir zu der Ungleichung

$$\sum_{p=1}^{\infty} \left| \sum_{q=1}^{\infty} u_{pq} \xi_q \right|^2 \leq \sum_{q=1}^{\infty} |\xi_q|^2; \quad (29)$$

damit ist die Beschränktheit des Operators U bewiesen. Da U unitär ist, gilt in (29) sicher das Gleichheitszeichen.

Wir betrachten nun eine Matrix von Elementen p_{ik} , die einem Projektionsoperator P in einen Unterraum L zugeordnet ist. Berücksichtigen wir, daß P selbstadjungiert ist und $P^2 = P$ gilt, so erhalten wir die Bedingungen

$$p_{kt} = \overline{p_{tk}}, \quad \sum_{s=1}^{\infty} p_{ts} p_{sk} = p_{tk}. \quad (30)$$

Wie man analog zu der Untersuchung bei unitären Transformationen zeigen kann, sind diese Forderungen nicht nur notwendig, sondern auch hinreichend dafür, daß der Matrix der p_{ik} ein Projektionsoperator entspricht. Wir wählen ein vollständiges normiertes Orthogonalsystem, so daß ein Teil der Einheitsvektoren ein vollständiges System in dem Unterraum L , der andere Teil ein solches System in dem komplementären Raum $H \ominus L$ bildet. Die Einheitsvektoren des ersten Teils bleiben bei Anwendung von P unverändert, die des zweiten Teils werden annulliert. Bei dem gewählten System von Einheitsvektoren ist also dem Projektionsoperator P eine reine Diagonalmatrix zugeordnet, deren Hauptdiagonale nur aus den Elementen 1 und 0 besteht. Mit anderen Worten: *Eine Projektionsmatrix ist unitär äquivalent einer reinen Diagonalmatrix, deren Hauptdiagonale nur aus den Elementen 1 und 0 besteht.*

165. Selbstadjungierte Matrizen. Selbstadjungierte Matrizen sind durch die Bedingungen (13) und (6) charakterisiert. Die Eigenwerte und Eigenelemente dieser Matrizen werden durch die Bedingung definiert, daß das unendliche System

$$\sum_{s=1}^{\infty} a_{ts} \xi_s = \lambda \xi_t \quad (i = 1, 2, \dots) \quad (31)$$

in l_2 eine von 0 verschiedene Lösung hat. Wenn die Eigenelemente ψ_k ein vollständiges normiertes Orthogonalsystem bilden und wir sie als Einheitsvektoren auffassen, wird dem Operator A die Matrix mit den Elementen

$$a_{pq} = (A\psi_q, \psi_p) = \begin{cases} 0 & \text{für } p \neq q, \\ \lambda_p & \text{für } p = q \end{cases} \quad (32)$$

zugeordnet, d. h., wir erhalten eine reine Diagonalmatrix mit den Zahlen λ_p in der Hauptdiagonale. Im allgemeinen besitzt eine selbstadjungierte Matrix dann und nur dann ein reines Punktspektrum, wenn sie unitär äquivalent einer reinen Diagonalmatrix mit reellen Elementen ist. Bei der oben getroffenen Wahl der Einheitsvektoren ψ_k gilt

$$(Ax, y) = \sum_{s=1}^{\infty} \lambda_s \xi_s \overline{\eta_s}, \quad (Ax, x) = \sum_{s=1}^{\infty} \lambda_s |\xi_s|^2. \quad (33)$$

Im allgemeinen Fall existiert für eine selbstadjungierte Matrix a_{ik} eine Zerlegung der Einheit $\mathcal{E}_\lambda$, d. h. eine nichtfallende Projektionsmatrix von Elementen $l_{ik}(\lambda)$, für die

$$l_{ik}(a) = 0, \quad l_{ik}(b) = \begin{cases} 0 & \text{für } i \neq k, \\ 1 & \text{für } i = k, \end{cases} \quad (i, k = 1, 2, \dots) \quad (34)$$

ist, und es gelten die Gleichungen

$$x'_i = \sum_{s=1}^{\infty} a_{is} \xi_s = (Ax, \varphi_i) = \int_a^b \lambda d(\mathcal{E}_\lambda x, \varphi_i) = \int_a^b \lambda d \left(\sum_{s=1}^{\infty} l_{is}(\lambda) \xi_s \right),$$

also

$$a_{is} = \int_a^b \lambda d l_{is}(\lambda) \quad (35)$$

und

$$(Ax, y) = \int_a^b \lambda d \left(\sum_{s,t=1}^{\infty} l_{st}(\lambda) \xi_t \overline{\eta_s} \right), \quad (Ax, x) = \int_a^b \lambda d \left(\sum_{s,t=1}^{\infty} l_{st}(\lambda) \xi_t \overline{\xi_s} \right). \quad (36)$$

Dabei erhalten wir, wenn wir die Eigenschaften der Zerlegung der Einheit berücksichtigen,

$$\sum_{s=1}^{\infty} l_{is}(\lambda_1) l_{sk}(\lambda_2) = \sum_{s=1}^{\infty} l_{is}(\lambda_2) l_{sk}(\lambda_1) = l_{ik}(\lambda_1) \quad \text{für } \lambda_1 \leq \lambda_2 \quad (37)$$

und allgemein

$$\sum_{s=1}^{\infty} \Delta_1 l_{is}(\lambda) \cdot \Delta_2 l_{sk}(\lambda) = \Delta_1 \Delta_2 l_{ik}(\lambda); \quad (38)$$

dabei steht auf der rechten Seite die Differenz der Werte von $l_{ik}(\lambda)$ in den Randpunkten des Intervalls $\Delta_1 \Delta_2$, das den Durchschnitt zweier Intervalle Δ_1 und Δ_2 darstellt. Ist $f(\lambda)$ im Intervall $[a, b]$ eine stetige Funktion, so ist dem Operator $f(A)$ die Matrix mit den Elementen

$$\{f(A)\}_{ik} = \int_a^b f(\lambda) d l_{ik}(\lambda) \quad (39)$$

zugeordnet. Nun ist [43]

$$\sum_{s=1}^{\infty} \int_a^b f_1(\lambda) d l_{is}(\lambda) \int_a^b f_2(\lambda) d l_{sk}(\lambda) = \int_a^b f_1(\lambda) f_2(\lambda) d l_{ik}(\lambda). \quad (40)$$

Da die Bilinearform

$$(\mathcal{E}_\lambda x, y) = \sum_{s,t=1}^{\infty} l_{st}(\lambda) \xi_t \overline{\eta_s} \quad (41)$$

eine Funktion von beschränkter Variation bezüglich λ ist, sind die Funktionen $l_{st}(\lambda)$ ebenfalls von beschränkter Variation bezüglich λ . Für $y=x$ ergibt (41) eine nichtfallende Funktion von λ , woraus unmittelbar folgt, daß die Funktionen $l_{ss}(\lambda)$ nicht abnehmen. Fassen wir die Integrale (39) als Lebesgue-Stieltjessche Integrale auf, so gilt (39) für die in [155] angegebene umfangreiche Klasse von Funktionen $f(\lambda)$. Man braucht nur vorauszusetzen, daß $f(\lambda)$ beschränkt und eine B -Funktion ist [47]. Dann ist sie auch bezüglich einer beliebigen, nichtfallenden Funktion meßbar. Im Fall eines rein kontinuierlichen Spektrums sind alle Funktionen $l_{ik}(\lambda)$ stetig. Auch die umgekehrte Behauptung ist richtig. Im Fall eines gemischten Spektrums betrachten wir den Unterraum L , in dem der Operator A ein reines Punktspektrum hat, und den komplementären Unterraum $H \ominus L$, in dem A ein rein kontinuierliches Spektrum hat. Wir führen in diesen Unterräumen vollständige normierte Orthogonalsysteme ein. Bezeichnen wir mit $(\xi'_1, \xi'_2, \dots)$ Elemente aus L und mit $(\xi''_1, \xi''_2, \dots)$ Elemente aus $H \ominus L$, so können wir die Bilinearform und die quadratische Form in der Gestalt

$$\begin{aligned} \sum_{i,k=1}^{\infty} a_{ik} \xi_k \bar{\eta}_i &= \sum_k \lambda_k \xi_k \bar{\eta}_k + \int_a^b \lambda d \left(\sum_{s,t=1}^{\infty} l_{st}(\lambda) \xi'_s \bar{\eta}''_t \right), \\ \sum_{i,k=1}^{\infty} a_{ik} \xi_k \bar{\xi}_i &= \sum_k \lambda_k |\xi'_k|^2 + \int_a^b \lambda d \left(\sum_{s,t=1}^{\infty} l_{st}(\lambda) \xi'_s \bar{\xi}''_t \right) \end{aligned} \quad (42)$$

schreiben, wobei die $l_{st}(\lambda)$ stetig sind.

Wir betrachten die Resolvente der Matrix A , d. h. die Matrix mit den Elementen $\{R(\lambda)\}_{ik}$, die durch die Gleichung

$$\{R(\lambda)\}_{ik} = \{(A - \lambda E)^{-1}\}_{ik}$$

definiert wird. Es gilt

$$\{R(\lambda)\}_{ik} = \int_a^b \frac{dl_{ik}(\mu)}{\mu - \lambda}; \quad (43)$$

dabei wird angenommen, daß λ nicht zu dem Spektrum von A gehört. Wegen (39) gilt für ganze positive Potenzen von A die Darstellung

$$\{A^m\}_{ik} = \int_a^b \lambda^m dl_{ik}(\lambda). \quad (44)$$

Wenn $\lambda=0$ nicht zum Spektrum von A gehört, d. h., wenn alle $l_{ik}(\lambda)$ in einer gewissen Umgebung von $\lambda=0$ konstant sind, existiert eine beschränkte inverse Matrix A^{-1} , und für ihre Potenzen gelten die Gleichungen

$$\{A^{-m}\}_{ik} = \int_a^b \lambda^{-m} dl_{ik}(\lambda). \quad (45)$$

166. Das kontinuierliche Spektrum. Bekanntlich läßt sich eine weitere Zerlegung von $H \ominus L$ in bezüglich A invariante Unterräume durchführen, in denen dann A ein einfaches kontinuierliches Spektrum besitzt. Es sei H_1 ein solcher Unterraum. In ihm führen wir ein vollständiges Orthogonalsystem ein, und wir nehmen an, daß sich die weitere Untersuchung auf H_1 bezieht; wir fassen dabei H_1 als einen gewissen Hilbertschen Raum auf, in dem die Einheitsvektoren $\varphi_1, \varphi_2, \dots$ eingeführt sind, so daß jedes Element x durch seine Komponenten $(\xi_1, \xi_2, \dots)$ bestimmt wird. Es sei x ein Element aus H_1 , für das die lineare abgeschlossene Hülle von $\mathfrak{E}_a x$ ($a \leq \lambda \leq b$) der gesamte Raum H_1 ist. Wir bezeichnen mit $p_k(\lambda)$ die Komponenten des

Elements $\mathcal{E}_\lambda x$. Jedem Element $y(\eta_1, \eta_2, \dots)$ aus H_1 läßt sich eine Funktion

$$\varphi_y(\lambda) = (y, \mathcal{E}_\lambda x) = \sum_{k=1}^{\infty} \overline{p_k(\lambda)} \eta_k \quad (46)$$

zuordnen. Mit Hilfe der Gleichung (259) aus [147] können wir das bilineare Funktional in der Gestalt

$$(Ay, z) = \sum_{i,k=1}^{\infty} a_{ik} \eta_k \bar{\zeta}_i = \int_a^b \lambda \frac{\sum_{k=1}^{\infty} \overline{p_k(\lambda)} \eta_k \cdot \sum_{i=1}^{\infty} p_i(\lambda) \bar{\zeta}_i}{d\varrho(\lambda)}$$

schreiben, wobei

$$\varrho(\lambda) = \|\mathcal{E}_\lambda x\|^2 = \sum_{s=1}^{\infty} |p_s(\lambda)|^2 \quad (47)$$

ist. Folglich gilt für die Elemente der Matrix, die die Transformation A in H_1 definiert, bei dem angenommenen System der Einheitsvektoren die Beziehung

$$a_{ki} = \int_a^b \lambda \frac{\overline{dp_k(\lambda)} dp_i(\lambda)}{d\varrho(\lambda)}. \quad (48)$$

Ferner ist jedem Element y aus H_1 im Intervall $[a, b]$ eine Funktion $y(\lambda)$ aus L_2 bezüglich $\varrho(\lambda)$ zugeordnet, so daß

$$\varphi_y(\lambda) = \sum_{k=1}^{\infty} \overline{p_k(\lambda)} \eta_k = \int_a^\lambda y(\mu) d\varrho(\mu) \quad (49)$$

gilt, und umgekehrt entspricht jeder Funktion $y(\lambda)$ aus L_2 ein bestimmtes Element y aus H_1 . Bei dieser Zuordnung bleiben die Normen und die Skalarprodukte erhalten. Wenn wir mit $\varphi_k(\lambda)$ die dem Einheitsvektor φ_k zugeordnete Funktion $y(\lambda)$ bezeichnen, finden wir

$$\overline{p_k(\lambda)} = \int_a^\lambda \varphi_k(\mu) d\varrho(\mu); \quad (50)$$

die $\varphi_k(\lambda)$, $k=1, 2, \dots$, bilden ein vollständiges normiertes Orthogonalsystem in L_2 . Dem Operator A in H_1 entspricht die Multiplikation mit λ in L_2 , und für $a_{ik} = (A\varphi_k, \varphi_i)$ erhalten wir statt (48) die Beziehung

$$a_{ik} = \int_a^b \lambda \varphi_k(\lambda) \overline{\varphi_i(\lambda)} d\varrho(\lambda). \quad (51)$$

Wir nehmen an, daß wir an Stelle von $\varphi_k(\lambda)$ ein anderes vollständiges normiertes Orthogonalsystem $\psi_k(\lambda)$ in L_2 gewählt haben; dabei ist diesen $\psi_k(\lambda)$, $k=1, 2, \dots$, in H_1 ein gewisses vollständiges System von Einheitsvektoren ψ_k zugeordnet. In H_1 führen wir die unitäre Transformation U ein, die die Einheitsvektoren φ_k in die Einheitsvektoren ψ_k überführt. Es ist also $U\varphi_k = \psi_k$. Dieser unitären Transformation U in H_1 entspricht eine Matrix, die von der Wahl der Einheitsvektoren abhängt. Wenn wir als Einheitsvektoren die φ_k oder die ψ_k wählen, erhalten wir ein und dieselbe Matrix mit den Elementen $c_{ik} = (U\varphi_k, \varphi_i) = (\psi_k, \varphi_i)$ bzw.

$$c_{ik} = (U\psi_k, \varphi_i) = (\psi_k, U^*\varphi_i) = (\psi_k, U^{-1}\varphi_i) = (\psi_k, \varphi_i).$$

Da der Übergang von H_1 zu L_2 das Skalarprodukt invariant läßt, ergibt sich

$$c_{ik} = \int_a^b \psi_k(\lambda) \overline{\varphi_i(\lambda)} d\rho(\lambda). \quad (52)$$

In den neuen Einheitsvektoren werden die Elemente der dem Operator A entsprechenden Matrix durch Gleichungen definiert, die den Beziehungen (51) analog sind:

$$b_{ik} = \int_a^b \lambda \psi_k(\lambda) \overline{\psi_i(\lambda)} d\rho(\lambda). \quad (53)$$

Sind ξ_k ($k=1, 2, \dots$) die Komponenten eines gewissen Elements in den Einheitsvektoren φ_k und ξ'_k ($k=1, 2, \dots$) die Komponenten desselben Elements in den ψ_k , so ist $\xi_k = (x, \varphi_k)$ und $\xi'_k = (x, \psi_k) = (x, U\varphi_k) = (U^{-1}x, \varphi_k)$; daraus erkennt man, daß sich $(\xi'_1, \xi'_2, \dots)$ durch $(\xi_1, \xi_2, \dots)$ mit Hilfe der Matrix ausdrücken läßt, die zu der unitären Matrix der c_{ik} invers ist. Es ergibt sich daher, wenn wir mit A , B und C die Matrizen mit den Elementen a_{ik} , b_{ik} bzw. c_{ik} bezeichnen, die Matrixgleichung

$$B = C^{-1}AC. \quad (54)$$

Benutzen wir (263) aus [147], (49) und (50), so erhalten wir

$$\{\mathcal{E}_\mu\}_{ik} = l_{ik}(\mu) = \int_a^\mu \varphi_k(\lambda) \overline{\varphi_i(\lambda)} d\rho(\lambda). \quad (55)$$

Unter Berücksichtigung von (39) können wir, wenn $f(\lambda)$ eine beliebige beschränkte und im Intervall $[a, b]$ definierte B -Funktion ist, die Elemente der Matrix $f(A)$ in bezug auf die Einheitsvektoren φ_k durch

$$\{f(A)\}_{ik} = \int_a^b f(\lambda) \varphi_k(\lambda) \overline{\varphi_i(\lambda)} d\rho(\lambda) \quad (56)$$

definieren. Für $f(\lambda) = \frac{1}{\lambda - \mu}$ erhalten wir die Resolvente des Operators A :

$$\{R(\mu)\}_{ik} = \int_a^b \frac{\varphi_k(\lambda) \overline{\varphi_i(\lambda)}}{\lambda - \mu} d\rho(\lambda). \quad (57)$$

Ist $f(\lambda)$ eine reelle Funktion, so ist $f(A)$ ein selbstadjungierter Operator. Analog zu (55) ergibt sich für die Elemente der Spektralschar $\mathcal{E}'_\mu$ des Operators $f(A)$

$$\{\mathcal{E}'_\mu\}_{ik} = \int_{C_\mu} \varphi_k(\lambda) \overline{\varphi_i(\lambda)} d\rho(\lambda); \quad (58)$$

dabei ist C_μ die Menge der Werte λ , für die $f(\lambda) \leq \mu$ gilt. Den Beweis dieser Gleichung wollen wir übergehen. Der Charakter des Spektrums von $f(A)$ hängt von den Eigenschaften von $f(\lambda)$ ab.

Wir sind oben von einem vorgegebenen selbstadjungierten Operator in H und von einem bestimmten Element x ausgegangen, für das die lineare abgeschlossene Hülle von $\mathcal{E}_\lambda x$ gleich H_1 ist. Durch die Einführung von Einheitsvektoren φ_k kamen wir zu l_2 und zu unendlichen Matrizen und erhielten die angegebenen Gleichungen. Man kann nun umgekehrt eine willkürliche, stetige, für $\lambda = a$ verschwindende und im Intervall $[a, b]$ nichtfallende Funktion $\rho(\lambda)$ sowie ein vollständiges normiertes Orthogonalsystem $\varphi_k(\lambda)$ wählen. Dann werden durch (51) Elemente a_{ik} bestimmt,

die offenbar der Bedingung $\overline{a_{ik}} = a_{ki}$ genügen. Wir können leicht zeigen, daß der Matrix mit den Elementen a_{ik} ein beschränkter Operator in l_2 zugeordnet ist. Bezeichnen wir nämlich mit N den größten Wert des absoluten Betrags $|\lambda|$ im Intervall $[a, b]$, so erhalten wir wegen (51)

$$\left| \sum_{i,k=1}^m a_{ik} \xi_k \bar{\xi}_i \right| \leq N \int_a^b \left| \sum_{k=1}^m \varphi_k(\lambda) \xi_k \right|^2 d\rho(\lambda)$$

oder, wenn wir berücksichtigen, daß die $\varphi_k(\lambda)$ orthogonal und normiert sind,

$$\left| \sum_{i,k=1}^m a_{ik} \xi_k \bar{\xi}_i \right| \leq N \sum_{k=1}^m |\xi_k|^2,$$

woraus die Beschränktheit des zugeordneten Operators folgt. Aus $\overline{a_{ik}} = a_{ki}$ ergibt sich die Selbstadjungiertheit des Operators. Durch die Gleichungen (55) werden die Elemente eines von dem Parameter μ abhängigen Projektionsoperators bestimmt, der eine Zerlegung der Einheit darstellt; dabei gilt offenbar

$$a_{ik} = \int_a^b \lambda d\{\mathcal{E}_\lambda\}_{ik},$$

d. h., $\mathcal{E}_\lambda$ ist die Spektralschar des Operators A . Gehen wir in (50) zu den adjungierten Größen über, so erhalten wir die Komponenten $p_k(\lambda)$ des Elements $\mathcal{E}_\lambda x$ und für $\lambda = b$ die Komponenten des Elements x selbst. Mit Hilfe der Parsevalschen Gleichung folgt aus (50), daß $\rho(\lambda)$ durch (47) bestimmt ist. Im allgemeinen Fall eines selbstadjungierten Operators A mit kontinuierlichem Spektrum bilden wir die paarweise orthogonalen invarianten Unterräume $H_1, H_2, \dots$; in jedem von ihnen hat A ein einfaches Spektrum. Führen wir in jedem H_k zugehörige Einheitsvektoren ein, so erhalten wir für jedes H_k Gleichungen der oben angegebenen Gestalt. Den endgültigen Ausdruck können wir dann z. B. für die Bilinearform (Ax, y) erhalten, indem wir die Ausdrücke für die Bilinearformen in jedem der Unterräume H_k addieren.

Den Begriff des einfachen Spektrums kann man leicht verallgemeinern, wenn man auf die Stetigkeit der Spektralschar $\mathcal{E}_\lambda$ verzichtet. Nach wie vor muß jedoch ein Element x existieren, mit welchem $\mathcal{E}_\lambda x$ den gesamten Raum H aufspannt. Dann ist die durch (47) definierte nichtfallende Funktion $\rho(\lambda)$ nicht unbedingt stetig. Wir können x als normiert annehmen; außer $\rho(a) = 0$ gilt dann noch $\rho(b) = 1$. Hat etwa A ein reines Punktspektrum, haben alle Eigenwerte den Rang 1 und nehmen wir für x ein beliebiges Element, für das alle Fourierkoeffizienten bezüglich des vollständigen Systems der Eigenwerte von 0 verschieden sind, so können wir behaupten, daß $\mathcal{E}_\lambda x$ den gesamten Raum H aufspannt und infolgedessen das Spektrum von A einfach ist. Beim Vorhandensein mehrfacher Eigenwerte kann das Spektrum, wie man leicht erkennt, nicht einfach sein. Wir erhalten den allgemeinen Fall eines einfachen Spektrums, wenn wir bei der Zerlegung von H in den Unterraum der Eigenelemente und den Unterraum des kontinuierlichen Spektrums in beiden Unterräumen einfache Spektren haben. Im ersten Unterraum ist das Spektrum dann und nur dann einfach, wenn alle Eigenwerte den Rang 1 haben. Ist das einfache Spektrum nicht kontinuierlich, so hat $\mathcal{E}_\mu$ Sprungstellen, und die durch (47) definierte Funktion $\rho(\lambda)$ muß in den Unstetigkeitsstellen von $\mathcal{E}_\lambda$ ebenfalls Unstetigkeiten besitzen; denn wäre die Funktion $\rho(\lambda)$ in einer Unstetigkeitsstelle $\lambda = \lambda'$ der Spektralschar $\mathcal{E}_\lambda$ stetig, so hätten wir $(\mathcal{E}_{\lambda'} - \mathcal{E}_{\lambda'-0})x = 0$, und alle Elemente des durch $\mathcal{E}_\lambda x$ aufgespannten Raumes wären orthogonal zu den Eigenelementen, die dem Eigenwert $\lambda = \lambda'$ entsprechen; daraus würde folgen, daß $\mathcal{E}_\lambda x$ nicht den gesamten Raum H aufspannen kann.

167. Jacobische Matrizen. In dem unendlichdimensionalen Raum H habe der selbstadjungierte Operator A ein einfaches kontinuierliches Spektrum. Orthogonalisieren wir die Potenzen von λ bezüglich $\varrho(\lambda)$ im Intervall $[a, b]$, so erhalten wir als vollständiges System $\varphi_k(\lambda)$ aus [166] ein System reeller Polynome $P_k(\lambda)$, $k=0, 1, 2, \dots$, vom Grad k :

$$\int_a^b P_i(\lambda) P_k(\lambda) d\varrho(\lambda) = \begin{cases} 0 & \text{für } i \neq k, \\ 1 & \text{für } i = k. \end{cases} \quad (59)$$

Der Koeffizient der höchsten auftretenden Potenz in jedem Polynom $P_k(\lambda)$ kann positiv gewählt werden. In den bisherigen Bezeichnungen haben wir die Funktionen $\varphi_k(\lambda)$ mit $k=1$ beginnend numeriert. Wir beginnen jetzt bei der Numerierung der $P_k(\lambda)$ mit $k=0$, da k den Grad von $P_k(\lambda)$ angibt; $P_k(\lambda)$ ersetzt daher die Funktion $\varphi_{k+1}(\lambda)$. Bei dieser Wahl der Einheitsvektoren werden die Elemente der dem Operator A entsprechenden Matrix gemäß (51) durch die Gleichungen

$$a_{ik} = \int_a^b \lambda P_i(\lambda) P_k(\lambda) d\varrho(\lambda) \quad (i, k=0, 1, 2, \dots) \quad (60)$$

definiert. Wir nehmen $k-i > 1$ an. Dann ist das Produkt $\lambda P_i(\lambda)$ ein Polynom von kleinerem als k -tem Grad. Dieses Produkt läßt sich linear durch die $P_s(\lambda)$ mit $s < k$ ausdrücken. Damit ist das Integral (60) wegen (59) gleich 0; (60) verschwindet auch für $i-k > 1$, denn es ist $a_{ki} = a_{ik}$. Es gilt also $a_{ik} = 0$ für $|i-k| > 1$. Wir führen nun die Bezeichnungen

$$a_k = \int_a^b \lambda P_k^2(\lambda) d\varrho(\lambda), \quad b_k = \int_a^b \lambda P_k(\lambda) P_{k+1}(\lambda) d\varrho(\lambda) \quad (k=0, 1, 2, \dots) \quad (61)$$

ein. Die Zahl b_k tritt in der linearen Darstellung des Produktes $\lambda P_k(\lambda)$ durch die $P_s(\lambda)$, $s=0, 1, 2, \dots, k+1$, auf:

$$\lambda P_k(\lambda) = b_k P_{k+1}(\lambda) + \sum_{s=0}^k c_s^{(k)} P_s(\lambda). \quad (62)$$

Da die Koeffizienten der höchsten in den Polynomen $P_m(\lambda)$ auftretenden Potenzen positiv sind, ist $b_k > 0$. Aus (60) und aus der obigen Überlegung folgt

$$a_{kk} = a_k, \quad a_{k,k+1} = a_{k+1,k} = b_k, \quad a_{ik} = 0 \quad \text{für } |i-k| > 1. \quad (63)$$

Bei dem gewählten System der Einheitsvektoren hat also die Transformationsmatrix die Gestalt

$$\begin{pmatrix} a_0 & b_0 & 0 & 0 & 0 & \dots \\ b_0 & a_1 & b_1 & 0 & 0 & \dots \\ 0 & b_1 & a_2 & b_2 & 0 & \dots \\ 0 & 0 & b_2 & a_3 & b_3 & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \end{pmatrix} \quad (64)$$

und es gilt $b_k > 0$. Eine reelle, selbstadjungierte und den Bedingungen (63) genügende Matrix heißt eine *Jacobische Matrix*. Bei geeigneter Wahl der Einheitsvektoren ist also die Matrix eines selbstadjungierten Operators mit einfachem kontinuierlichem Spektrum eine *Jacobische Matrix*.

Mit Hilfe der Gleichungen (59) und der Bezeichnungen (61) kann man leicht die Koeffizienten in der Entwicklung (62) berechnen, indem man beide Seiten mit $P_m(\lambda) d\varrho(\lambda)$ multipliziert und nach λ integriert. Für $m < k-1$ ist das Integral über das Produkt $\lambda P_k(\lambda) P_m(\lambda) d\varrho(\lambda)$, wie wir eben gesehen haben, gleich 0; folglich ist

$c_m^{(k)} = 0$ für $m < k - 1$. Bei der Berechnung der übrigen Koeffizienten benutzen wir (61) und erhalten die zwischen den Polynomen $P_m(\lambda)$ geltenden Beziehungen

$$\lambda P_k(\lambda) = b_k P_{k+1}(\lambda) + a_k P_k(\lambda) + b_{k-1} P_{k-1}(\lambda); \quad (65)$$

dabei ist

$$P_{-1}(\lambda) = 0, \quad P_0(\lambda) = 1. \quad (66)$$

Die zweite Gleichung (66) ergibt sich daraus, daß wir, wie oben gezeigt wurde, $\varrho(b) = 1$ annehmen können; die erste Gleichung (66) erhält man aus der Definition. Man kann auch wie in [166] von einer stetigen nichtfallenden Funktion $\varrho(\lambda)$ ausgehen, ein System von bezüglich $\varrho(\lambda)$ orthogonalen Polynomen konstruieren und mit Hilfe von (60) die Elemente einer Jacobischen Matrix bilden. Nach dem in [163] Dargestellten müssen die Elemente der Matrix (64) ihrem absoluten Betrag nach durch ein und dieselbe Zahl beschränkt sein. Das erhält man auch leicht aus (61). Diese Überlegungen führen uns zu dem folgenden Satz:

Satz. Jede selbstadjungierte, zu einem beschränkten Operator mit einfachem kontinuierlichem Spektrum gehörige Matrix ist unitär äquivalent einer gewissen Jacobischen Matrix der Gestalt (64), deren Elemente beschränkt sind und für die $b_k > 0$ gilt. Wir können mit Hilfe von (60) alle diese Matrizen erhalten, wenn $[a, b]$ ein beliebiges endliches Intervall und $\varrho(\lambda)$ eine in ihm nichtfallende stetige Funktion ist, die der Bedingung $\varrho(a) = 0$ und der (unwesentlichen) Bedingung $\varrho(b) = 1$ genügt, und die $P_i(\lambda)$ ein bezüglich $\varrho(\lambda)$ normiertes Orthogonalsystem von Polynomen bilden.

Wir gehen von einer vorgegebenen Jacobischen Matrix aus und nehmen an, daß in ihr die Elemente $a_{ik} = a_{ki}$ für $|i - k| = 1$ von 0 verschieden sind. Wenn wir von dem ursprünglichen System von Einheitsvektoren zu einem neuen System übergehen, indem wir jeden Einheitsvektor mit einem Ausdruck der Gestalt $e^{i\omega_k}$ multiplizieren, erhalten wir bei geeigneter Wahl der ω_k , wie sich leicht zeigen läßt, eine zu der gegebenen Matrix unitär äquivalente Jacobische Matrix, in der die Elemente a_{ik} für $|i - k| = 1$ positiv sind. Wir können also annehmen, daß die gegebene Jacobische Matrix die Gestalt (64) hat, die a_k offenbar reell und die b_k positiv sind. Wenn wir Satz 2 aus [163] und eine der Folgerungen aus Satz 1 von [163] berücksichtigen, sehen wir, daß die Matrix (64) dann und nur dann eine lineare beschränkte Transformation realisiert, wenn die Zahlen a_k und b_k durch ein und dieselbe, nicht von k abhängige Zahl N beschränkt sind:

$$|a_k| \leq N, \quad |b_k| \leq N. \quad (67)$$

Wir werden im folgenden diese Eigenschaft voraussetzen. Es sei

$$\psi_0, \psi_1, \psi_2, \dots \quad (68)$$

das zugrunde liegende System von Einheitsvektoren. Mit A bezeichnen wir den der Matrix (64) entsprechenden selbstadjungierten Operator. Dann gilt

$$A\psi_k = b_{k-1}\psi_{k-1} + a_k\psi_k + b_k\psi_{k+1} \quad (k=0, 1, \dots; \psi_{-1}=0). \quad (69)$$

Führen wir die durch (65) und (66) definierten Polynome $P_k(\lambda)$ ein und benutzen die letzte Gleichung, so können wir jeden Einheitsvektor ψ_k unmittelbar durch den ersten Einheitsvektor ψ_0 mit Hilfe der Beziehung

$$\psi_k = P_k(A) \psi_0 \quad (70)$$

ausdrücken.

Es sei $\mathcal{E}_\lambda$ die Spektralschar des Operators A , d. h. der Matrix (64). Aus (70) folgt unmittelbar, daß die Elemente $\mathcal{E}_\lambda \psi_0$ den gesamten Raum H aufspannen; denn mit (70) gilt, wenn a die untere und b die obere Grenze des Operators (64) ist,

$$\psi_k = \int_a^b P_k(\lambda) d\mathcal{E}_\lambda \psi_0,$$

d. h., ψ_k läßt sich als Grenzelement von Linearkombinationen der Elemente $\mathcal{E}_\lambda \psi_0$ darstellen, und jedes Element läßt sich nach den Einheitsvektoren ψ_k entwickeln. Also entspricht einer Jacobischen Matrix ein einfaches (nicht notwendig kontinuierliches) Spektrum, und als Ausgangselement x kann man den ersten Einheitsvektor ψ_0 nehmen. Auf Grund von (70) können wir

$$\begin{aligned}(\psi_i, \psi_k) &= (P_i(A) \psi_0, P_k(A) \psi_0) = (P_i(A) P_k(A) \psi_0, \psi_0), \\(A \psi_i, \psi_k) &= (A P_i(A) P_k(A) \psi_0, \psi_0)\end{aligned}$$

schreiben, womit man durch Einführung der Funktion

$$\varrho(\lambda) = (\mathcal{E}_\lambda \psi_0, \psi_0) = \|\mathcal{E}_\lambda \psi_0\|^2 \quad (71)$$

die Beziehungen

$$\begin{aligned}(\psi_i, \psi_k) &= \int_a^b P_i(\lambda) P_k(\lambda) d\varrho(\lambda) = \begin{cases} 0 & \text{für } i \neq k, \\ 1 & \text{für } i = k, \end{cases} \\(A \psi_i, \psi_k) &= \int_a^b \lambda P_i(\lambda) P_k(\lambda) d\varrho(\lambda)\end{aligned}$$

erhält. Daraus folgt, daß die Polynome $P_i(\lambda)$ ein normiertes Orthogonalsystem bezüglich $\varrho(\lambda)$ bilden und daß die Elemente der Matrix (64) vermöge (60) durch sie ausgedrückt werden können. Besitzt die Spektralschar $\mathcal{E}_\lambda$ Unstetigkeiten, so ist, wie wir in [166] gesehen haben, jeder solchen Unstetigkeit ein Eigenwert vom Rang 1 zugeordnet. Daraus erkennt man, daß $\mathcal{E}_\lambda$ nicht auf endlich viele Sprungstellen reduziert werden kann; dasselbe kann man von $\varrho(\lambda)$ behaupten. Umgekehrt lassen sich mit Hilfe von (60) beliebige Jacobische Matrizen bilden, wenn für $\varrho(\lambda)$ eine nicht unbedingt stetige, aber beliebige nichtfallende Funktion gewählt wird, die sich nur nicht auf endlich viele Sprungstellen reduzieren läßt.

168. Differentiallösungen. Wir betrachten einen selbstadjungierten Operator (eine selbstadjungierte Matrix) mit rein kontinuierlichem Spektrum. Wie wir gesehen haben, läßt sich eine Folge paarweise orthogonaler normierter Elemente $y^{(s)}$ ($s = 1, 2, 3, \dots$) derart bilden, daß die $\mathcal{E}_\lambda y^{(s)}$ paarweise orthogonale Unterräume bilden, deren orthogonale Summe der gesamte Raum H ist. Die Anzahl der Elemente $y^{(s)}$ kann endlich oder auch unendlich sein. Es seien $p_k^{(s)}(\lambda)$, $k = 1, 2, \dots$, die Komponenten des Elements $\mathcal{E}_\lambda y^{(s)}$. Die Funktionen $p_k^{(s)}(\lambda)$ sind Funktionen von beschränkter Variation und genügen für jedes s in einem beliebigen, im Intervall $[a, b]$ enthaltenen Intervall Δ den Gleichungen

$$\sum_{k=1}^{\infty} a_{ik} \Delta p_k^{(s)}(\lambda) = \int_{\Delta} \lambda d p_i^{(s)}(\lambda) \quad (i = 1, 2, \dots). \quad (72)$$

Es lassen sich die folgenden Orthogonalitätseigenschaften der Lösungen $p_k^{(s)}(\lambda)$ bestätigen [151]:

$$\left. \begin{aligned} \sum_{k=1}^{\infty} \Delta_1 p_k^{(s)}(\lambda) \cdot \overline{\Delta_2 p_k^{(t)}(\lambda)} &= 0 \\ (s \neq t; \Delta_1 \text{ und } \Delta_2 \text{ beliebige Intervalle}), \end{aligned} \right\} \quad (73)$$

$$\left. \begin{aligned} \sum_{k=1}^{\infty} \Delta_1 p_k^{(s)}(\lambda) \cdot \overline{\Delta_2 p_k^{(s)}(\lambda)} &= 0 \\ (\Delta_1 \text{ und } \Delta_2 \text{ ohne gemeinsame innere Punkte}). \end{aligned} \right\} \quad (73_1)$$

Aus den grundlegenden Gleichungen in [149] ergeben sich die folgenden Gleichungen, die die konstruierten Differentiallösungen enthalten:

$$\sum_s \int_a^b \frac{\overline{dp_k^{(s)}(\lambda)} dp_i^{(s)}(\lambda)}{d\varrho_s(\lambda)} = \begin{cases} 0 & \text{für } k \neq i, \\ 1 & \text{für } k = i, \end{cases} \quad (74)$$

$$a_{ik} = \sum_s \int_a^b \lambda \frac{\overline{dp_k^{(s)}(\lambda)} dp_i^{(s)}(\lambda)}{d\varrho_s(\lambda)}, \quad (75)$$

$$l_{ik}(\lambda) = \sum_s \int_a^\lambda \frac{\overline{dp_k^{(s)}(\mu)} dp_i^{(s)}(\mu)}{d\varrho_s(\mu)}; \quad (76)$$

dabei sind die $l_{ik}(\lambda)$ die Elemente der Spektralmatrix. Konstruieren wir auf irgendeine Weise die Lösungen des Systems (72), die den Orthogonalitätsbedingungen (73) und (73₁) genügen, und weisen wir für beliebiges λ die Gültigkeit von (76) nach, so können wir sicher sein, daß das erhaltene Lösungssystem abgeschlossen ist und die übrigen Gleichungen gelten. Es seien y und z Elemente aus l_2 , und es sei

$$y^{(s)}(\lambda) = \sum_{k=1}^{\infty} \eta_k p_k^{(s)}(\lambda), \quad z^{(s)}(\lambda) = \sum_{k=1}^{\infty} \zeta_k p_k^{(s)}(\lambda).$$

Die Gleichung (272) aus [149] läßt sich dann auf die Gestalt

$$(Ay, z) = \sum_s \int_a^b \lambda \frac{dy^{(s)}(\lambda) \overline{dz^{(s)}(\lambda)}}{d\varrho_s(\lambda)} \quad (76_1)$$

bringen, die auch unmittelbar aus (75) folgt. Analog lassen sich auch die übrigen Gleichungen aus [149] schreiben. Wir können leicht zeigen: Ist eine bestimmte Differentiallösung $v_k(\lambda)$, $k=1, 2, \dots$, zu jeder Differentiallösung des eben betrachteten abgeschlossenen Systems orthogonal, so sind alle $v_k(\lambda)$ konstant. Dieser Fall ergibt eine triviale Lösung des Systems (72), weil dann $\Delta v_k(\lambda)=0$ und $dv_k(\lambda)=0$ gilt. Die Differentiallösungen $p_k^{(s)}(\lambda)$ sind offenbar auch zu allen Eigenelementen des Operators A orthogonal.

Wir kehren nun zum System (72) zurück; $p_k(\lambda)$, $k=1, 2, 3, \dots$, seien gewisse Differentiallösungen dieses Systems:

$$\sum_{k=1}^{\infty} a_{ik} \Delta p_k(\lambda) = \int_a^\lambda \lambda dp_i(\lambda) \quad (i=1, 2, \dots).$$

Wir nehmen an, jede Funktion $p_k(\lambda)$ habe eine stetige Ableitung $p'_k(\lambda)$. Dann werden aus den auftretenden Stieltjesschen Integralen gewöhnliche Integrale stetiger Funktionen, und wir erhalten, wenn wir auf diese den Mittelwertsatz anwenden und mit λ' und λ'' die Endpunkte des Intervalls Δ bezeichnen,

$$\sum_{k=1}^{\infty} a_{ik} [p_k(\lambda'') - p_k(\lambda')] = \lambda_i p'_i(\lambda_i) (\lambda'' - \lambda') \quad (77)$$

mit $\lambda' < \lambda_i < \lambda''$. Wenn wir auf die linke Seite dieser Beziehung den Mittelwertsatz anwenden, beide Seiten durch $\lambda'' - \lambda'$ dividieren und λ' sowie λ'' gegen einen gemeinsamen Grenzwert λ streben lassen, erhalten wir

$$\sum_{k=1}^{\infty} a_{ik} p'_k(\lambda) = \lambda p'_i(\lambda) \quad (i = 1, 2, \dots). \quad (78)$$

Wir setzen dabei voraus, daß wir in der unendlichen Summe auf der linken Seite von (77) gliedweise zur Grenze übergehen können. Das ist sicher möglich, wenn diese Summe endlich ist, d. h., wenn die Matrix der a_{ik} in jeder Zeile und damit auch in jeder Spalte nur endlich viele, von 0 verschiedene Elemente besitzt. Aus den Gleichungen (78) ist ersichtlich, daß unter den angegebenen Bedingungen im Fall eines kontinuierlichen Spektrums die $p'_k(\lambda)$, $k = 1, 2, \dots$, für jedes λ denselben Gleichungen genügen, die wir für die Eigenelemente aufgestellt haben. Die $p'_k(\lambda)$ liegen aber nicht in l_2 , d. h., die aus den Ausdrücken $|p'_k(\lambda)|^2$ gebildete Summe ist gleich ∞ , denn im Fall eines rein kontinuierlichen Spektrums existieren keine Eigenwerte. Im Fall eines gemischten Spektrums können die Differentiallösungen den gewöhnlichen Lösungen aus l_2 , welche Eigenelemente ergeben, überlagert werden.

169. Beispiele. 1. Wir setzen im Intervall $[-1, 1]$

$$q(\lambda) = \frac{2}{\pi} \int_{-1}^{\lambda} \sqrt{1 - \lambda^2} d\lambda.$$

Damit erhält die Bedingung (59) für die Polynome $P_i(\lambda)$ die Gestalt

$$\frac{2}{\pi} \int_{-1}^1 \sqrt{1 - \lambda^2} P_i(\lambda) P_k(\lambda) d\lambda = \begin{cases} 0 & \text{für } i \neq k, \\ 1 & \text{für } i = k. \end{cases} \quad (79)$$

Diesen Bedingungen genügen die Polynome

$$P_n(\lambda) = \frac{\sin(n+1)\theta}{\sin\theta} \quad \text{mit} \quad \cos\theta = \lambda.$$

Durch Verwendung der Moivreschen Formel können wir leicht zeigen, daß der Bruch tatsächlich ein Polynom n -ten Grades in $\cos\theta$ ist. Die Bedingungen (79) lassen sich sofort verifizieren, wenn man mit Hilfe von $\lambda = \cos\theta$ eine neue Veränderliche λ einführt. Die in der Matrix (64) auftretenden Zahlen a_k und b_k werden durch die Gleichungen

$$a_k = \frac{2}{\pi} \int_{-1}^1 \lambda \sqrt{1 - \lambda^2} P_k^2(\lambda) d\lambda, \quad b_k = \frac{2}{\pi} \int_{-1}^1 \lambda \sqrt{1 - \lambda^2} P_k(\lambda) P_{k+1}(\lambda) d\lambda$$

definiert. Durch Einführung der Veränderlichen θ und Auswertung der erhaltenen Integrale ergeben sich für jedes k die Werte $a_k = 0$ und $b_k = 1/2$, d. h., die Elemente der entsprechenden Matrix sind $a_{k,k+1} = a_{k+1,k} = 1/2$ und $a_{ik} = 0$ für die übrigen Fälle. Diese Matrix besitzt ein einfaches rein kontinuierliches Spektrum. Die einzige Differentiallösung des Systems wird auf Grund von (50) durch die Beziehungen

$$p_n(\lambda) = \frac{2}{\pi} \int_{-1}^{\lambda} \sqrt{1 - \lambda^2} P_{n-1}(\lambda) d\lambda = -\frac{2}{\pi} \int_{\pi}^{\theta} \sin\theta \sin n\theta$$

definiert; dabei ist $p'_n(\lambda) = \frac{2}{\pi} \sqrt{1 - \lambda^2} P_{n-1}(\lambda) = \frac{2}{\pi} \sin n\theta$. Sehen wir von dem Faktor $\frac{2}{\pi}$

ab, so erhalten wir für das System

$$\frac{1}{2} x_2 = \lambda x_1, \quad \frac{1}{2} x_1 + \frac{1}{2} x_3 = \lambda x_2, \quad \dots, \quad \frac{1}{2} x_{n-1} + \frac{1}{2} x_{n+1} = \lambda x_n, \quad \dots$$

die Lösung $x_n = \sin(n \arccos \lambda)$; dabei ist $-1 \leq \lambda \leq 1$.

2. Wir betrachten im Intervall $[-\pi, \pi]$ das vollständige normierte Orthogonalsystem

$\frac{1}{\sqrt{2\pi}} e^{ikx}$ ($k=0, \pm 1, \pm 2, \dots$). Die Gleichungen (51) ergeben mit $\varrho(\lambda) = \lambda$ für die entsprechenden Matrizen die Elemente

$$a_{pq} = \begin{cases} \frac{1}{2\pi} \int_{-\pi}^{\pi} \lambda e^{i(p-q)\lambda} d\lambda = \frac{(-1)^{p-q}}{i(p-q)} & \text{für } p \neq q, \\ 0 & \text{für } p = q. \end{cases}$$

dabei laufen die Größen p und q von $-\infty$ bis ∞ . Nach (50) ist

$$p_k(\lambda) = \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\lambda} e^{-ik\lambda} d\lambda, \quad p'_k(\lambda) = \frac{1}{\sqrt{2\pi}} e^{-ik\lambda},$$

und aus (78) ergeben sich die Beziehungen

$$\sum'_{k=-\infty}^{\infty} \frac{(-1)^{s-k}}{i(s-k)} \frac{1}{\sqrt{2\pi}} e^{-ik\lambda} = \lambda \frac{1}{\sqrt{2\pi}} e^{-is\lambda}.$$

Der Strich bei dem Summenzeichen bedeutet dabei, daß der Wert $k=s$ auszuschließen ist. Die letzte Gleichung können wir auf die Form

$$\sum'_{k=-\infty}^{\infty} \frac{(-1)^{s-k}}{i(s-k)} e^{i(s-k)\lambda} = \lambda \quad \text{oder} \quad \sum'_{j=-\infty}^{\infty} \frac{(-1)^j}{ij} e^{ij\lambda} = \lambda$$

bringen; dabei ist der Wert $j=0$ auszuschließen. Die letzte Formel stellt die gewöhnliche Entwicklung von λ in eine Fourierreihe dar; die Reihe ist in den Randpunkten des Intervalls, d. h. für $\lambda = \pm \pi$, divergent, da die Entwicklung in komplexer Form geschrieben wurde. Wir verwenden nun die Gleichung (56) und nehmen $f(\lambda) = -\pi - \lambda$ für $\lambda < 0$ und $f(\lambda) = \pi - \lambda$ für $\lambda > 0$ an. Auf diese Weise erhalten wir die Matrix der

$$b_{pq} = \frac{1}{2\pi} \int_{-\pi}^0 (-\pi - \lambda) e^{i(p-q)\lambda} d\lambda + \frac{1}{2\pi} \int_0^{\pi} (\pi - \lambda) e^{i(p-q)\lambda} d\lambda$$

oder

$$b_{pq} = \begin{cases} \frac{i}{p-q} & \text{für } p \neq q, \\ 0 & \text{für } p = q. \end{cases}$$

Da im Intervall $[-\pi, \pi]$ die Ungleichung $|f(\lambda)| \leq \pi$ gilt, erhalten wir für eine quadratische Form die Abschätzung [155]

$$\left| \sum'_{p,q=-s}^t \frac{\xi_p \bar{\xi}_q}{p-q} i \right| \leq \pi \sum_{k=-s}^t |\xi_k|^2. \quad (81)$$

In dieser Abschätzung kann der Faktor i auf der linken Seite offenbar fortgelassen werden. Setzen wir $\xi_p = 0$ für $p \leq 0$ und nehmen wir die übrigen ξ_p als reell an, so erhalten wir die von HILBERT angegebene Abschätzung

$$\left| \sum'_{p,q=1}^n \frac{\xi_p \bar{\xi}_q}{p-q} \right| \leq \pi \sum_{k=1}^n \xi_k^2. \quad (82)$$

Es läßt sich leicht zeigen, daß der Matrix mit den Elementen $a_{pq} = 1/(p-q)$ für $p \neq q$ und $a_{pp} = 0$ nicht mehr ein beschränkter Operator zugeordnet ist. In der Tat, nehmen wir $\xi_k = 1/\sqrt{n}$ für $1 \leq k \leq n$ und $\xi_k = 0$ für $k > n$ an, so ist die Norm des Elements $(\xi_1, \xi_2, \xi_3, \dots)$ gleich 1, und die zugeordnete quadratische Form lautet

$$\begin{aligned} \sum_{p,q=1}^n \frac{\xi_p \xi_q}{|p-q|} &= \frac{2}{n} \sum_{q < p}^{1,2,\dots,n} \frac{1}{p-q} = \frac{2}{n} \sum_{p=2}^n \left(\frac{1}{1} + \frac{1}{2} + \dots + \frac{1}{p-2} + \frac{1}{p-1} \right) \\ &= \frac{2}{n} \left(\frac{n-1}{1} + \frac{n-2}{2} + \dots + \frac{n-(n-1)}{n-1} \right) \\ &= 2 \left(1 + \frac{1}{2} + \dots + \frac{1}{n-1} - \frac{n-1}{n} \right); \end{aligned}$$

der letzte Ausdruck wächst mit n unbegrenzt an, weil die Summe $1 + \frac{1}{2} + \dots + \frac{1}{n-1}$ unbegrenzt wächst und der Bruch $(n-1)/n$ gegen 1 strebt. Im Fall (80) liegt keine absolute Konvergenz der unendlichen Doppelreihe vor; man kann aber bestätigen, daß für je zwei Elemente aus l_2 der Grenzwert

$$\lim_{\substack{n \rightarrow \infty \\ m \rightarrow \infty}} \sum'_{p=1}^m \sum'_{q=1}^n \frac{\xi_p \bar{\eta}_q}{p-q} i = \sum_{q=1}^{\infty} \sum'_{p=1}^{\infty} \left[\frac{\xi_p}{p-q} i \right] \bar{\eta}_q$$

existiert. Die Striche bedeuten, daß die Summanden, in denen $p=q$ ist, ausgelassen werden.

3. Wir betrachten nun an Stelle des Systems $e^{ik\lambda}$ das reelle vollständige normierte Orthogonalsystem

$$\varphi_k(\lambda) = \frac{1}{\sqrt{2\pi}} (\sin k\lambda + \cos k\lambda) \quad (k=0, \pm 1, \pm 2, \dots).$$

Durch Anwendung von Gleichung (56) kommen wir zu der Matrix mit den Elementen

$$b_{pq} = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\lambda) (\sin p\lambda + \cos p\lambda) (\sin q\lambda + \cos q\lambda) d\lambda$$

oder

$$b_{pq} = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\lambda) \cos(p-q)\lambda d\lambda + \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\lambda) \sin(p+q)\lambda d\lambda.$$

Wenn wir wie oben $f(\lambda) = -\pi - \lambda$ für $\lambda < 0$ und $f(\lambda) = \pi - \lambda$ für $\lambda > 0$ annehmen, erhalten wir die Matrix mit den Elementen

$$b_{pq} = \begin{cases} \frac{1}{p+q} & \text{für } p+q \neq 0, \\ 0 & \text{für } p+q = 0. \end{cases}$$

Analog zu (81) gelangen wir zu der Ungleichung

$$\left| \sum'_{p,q=-s}^t \frac{\xi_p \xi_q}{p+q} \right| \leq \pi \sum_{k=-s}^t |\xi_k|^2; \quad (83)$$

der Strich bedeutet hier, daß diejenigen Summanden auszuschließen sind, für die $p+q=0$ gilt. Der Ungleichung (82) entspricht

$$\left| \sum'_{p,q=1}^n \frac{\xi_p \xi_q}{p+q} \right| \leq \pi \sum_{k=1}^n \xi_k^2. \quad (84)$$

Wir können alle Zahlen positiv wählen, und die Doppelreihe konvergiert absolut für jedes Element aus l_2 . Es ist also

$$\sum_{p,q=1}^{\infty} \frac{\xi_p \xi_q}{p+q} \leq \pi \sum_{k=1}^{\infty} \xi_k^2. \quad (85)$$

170. Schwache Konvergenz im Raum l_2 . Wir nehmen an, eine Folge von Elementen $x^{(n)}(\xi_1^{(n)}, \xi_2^{(n)}, \dots)$, $n=1, 2, \dots$, konvergiere stark gegen das Element $x(\xi_1, \xi_2, \dots)$, d. h., es sei $x^{(n)} \rightarrow x$. Diese Aussage läßt sich in der Form

$$\sum_{k=1}^{\infty} |\xi_k - \xi_k^{(n)}|^2 \rightarrow 0 \quad \text{für } n \rightarrow \infty \quad (86)$$

darstellen. Daraus folgt

$$\sum_{k=1}^{\infty} |\xi_k^{(n)}|^2 \rightarrow \sum_{k=1}^{\infty} |\xi_k|^2, \quad (87)$$

und die $\|x^{(n)}\|$ sind durch eine nicht von n abhängige Zahl l beschränkt.

Aus (86) folgt sofort

$$\xi_n^{(k)} \rightarrow \xi_k \quad (k=1, 2, \dots), \quad (88)$$

aber aus (88) folgt nicht (86).

Wir zeigen, daß die Bedingung (88) zusammen mit der Beschränktheit der Norm von $x^{(n)}$,

$$\sum_{k=1}^{\infty} |\xi_k^{(n)}|^2 \leq l^2, \quad (89)$$

der schwachen Konvergenz $x^{(n)} \rightarrow x$ äquivalent ist. Gilt die schwache Konvergenz, so müssen die $\|x^{(n)}\|$ beschränkt sein, d. h., es muß (89) für gewisse l sein und außerdem $(x^{(n)}, \varphi_k) \rightarrow (x, \varphi_k)$ gelten, wobei die φ_k die schon oben erwähnten Einheitsvektoren sind. Hieraus ergibt sich (88).

Sind umgekehrt die Bedingungen (88) und (89) erfüllt, so folgt die schwache Konvergenz der $x^{(n)}$ gegen x sofort aus [132].

Wir können somit den folgenden Satz formulieren:

Satz. Die Bedingungen (88) und (89) sind notwendig und hinreichend für die Existenz eines schwachen Grenzelements für die Folge der Elemente $x^{(n)}(\xi_1^{(n)}, \xi_2^{(n)}, \dots)$. Sind sie erfüllt, so besitzt das Grenzelement die Komponenten $\xi_1, \xi_2, \dots$

171. Vollstetige Operatoren im Raum l_2 . Wir leiteten in [108] eine hinreichende Bedingung dafür her, daß eine unendliche Matrix einen vollstetigen Operator in l_2 definiert, nämlich: *Konvergiert die Doppelreihe*

$$\sum_{n,m=1}^{\infty} |a_{nm}|^2, \quad (90)$$

so definiert die Matrix der a_{nm} einen vollstetigen Operator in l_2 .

Die Konvergenz der Reihe (90) ist nur hinreichend dafür, daß der durch die Matrix der a_{nm} definierte Operator vollstetig ist. Wir können zeigen, daß notwendig und hinreichend ist, daß der Grenzübergang in Formel (8) gleichmäßig für alle Elemente x und y gilt, deren Norm nicht größer als 1 ist. Die Gleichung $(E - \mu A)x = y$ hat in l_2 die Gestalt

$$\xi_n - \mu \sum_{m=1}^{\infty} a_{nm} \xi_m = \eta_n \quad (n=1, 2, \dots), \quad (91)$$

wobei $(\eta_1, \eta_2, \dots)$ das gegebene und $(\xi_1, \xi_2, \dots)$ das gesuchte Element aus l_2 ist. Ist A ein vollstetiger Operator, so gilt für das System (91) das in [135] Gesagte. Nun sei A ein selbstadjungierter vollstetiger Operator und ψ_k ($k=1, 2, \dots$) ein abgeschlossenes normiertes Orthogonalsystem seiner Elemente. Ferner sei ein unitärer Operator U in l_2 durch die Bedingungen $\varphi_k = U\psi_k$ definiert, wobei die φ_k die obigen Einheitsvektoren aus l_2 sind. Nehmen wir die ψ_k als neue Einheitsvektoren in l_2 , so geht A in diesen neuen Vektoren in den Operator $B = UAU^{-1}$ über. Seine Komponenten werden durch $\{B\}_{nm} = (B\psi_m, \psi_n) = \lambda_m(\psi_m, \psi_n)$ definiert, denn es ist $B\psi_m = \lambda_m\psi_m$. Folglich gilt

$$\{B\}_{nm} = \begin{cases} \lambda_m & \text{für } m=n, \\ 0 & \text{für } m \neq n, \end{cases} \quad (92)$$

d. h., in den Einheitsvektoren ψ_k entspricht dem Operator B eine Diagonalmatrix, in deren Hauptdiagonale die Eigenwerte von B stehen. Dies gilt auch für jeden linearen selbstadjungierten oder unitären Operator mit reinem Punktspektrum [146].

Der Operator A ist wegen $A = U^{-1}BU$ dem Operator B unitär äquivalent. Berücksichtigen wir das oben Gesagte, so können wir behaupten, daß die den vollstetigen selbstadjungierten Operatoren in l_2 zugeordneten Matrizen denjenigen Diagonalmatrizen unitär äquivalent sind, deren Diagonalelemente λ_m den Bedingungen aus [136] genügen.

172. Integraloperatoren im Raum L_2 . Die früher im Raum L_p betrachteten Integraloperatoren wollen wir jetzt im Raum L_2 genauer betrachten. Es sei

$$\varphi(x) = \int_a^b K(x, y) f(y) dy, \quad (93)$$

wobei $K(x, y)$ eine im Intervall Δ_0 ($a \leq x \leq b, a \leq y \leq b$) meßbare Funktion ist. Daher ist $K(x, y)$ für fast alle x aus $[a, b]$ bezüglich y meßbar und umgekehrt. Wir nehmen ferner an, daß $K(x, y)$ für fast alle x als Funktion von y zu L_2 gehört und umgekehrt, d. h.

$$K^2(x) = \int_a^b |K(x, y)|^2 dy < \infty, \quad (94)$$

$$K_1^2(y) = \int_a^b |K(x, y)|^2 dx < \infty, \quad (95)$$

wobei $K(x)$ und $K_1(y)$ meßbare nichtnegative Funktionen sind [67, 68]. Aus (94) folgt, daß zu jedem $f(y) \in L_2$ das Integral (93) für fast alle x existiert und die Funktion $\varphi(x)$ meßbar ist [67, 68]. Die Transformation (93) ist bei Gültigkeit von (94) genau dann beschränkt, wenn für jedes $f(x) \in L_2$ eine positive Zahl N existiert, für die

$$\int_a^b |\varphi(x)|^2 dx = \int_a^b \left| \int_a^b K(x, y) f(y) dy \right|^2 dx \leq N^2 \int_a^b |f(y)|^2 dy \quad (96)$$

gilt.

Wir zeigen nun eine einfache hinreichende Bedingung für die Beschränktheit eines dem Kern $K(x, y)$ entsprechenden Operators, die derjenigen für die Beschränktheit einer Matrix vollkommen analog ist. Es gibt eine positive Zahl l , für die die Ungleichungen

$$\int_a^b |K(x, y)| dy \leq l \quad \text{und} \quad \int_a^b |K(x, y)| dx \leq l \quad (97)$$

gelten. Es genügt, die Beschränktheit des zugeordneten bilinearen Funktional zu beweisen. Ersetzen wir in dem iterierten Integral, durch das dieses Funktional ausgedrückt wird, alle Funktionen durch ihre Absolutbeträge, so läßt sich das Doppelintegral in folgender Form schreiben:

$$\begin{aligned} & \int_a^b \int_a^b |K(x, y)| \cdot |f_1(y)| \cdot |f_2(x)| \, dx dy \\ & \leq \frac{1}{2} \int_a^b \int_a^b |K(x, y)| [|f_1(y)|^2 + |f_2(x)|^2] \, dx dy \\ & = \frac{1}{2} \int_a^b \left[\int_a^b |K(x, y)| \, dx \right] |f_1(y)|^2 \, dy + \frac{1}{2} \int_a^b \left[\int_a^b |K(x, y)| \, dy \right] |f_2(x)|^2 \, dx \\ & \leq \frac{l}{2} \left[\int_a^b |f_1(y)|^2 \, dx + \int_a^b |f_2(x)|^2 \, dx \right]. \end{aligned}$$

Der letzte Ausdruck ist dabei gleich l , wenn $\|f_1\| = \|f_2\| = 1$ ist. Mit genau derselben Methode läßt sich eine noch allgemeinere Bedingung für die Beschränktheit des Operators (93) nachweisen: Es existieren eine positive Zahl l und eine positive, in $[a, b]$ stetige Funktion $\omega(x)$, für die die Ungleichungen

$$\int_a^b |K(x, y)| \, \omega(y) \, dy \leq l \omega(x), \quad \int_a^b |K(x, y)| \, \omega(x) \, dx \leq l \omega(y) \quad (98)$$

gelten.

173. Der adjungierte Operator. Im Fall eines beschränkten Operators braucht das Integral

$$\int_a^b K(x) \, \tau(x) \, dx, \quad (99)$$

wobei die nichtnegative Funktion $K(x)$ durch (94) definiert ist, für gewisse $\tau(x)$ aus L_2 keinen Sinn zu haben. Die Menge der $\tau(x)$ aus L_2 , für welche das Integral (99) sinnvoll ist, ist offenbar eine lineare Mannigfaltigkeit l in L_2 .

Satz 1. Die lineare Mannigfaltigkeit l ist in L_2 überall dicht.

Beweis. Wir müssen zeigen, daß die abgeschlossene Hülle von l den ganzen Raum L_2 ergibt. Dazu nehmen wir das Gegenteil an. Dann existiert in L_2 ein vom Nullelement verschiedenes Element $\pi(x)$, das zu dem durch die Abschließung von l erhaltenen Unterraum und somit zu allen $\tau(x)$ aus l orthogonal ist. Es genügt also zu zeigen, daß das Element $\pi(x) = \pi_1(x) + i\pi_2(x)$ äquivalent 0 ist, wenn es orthogonal zu allen $\tau(x)$ aus l ist, d. h. wenn

$$\int_a^b \tau(x) \, \overline{\pi(x)} \, dx = 0 \quad (100)$$

gilt. Wir wählen dazu speziell ein $\tau(x)$ aus l . Ferner sei m eine beliebige endliche positive Zahl, e_m die Menge der x , für die $K(x) \leq m$ ist, und e'_m eine beliebige Unter-
menge von e_m , deren Maß höchstens gleich m ist. Wir definieren $\tau(x)$ so, daß $\tau(x) = 1$ für $x \in e'_m$ und $\tau(x) = 0$ für alle anderen x ist. Dieses $\tau(x)$ gehört zu l , und wir erhalten,

wenn wir (100) berücksichtigen,

$$\int_{e'_m} \overline{\pi(x)} dx = \int_{e'_m} [\pi_1(x) - i\pi_2(x)] dx = 0. \quad (101)$$

Da diese Gleichung für jede Untermenge e'_m gilt, ist z. B.

$$\int_{e'_m} \pi_1^+(x) dx = 0, \quad (102)$$

wobei $\pi_1^+(x)$ der positive Teil von $\pi_1(x)$ ist; $\pi_1^+(x)$ ist also auf e'_m äquivalent 0. Lassen wir m unbegrenzt wachsen und beachten wir (94), so sehen wir, daß $\pi_1^+(x)$ auf $[a, b]$ ebenfalls äquivalent 0 ist. Analoges läßt sich auch für $\pi_1^-(x)$, $\pi_2^+(x)$ und $\pi_2^-(x)$ behaupten. Damit ist der Satz bewiesen.

Beim Beweis dieses Satzes verwendeten wir nur, daß $K(x)$ eine beliebige gegebene, nichtnegative, fast überall endliche und auf $[a, b]$ meßbare Funktion ist.

Im folgenden bezeichnen wir mit l_1 die analoge lineare Mannigfaltigkeit für das Produkt $K_1(x) \tau(x)$; sie ist ebenfalls in L_2 überall dicht.

Satz 2. *Definiert die Formel (93) einen beschränkten Operator, so ist der adjungierte Operator ein Integraloperator mit dem Kern*

$$K^*(x, y) = \overline{K(y, x)}. \quad (103)$$

Beweis. Bezeichnen wir den Operator (93) mit A und benutzen wir die Definitionsgleichung $(Ax, y) = (x, A^*y)$ für den adjungierten Operator, so ist

$$\int_a^b \left[\int_a^b K(x, y) \tau(y) dy \right] \overline{g(x)} dx = \int_a^b \tau(y) \overline{g^*(y)} dy, \quad (104)$$

wobei $g^*(x) = A^*g(x)$ ist, und wir nehmen $\tau(y) \in l_1$ an. Wegen

$$\int_a^b |K(x, y) \overline{g(x)}| dx \leq \left(\int_a^b |K(x, y)|^2 dx \right)^{\frac{1}{2}} \left(\int_a^b |g(x)|^2 dx \right)^{\frac{1}{2}} = K_1(y) \cdot \|g\|$$

und $\tau(x) \in l_1$ existiert eines der iterierten Integrale für $|K(x, y) \overline{g(x)} \tau(y)|$. Folglich können wir in dem auf der linken Seite von (104) stehenden Integral die Reihenfolge der Integration ändern; die Formel (104) läßt sich dann in der Gestalt

$$\int_a^b \tau(y) \left[\int_a^b K(x, y) \overline{g(x)} dx - \overline{g^*(y)} \right] dy = 0$$

schreiben. Wiederholen wir den Beweis von Satz 1, so können wir uns davon überzeugen, daß die Differenz in der eckigen Klammer äquivalent 0 ist, und gehen wir zu den konjugierten Größen über, so ist

$$g^*(y) = \int_a^b \overline{K(x, y)} g(x) dx. \quad (105)$$

Damit ist der Satz bewiesen.

Die Gleichung $(Ax, y) = (x, A^*y)$ läßt sich für Integraloperatoren auf Grund von Satz 2 auf die Form

$$\int_a^b \left[\int_a^b K(x, y) f(y) dy \right] \overline{g(x)} dx = \int_a^b \left[\int_a^b K(x, y) \overline{g(x)} dx \right] f(y) dy \quad (106)$$

bringen; die Reihenfolge der Integration kann also geändert werden. Das entsprechende Doppelintegral braucht nicht zu existieren. Genügt der Kern außer den erwähnten Bedingungen noch der Bedingung

$$K(x, y) = \overline{K(y, x)}, \quad (107)$$

so stimmt der Operator (105) mit (93) überein, d. h., der Operator (93) ist selbstadjungiert.

174. Vollstetige Operatoren. Wir sahen oben, daß für eine im Quadrat $\Delta_0 [a < x < b; a < y < b]$ meßbare und der Bedingung

$$\iint_{\Delta_0} |K(x, y)|^2 dx dy < \infty \quad (108)$$

genügende Funktion $K(x, y)$ der Operator (93) in L_2 vollstetig ist [108]. Für die Integralgleichung

$$\int_a^b K(x, y) f(y) dy = \lambda f(x) + \psi(x), \quad (109)$$

in der $\psi(x)$ die gegebene und $f(x)$ die gesuchte Funktion aus $L_2 [a, b]$ ist, gilt das in [135] Gesagte.

Ist der Operator (93) selbstadjungiert, d. h. ist $\overline{K(x, y)}$ äquivalent $K(y, x)$, so ist auf (109) das in [136] Gesagte anwendbar.

Wir wollen nun zeigen, daß das Integral (108) gleich der absoluten Norm des Operators (93) ist. Dazu benötigen wir den folgenden

Hilfssatz. Ist $\varphi_n(x)$, $n=1, 2, \dots$, ein vollständiges normiertes Orthogonalsystem auf dem Intervall $[a, b]$, so ist $\varphi_{mn}(x, y) = \varphi_m(x) \varphi_n(y)$ ein im Quadrat Δ_0 vollständiges normiertes Orthogonalsystem.

Beweis. Nach Voraussetzung ist

$$\int_a^b \varphi_i(x) \overline{\varphi_k(x)} dx = \begin{cases} 0 & \text{für } i \neq k, \\ 1 & \text{für } i = k. \end{cases}$$

Die Funktionen $\varphi_{mn}(x, y)$ gehören offenbar zu $L_2(\Delta_0)$, und es ist

$$\iint_{\Delta_0} \varphi_{mn}(x, y) \overline{\varphi_{pq}(x, y)} dx dy = \int_a^b \varphi_m(x) \overline{\varphi_p(x)} dx \cdot \int_a^b \varphi_n(y) \overline{\varphi_q(y)} dy$$

auf Grund des Satzes von FUBINI. Aus dieser Beziehung folgt, daß die $\varphi_{mn}(x, y)$ in Δ_0 orthogonal und normiert sind. Um die Vollständigkeit dieses Systems zu beweisen, zeigen wir, daß die Funktion f äquivalent 0 in Δ_0 ist, wenn sie zu allen φ_{mn} orthogonal ist [58].

Es sei

$$\iint_{\Delta_0} f(x, y) \overline{\varphi_m(x) \varphi_n(y)} dx dy = 0,$$

also

$$\int_a^b \left[\int_a^b f(x, y) \varphi_n(y) dy \right] \overline{\varphi_m(x)} dx = 0,$$

und wir erhalten wegen der Vollständigkeit des Systems der $\varphi_m(x)$, wenn wir zu den konjugierten Größen übergehen,

$$\int_a^b \overline{f(x, y) \varphi_n(y)} dy = 0 \quad \text{fast überall bezüglich } x \text{ in } [a, b].$$

Mit Hilfe derselben Überlegungen gelangen wir zu der Beziehung $f(x, y) = 0$ fast überall in Δ_0 . Damit ist der Hilfssatz bewiesen.

Ist b_{mn} der Fourierkoeffizient des Kerns $K(x, y)$, der zu $L_2(\Delta_0)$ gehört, so ist wegen (108)

$$b_{mn} = \iint_{\Delta_0} K(x, y) \overline{\varphi_m(x)} \varphi_n(y) dx dy.$$

Das Quadrat der absoluten Norm des Operators A , der zu diesem Kern gehört [138], ist gleich

$$\begin{aligned} N^2(A) &= \sum_{m,n} |(A\varphi_n, \varphi_m)|^2 = \sum_{m,n} \left| \int_a^b \left[\int_a^b K(x, y) \varphi_n(y) dy \right] \overline{\varphi_m(x)} dx \right|^2 \\ &= \sum_{m,n} \left| \iint_{\Delta_0} K(x, y) \varphi_m(x) \varphi_n(y) dx dy \right|^2 = \sum_{m,n} |b_{mn}|^2. \end{aligned}$$

Die letzte Summe ist wegen der Parsevalschen Gleichung gleich dem Integral (108). Ist (108) erfüllt und ist A ein selbstadjungierter Operator, so gilt

$$\iint_{\Delta_0} |K(x, y)|^2 dx dy = \sum_k \lambda_k^2, \quad (110)$$

wobei die λ_k die Eigenwerte sind.

Die Ergebnisse aus [135] und [136] bleiben auch für ein unendliches Intervall und für einen mehrdimensionalen Operator

$$\varphi(x) = \int_D K(x, y) f(y) dy \quad (111)$$

gültig; dabei sind $x(x_1, x_2, \dots, x_n)$ und $y(y_1, y_2, \dots, y_n)$ Punkte des n -dimensionalen Raumes R_n , ferner ist $dy = dy_1 dy_2 \dots dy_n$ und D ein Bereich des R_n .

175. Die Spektralschar. Wir bezeichnen den Operator (93) mit K , nehmen an, daß er vollstetig und selbstadjungiert ist, und zeigen für ihn die Konstruktion der Spektralschar $\mathcal{E}_\lambda$ und der Resolvente $R_\lambda = (A - \lambda E)^{-1}$.

Statt $\mathcal{E}_\lambda$ führen wir zunächst eine andere Funktion ein, die uns die Möglichkeit gibt, sie als Integraloperator darzustellen; wir setzen nämlich

$$\theta_\lambda = \begin{cases} \mathcal{E}_\lambda & \text{für } \lambda < 0, \\ 0 & \text{für } \lambda = 0, \\ \mathcal{E}_\lambda - E & \text{für } \lambda > 0. \end{cases} \quad (112)$$

Da ein reines Punktspektrum vorliegt, können wir behaupten, daß $\mathcal{E}_\lambda$ in den Raum derjenigen Eigenfunktionen $\varphi_k(x)$ projiziert, für die λ_k höchstens gleich λ ist. Die Projektion der Funktion $f(x)$ auf den eindimensionalen Unterraum der Eigenfunktion $\varphi_k(x)$ ist das Produkt

$$\alpha_k \varphi_k(x) = \int_a^b \varphi_k(x) \overline{\varphi_k(y)} f(y) dy,$$

wobei α_k der Fourierkoeffizient von $f(x)$ ist. Somit ist der Projektionsoperator auf den eindimensionalen Unterraum der Integraloperator mit dem Kern $\varphi_k(x) \overline{\varphi_k(y)}$, und wir können

$$\theta_\lambda f(x) = \int_a^b \sum_{\alpha} \varphi_k(x) \overline{\varphi_k(y)} f(y) dy \quad \text{für } \lambda < 0$$

schreiben, wobei über diejenigen k summiert wird, für die $\lambda_k \leq \lambda$ ist. Diese Summe enthält endlich viele Summanden wegen der Eigenschaft des Spektrums eines vollstetigen Operators. Somit ist der Operator θ_λ für $\lambda < 0$ ein Integraloperator mit dem Kern

$$\theta(x, y; \lambda) = \sum_{\lambda_k \leq \lambda} \varphi_k(x) \overline{\varphi_k(y)} \quad \text{für } \lambda < 0. \quad (113)$$

Berücksichtigen wir (112) und das oben über $\mathfrak{E}_\lambda$ Gesagte, so ist θ_λ für $\lambda > 0$ ein Integraloperator mit dem Kern

$$\theta(x, y; \lambda) = - \sum_{\lambda_k > \lambda} \varphi_k(x) \overline{\varphi_k(y)} \quad \text{für } \lambda > 0; \quad (114)$$

auch hier treten nur endlich viele Summanden auf. Durchläuft λ die Eigenwerte, so erfährt der Kern einen Sprung. Aus (112) folgt unmittelbar $\theta_\lambda^2 = \theta_\lambda$ für $\lambda < 0$ und $\theta_\lambda^2 = -\theta_\lambda$ für $\lambda > 0$. Wir können also

$$\int_a^b \theta(x, t; \lambda) \theta(t, y; \lambda) dt = \pm \theta(x, y; \lambda) \quad (115)$$

schreiben; dabei gilt das positive Vorzeichen für $\lambda < 0$ und das negative Vorzeichen für $\lambda > 0$.

Die Funktion $R_l f(x)$ ist offenbar eine Lösung der Gleichung (109) für $\lambda = l$, wobei aber vorausgesetzt ist, daß l von 0 verschieden sein und mit keinem der λ_k übereinstimmen soll. Für die Resolvente läßt sich auch eine andere Beziehung angeben, wenn man von der Formel

$$R_l f(x) = \int_{-\infty}^{\infty} \frac{d\mathfrak{E}_\lambda f(x)}{\lambda - l} \quad (116)$$

ausgeht, in der tatsächlich über ein endliches Intervall integriert wird, das das Spektrum des Operators enthält.

Es sei $\lambda = 0$ kein Eigenwert. Ersetzen wir $\mathfrak{E}_\lambda$ durch θ_λ mit Hilfe von (112) und berücksichtigen wir, daß θ_λ beim Durchgang durch $\lambda = 0$ den zusätzlichen Sprung $-E$ hat, so erhält (116) die Gestalt

$$R_l f(x) = -\frac{1}{l} f(x) + \lim_{\substack{\epsilon_1 \rightarrow +0 \\ \epsilon_2 \rightarrow +0}} \left[\int_{-\infty}^{-\epsilon_1} \frac{d_\lambda \left[\int_a^b \theta(x, y; \lambda) f(y) dy \right]}{\lambda - l} + \int_{\epsilon_2}^{\infty} \frac{d_\lambda \left[\int_a^b \theta(x, y; \lambda) f(y) dy \right]}{\lambda - l} \right]. \quad (117)$$

176. Die Spektralschar (Fortsetzung). Die Spektralschar wurde für ganz allgemeine Integraloperatoren von CARLEMAN¹⁾ eingeführt. Vom modernen Gesichtspunkt aus haben STONE²⁾ und ACHIESER³⁾ die Theorie dargestellt. Sie untersuchten Integraloperatoren von allgemeinerem Typ als selbstadjungierte beschränkte Operatoren, über die wir jetzt sprechen werden. Beschränkte selbstadjungierte Operatoren wurden von HILBERT, HELLINGER u. a. untersucht.

Wir geben nun kurz einige Resultate für den letzten Fall an. Der Kern $K(x, y)$ eines beschränkten selbstadjungierten Operators K läßt sich durch Kerne $K^{(n)}(x, y)$, $n = 1, 2, \dots$, approximieren, denen vollstetige selbstadjungierte Operatoren entsprechen. Wir können also zeigen, daß der durch (112) definierte Operator θ_λ , wobei $\mathfrak{E}_\lambda$ die Spektralschar von K ist, für $\lambda \neq 0$ ein Integraloperator ist und die Formeln

$$\theta_\lambda f(x) = \int_a^b \theta(x, y; \lambda) f(y) dy, \quad (118)$$

$$\int_a^b K(x, y) f(y) dy = \int_{-\infty}^{\infty} \lambda d_\lambda \left[\int_a^b \theta(x, y; \lambda) f(y) dy \right] \quad (119)$$

¹⁾ T. CARLEMAN, Sur les équations intégrales singulières à noyau réel et symétrique, A.-B. Lundequistska Bokhandeln, Uppsala 1923.

²⁾ M. H. STONE, Linear transformations in Hilbert space and their applications to analysis, Amer. Math. Soc. Coll. Publ., Vol. XV, New York 1932.

³⁾ N. I. ACHIESER, Integraloperatoren mit Carlemanschen Kernen (russ.), Uspechi mat. nauk 2: 5 (1947), 93–131.

und ebenfalls (115) gelten. In (119) müssen wir das rechts stehende Integral als uneigentliches Integral in bezug auf $\lambda = 0$ auffassen.

Wir wollen annehmen, daß der Operator (97) kein Punktspektrum hat, und leiten für ihn aus [149] die allgemeinen Formeln her. Es seien $\omega_k(x)$ Elemente aus L_2 , die den y_k aus [149] entsprechen, wobei die $\omega_k(x)$ ein normiertes Orthogonalsystem bilden mögen. Mit Hilfe von $\theta(x, y; \lambda)$ erhalten wir ein Fundamentalsystem von Differentiallösungen

$$\pi_k(x, \lambda) = \begin{cases} \int_a^b \theta(x, y; \lambda) \omega_k(y) dy & \text{für } \lambda < 0, \\ \int_a^b \theta(x, y; \lambda) \omega_k(y) dy + \omega_k(x) & \text{für } \lambda > 0. \end{cases}$$

Im Punkt $\lambda = 0$ hat der Operator θ_λ den Sprung $-E$.

Es gilt

$$\varrho_k(\lambda) = \int_a^b |\pi_k(x, \lambda)|^2 dx; \quad (120)$$

sind $\varphi(x)$ und $\psi(x)$ zwei beliebige Elemente aus L_2 und setzen wir

$$g_k(\lambda) = \int_a^b \pi_k(x, \lambda) \overline{\varphi(x)} dx, \quad (121)$$

$$h_k(\lambda) = \int_a^b \pi_k(x, \lambda) \overline{\psi(x)} dx, \quad (122)$$

so erhalten die Formeln aus [149] die Gestalt

$$\int_a^b \varphi(x) \overline{\psi(x)} dx = \sum_k \int_m^M \frac{\overline{dg_k(\lambda)} dh_k(\lambda)}{d\varrho_k(\lambda)}, \quad (123)$$

$$\int_a^b \int_a^b K(x, y) \varphi(y) \overline{\psi(x)} dx dy = \sum_k \int_m^M \lambda \frac{\overline{dg_k(\lambda)} dh_k(\lambda)}{d\varrho_k(\lambda)}, \quad (124)$$

$$\int_a^b \mathcal{E}_\lambda \varphi(x) \overline{\psi(x)} dx = \sum_k \int_m^\lambda \frac{\overline{dg_k(\mu)} dh_k(\mu)}{d\varrho_k(\mu)}. \quad (125)$$

Dabei sind m und M die Grenzen des Operators, und das in (124) links stehende Integral ist als iteriertes Integral mit beliebiger Reihenfolge der Integration aufzufassen. Gehört das Element $\varphi(x)$ zu der linearen Mannigfaltigkeit l , so existiert (124) als Doppelintegral. Die übrigen Formeln aus [149] haben die Gestalt

$$\varphi(x) = \sum_k \int_m^M \frac{\overline{dg_k(\lambda)}}{d\varrho_k(\lambda)} d\pi_k(x, \lambda), \quad (126)$$

$$\int_a^b K(x, y) \varphi(y) dy = \sum_k \int_m^M \lambda \frac{\overline{dg_k(\lambda)}}{d\varrho_k(\lambda)} d\pi_k(x, \lambda), \quad (127)$$

$$\mathcal{E}_\lambda \varphi(x) = \sum_k \int_m^\lambda \frac{\overline{dg_k(\mu)}}{d\varrho_k(\mu)} d\pi_k(x, \mu). \quad (128)$$

Haben die Summen unendlich viele Summanden, so müssen die Reihen im Mittel gegen die links stehenden Größen konvergieren.

Die Differentiallösungen $\pi(x, \lambda)$ genügen der Beziehung

$$\int_a^b K(x, y) \pi(y, \lambda) dy = \int_m^\lambda \mu d\pi(x, \mu), \quad (129)$$

wobei wir, wie immer, $\pi(x, m) = 0$ annehmen. Die Orthogonalität dieser Lösungen läßt sich durch

$$\int_a^b \Delta_1 \pi_p(x, \lambda) \cdot \overline{\Delta_2 \pi_q(x, \lambda)} dx = 0 \quad (p \neq q)$$

ausdrücken.

Zur Herleitung der oben erwähnten Formeln können wir von einem beliebigen abgeschlossenen Orthogonalsystem von Differentiallösungen $\pi_k(x, \lambda)$ ausgehen. Im folgenden werden wir Beispiele für Integraloperatoren im Raum L_2 untersuchen.

177. Unitäre Transformationen im Raum L_2 . Nicht jede unitäre Transformation $\varphi(x) = Uf(x)$ in L_2 läßt sich in Integralform darstellen.

Ein Beispiel dafür ist die identische Transformation $\varphi(x) = f(x)$, die sich nur in Integralform darstellen läßt, wenn zu der Stammfunktion übergegangen wird:

$$\int_0^x \varphi(y) dy = \int_{-\infty}^{\infty} K(x, y) f(y) dy.$$

Als Grundintervall wird dabei $(-\infty, \infty)$ angenommen; ferner setzen wir $K(x, y) = 1$ für $0 \leq y \leq x$ sowie $K(x, y) = 0$ für $y < 0$ und $y > x$, wenn $x > 0$ ist, und analog für $x < 0$. Ein ähnliches Resultat gilt auch für den beschränkten Operator A . Das Grundintervall sei wieder $(-\infty, \infty)$. Wir wählen ein beliebiges, aber festes x und untersuchen für $Af(x)$ die Stammfunktion:

$$l(f) = \int_0^x [Af(t)] dt;$$

sie ist distributiv, und wegen der Schwarzischen Ungleichung gilt die Abschätzung

$$|l(f)| \leq \left[\int_0^x |Af(t)|^2 dt \right]^{1/2} \left[\int_0^x dt \right]^{1/2} \leq \|A\| \cdot \sqrt{x} \cdot \|f\| \quad (x > 0).$$

Wir können $l(f)$ demzufolge als lineares beschränktes und von x abhängiges Funktional auffassen. Auf Grund des Satzes aus [123] gilt

$$\int_0^x [Af(t)] dt = \int_{-\infty}^{\infty} K(x, y) f(y) dy;$$

dabei gehört $K(x, y)$ für jedes $x \in (-\infty, \infty)$ bezüglich y zu $L_2(-\infty, \infty)$, und es ist

$$\int_{-\infty}^{\infty} |K(x, y)|^2 dy \leq \|A\|^2 |x|.$$

Wir beweisen nun einen Satz, der dadurch, daß wir zur Stammfunktion übergehen, die allgemeine analytische Form für unitäre Operatoren angibt. Dieser Satz wurde zuerst von BOCHNER¹⁾ für das Intervall $0 \leq x < \infty$ bewiesen. Der Beweis hängt jedoch nicht von der Wahl des Intervalls ab. Der Bestimmtheit wegen nehmen wir das Intervall $(-\infty, \infty)$ und bezeichnen mit L_2 die Klasse der Funktionen auf $(-\infty, \infty)$.²⁾

¹⁾ S. BOCHNER, Inversion formulae and unitary transformations, Ann. Math. **35** (1934), 111–115.

²⁾ Vgl. F. RIESZ und B. SZ.-NAGY, Vorlesungen über Funktionalanalysis (vgl. das Literaturverzeichnis).

Satz. Die Funktionen $K(x, y)$ und $L(x, y)$ mögen für jedes feste x aus $(-\infty, \infty)$ bezüglich y zu L_2 gehören, wobei für alle a und b aus $(-\infty, \infty)$ die Beziehungen

$$\left. \begin{aligned} \int_{-\infty}^{\infty} K(a, y) \overline{K(b, y)} dy \\ \int_{-\infty}^{\infty} L(a, y) \overline{L(b, y)} dy \end{aligned} \right\} = \begin{cases} \min(|a|, |b|) & \text{für } ab > 0, \\ 0 & \text{für } ab \leq 0 \end{cases} \quad (130)$$

und

$$\int_0^b K(a, y) dy = \int_0^a \overline{L(b, y)} dy \quad (131)$$

gelten. Dann definieren die Formeln

$$\int_0^a \varphi(y) dy = \int_{-\infty}^{\infty} \overline{L(a, y)} f(y) dy, \quad (132)$$

$$\int_0^a f(y) dy = \int_{-\infty}^{\infty} \overline{K(a, y)} \varphi(y) dy \quad (133)$$

die unitäre Transformation $\varphi(x) = Uf(x)$ und ihre Inverse.

Existiert umgekehrt eine unitäre Transformation $\varphi(x) = Uf(x)$, so gibt es Funktionen $K(x, y)$ und $L(x, y)$, welche die obigen Eigenschaften haben und mit deren Hilfe U und U^{-1} durch (132) bzw. (133) ausgedrückt werden können.

Beweis. Zunächst beweisen wir den ersten Teil des Satzes. Wir definieren eine Funktion $f_a(x)$ durch

$$f_a(x) = \begin{cases} 1 & \text{für } 0 < x \leq a, \\ 0 & \text{für } x \leq 0 \text{ und } x > a \quad (a > 0), \end{cases}$$

$$f_a(x) = \begin{cases} 1 & \text{für } a \leq x < 0, \\ 0 & \text{für } x < a \text{ und } x \geq 0 \quad (a < 0) \end{cases}$$

und $f_0(x) \equiv 0$ sowie zwei Operatoren U_0 und V_0 durch

$$K(a, x) = U_0 f_a(x), \quad L(a, x) = V_0 f_a(x). \quad (134)$$

Bilden wir alle möglichen Linearkombinationen der Funktionen $f_a(x)$ für verschiedene a , so erhalten wir die lineare Mannigfaltigkeit l der stückweise konstanten Funktionen, die in endlich vielen endlichen Intervallen konstant und außerhalb dieser Intervalle gleich 0 sind. Dabei spielen die Werte der Funktionen in den Endpunkten der Intervalle keine Rolle, da wir äquivalente Funktionen miteinander identifizieren. Wir erweitern die Operatoren U_0 und V_0 auf die lineare Mannigfaltigkeit l , indem wir die Distributivität von U_0 und V_0 zugrunde legen. Es ist leicht zu sehen, daß diese Erweiterung eindeutig ist. Wir bezeichnen mit Uf und Vf die auf l erhaltenen distributiven Operatoren. Die Formeln (130), (131) und (134) ergeben

$$(U_0 f_a, U_0 f_b) = (f_a, f_b), \quad (V_0 f_a, V_0 f_b) = (f_a, f_b), \quad (135)$$

$$(U_0 f_a, f_b) = (f_a, V_0 f_b). \quad (136)$$

Benutzen wir die Distributivität, so erhalten wir analoge Formeln für U und V auf l :

$$(Uf, Ug) = (f, g), \quad (Vf, Vg) = (f, g), \quad (137)$$

$$(Uf, g) = (f, Vg); \quad (138)$$

dabei gehören f und g zu l . Aus (137) folgt, daß die Operatoren U und V auf l die Norm invariant lassen, und wir erhalten, da die lineare Mannigfaltigkeit l in L_2 dicht ist [173], eine eindeutige Erweiterung der Operatoren U und V (bezüglich ihrer Stetigkeit) auf ganz L_2 . Infolge der Stetigkeit des Skalarprodukts gelten auf L_2 die Formeln (137) und (138), und die Operatoren U und V lassen die Norm sowie das Skalarprodukt auf L_2 invariant. Aus (138) folgt $V = U^*$. Ersetzen wir f in (138) durch Vf und verwenden wir (137), so ergibt sich $UU^* = E$ und analog, wenn wir g durch Ug ersetzen, $U^*U = E$. Also ist U eine unitäre Transformation und V ihre Inverse [137]. Es muß nun noch die Gültigkeit von (132) und (133) nachgewiesen werden. Die erste Formel ergibt sich aus (138), wenn $g(x) = f_a(x)$ gesetzt wird, und die zweite, wenn $f(x) = f_a(x)$ gesetzt und zu den konjugierten Größen übergegangen wird.

Wir kommen nun zum Beweis des zweiten Teils. Gegeben sei ein unitärer Operator U , und es sei $V = U^{-1} = U^*$. Ferner konstruieren wir die Funktionen

$$K(a, x) = Uf_a(x), \quad L(a, x) = U^{-1}f_a(x). \quad (139)$$

Setzen wir wie oben $\varphi(x) = Uf(x)$ und beachten wir, daß U unitär ist, so erhalten wir

$$(\varphi, f_a) = (Uf, f_a) = (f, U^{-1}f_a), \quad (f, f_a) = (U^{-1}\varphi, f_a) = (\varphi, Uf_a);$$

dies führt wegen (139) auf (132) und (133). Die Formeln (137) und (138) gelten für die obengenannten U und V . Setzen wir in ihnen $f(x) = f_a(x)$ und $g(x) = f_b(x)$, so ergibt sich (130) und (131). Damit ist der Satz vollständig bewiesen.

178. Fouriertransformationen. WATSON¹⁾ untersuchte im Intervall $0 \leq x < \infty$ Kerne der Form

$$K(a, x) = \frac{\overline{\chi(ax)}}{x}, \quad L(a, x) = \frac{\chi(ax)}{x};$$

dabei sei $\chi(0) = 0$, und $\frac{\chi(x)}{x}$ gehöre zu $L_2(0, \infty)$. Die Bedingung (130) erhält damit die Gestalt

$$\int_0^\infty \frac{\chi(ax) \overline{\chi(bx)}}{x^2} dx = \min(a, b) \quad (a > 0, b > 0),$$

und (131) ist von selbst erfüllt.

Wir wollen nun die Fouriertransformation betrachten, für die $(-\infty, \infty)$ das Grundintervall ist und

$$K(a, x) = \frac{1}{\sqrt{2\pi}} \frac{e^{-iax} - 1}{-ix}, \quad L(a, x) = \frac{1}{\sqrt{2\pi}} \frac{e^{iax} - 1}{ix} \quad (140)$$

gilt. Der Absolutbetrag des Zählers ist höchstens gleich 2, und die beiden Funktionen gehören zu L_2 . Die Bedingung (131) läßt sich leicht nachprüfen; daher werden wir uns jetzt mit (130) beschäftigen. Beide Bedingungen führen wie oben auf die eine Formel

$$I = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{(e^{-iax} - 1)(e^{ibx} - 1)}{x^2} dx = \begin{cases} \min(|a|, |b|) & \text{für } ab > 0, \\ 0 & \text{für } ab \leq 0. \end{cases} \quad (141)$$

¹⁾ G. N. WATSON, General transforms, Proc. London Math. Soc. (2) **35** (1933), 156–199. (Ann. d. Red.).

Es ergibt sich leicht

$$\int_{-\infty}^{\infty} \frac{\sin^2 \alpha x}{x^2} dx = \pi |\alpha| \quad (\alpha \text{ reell}), \quad (142)$$

wenn wir nach dem Parameter α differenzieren. Das Integral I ist auch in der Form

$$I = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{\sin^2 \frac{a}{2} x + \sin^2 \frac{b}{2} x - \sin^2 \frac{a-b}{2} x}{x^2} dx$$

darstellbar, woraus wir, wenn wir (142) verwenden,

$$I = \frac{1}{2} (|a| + |b| - |a-b|)$$

erhalten. Hieraus ergibt sich sofort (141). Wir gelangen also zu einer unitären Fouriertransformierten $F(x) = Tf(x)$ der Gestalt

$$\begin{aligned} \int_0^a F(x) dx &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \frac{e^{-iax} - 1}{-ix} f(x) dx, \\ \int_0^a f(x) dx &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \frac{e^{iax} - 1}{ix} F(x) dx. \end{aligned} \quad (143)$$

Die Fouriertransformierte kann auch noch die in [II, 173] angegebene Form haben.

Wir setzen zunächst voraus, daß $f(x)$ außerhalb eines endlichen Intervalls $[-n, n]$ verschwindet. Nach Voraussetzung gehört $f(x)$ zu $L_2[-n, n]$ und somit auch zu $L_1[-n, n]$, so daß für jedes reelle y das Integral

$$F_1(y) = \frac{1}{\sqrt{2\pi}} \int_{-n}^n e^{-iyx} f(x) dx \quad (144)$$

existiert. Es ist leicht zu zeigen, daß $F_1(y)$ stetig ist und Ableitungen besitzt. Es gilt $|e^{-iyx} f(x)| = |f(x)|$, und wir können über ein endliches Intervall nach y unter dem Integralzeichen integrieren:

$$\int_0^a F_1(y) dy = \frac{1}{\sqrt{2\pi}} \int_{-n}^n \frac{e^{-iax} - 1}{-ix} f(x) dx.$$

Vergleichen wir diese Formel mit (143) und beachten wir, daß a beliebig und $f(x)$ nach Voraussetzung außerhalb des Intervalls $[-n, n]$ gleich 0 ist, so sehen wir, daß $F_1(y)$ äquivalent $F(y)$ ist, d. h., in diesem Fall kann die Fouriertransformierte in der Gestalt

$$F(y) = \frac{1}{\sqrt{2\pi}} \int_{-n}^n e^{-iyx} f(x) dx \quad (145)$$

geschrieben werden.

Im allgemeinen hat das Integral

$$\frac{1}{\sqrt{2\pi}} \int_{-n}^n e^{-iyx} f(x) dx \quad (146)$$

keinen Sinn, da aus der Zugehörigkeit von $f(x)$ zu $L_2(-\infty, \infty)$ noch nicht folgt, daß $f(x)$ auch zu $L_1(-\infty, \infty)$ gehört. Wir betrachten die Funktion $f_{nm}(x)$, die für $-n \leq x \leq m$ gleich $f(x)$ und für $x > m$ und $x < -n$ gleich 0 sein soll. Die Fouriertransformierte dieser Funktion ist, wie wir wissen,

$$F_{nm}(y) = \frac{1}{\sqrt{2\pi}} \int_{-n}^m e^{-iyx} f_{nm}(x) dx. \quad (147)$$

Nun ist $f_{nm}(x) \Rightarrow f(x)$ in L_2 für $n \rightarrow \infty, m \rightarrow \infty$ und folglich $Tf_{nm}(x) \Rightarrow Tf(x)$. Bezeichnen wir mit lm den Grenzwert im Mittel (in L_2), so hat jede Funktion aus L_2 die Fouriertransformierte

$$F(y) = Tf = \lim_{\substack{n \rightarrow \infty \\ m \rightarrow \infty}} \frac{1}{\sqrt{2\pi}} \int_{-n}^m e^{-iyx} f(x) dx. \quad (148)$$

Gehört die Funktion $f(x)$ nicht nur zu L_2 , sondern auch zu L_1 im Intervall $(-\infty, \infty)$, so existiert das Integral (146) für jedes reelle y und ist der Grenzwert des Integrals (147) für $m, n \rightarrow \infty$. Existieren der Grenzwert im Mittel und der gewöhnliche Grenzwert, so stimmen sie überein, so daß in diesem Fall die Fouriertransformierte in der Gestalt

$$F(y) = Tf = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-iyx} f(x) dx \quad (f(x) \text{ aus } L_2 \text{ und } L_1) \quad (149)$$

geschrieben werden kann.

Das Vorhergehende gilt auch für die adjungierte (inverse) Transformation. Statt (148) haben wir, wenn wir $m = n$ setzen,

$$f(y) = T^*F = \lim_{n \rightarrow \infty} \frac{1}{\sqrt{2\pi}} \int_{-n}^n e^{iyx} F(x) dx \quad (150)$$

und statt (149)

$$f(y) = T^*F = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{iyx} F(x) dx \quad (F(x) \text{ aus } L_2 \text{ und } L_1). \quad (151)$$

Für eine Fouriertransformierte gilt die Faltungsformel, die wir nun einführen wollen (vgl. [IV/1, 52]). Es seien $g(t)$ und $f(t)$ zwei Funktionen aus L_2 . Die Funktion $g(x-t)$ gehört für jedes reelle x als Funktion von t ebenfalls zu L_2 . Es gilt nun

$$\begin{aligned} T[g(x-t)] &= \lim_{a \rightarrow \infty} \frac{1}{\sqrt{2\pi}} \int_{-a}^a g(x-t) e^{-iyt} dt \\ &= \lim_{a \rightarrow \infty} \frac{1}{\sqrt{2\pi}} \int_{x-a}^{x+a} g(u) e^{-iy(x-u)} du = e^{-iyx} \lim_{a \rightarrow \infty} \frac{1}{\sqrt{2\pi}} \int_{x-a}^{x+a} g(t) e^{iyt} dt, \end{aligned}$$

also $T[g(x-t)] = e^{-iyx} T^*[g(t)]$. Da eine unitäre Transformation das Skalarprodukt invariant läßt, gilt

$$\int_{-\infty}^{\infty} g(x-t) f(t) dt = \int_{-\infty}^{\infty} e^{-iyx} T^*[g(t)] \overline{T[f(t)]} dy$$

mit

$$T^*[g(t)] = \lim_{n \rightarrow \infty} \frac{1}{\sqrt{2\pi}} \int_{-n}^n g(t) e^{iyt} dt, \quad \overline{T[f(t)]} = \lim_{n \rightarrow \infty} \frac{1}{\sqrt{2\pi}} \int_{-n}^n f(t) e^{iyt} dt = T^*f$$

sowie

$$\int_{-\infty}^{\infty} g(x-t) f(t) dt = \int_{-\infty}^{\infty} G_1(y) F_1(y) e^{-iyx} dy \quad (152)$$

mit $G_1(y) = T^*g$ und $F_1(y) = T^*f$.

Der Hauptsatz läßt sich genauso wie für den Fall einer Funktion mehrerer Veränderlicher beweisen. Dabei wird eine unitäre Transformation T durch

$$Tf = \lim_{m_k \rightarrow \infty} \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{-m_1}^{m_1} \dots \int_{-m_n}^{m_n} f(x_1, \dots, x_n) e^{-i(x_1 y_1 + \dots + x_n y_n)} dx_1 \dots dx_n \quad (153)$$

und ihre inverse Transformation durch

$$T^*F = \lim_{m_k \rightarrow \infty} \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{-m_1}^{m_1} \dots \int_{-m_n}^{m_n} F(y_1, \dots, y_n) e^{i(x_1 y_1 + \dots + x_n y_n)} dy_1 \dots dy_n \quad (154)$$

definiert.

Wir kehren zum Fall einer Veränderlichen zurück und erwähnen noch einige Eigenschaften der Fouriertransformierten. Ersetzen wir in (148) für $m=n$ die Größe x durch $-x$, vergleichen das Ergebnis mit (150) und berücksichtigen die Beziehung $T^* = T^{-1}$, so erhalten wir $T^2 f(x) = f(-x)$ und analog $T^{*2} F(y) = F(-y)$. Ist $f(x)$ eine gerade Funktion, so ergibt die Transformation eine Funktion, die einer geraden Funktion äquivalent ist, und wir erhalten

$$F(y) = \lim_{n \rightarrow \infty} \sqrt{\frac{2}{\pi}} \int_0^n f(x) \cos xy dx, \quad f(x) = \lim_{n \rightarrow \infty} \sqrt{\frac{2}{\pi}} \int_0^n F(y) \cos xy dy.$$

Diese Formeln vermitteln eine unitäre Transformation für Funktionen aus L_2 auf dem Intervall $(0, \infty)$. Ändern wir das Vorzeichen und multiplizieren wir mit i (diese Operationen bilden offenbar unitär ab), so ergeben sich bei ungeraden Funktionen auf dem Intervall $(0, \infty)$ die zueinander inversen unitären Fouriertransformierten

$$F(y) = \lim_{n \rightarrow \infty} \sqrt{\frac{2}{\pi}} \int_0^n f(x) \sin xy dx, \quad f(x) = \lim_{n \rightarrow \infty} \sqrt{\frac{2}{\pi}} \int_0^n F(y) \sin xy dy.$$

179. Fouriertransformation und Hermite'sche Funktionen. Wir zeigen nun, daß eine Fouriertransformierte die vier Eigenwerte $\pm 1, \pm i$ hat; ihnen entspricht auf dem Intervall $(-\infty, \infty)$ ein vollständiges normiertes Orthogonalsystem von Eigenfunktionen, den Hermite'schen Funktionen [III/2, 157]. Wir erwähnen zunächst ihre Hauptformeln. Die Hermite'schen Polynome werden durch

$$H_n(x) = (-1)^n e^{x^2} \frac{d^n}{dx^n} (e^{-x^2})$$

und die Hermite'schen Funktionen durch

$$\psi_n(x) = e^{-x^2/2} H_n(x)$$

definiert. Sie sind in $(-\infty, \infty)$ orthogonal. Die normierten Hermiteschen Funktionen lauten

$$\varphi_n(x) = \frac{1}{2^{n/2} \sqrt{n!} \sqrt{\pi}} \psi_n(x)$$

und bilden ein vollständiges normiertes Orthogonalsystem. Wir zeigen, daß $\varphi_n(x)$ eine dem Eigenwert $(-i)^n$ entsprechende Eigenfunktion des Operators T ist, d. h.

$$T\varphi_n = (-i)^n \varphi_n(y). \quad (155)$$

Wir müssen also die Gültigkeit der Beziehung

$$I_n = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-iyx+(x^2/2)} \frac{d^n}{dx^n} (e^{-x^2}) dx = (-i)^n e^{y^2/2} \frac{d^n}{dy^n} (e^{-y^2})$$

nachweisen. Partielle Integration ergibt

$$I_n = \frac{(-1)^n}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-x^2} \frac{d^n}{dx^n} (e^{-iyx+(x^2/2)}) dx,$$

wenn wir berücksichtigen, daß die integralfreien Glieder verschwinden. Nun setzen wir vor das Integral den Faktor $e^{y^2/2}$ und multiplizieren den Integranden mit $e^{-y^2/2}$. Dann folgt

$$\begin{aligned} I_n &= \frac{(-1)^n}{\sqrt{2\pi}} e^{y^2/2} \int_{-\infty}^{\infty} e^{-x^2} \frac{d^n}{dx^n} e^{(x-iy)^2/2} dx \\ &= \frac{(-1)^n i^n}{\sqrt{2\pi}} e^{y^2/2} \int_{-\infty}^{\infty} e^{-x^2} \frac{d^n}{dy^n} e^{(x-iy)^2/2} dx \\ &= (-i)^n e^{y^2/2} \frac{d^n}{dy^n} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-(x^2/2)-(y^2/2)-ixy} dx. \end{aligned}$$

Durch Differentiation nach dem Parameter y sehen wir, daß das letzte Integral gleich $e^{-y^2/2}$ und somit die Formel (155) bewiesen ist. Da das System der Hermiteschen Funktionen vollständig ist, bilden die Punkte $\lambda = \pm 1$ und $\lambda = \pm i$ das ganze Spektrum des Operators T .

180. Die Operation der Multiplikation. Wir untersuchen nun die Operation der Multiplikation mit der unabhängigen Veränderlichen auf einem endlichen Intervall, dessen linker Endpunkt der Punkt $x=0$ sei. Wir betrachten also den Raum L_2 auf dem Intervall $[0, a]$ und die Multiplikation mit der unabhängigen Veränderlichen:

$$Af(x) = xf(x). \quad (156)$$

Es gilt

$$(Af, g) = \int_0^a xf(x) \overline{g(x)} dx \quad \text{und} \quad (Af, f) = \int_0^a x|f(x)|^2 dx,$$

d. h., A ist ein selbstadjungierter Operator, dessen Norm höchstens gleich a ist. Ist $f(x)$ nur in einer kleinen Umgebung von $x=a$ von 0 verschieden, so ist, wie man leicht sieht, die Norm von A genau gleich a . Die Grenzen der quadratischen Form (Af, f) sind für $\|f\|=1$ gleich $m=0$ bzw. $M=a$. Die Gleichung für die Eigenwerte und die Eigenelemente hat die Gestalt $xf(x) = \lambda f(x)$ oder $(x-\lambda)f(x) = 0$, so daß $f(x)$ äquivalent 0 ist. Es gibt also keine Eigenwerte, und das Spektrum ist rein

kontinuierlich. Die Resolvente hat offenbar die Gestalt $R_\lambda f(x) = \frac{f(x)}{x-\lambda}$. Liegt λ außerhalb von $[0, a]$, so gehört $R_\lambda f(x)$ zu L_2 ; ist $\lambda \in [0, a]$, so gehört $R_\lambda f(x)$ nicht für jedes $f(x)$ zu L_2 . Dabei bildet der Operator $(A - \lambda E) f(x) = (x - \lambda) f(x)$ den Raum L_2 eineindeutig auf die lineare Mannigfaltigkeit M_λ derjenigen Funktionen $\varphi(x) = (x - \lambda) f(x)$ ab, für die $\frac{\varphi(x)}{x-\lambda}$ zu L_2 gehört. Wir bestimmen nun die Spektralschar $\mathcal{E}_\lambda$ mit $\lambda \in [0, a]$. Mit Berücksichtigung von

$$\lim_{\tau \rightarrow +0} \int_{-\infty}^{\lambda} \frac{2\tau i}{(\sigma - x)^2 + \tau^2} d\sigma = 2i \lim_{\tau \rightarrow +0} \left(\arctan \frac{\lambda - x}{\tau} + \frac{\pi}{2} \right) = \begin{cases} 0 & \text{für } \lambda < x, \\ 2\pi i & \text{für } \lambda > x \end{cases}$$

erhalten wir für alle Elemente $f(x)$ und $\psi(x)$ [144]

$$(\mathcal{E}_\lambda f, \psi) = \lim_{\tau \rightarrow +0} \frac{1}{2\pi i} \int_{-\infty}^{\lambda} \left[\int_0^a \frac{2\tau i}{(x - \sigma)^2 + \tau^2} f(x) \overline{\psi(x)} dx \right] d\sigma = \int_0^{\lambda} f(x) \overline{\psi(x)} dx ;$$

daraus folgt

$$\mathcal{E}_\lambda f(x) = \begin{cases} f(x) & \text{für } x \leq \lambda, \\ 0 & \text{für } x > \lambda. \end{cases} \quad (157)$$

Wählen wir $f(x) \equiv 1$, so erhalten wir die Differentiallösung $\pi(x, \lambda) = 1$ für $x \leq \lambda$ und $\pi(x, \lambda) = 0$ für $x > \lambda$. Mit Hilfe von (157) und der Eigenschaft 11 aus [52] können wir leicht zeigen, daß keine zu dieser Lösung orthogonale Lösungen existieren.

Wir betrachten ferner den allgemeinen selbstadjungierten Operator

$$Bf(x) = \omega(x) f(x), \quad (158)$$

wobei $\omega(x)$ eine reelle, meßbare und im Intervall $[0, a]$ beschränkte Funktion ist. Die Gleichung für die Eigenwerte und Eigenelemente hat hier die Form $(\omega(x) - \lambda) f(x) = 0$. Ist K_λ die Menge derjenigen x , für die $\omega(x) = \lambda$ gilt, und ist das Maß von K_λ gleich 0, so ist λ kein Eigenwert. Ist jedoch das Maß von K_λ positiv, so ist λ ein Eigenwert, und jedes auf der Menge K_λ abgeschlossene Orthogonalsystem von Funktionen ist ein abgeschlossenes System von Eigenfunktionen, die diesem Eigenwert entsprechen. Dabei muß angenommen werden, daß diese Funktionen außerhalb von K_λ gleich 0 sind. Ist für jedes λ das Maß von K_λ gleich 0, so hat der Operator (158) ein rein kontinuierliches Spektrum. Seine Spektralschar lautet analog zu (157)

$$\mathcal{E}_\lambda f(x) = \begin{cases} f(x) & \text{für } \omega(x) \leq \lambda, \\ 0 & \text{für } \omega(x) > \lambda. \end{cases} \quad (159)$$

Das eben Gesagte läßt sich leicht auf Funktionen mehrerer Veränderlicher übertragen. Wir können z. B. für die Funktionen $f(x_1, x_2, \dots, x_n)$, die auf einem endlichen Intervall $a_s \leq x_s \leq b_s$ ($s = 1, 2, \dots, n$) zu L_2 gehören, einen selbstadjungierten Operator A definieren, der die Multiplikation mit der unabhängigen Veränderlichen x_k vermittelt, nämlich $Af = x_k f$. Dieser Operator hat ein rein kontinuierliches Spektrum auf dem Intervall $a_k \leq \lambda \leq b_k$, und seine Spektralschar lautet

$$\mathcal{E}_\lambda f(x_1, x_2, \dots, x_n) = \begin{cases} f(x_1, x_2, \dots, x_n) & \text{für } x_k \leq \lambda, \\ 0 & \text{für } x_k > \lambda. \end{cases} \quad (160)$$

Wir kehren nun wieder zu Funktionen einer Veränderlichen zurück. Der Operator der Multiplikation mit der unabhängigen Veränderlichen auf einem unendlichen Intervall ist nicht beschränkt; wir werden ihn im folgenden untersuchen. Der

Operator aus (158) ist beschränkt und linear, wenn wir $\omega(x)$ auf einem unendlichen Intervall beschränkt annehmen. Legen wir den Raum $L_2(-\infty, \infty)$ zugrunde und ist $\omega(x)$ auf dem Intervall $(-\infty, \infty)$ eine reelle, meßbare und beschränkte Funktion, so definiert (158) einen selbstadjungierten beschränkten Operator. Ist $\omega(x)$ im abgeschlossenen Intervall $[-\infty, \infty]$ stetig, so stimmen die Grenzen des Operators mit dem kleinsten und dem größten Wert der Funktion $\omega(x)$ überein.

181. Differenzkerne. Benutzen wir den Operator (158) auf dem Intervall $(-\infty, \infty)$ und gehen wir mit Hilfe der Transformation T zu den unitär äquivalenten Operatoren über, so können wir beschränkte selbstadjungierte Integraloperatoren mit Differenzkernen folgendermaßen leicht konstruieren:

Der zu (158) unitär äquivalente Operator $B' = T^* B T$ ist offenbar durch die Beziehung

$$B'f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \left[\omega(y) \int_{-\infty}^{\infty} f(t) e^{-ity} dt \right] e^{ixy} dy$$

definiert. Hier und im folgenden schreiben wir statt $\lim_{n \rightarrow \infty}$ stets nur das Integral mit unendlichen Grenzen. Nehmen wir an, daß $\omega(y)$ nicht nur beschränkt, sondern auch auf dem Intervall $(-\infty, \infty)$ summierbar und daß $f(t) \in L_2$ ebenfalls summierbar ist, so können wir die Reihenfolge der Integration ändern, und wir erhalten

$$B'f(x) = \int_{-\infty}^{\infty} \left[\frac{1}{2\pi} \int_{-\infty}^{\infty} \omega(y) e^{-iy(x-t)} dy \right] f(t) dt$$

oder

$$B'f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} g(x-t) f(t) dt, \quad (161)$$

wenn wir die Funktion

$$g(u) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \omega(y) e^{iyu} dy = T^* \omega \quad (162)$$

einführen.

Die Spektralschar von B' läßt sich, wie wir wissen, durch $\mathcal{E}'_\lambda = T^* \mathcal{E}_\lambda T$ ausdrücken, wobei $\mathcal{E}_\lambda$ die Spektralschar von B ist. Genügt wie in den früheren Beispielen auch hier der Kern der Bedingung (97) aus [172], d. h., ist

$$\int_{-\infty}^{\infty} |g(u)| du < \infty, \quad (163)$$

so können wir die Formel (161) nicht nur auf diejenigen $f(x)$, die auf $(-\infty, \infty)$ summierbar sind, sondern auf ganz L_2 anwenden. Dazu betrachten wir einige Beispiele.

Es sei

$$Bf(x) = \frac{2}{1+x^2} f(x). \quad (164)$$

Die Grenzen des Operators sind $m=0$ und $M=2$. Für jedes λ aus $[0, 2]$ hat die Gleichung $\frac{2}{1+x^2} = \lambda$ nicht mehr als zwei Wurzeln, und der Operator (164) besitzt ein rein kontinuierliches Spektrum.

Der Kern des Operators B' ist durch die Beziehung

$$g(u) = \sqrt{\frac{2}{\pi}} \int_{-\infty}^{\infty} \frac{e^{iyu}}{1+y^2} dy = \sqrt{\frac{2}{\pi}} \int_{-\infty}^{\infty} \frac{\cos yu}{1+y^2} dy = \sqrt{2\pi} e^{-|u|}$$

bestimmt, und es ist

$$B'f(x) = \int_{-\infty}^{\infty} e^{-|x-y|} f(y) dy. \quad (165)$$

Der erhaltene Kern genügt der Beziehung (97) aus [172]:

$$\int_{-\infty}^{\infty} |K(x, y)| dy = \int_{-\infty}^x e^{y-x} dy + \int_x^{\infty} e^{x-y} dy = 2.$$

Wegen (159) ist die Spektralschar des Operators (164) folgendermaßen definiert:

$$\mathcal{E}_\lambda f(x) = \begin{cases} f(x) & \text{für } \frac{2}{1+x^2} \leq \lambda, \\ 0 & \text{für } \frac{2}{1+x^2} > \lambda, \end{cases} \quad \text{d. h.} \quad \mathcal{E}_\lambda f(x) = \begin{cases} f(x) & \text{für } |x| \geq \mu, \\ 0 & \text{für } |x| < \mu \end{cases}$$

mit $\mu = \sqrt{2-\lambda}$. Ferner ist

$$\mathcal{E}'_\lambda f(x) = T^* \mathcal{E}_\lambda T f(x) = \frac{1}{2\pi} \left(\int_{-\mu}^x + \int_x^\mu \right) \left[\int_{-\infty}^{\infty} f(t) e^{-itv} dt \right] e^{ixv} dy,$$

also

$$\mathcal{E}'_\lambda f(y) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \left[\int_{-\mu}^y f(t) e^{-itv} dt \right] e^{ixv} dy - \frac{1}{2\pi} \int_{-\mu}^y \left[\int_{-\infty}^{\infty} f(t) e^{itv} dt \right] e^{-ixv} dy.$$

Diese uneigentlichen Integrale sind im Sinne der mittleren quadratischen Approximation zu verstehen. Ändern wir im letzten Integral die Reihenfolge der Integration, wozu wir ja die Möglichkeit haben, und berücksichtigen wir die Beziehung $T^*T = E$, so erhalten wir

$$\mathcal{E}'_\lambda f(x) = f(x) - \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{\sin \mu(x-t)}{x-t} f(t) dt. \quad (166)$$

Der Operator B' hat ein rein kontinuierliches Spektrum, und $\mathcal{E}'_\lambda f(x)$ strebt also für $\lambda \rightarrow 0$ (im Mittel) gegen 0. Somit ist

$$\lim_{\mu \rightarrow \infty} \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{\sin \mu(x-t)}{x-t} f(t) dt = f(x). \quad (167)$$

Wir bilden nun für den Operator (165) die Differentiallösungen. Die homogene Gleichung $B'f(x) = \lambda f(x)$ besitzt die nicht zu L_2 gehörenden Lösungen $\cos \mu x$ und $\sin \mu x$, d. h., es ist

$$\begin{aligned} \int_{-\infty}^{\infty} e^{-|x-y|} \cos \mu y dy &= \lambda \cos \mu x, \\ \int_{-\infty}^{\infty} e^{-|x-y|} \sin \mu y dy &= \lambda \sin \mu x. \end{aligned}$$

Multiplizieren wir beide Seiten mit $e^{-\mu} \frac{d\mu}{d\lambda}$ und integrieren über λ von 0 bis λ bzw. (was dasselbe ist) über μ von ∞ bis μ , so erhalten wir die Differentiallösungen

$$\begin{aligned} \pi_1(x, \lambda) &= \int_{\infty}^{\mu} e^{-\mu} \cos \mu x d\mu = -e^{-\mu} \left(\frac{\cos \mu x}{1+x^2} - \frac{x \sin \mu x}{1+x^2} \right), \\ \pi_2(x, \lambda) &= \int_{\infty}^{\mu} e^{-\mu} \sin \mu x d\mu = -e^{-\mu} \left(\frac{x \cos \mu x}{1+x^2} + \frac{\sin \mu x}{1+x^2} \right). \end{aligned} \quad (168)$$

Diese Funktionen gehören zu L_2 . Aus ihrer Konstruktion ist ersichtlich, daß sie der Gleichung (129) genügen und für $\lambda=0$ (also $\mu=\infty$) verschwinden. Der Faktor $e^{-\mu}$ ist notwendig, damit wir von $\mu=\infty$ integrieren können und auf diese Art eine Lösung erhalten, die bis zu $\lambda=0$ stetig ist. Die Lösungen (168) sind zueinander orthogonal [176] im Grundintervall $(-\infty, \infty)$, da die eine eine gerade und die andere eine ungerade Funktion ist.

Aus den Formeln (120) und (121) aus [176] ergibt sich mit Hilfe einer einfachen Rechnung

$$\varrho_1(\lambda) = \varrho_2(\lambda) = \frac{\pi}{2} e^{-2\mu}$$

bzw.

$$g_1(\lambda) = \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\mu} e^{-\mu} \cos \mu y d\mu \right] \overline{\varphi(y)} dy, \quad g_2(\lambda) = \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\mu} e^{-\mu} \sin \mu y d\mu \right] \overline{\varphi(y)} dy.$$

Die Abgeschlossenheit des Lösungssystems (168) läßt sich auf Grund von (128) aus [176] nachweisen. Wenden wir (123) auf die reelle Funktion $\varphi(x)$ aus L_2 und auf die Funktion $\varphi(x)$ an, die gleich 1 auf dem Intervall $[0, x]$ und gleich 0 außerhalb dieses Intervalls ist, so finden wir nach elementaren Umformungen

$$\int_0^x \varphi(x) dx = \frac{1}{\pi} \int_0^{\infty} \left[\frac{\sin \mu x}{\mu} \int_{-\infty}^{\infty} \varphi(y) \cos \mu y dy + \left(\frac{1}{\mu} - \frac{\cos \mu x}{\mu} \right) \int_{-\infty}^{\infty} \varphi(y) \sin \mu y dy \right] d\mu.$$

Dies führt unter gewissen Zusatzbedingungen auf die gewöhnliche Fouriersche Formel.

Die Lösungen (168) ergeben sich auch, wie man leicht zeigen kann, nach Anwendung des Operators $\mathfrak{S}'_\lambda$ auf die Funktionen $\pi_k(x, 0)$, d. h. auf $\frac{1}{1+x^2}$ und $\frac{x}{1+x^2}$. Würden wir zur Integration nicht den Faktor $e^{-\mu}$ hinzufügen und von $\mu=0$ integrieren, so erhielten wir die einfachen Differentiallösungen $\frac{\sin \mu x}{x}$ und $\frac{1-\cos \mu x}{x}$, die für $\lambda=0$ ihren Sinn verlieren, wobei ihre Normen für $\lambda \rightarrow 0$ unbeschränkt wachsen.

Wir betrachten nun den allgemeinen Fall (161), wobei wir annehmen, daß $g(y)$ eine reelle gerade Funktion ist, die der Bedingung (163) genügt. Dabei ist der Operator (161) in ganz L_2 definiert und beschränkt sowie selbstadjungiert. Wir setzen

$$G_1(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} g(u) e^{itu} du, \quad F_1(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(u) e^{itu} du \quad (169)$$

mit

$$|G_1(t)| \leq \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} |g(u)| du,$$

d. h., $G_1(t)$ ist eine beschränkte Funktion und $F_1(t)$ gehört zu L_2 , so daß auch $G_1(t) F_1(t)$ in L_2 liegt. Wir können zeigen, daß auch in diesem Fall die Formel (152) gilt, der wir die Gestalt

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} g(x-t) f(t) dt = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} G_1(t) F_1(t) e^{-ixt} dt = T[G_1(t) F_1(t)]$$

geben können. Hieraus folgt $G_1(t) F_1(t) = T^*[\varphi(x)]$. Wir sehen also, wenn wir die zweite Formel (169) beachten, daß der Operator (161) dem Operator der Multiplikation mit der beschränkten Funktion $G_1(t)$ unitär äquivalent ist.

Wir erwähnen nun einen Typ von Kernen, die auf Differenzkerne führen. Es sei $K(x, y)$ ein reeller symmetrischer Kern auf dem Intervall $(0, \infty)$, der eine homogene Funktion (-1) -ten Grades sein soll. Führen wir in dem mit diesem Kern gebildeten

Integraloperator

$$\varphi(x) = \int_0^{\infty} K(x, y) f(y) dy \quad (170)$$

statt x und y neue unabhängige Veränderliche e^s und e^t und statt $\varphi(x)$ und $f(y)$ neue Funktionen $\varphi_1(s) = e^{s/2}\varphi(e^s)$ und $f_1(t) = e^{t/2}f(e^t)$ ein, so gelangen wir zu dem Integraloperator

$$\varphi_1(s) = \int_{-\infty}^{\infty} K_1(s, t) f_1(t) dt$$

mit einem von $|s - t|$ abhängigen Kern. Wegen der Homogenität ist nämlich

$$K(x, y) = x^{-1} K\left(1, \frac{y}{x}\right),$$

und setzen wir $K(1, z) = \omega(z)$, so können wir

$$K_1(s, t) = e^{(s+t)/2} \omega(e^{t-s}) = e^{-s/2} \omega(e^{t-s})$$

schreiben. Dabei ist dieser Ausdruck wegen der Symmetrie von $K(x, y)$ eine gerade Funktion von $t - s$. Wegen $ds = \frac{dx}{x}$ sehen wir, daß bei der genannten Variablentransformation der Raum $L_2(0, \infty)$ in den Raum $L_2(-\infty, \infty)$ übergeht. Die Norm des Operators (170) können wir unmittelbar mit Hilfe des folgenden einfachen Satzes bestimmen.

Satz. Ist $K(x, y)$ nichtnegativ, homogen vom Grad -1 und ist

$$\int_0^{\infty} K(x, 1) x^{-1/2} dx = \int_0^{\infty} K(1, y) y^{-1/2} dy = k, \quad (171)$$

so gilt

$$|I| = \left| \int_0^{\infty} \int_0^{\infty} K(x, y) f(x) g(y) dx dy \right| \leq k \|f\| \cdot \|g\|. \quad (172)$$

Die Integrale in (171) sind wegen der Homogenität des Kerns einander gleich. Schreiben wir den Integranden von (172) in der Form $f(x) \sqrt{K}\left(\frac{x}{y}\right)^{1/4} g(y) \sqrt{K}\left(\frac{y}{x}\right)^{1/4}$ und wenden wir die Schwarzsche Ungleichung an, so erhalten wir $|I| \leq \sqrt{A} \sqrt{B}$, wobei

$$A = \int_0^{\infty} |f(x)|^2 \left[\int_0^{\infty} K(x, y) \left(\frac{x}{y}\right)^{1/2} dy \right] dx = k \|f\|^2$$

und analog $B = \|g\|^2$ ist. Daraus folgt (172). Aus (172) sieht man, daß die Norm eines Operators mit dem Kern $K(x, y)$ nicht größer als k ist. Setzen wir insbesondere $K(x, y) = \frac{1}{x+y}$, so gelangen wir wegen

$$\int_0^{\infty} \frac{x^{-1/2}}{1+x} dx = \pi$$

zu der Beziehung

$$\left| \int_0^{\infty} \int_0^{\infty} \frac{f(x) \overline{g(y)}}{x+y} dx dy \right| \leq \pi \|f\| \cdot \|g\|.$$

Mit Hilfe der oben angegebenen Variablensubstitution läßt sich zeigen, daß ein Operator mit dem Kern $\frac{1}{x+y}$ auf dem Intervall $[0, \pi]$ ein kontinuierliches Spektrum besitzt.

182. Schwache Konvergenz. Wir untersuchten früher schon die schwache Konvergenz in L_p . Nun wollen wir die Hauptergebnisse für den Fall $p=2$ wiederholen. Wird der Raum $L_2(\mathfrak{E})$ betrachtet, wobei $\mathfrak{E}$ eine beliebige, aber feste meßbare Menge ist, so läßt sich die schwache Konvergenz $\varphi_n(x) \rightarrow \varphi(x)$ durch die Beziehung

$$\lim_{n \rightarrow \infty} \int_{\mathfrak{E}} \psi(x) \varphi_n(x) dx = \int_{\mathfrak{E}} \psi(x) \varphi(x) dx \quad (173)$$

für beliebige Funktionen $\psi(x) \in L_2(\mathfrak{E})$ definieren. Für die schwache Konvergenz sind die folgenden Bedingungen notwendig und hinreichend; a) Die Normen $\|\varphi_n\|$ in $L_2(\mathfrak{E})$ sind beschränkt; b) die Bedingung (173) ist auf der Menge derjenigen Elemente $\psi(x) \in L_2(\mathfrak{E})$ erfüllt, deren lineare Hülle in $L_2(\mathfrak{E})$ dicht ist.

Im eindimensionalen Fall, wenn $\mathfrak{E}$ ein endliches oder unendliches Intervall ist, läßt sich die zweite Forderung durch

$$\lim_{n \rightarrow \infty} \int_c^{\xi} \varphi_n(x) dx = \int_c^{\xi} \varphi(x) dx$$

ersetzen; dabei ist c eine beliebige, aber feste Zahl und ξ eine beliebige Zahl aus einem solchen Intervall.

Es läßt sich leicht beweisen, daß zwei Funktionen $\varphi(x) \in L_2(\mathfrak{E})$ und $\omega(x) \in L_2(\mathfrak{E})$ äquivalent sind, wenn die Folge der Funktionen $\varphi_n(x) \in L_2(\mathfrak{E})$ gegen die Funktion $\varphi(x)$ schwach konvergiert und fast überall auf $\mathfrak{E}$ gegen die Funktion $\omega(x)$ konvergiert.

183. Andere Realisierungen des Raumes H . Außer l_2 und L_2 gibt es noch eine Reihe anderer Realisierungen des Hilbertschen Raumes, die auch angewendet werden. Es sei $\mathfrak{E}$ eine meßbare Menge des n -dimensionalen Raumes und L_2 der Raum der auf $\mathfrak{E}$ meßbaren und quadratisch summierbaren Funktionen, wobei auf Grund der Meßbarkeit das Lebesguesche Maß oder eine andere beliebige normale Mengenfunktion gewählt werden kann. Im letzten Fall erhalten wir ein Lebesgue-Stieltjesches Integral. Dann definieren wir den Raum $L_{2,m}$ folgendermaßen: Die Folge der m Funktionen $f_1, f_2, \dots, f_m$ aus L_2 ist ein *Element des Raumes $L_{2,m}$* . Ein Element ist Nullelement, wenn jede der Funktionen f_k ($k=1, 2, \dots, m$) äquivalent 0 ist. Die Multiplikation eines Elements mit einer Zahl und die Addition von Elementen wird mit Hilfe der Formeln

$$a(f_1, f_2, \dots, f_m) = (af_1, af_2, \dots, af_m),$$

$$(f_1, f_2, \dots, f_m) + (g_1, g_2, \dots, g_m) = (f_1 + g_1, f_2 + g_2, \dots, f_m + g_m)$$

vorgenommen; das Skalarprodukt wird durch

$$(x, y) = \int_{\mathfrak{E}} (f_1 \bar{g}_1 + f_2 \bar{g}_2 + \dots + f_m \bar{g}_m) d\omega$$

definiert, wobei $d\omega$ im Fall eines Lebesgueschen Integrals ein Element des R_n und im Fall eines Lebesgue-Stieltjeschen Integrals das Differential einer normalen Mengenfunktion ist. Wir können leicht zeigen, daß $L_{2,m}$ die Realisierung eines separablen Hilbertschen Raumes ist. Der lineare Operator $y = Ax$ aus $L_{2,m}$ besteht aus m^2 linearen Operatoren A_{ik} ($i, k=1, 2, \dots, m$) aus L_2 , mit deren Hilfe sich die Komponenten von y durch die von x ausdrücken lassen

$$g_i = \sum_{k=1}^m A_{ik} f_k.$$

Die erwähnte Realisierung eines Hilbertschen Raumes ist die Realisierung einer abstrakten Konstruktion des Hilbertschen Raumes H durch gegebene Hilbertsche

Räume $H_1, H_2, \dots, H_m$. Als Element $x \in H$ bezeichnen wir die Folge der Elemente $x_1, x_2, \dots, x_m$ ($x_k \in H_k$; $k = 1, 2, \dots, m$). Das Element x ist das Nullelement, wenn jedes x_k das Nullelement aus H_k ist. Die Multiplikation mit einer Zahl, die Addition von Elementen und das Skalarprodukt werden durch

$$\begin{aligned} \alpha(x_1, x_2, \dots, x_m) &= (\alpha x_1, \alpha x_2, \dots, \alpha x_m), \\ (x_1, x_2, \dots, x_m) + (y_1, y_2, \dots, y_m) &= (x_1 + y_1, x_2 + y_2, \dots, x_m + y_m), \\ (x, y) &= \sum_{k=1}^m (x_k, y_k) \end{aligned}$$

definiert.

Jeder Raum $W_2^{(l)}(D)$ [112] von Funktion $\varphi(x)$, welche zu $L_2(D)$ gehören, wobei D ein Gebiet des n -dimensionalen Raumes ist, und welche verallgemeinerte Ableitungen bis zur l -ten Ordnung besitzen, die ebenfalls zu $L_2(D)$ gehören, ist ein vollständiger Hilbertscher Raum mit dem Skalarprodukt

$$(\varphi, \psi) = \int_D [\varphi(x) \overline{\psi(x)} + \sum_{1 \leq \alpha \leq l} D^\alpha \varphi(x) \overline{D^\alpha \psi(x)}] dx, \quad (174)$$

wobei über alle Ableitungen bis einschließlich l -ter Ordnung summiert wird. Dabei wird vorausgesetzt, daß das Gebiet D bezüglich irgendeines seiner Punkte sternförmig ist, so daß der Raum $W_2^{(l)}(D)$ separabel ist [111].

Wir untersuchen nun speziell den Raum $W_2^{(l)}(D)$. Die zu ihm gehörenden Funktionen $\varphi(x)$ haben auf der Oberfläche S von D (S soll hinreichend glatt sein) Grenzwerte. Es läßt sich dann zeigen, daß die Menge der Funktionen $\varphi(x) \in W_2^{(l)}(D)$, die der Randbedingung

$$\varphi(x)|_S = 0$$

genügen, ein vollständiger Hilbertscher Raum mit dem Skalarprodukt

$$(\varphi, \psi) = \int_D \sum_{k=1}^n \frac{\partial \varphi(x)}{\partial x_k} \frac{\partial \overline{\psi(x)}}{\partial x_k} dx \quad (175)$$

ist.

Ein vollständiger Hilbertscher Raum ist auch die Menge der Funktionen aus $W_2^{(l)}(D)$ mit dem Skalarprodukt (175) und ohne jede Randbedingung, wenn man die Funktionen, deren Differenz gleich einer Konstanten ist, als ein und dasselbe Element des Raumes auffaßt (vgl. [114], Satz 3).

§ 3. Nichtbeschränkte Operatoren

184. Abgeschlossene Operatoren. Wir gehen nun zu distributiven Operatoren über, die nicht in ganz H gegeben zu sein und für die wir nicht Beschränktheit (Endlichkeit der Norm) voraussetzen brauchen. In den folgenden Ausführungen ist A ein distributiver Operator, $D(A)$ sein Definitionsbereich, der immer eine lineare Mannigfaltigkeit sein soll, und $R(A)$ der Wertebereich von A , der wegen der Distributivität von A ebenfalls eine lineare Mannigfaltigkeit ist. Vermittelt A eine eindeutige Zuordnung zwischen $D(A)$ und $R(A)$, so ist auf $R(A)$ der inverse Operator A^{-1} definiert. Der Operator A^{-1} existiert genau dann, wenn die Gleichung $Ax = 0$ (auf $D(A)$) nur die Nulllösung hat [127].

Zwei Operatoren A und B *stimmen überein* (sind *gleich*), in Zeichen $A = B$, wenn ihre Definitionsbereiche übereinstimmen und für alle Elemente dieses Bereichs $Ax = Bx$ gilt. Der Operator B ist eine *Erweiterung des Operators* A , in Zeichen $A \subseteq B$, wenn $D(A)$ in $D(B)$ enthalten und $Ax = Bx$ für $x \in D(A)$ ist. Die Schreibweise $A \subseteq B$ läßt noch die Möglichkeit $A = B$ zu. Ist für $Ax = Bx$ mit $x \in D(A)$ die lineare Mannigfaltigkeit $D(B)$ echt größer als $D(A)$, so werden wir $A \subset B$ schreiben. Wir erwähnen noch, daß $(A + B)x = Ax + Bx$ für $x \in D(A)$ und $x \in D(B)$ sowie $(AB)x = A(Bx)$ für $x \in D(B)$ und $Bx \in D(A)$ sinnvoll ist. Da wir den Operator nicht als überall gegeben und der Norm nach beschränkt voraussetzen, können wir nichts über seine Stetigkeit aussagen. Analysieren wir jedoch die Grundeigenschaften für beschränkte Operatoren, so können wir uns davon überzeugen, daß man zur Herleitung vieler dieser Eigenschaften nicht die Stetigkeit der Operatoren braucht, sondern eine viel schwächere Eigenschaft, nämlich die Abgeschlossenheit. Diese überaus wichtige Eigenschaft linearer Operatoren wollen wir nun definieren und untersuchen.

Definition. Ein Operator A heißt *abgeschlossen*, wenn gilt: Gehört x_n ($n = 1, 2, \dots$) zu $D(A)$ und besitzen die Folgen der Elemente x_n und Ax_n Grenzelemente, $x_n \Rightarrow x_0$ und $Ax_n \Rightarrow y_0$, so ist $x_0 \in D(A)$ und $Ax_0 = y_0$.

Ist ein Operator nicht abgeschlossen, so ergibt sich die Frage, ob er abgeschlossene Erweiterungen besitzt. Haben wir zwei Folgen von Elementen x_n und x'_n aus $D(A)$, die das gleiche Grenzelement besitzen, für die aber die Grenzelemente der Folgen Ax_n und Ax'_n verschieden sind, so läßt der Operator A offenbar keine abgeschlossenen Erweiterungen zu. Sind aber bei gleichen Grenzelementen für x_n und x'_n die von Ax_n und Ax'_n , falls sie existieren, ebenfalls gleich, so gibt es abgeschlossene Erweiterungen, und unter ihnen existiert eine *minimale*, die im allgemeinen mit $\bar{A}$ bezeichnet wird. Wir beschreiben nun die Konstruktion von $\bar{A}$. Ist $x_n \in D(A)$, $x_n \Rightarrow x_0$ und $Ax_n \Rightarrow y_0$, so zählen wir x_0 zum Definitionsbereich von $\bar{A}$ und setzen $\bar{A}x_0 = y_0$. Wegen der obenerwähnten Bedingung ist $\bar{A}$ dadurch eindeutig bestimmt. Mit Hilfe der Dreiecksungleichung läßt sich zeigen, daß $\bar{A}$ abgeschlossen ist. Die eben beschriebene Operation der Erweiterung von A heißt die *Abschließung* von A . Ist B eine beliebige abgeschlossene Erweiterung von A , so ist $\bar{A} \subseteq B$.

Satz 1. Ist A ein abgeschlossener Operator, so ist $A + B$, wobei B ein auf $D(A)$ beschränkter Operator ist, ebenfalls abgeschlossen. Der Operator A^{-1} ist, falls er existiert, abgeschlossen, und die Menge der Lösungen von $Ax = 0$ ist ein Unterraum.

Der Beweis dieser Behauptungen folgt unmittelbar aus der obigen Definition der Abgeschlossenheit eines Operators.

Satz 2. Läßt der Operator A die Abschließung zu und hat er auf $R(A)$ ein beschränktes Inverses A^{-1} , so hat $\bar{A}$ das Inverse $(\bar{A})^{-1}$; es ist auf dem Unterraum $\overline{R(A)}$ definiert und beschränkt.

Beweis. Ist $R(A)$ ein Unterraum, so ist A^{-1} ein abgeschlossener Operator. Anderenfalls läßt sich der beschränkte Operator A^{-1} von der linearen Mannigfaltigkeit $R(A)$ auf den Unterraum $\overline{R(A)}$ fortsetzen. Den so erhaltenen Operator bezeichnen wir mit B , so daß also $R(B) = \overline{R(A)}$ ist. Dann hat die Gleichung $Bx = 0$ auf $\overline{R(B)}$ nur die Nulllösung. Wäre dies nicht der Fall, so würde eine Folge von Elementen $x_n \in D(A)$ existieren, für die $x_n \Rightarrow 0$, aber $Ax_n \Rightarrow y \neq 0$ wäre. Das widerspricht jedoch die Voraussetzung, daß A die Abschließung zulassen soll, denn für $x'_n = 0$

($n = 1, 2, \dots$) ist $Ax'_n = 0$. Der Operator $A_1 = B^{-1}$ ist somit die Abschließung von A . Damit ist der Satz bewiesen.

Bemerkung. In dem betrachteten Fall hängt die Abschließung von A eindeutig mit der Erweiterung bezüglich der Stetigkeit des beschränkten Operators A^{-1} zusammen.

Folgerung. Ist A ein abgeschlossener Operator und existiert auf $R(A)$ ein beschränkter inverser Operator A^{-1} , so ist $R(A)$ ein Unterraum.

185. Der adjungierte Operator. Wir beginnen hier mit der einfachen Bemerkung, daß das Element z das Nullelement ist, wenn es zu einer in H dichten linearen Mannigfaltigkeit l orthogonal ist. Zum Beweis nehmen wir an, es sei $(x, z) = 0$ für $x \in l$, und y sei ein beliebiges Element aus H . Da l in H dicht sein soll, existiert eine Folge von Elementen x_n aus l derart, daß $x_n \rightarrow y$ ist. Wegen der Eigenschaft von z ist $(x_n, z) = 0$ und nach Grenzübergang $(y, z) = 0$, d. h., z ist zu jedem Element aus H orthogonal, insbesondere auch zu sich selbst, so daß $(z, z) = \|z\|^2 = 0$, also $z = 0$ ist. Ist l nicht dicht in H , so existiert ein von 0 verschiedenes zu l orthogonales Element z .

Im folgenden seien die Operatoren stets distributiv.

Wir setzen ferner voraus, daß A auf der in H dichten linearen Mannigfaltigkeit $D(A)$ definiert ist, und bilden (Ax, y) mit $x \in D(A)$ und $y \in H$. Dann existieren solche Elemente y , daß (Ax, y) für jedes $x \in D(A)$ in der Form

$$(Ax, y) = (x, y^*) \quad (x \in D(A)) \quad (1)$$

dargestellt werden kann; dabei ist y^* ein gewisses Element aus H . Ist also $y = 0$, so gilt $(Ax, 0) = (x, 0)$ für jedes $x \in D(A)$. Ist für ein gewisses y die Darstellung (1) möglich, so ist y^* durch (1) eindeutig bestimmt; denn gelten für ein gewisses $y \in H$ und ein $x \in D(A)$ die Beziehungen $(Ax, y) = (x, y_1^*)$ und $(Ax, y) = (x, y_2^*)$, so erhalten wir nach Subtraktion $(x, y_1^* - y_2^*) = 0$, so daß also $y_1^* - y_2^*$ zu $D(A)$ orthogonal ist, und daraus folgt $y_1^* = y_2^*$. Die Menge der Elemente y , für welche die Darstellung (1) möglich ist, bildet offenbar eine lineare Mannigfaltigkeit l^* , auf der der distributive Operator definiert ist, der y in y^* überführt. Dies ist der zu A adjungierte Operator, den wir mit A^* bezeichnen wollen. Es ist also $y^* = A^*y$, die lineare Mannigfaltigkeit l^* stimmt mit $D(A^*)$ überein, und die Formel (1) erhält die Gestalt

$$(Ax, y) = (x, A^*y) \quad (x \in D(A), y \in D(A^*)). \quad (2)$$

Aus diesen Überlegungen folgt, daß A^* genau dann existiert, wenn die lineare Mannigfaltigkeit $D(A)$ in H dicht ist. Wie wir schon wissen, ist A^* distributiv. Für den beschränkten Operator A gilt jetzt die frühere Definition von A^* [124].

Wir beweisen nun eine Reihe von Eigenschaften des adjungierten Operators.

Satz 1. Der Operator A^* ist abgeschlossen.

Es sei $x_n \in D(A^*)$ und $x_n \rightarrow x_0$, $A^*x_n \rightarrow y_0$. Nach Definition von A^* gilt

$$(Ax, x_n) = (x, A^*x_n) \quad \text{für } x \in D(A)$$

oder nach Grenzübergang $(Ax, x_0) = (x, y_0)$, woraus wieder nach Definition von A^* folgt, daß x_0 zu $D(A^*)$ gehört und $A^*x_0 = y_0$ ist. Damit ist der Satz bewiesen.

Satz 2. Sind $D(A)$ und $D(B)$ in H dicht und ist $A \subset B$, so gilt $B^* \subseteq A^*$.

Die lineare Mannigfaltigkeit $D(B^*)$ wird von den Elementen y gebildet, die für alle $x \in D(B)$ der Beziehung $(Bx, x) = (x, y^*)$ genügen; dabei ist $y^* = B^*y$. Wegen $A \subset B$ folgt aus $(Bx, y) = (x, y^*)$ für $x \in D(B)$ die Beziehung $(Ax, y) = (x, y^*)$ für

$x \in D(A)$. Gehört also y zu $D(B^*)$, so ist $y \in D(A^*)$ und $B^*y = A^*y = y^*$; daraus ergibt sich $B^* \subseteq A^*$.

Satz 3. *Ist $D(A)$ in H dicht und läßt sich A abschließen, so gilt $(\bar{A})^* = A^*$.*

Es gilt $A \subseteq \bar{A}$ und folglich $(\bar{A})^* \subseteq A^*$. Somit bleibt nur noch zu zeigen, daß jedes Element y aus $D(A^*)$ auch zu $D((\bar{A})^*)$ gehört. Für $x \in D(A)$ ist nach Voraussetzung $(Ax, y) = (x, A^*y)$; es muß also $(\bar{A}x, y) = (x, A^*y)$ für $x \in D(\bar{A})$ nachgewiesen werden. Für $x \in D(\bar{A})$ existiert eine solche Folge von Elementen $x_n \in D(A)$, daß $x_n \rightarrow x$ und $Ax_n \rightarrow \bar{A}x$ strebt. Da nach Voraussetzung $(Ax_n, y) = (x_n, A^*y)$ und nach Grenzübergang $(\bar{A}x, y) = (x, A^*y)$ gilt, ist der Satz bewiesen.

Satz 4. *Existieren die Operatoren A^* und $(A^*)^* = A^{**}$, so gilt $A \subseteq A^{**}$.*

Die lineare Mannigfaltigkeit $D(A^{**})$ von Elementen z wird durch die Gleichung $(A^*y, z) = (y, z^{**})$ für $y \in D(A^*)$, $z^{**} = A^{**}z$ bestimmt. Nach Definition von A^* gilt $(A^*y, z) = (y, Az)$ mit $y \in D(A^*)$ und $z \in D(A)$, woraus $A \subseteq A^{**}$ folgt.

Da A^{**} wegen Satz 1 ein abgeschlossener Operator ist, folgt aus $A \subseteq A^{**}$, daß A abgeschlossene Erweiterungen zuläßt. Die Existenz von A^{**} ist also eine hinreichende Bedingung dafür, daß A abgeschlossene Erweiterungen gestattet. Im folgenden werden wir sehen, daß diese Bedingung auch notwendig ist. Die Existenz von A^{**} ist, wie wir wissen, damit gleichwertig, daß $D(A^*)$ in H dicht ist.

Satz 5. *Sind $D(A)$ und $R(A)$ in H dicht und existiert der inverse Operator A^{-1} , so existieren auch die Operatoren A^* , $(A^{-1})^*$ und $(A^*)^{-1}$, und es gilt*

$$(A^*)^{-1} = (A^{-1})^* . \quad (3)$$

Die Existenz von A^* und $(A^{-1})^*$ folgt unmittelbar daraus, daß $D(A)$ und $R(A)$ in H dicht sind. Es sei nun $x \in D(A^*)$ und $y \in D(A^{-1}) = R(A)$; dann ist

$$(x, y) = (x, AA^{-1}y) = (A^*x, A^{-1}y) ;$$

daraus ergibt sich $A^*x \in D((A^{-1})^*)$ und

$$(A^{-1})^* A^*x = x \quad (x \in D(A^*)) . \quad (4)$$

Die Gleichung $A^*x = 0$ hat somit (in $D(A^*)$) nur die Nulllösung, d. h., es existiert $(A^*)^{-1}$. Außerdem folgt aus (4)

$$(A^*)^{-1} \subseteq (A^{-1})^* . \quad (5)$$

Nun sei $x \in D(A)$ und $y \in D((A^{-1})^*)$; dann ist $(x, y) = (A^{-1}Ax, y) = (Ax, (A^{-1})^*y)$; daraus ergibt sich $(A^{-1})^*y \in D(A^*)$,

$$A^*(A^{-1})^*y = y \quad (y \in D((A^{-1})^*)) \quad \text{und} \quad (A^{-1})^* \subseteq (A^*)^{-1} .$$

Dies ergibt zusammen mit (5) die Beziehung (3). Damit ist der Satz bewiesen.

Mit dem Begriff des adjungierten Operators ist die Frage nach der Lösbarkeit der Gleichung

$$Ax = y \quad (6)$$

verknüpft. Einen abgeschlossenen Operator A mit einem in H dichten Definitionsbereich $D(A)$ nennen wir *normal lösbar*, wenn für die (nicht unbedingt eindeutige) Lösbarkeit von (6) notwendig und hinreichend ist, daß y zum Unterraum der Lösungen von

$$A^*z = 0 \quad (7)$$

orthogonal ist.

Satz 6. *Ein abgeschlossener Operator A mit in H dichtem Definitionsbereich $D(A)$ ist genau dann normal lösbar, wenn $R(A)$ ein Unterraum von H ist.*

Der Operator A^* ist abgeschlossen, und die Menge der Lösungen von (7) bildet einen Unterraum l . Alle Elemente von l sind orthogonal zu $R(A)$, denn für $y \in R(A)$ ist $y = Ax$ und

$$(y, z) = (Ax, z) = (x, A^*z) = (x, 0) = 0.$$

Somit ist l wegen der Stetigkeit des Skalarprodukts auch zu $\overline{R(A)}$ orthogonal. Wir zeigen nun, daß ein zu $R(A)$ orthogonales Element w in l liegen muß. Aus $(Ax, w) = 0$ folgt tatsächlich $(Ax, w) = (x, 0) = (x, A^*w)$, also $A^*w = 0$ und $w \in l$. Somit ist H die direkte Summe zweier orthogonaler Unterräume,

$$H = \overline{R(A)} \oplus l,$$

und A ist genau dann normal lösbar, wenn $R(A)$ mit $\overline{R(A)}$ übereinstimmt, d. h., wenn $R(A)$ ein Unterraum von H ist.

186. Der Graph eines Operators. Außer dem Raum H betrachten wir noch den Raum $\tilde{H}$, dessen Elemente die Paare $\{x, y\}$ mit $x, y \in H$ sind. Die Multiplikation dieser Elemente mit einer Zahl, die Addition und das Skalarprodukt sind dabei durch

$$\left. \begin{aligned} \alpha\{x, y\} &= \{\alpha x, \alpha y\}, \\ \{x_1, y_1\} + \{x_2, y_2\} &= \{x_1 + x_2, y_1 + y_2\}, \end{aligned} \right\} \quad (8)$$

$$(\{x_1, y_1\}, \{x_2, y_2\}) = (x_1, x_2) + (y_1, y_2) \quad (9)$$

definiert.

Die Gültigkeit aller Axiome aus [120] läßt sich leicht nachweisen. Ist A ein Operator aus H , so heißt die Menge $F(A)$ der Elemente $\{x, Ax\}$ von $\tilde{H}$ für $x \in D(A)$ der *Graph des Operators A* . Alle Elemente dieser Menge werden durch ihre Abszissen (das erste Element jedes Paares) eindeutig bestimmt. Sind umgekehrt alle Elemente der Menge F von Elementen aus $\tilde{H}$ durch ihre Abszissen eindeutig bestimmt, so existiert in H ein (nicht notwendig distributiver) Operator, dessen Graph genau die Menge F ist. Die Abgeschlossenheit des Operators A ist, wie wir leicht sehen, damit gleichwertig, daß $F(A)$ in $\tilde{H}$ abgeschlossen ist. Ist A ein auf einer linearen Mannigfaltigkeit definierter distributiver Operator, so ist $F(A)$ eine lineare Mannigfaltigkeit in $\tilde{H}$. Wir wollen wie vorher nur von distributiven Operatoren sprechen, die auf linearen Mannigfaltigkeiten definiert sind.

Wir definieren in ganz H einen Operator U durch

$$U\{x, y\} = \{iy, -ix\}; \quad (10)$$

U ist unitär, und es gilt $U^{-1} = U$. Mit einem Operator A auf H bilden wir jetzt das Skalarprodukt eines Elements der Menge $UF(A)$ mit einem Element $\{x, y\}$:

$$(\{iAz, -iz\}, \{x, y\}) = i[(Az, x) - (z, y)] \quad (z \in D(A)). \quad (11)$$

Es sei nun A ein in H abgeschlossener Operator und $D(A)$ in H dicht. Wir zeigen, daß sich dann $\tilde{H}$ in zwei orthogonale Unterräume zerlegen läßt:

$$\tilde{H} = UF(A) \oplus F(A^*). \quad (12)$$

Ist das Element $\{x, y\}$ orthogonal zu $UF(A)$, so folgt aus (11) die Beziehung $(Az, x) = (z, y)$ für $z \in D(A)$, also $x \in D(A^*)$ und $y = A^*x$, oder mit anderen Worten,

$\{x, y\} \in F(A^*)$. Umgekehrt ergibt sich aus (11), daß das Element $\{x, y\}$ zu $UF(A)$ orthogonal ist, wenn es zu $F(A^*)$ gehört. Zum Beweis von (12) genügt dann die Bemerkung, daß $UF(A)$ und $F(A^*)$ wegen der Abgeschlossenheit von A und A^* Unterräume von $\vec{H}$ sind.

Ist A nicht abgeschlossen, aber $D(A)$ wie oben in H dicht, so gilt statt (12)

$$\vec{H} = \overline{UF(A)} \oplus F(A^*). \quad (13)$$

Die Differenz $\vec{H} \ominus \overline{UF(A)}$ ist eine Menge $\mathfrak{M}$ von Elementen $\{x, y\}$, die zu $\overline{UF(A)}$ oder, was das gleiche ist, zu $UF(A)$ orthogonal sind; dies bedeutet wegen (11), daß $\mathfrak{M}$ aus den Paaren $\{x, y\}$ besteht, die der Bedingung $(Az, x) = (z, y)$ für $z \in D(A)$ genügen. Die Existenz des Operators A^* ist also damit gleichwertig, daß die Elemente von $\mathfrak{M}$ durch ihre Abszissen eindeutig bestimmt sind.

Das eben Bewiesene fassen wir in dem folgenden Hilfssatz zusammen.

Hilfssatz. Der Operator A^* existiert genau dann, wenn die Elemente der Menge

$$\vec{H} \ominus \overline{UF(A)}$$

durch ihre Abszissen eindeutig bestimmt sind.

Wir beweisen nun den folgenden Satz.

Satz 1. *Ist der Operator A auf einer dichten Menge definiert und läßt er sich abschließen, so existieren A^* und A^{**} , und es gilt*

$$A^{**} = \bar{A}. \quad (14)$$

Beweis. Zunächst nehmen wir an, A sei abgeschlossen. Aus (12) und $U^{-1} = U$ folgt $\vec{H} = F(A) \oplus UF(A^*)$, d. h., die Elemente der Menge $F(A) = \vec{H} \ominus UF(A^*)$ sind durch ihre Abszissen eindeutig bestimmt, und diese Menge definiert auf Grund des Hilfssatzes den Graphen des zu A^* adjungierten Operators, d. h. den Graphen des Operators A^{**} . Dies ist aber die Menge $F(A)$. Daraus folgt $A^{**} = A$.

Ist der Operator A nicht abgeschlossen, aber läßt er die Abschließung zu, so ist nach dem oben Bewiesenen $(\bar{A})^{**} = \bar{A}$. Andererseits gilt [185]

$$(\bar{A})^{**} = ((\bar{A})^*)^* = (A^*)^* = A^{**},$$

woraus sich dann (14) ergibt.

Folgerung. *Existieren die Operatoren A^* und A^{**} , so läßt sich A abschließen [185], und aus (14) folgt*

$$(A^{**})^* = A^{***} = A^*. \quad (15)$$

Satz 2. *Ist A ein abgeschlossener Operator und ist $D(A) = H$, so ist A auch beschränkt.*

Beweis. Aus den Voraussetzungen des Satzes und aus Satz 1 folgt, daß $D(A^*)$ in H dicht und $A = A^{**}$ ist. Wir zeigen zunächst, daß eine positive Zahl N existiert, für die die Ungleichung

$$||A^*x|| \leq N||x|| \quad (x \in D(A^*))$$

gilt, und betrachten dazu das Skalarprodukt

$$(Ay, x) = (y, A^*x) \quad (x \in D(A^*), y \in H).$$

Es definiert in H ein lineares (beschränktes) Funktional $l_x(y)$ für ein festes x aus $D(A^*)$. Ist $x_n \in D(A^*)$ und $x_n \rightarrow 0$, so strebt auch die Folge der Funktionale $l_{x_n}(y)$ auf jedem Element $y \in H$ gegen 0. Es existiert also eine positive Zahl N_1 , für die

$$|l_{x_n}(y)| = |(y, A^*x_n)| \leq N_1||y|| \quad (16)$$

ist [100]. Hätte der Operator A^* keine beschränkte Norm, so würde es eine Folge von Elementen $x_n \in D(A^*)$ geben, für welche $x_n \rightarrow 0$, aber $\|A^*x_n\| \rightarrow \infty$ gilt. Dies widerspricht jedoch (16), denn setzen wir $y = A^*x_n$ in (16), so erhalten wir $\|A^*x_n\|^2 \leq N_1 \|A^*x_n\|$, also $\|A^*x_n\| \leq N_1$. Somit ist $\|A^*\| \leq N$ auf $D(A^*)$. Dann läßt sich A^* auf ganz H ausdehnen, und wegen der Abgeschlossenheit von A^* ist $D(A^*) = H$. Der zu dem beschränkten Operator A^* adjungierte Operator $A^{**} = A$ ist selbst beschränkt. Damit ist der Satz bewiesen.

Bemerkung. Ist λ eine Zahl und existiert A^* , so existiert auch der Operator $(A - \lambda E)^* = A^* - \bar{\lambda}E$. Ist B ein beschränkter linearer Operator, der auf ganz H definiert ist, und gibt es zu A den adjungierten Operator A^* , so existiert $(A + B)^*$ und ist gleich $A^* + B^*$.

187. Symmetrische und selbstadjungierte Operatoren. Im folgenden wollen wir uns in der Hauptsache mit den sogenannten symmetrischen und selbstadjungierten Operatoren beschäftigen.

Definition. Ein Operator A heißt *symmetrisch*, wenn $D(A)$ in H dicht ist und

$$(Ax, y) = (x, Ay) \quad (17)$$

für alle x und y aus $D(A)$ gilt.

Aus (17) folgt, daß jedes $y \in D(A)$ auch zu $D(A^*)$ gehört und $A^*y = Ay$ für diese y gilt, d. h.

$$A \subseteq A^*. \quad (18)$$

Ein symmetrischer Operator A heißt *von unten halbbeschränkt*, wenn eine endliche Zahl

$$m_A = \inf (Ax, x) \quad \text{für } x \in D(A) \quad \text{und} \quad \|x\| = 1$$

existiert; daraus folgt

$$(Ax, x) \geq m_A(x, x) \quad (x \in D(A)). \quad (19)$$

Die Zahl m_A läßt sich nicht durch eine größere ersetzen.

Ist $m_A > 0$, so heißt A *positiv definit*; ist $m_A \geq 0$, so nennen wir A *positiv* (vgl. [126]).

Wir weisen noch darauf hin, daß A ein symmetrischer Operator ist, d. h. $(Ax, y) = (x, Ay)$ für $x, y \in D(A)$, wenn die lineare Mannigfaltigkeit $D(A)$ in H dicht und (Ax, x) für alle $x \in D(A)$ reell ist. Dies läßt sich genau wie Satz 2 aus [124] beweisen.

Definition. Ein symmetrischer Operator A heißt *selbstadjungiert*, wenn $A^* = A$ ist.

Für den Beweis der Selbstadjungiertheit eines symmetrischen Operators A genügt es auf Grund des oben Gesagten zu zeigen, daß $x \in D(A)$ ist, wenn x zu $D(A^*)$ gehört. Wegen (18) läßt sich ein symmetrischer Operator abschließen, und es gelten für ihn die Beziehungen

$$\bar{A} = A^{**} \subseteq A^*, \quad A^{***} = (\bar{A})^* = A^*. \quad (20)$$

Ein selbstadjungierter Operator ist offenbar abgeschlossen.

Ist λ eine reelle Zahl, so ist der Operator $A - \lambda E$ symmetrisch, wenn A symmetrisch ist, und selbstadjungiert, wenn A selbstadjungiert ist.

Satz 1. *Hat der selbstadjungierte Operator A ein Inverses A^{-1} , so ist der Wertebereich $R(A)$ von A in H dicht, und A^{-1} ist ein selbstadjungierter Operator auf $R(A)$.*

Beweis. Wäre die lineare Mannigfaltigkeit $R(A)$ nicht dicht in H , so würde ein vom Nullelement verschiedenes Element z existieren, das zu $R(A)$ orthogonal ist, d. h. $(Ax, z) = 0$ ($x \in D(A)$) oder $(Ax, z) = (x, 0)$, also $z \in D(A^*)$ und $A^*z = Az = 0$ für ein vom Nullelement verschiedenes Element z . Dies widerspricht jedoch der Existenz des inversen Operators A^{-1} . Also ist $R(A)$ in H dicht.

Auf Grund von Satz 5 aus [185] folgt $(A^{-1})^* = (A^*)^{-1}$ oder wegen der Selbstadjungiertheit von A

$$(A^{-1})^* = A^{-1};$$

d. h., A^{-1} ist tatsächlich selbstadjungiert.

Satz 2. *Gibt es zu einem symmetrischen Operator A eine Zahl λ derart, daß sowohl die Elemente der Form $(A - \lambda E)x$ als auch die der Form $(A - \lambda E)x$ ganz H ausschöpfen, so ist A ein auf $D(A)$ selbstadjungierter Operator.*

Beweis. Wir müssen zeigen, daß y zu $D(A)$ gehört, wenn $y \in D(A^*)$ ist. Es gilt $(Ax, y) = (x, y^*)$ für $x \in D(A)$, woraus

$$((A - \lambda E)x, y) = (x, y^* - \lambda y)$$

folgt. Nach Voraussetzung existiert wenigstens ein Element $z \in D(A)$, für das $y^* - \lambda y = (A - \lambda E)z$ gilt, und wegen der Symmetrie von A ist

$$((A - \lambda E)x, y) = (x, (A - \lambda E)z) = ((A - \lambda E)x, z).$$

Die Elemente der Form $(A - \lambda E)x$ spannen ganz H auf, und aus der letzten Gleichung folgt $y = z \in D(A)$. Damit ist der Satz bewiesen.

Folgerung. *Ist $R(A) = H$ für einen symmetrischen Operator A , so ist A selbstadjungiert.*

Zum Beweis dieser Folgerung benutzt man Satz 2 für $\lambda = 0$.

In [189] werden wir zeigen, daß eine Umkehrung von Satz 2 in einem gewissen Sinn gilt, nämlich: *Ist A ein selbstadjungierter Operator und λ eine nicht reelle Zahl, so besitzt der Operator $A - \lambda E$ ein auf ganz H definiertes Inverses.*

Satz 3. *Hat der symmetrische Operator A ein auf $R(A)$ beschränktes Inverses A^{-1} , so ist $R(A^*) = H$.*

Beweis. Da die Abschließung eines symmetrischen Operators wieder auf einen symmetrischen Operator führt und die Beziehung $A^* = (\bar{A})^*$ gilt, können wir zum Beweis annehmen, A sei abgeschlossen. Mit Hilfe von Satz 2 aus [184] können wir schließen, daß $R(A)$ Unterraum von H ist. Wir müssen also zeigen, daß für jedes feste $y^* \in H$ und jedes $x \in D(A)$ das Skalarprodukt (x, y^*) in der Form (Ax, y) dargestellt werden kann. Mit $Ax = z$ erhalten wir $(x, y^*) = (A^{-1}z, y^*)$, und den Ausdruck $(A^{-1}z, y^*)$ können wir als lineares (beschränktes) Funktional $l_{y^*}(z)$ auf dem Unterraum $R(A)$ auffassen, da der Operator A^{-1} auf $R(A)$ beschränkt ist. Das Funktional $l_{y^*}(z)$ können wir in der Form (z, y) mit $y \in R(A)$ darstellen [123], so daß sich $(x, y^*) = (z, y) = (Ax, y)$ ergibt. Diese Gleichung bedeutet auch, daß sich jedes $y^* \in H$ in der Form A^*y schreiben läßt. Damit ist der Satz bewiesen.

Folgerung. *Ist A ein selbstadjungierter, positiv definiter Operator, so existiert A^{-1} , ist beschränkt und auf ganz H definiert.*

Da A positiv definit ist, d. h., da $(Ax, x) \geq \alpha(x, x)$ ($\alpha > 0$) gilt, haben wir $\alpha\|x\| \leq \|Ax\|$, woraus auch folgt, daß auf $R(A)$ der Operator A^{-1} existiert und seine

Norm $\|A^{-1}\|$ nicht größer als α^{-1} ist. Auf Grund von Satz 3 ist $R(A^*)=H$ und $R(A^*)=R(A)$, so daß A^{-1} auf ganz H definiert ist.

Wir erwähnen hier noch zwei Tatsachen: Für jede symmetrische Erweiterung $\tilde{A}$ eines symmetrischen Operators A gilt $\tilde{A} \subseteq A^*$; ein selbstadjungierter Operator läßt keine symmetrischen Erweiterungen zu. Diese Behauptungen folgen aus den Definitionen des symmetrischen und des selbstadjungierten Operators.

Satz 4. *Ist A ein abgeschlossener linearer Operator mit einem in H dichten Definitionsbereich $D(A)$, so ist das Produkt A^*A ein selbstadjungierter positiver Operator.*

Beweis. Daß A^*A positiv ist, folgt aus

$$(A^*Ax, x) = (Ax, Ax) \geq 0 \quad (x \in D(A^*A)).$$

Die Symmetrie von A^*A auf $D(A^*A)$ läßt sich aus der Gleichung

$$(A^*Ax, y) = (Ax, Ay) = (x, A^*Ay)$$

ablesen.

Wir zeigen, daß

$$(A^*A + E)x = y \tag{21}$$

für jedes $y \in H$ (eindeutig) lösbar ist. Dazu betrachten wir den in [186] eingeführten Raum $\tilde{H}$ und seine Zerlegung

$$\tilde{H} = F(A) \oplus UF(A^*) ;$$

aus ihr ergibt sich, daß sich das Element $\{y, 0\}$ eindeutig in der Form

$$\{y, 0\} = \{x, Ax\} + i\{A^*z, -z\}$$

darstellen läßt. Somit ist $y = x + iA^*z$ und $z = -iAx$, also

$$y = x + A^*Ax \quad (x \in D(A^*A)),$$

d. h., die Gleichung (21) hat für jedes $y \in H$ eine Lösung.

Wir beweisen nun, daß $D(A^*A)$ in H dicht ist. Wäre das Gegenteil der Fall, so würde ein zu $D(A^*A)$ orthogonales Element $z \neq 0$ existieren, das man wegen des oben Bewiesenen in der Form

$$z = (A^*A + E)x_0 \quad (x_0 \in D(A^*A))$$

darstellen könnte. Für jedes $x \in D(A^*A)$ hätten wir dann

$$0 = (z, x) = ((A^*A + E)x_0, x) = (x_0, (A^*A + E)x)$$

oder mit $x = x_0$

$$\|x_0\|^2 + (x_0, A^*Ax_0) = \|x_0\|^2 + \|Ax_0\|^2 = 0,$$

also $x_0 = 0$. Dies widerspricht aber der Voraussetzung, daß z vom Nullelement verschieden sein soll. Somit ist $D(A^*A)$ in H dicht, und A^*A sowie $A^*A + E$ sind folglich symmetrisch. Wegen $R(A^*A + E) = H$ ist der Operator $A^*A + E$ selbstadjungiert (Satz 2). Also ist auch A^*A selbstadjungiert. Damit ist der Satz bewiesen.

188. Beispiele für nichtbeschränkte Operatoren. Wir untersuchen jetzt verschiedene Differentialoperatoren vom Standpunkt der Theorie der Operatoren und betrachten sie im komplexen Hilbertschen Raum L_2 der komplexwertigen Funktionen reellen Arguments. Die Formel für partielle Integration, die im folgenden eine fundamentale Rolle spielt, hat im Komplexen die gleiche Gestalt wie im Reellen; es gilt nämlich für zwei beliebige Funktionen $\varphi(x)$ und $\psi(x)$ aus $W_2^{(1)}(D)$, wenn der Bereich D eine stückweise

glatte Berandung S hat,

$$\int_D \frac{\partial \varphi(x)}{\partial x_i} \overline{\psi(x)} dx = - \int_D \varphi(x) \frac{\partial \overline{\psi(x)}}{\partial x_i} dx + \int_D \varphi(x) \overline{\psi(x)} \cos(n, x_i) dS, \quad (22)$$

wobei n die äußere Normale auf S ist (vgl. [113]).

Wir beginnen mit dem einfachsten Differentialoperator $D = i \frac{d}{dx}$.

1. Der Operator $D = i \frac{d}{dx}$ im Raum $L_2[0, 1]$. Wie wir wissen, ist in der abstrakten Theorie ein Operator A durch seinen Definitionsbereich $D(A)$ und durch die Anwendungsvorschrift auf Elemente aus $D(A)$ bestimmt.

Der Operator D läßt sich auf natürliche Weise auf allen Funktionen aus $L_2[0, 1]$ definieren, die eine verallgemeinerte Ableitung aus $L_2[0, 1]$ besitzen. Jedoch hat der auf dieser umfangreichen Funktionenklasse definierte Operator D nicht die Eigenschaften, die er z. B. besitzt, wenn man ihn auf finiten, stetig differenzierbaren Funktionen betrachtet. Deshalb beginnen wir in diesem und in den nächsten Beispielen mit der Betrachtung des Differentialoperators auf einer Menge stetig differenzierbarer Funktionen, welche gewissen Randbedingungen unterworfen sind. Dann untersuchen wir seine Eigenschaften, z. B. ob er symmetrisch, positiv definit, invertierbar usw. ist, und schließlich fragen wir nach seiner Erweiterung unter Beibehaltung dieser oder jener Eigenschaften des ursprünglichen Operators.

Die Wahl des Definitionsbereichs des ursprünglichen Differentialoperators ist nicht eindeutig; deshalb werden wir in den folgenden Beispielen jedesmal besonders darauf hinweisen.

Wir bezeichnen mit A den auf der Menge $C^{(1)}[0, 1]$ aller finiten, auf $[0, 1]$ stetig differenzierbaren Funktionen betrachteten Operator D (vgl. die Bezeichnungen in [113]). Der Wert von A auf $\varphi(x) \in D(A)$ wird durch

$$A\varphi = i \frac{d\varphi(x)}{dx} \quad (23)$$

gegeben, und $D(A)$ ist in $L_2[0, 1]$ dicht.

Der Operator A ist symmetrisch, denn aus (23) folgt für $\varphi(x) \in D(A)$ und $\psi(x) \in D(A)$

$$(A\varphi, \psi) = \int_0^1 i \frac{d\varphi(x)}{dx} \overline{\psi(x)} dx = - \int_0^1 i \varphi(x) \frac{d\overline{\psi(x)}}{dx} dx = (\varphi, A\psi).$$

Aus der Symmetrie von A folgt, daß A die Abschließung gestattet, den adjungierten Operator besitzt und daß $A \subseteq \bar{A} \subseteq A^*$ ist.

Wir untersuchen nun, aus welchen Funktionen die linearen Mannigfaltigkeiten $D(\bar{A})$ und $D(A^*)$ bestehen. Ist $\varphi_m(x) \in D(A)$ und $\varphi_m(x) \rightarrow \varphi(x)$, $A\varphi_m = \psi_m \rightarrow \psi(x)$, dann gilt $\varphi(x) \in D(\bar{A})$ und $\psi(x) = \bar{A}\varphi(x)$.

Aus der Theorie der verallgemeinerten Ableitungen folgt, daß diese Abschließung von A den Bereich $D(A)$ auf $D(\bar{A}) = \dot{W}_2^{(1)}[0, 1]$ erweitert [113].

Jede Funktion $\varphi(x)$ aus $\dot{W}_2^{(1)}[0, 1]$ ist absolut stetig, in den Endpunkten des Intervalls $[0, 1]$ gleich 0 und besitzt eine verallgemeinerte Ableitung aus $L_2[0, 1]$. Man könnte zeigen, daß jede solche Funktion in $\dot{W}_2^{(1)}[0, 1]$ auftritt. Der Operator $\bar{A}$ auf $\varphi(x) \in D(\bar{A})$ wird ebenfalls durch (23) gegeben, nur mit dem Unterschied, daß jetzt das Symbol $\frac{d}{dx}$ nicht die gewöhnliche, sondern die verallgemeinerte Differentiation bezeichnet. Wir können nun feststellen, aus welchen Funktionen der Bereich $D(A^*)$ besteht. Die Funktion $\varphi(x)$ gehört zu $D(A^*)$, wenn es eine Funktion $\psi(x) \in L_2[0, 1]$ gibt, für die $(A\omega, \varphi) = (\omega, \psi)$, d. h.

$$\int_0^1 i \frac{d\omega(x)}{dx} \overline{\varphi(x)} dx = \int_0^1 \omega(x) \overline{\psi(x)} dx \quad (24)$$

für alle $\omega(x) \in D(A)$ gilt. Dies bedeutet [109], daß $\varphi(x)$ die verallgemeinerte Ableitung

$$\frac{d\varphi(x)}{dx} = -i\psi(x)$$

besitzt, d. h., daß $\varphi(x)$ zu $W_2^{(1)}[0, 1]$ gehört und $i \frac{d\varphi(x)}{dx} = \psi(x)$ ist, wobei jede Funktion $\varphi(x) \in W_2^{(1)}[0, 1]$ der Bedingung (24) mit $\psi(x) = i \frac{d\varphi(x)}{dx}$ genügt. Damit zeigten wir $D(A^*) = W_2^{(1)}[0, 1]$ und $A^*\varphi = i \frac{d\varphi(x)}{dx}$. Der Raum $W_2^{(1)}[0, 1]$ ist offenbar umfassender als $\dot{W}_2^{(1)}[0, 1]$. Man kann sich leicht davon überzeugen, daß A^* auf $D(A^*)$ nicht symmetrisch ist.

Wir zeigen nun, daß auf $R(A)$ ein beschränktes Inverses existiert [woraus dann folgt, daß $R(\bar{A})$ ein Unterraum von H ist]. Es sei $\varphi(x) \in D(A)$; dann ist $\varphi(x) = \frac{1}{i} \int_0^x A\varphi(x) dx$, und auf Grund der Schwarzischen Ungleichung gilt

$$\|\varphi\|^2 = \int_0^1 \left| \int_0^x A\varphi(x) dx \right|^2 dx \leq \|A\varphi\|^2.$$

Daraus folgt, daß A^{-1} auf $R(A)$ existiert und daß $\|A^{-1}\| \leq 1$ ist.

Wir untersuchen nun, ob verschiedene selbstadjungierte Erweiterungen des Operators A möglich sind. Wir wissen, daß für jede symmetrische Erweiterung $\bar{A}$ die Beziehung $A \subseteq \bar{A} \subseteq A^*$ gilt, d. h., daß bei jeder Erweiterung Elemente aus $D(A^*)$ zu $D(A)$ hinzugefügt werden müssen und auf diesen Elementen z die Gültigkeit von $\bar{A}z = A^*z$ gefordert werden muß. Dabei ist zu beachten, daß bei der Erweiterung die Symmetrie des Operators erhalten bleibt.

Den Raum H stellen wir in der Form

$$H = R(\bar{A}) \oplus U$$

dar. Auf Grund der Sätze aus [185] besteht U aus den Nullstellen $u(x)$ des adjungierten Operators. Nun ist $A^*u = i \frac{du(x)}{dx}$, so daß $u(x) = \text{const}$ ist. Wir versuchen, A so zu erweitern, daß $R(A)$ zu H erweitert wird und der Operator dabei symmetrisch bleibt. Dazu müssen wir alle Lösungen von $A^*\varphi = \text{const}$ nehmen und unter ihnen diejenigen auswählen, für die der Operator D symmetrisch ist.

Offenbar hat $\varphi(x)$ die Form $\varphi(x) = c_1(x + c)$. Die Konstanten c und c_1 wählen wir so, daß D auf $\varphi(x)$ symmetrisch ist, d. h.

$$0 = (D\varphi, \varphi) - (\varphi, D\varphi) = i\varphi(x) \overline{\varphi(x)} \Big|_{x=0}^{x=1} = i|c_1|^2 [(1+c)(1+\bar{c}) - c\bar{c}] = i|c_1|^2 [1+c+\bar{c}].$$

Daraus folgt c_1 beliebig, $c = -\frac{1}{2} + \beta i$, wobei β eine beliebige reelle Zahl ist. Zu $D(\bar{A})$ fügen wir die Elemente

$$\varphi(x) = c_1 \left(x - \frac{1}{2} + \beta i \right)$$

hinzu, wobei β eine feste reelle und c_1 eine beliebige komplexe Zahl ist. Die so erhaltene Menge bezeichnen wir mit $D(\bar{A})$ und den Operator D auf ihr mit $\bar{A}$. Es läßt sich leicht zeigen, daß $\bar{A}$ eine symmetrische Erweiterung von A ist.

Da nach Konstruktion des Operators $\bar{A}$ sein Wertebereich mit ganz H übereinstimmt, ist $\bar{A}$ eine selbstadjungierte Erweiterung von A [187]. Die hinzugefügten Elemente dieser Erweiterung genügen, wie man leicht sieht, der Randbedingung

$$\varphi(1) = \frac{\frac{1}{2} + \beta i}{-\frac{1}{2} + \beta i} \varphi(0), \quad \text{d. h.} \quad \varphi(1) = e^{i\theta} \varphi(0), \quad 0 < \theta < 2\pi. \quad (25)$$

Dieser Bedingung genügen offenbar auch die Elemente von $D(\tilde{A})$. Andererseits gilt für jedes Element $\varphi(x) \in D(A^*)$, das die Bedingung (25) erfüllt, die Identität

$$\int_0^1 i \frac{d\varphi(x)}{dx} \overline{\omega(x)} dx = \int_0^1 \varphi(x) i \frac{d\overline{\omega(x)}}{dx} dx$$

für ein beliebiges $\omega(x) \in D(\tilde{A})$. Daher ist $\varphi(x) \in D(\tilde{A}^*)$ und $\tilde{A}^*\varphi = i \frac{d\varphi(x)}{dx}$. Der Operator $\tilde{A}$ ist eine selbstadjungierte Erweiterung von A , also $\tilde{A}^* = \tilde{A}$, und folglich besteht $D(\tilde{A})$ aus allen Elementen von $D(A^*)$, die (25) genügen. Der Raum $\dot{W}_2^{(1)}[0, 1] = D(\tilde{A})$ enthält demzufolge alle absolut stetigen Funktionen $\varphi(x)$, die in den Punkten 0 und 1 verschwinden und die Ableitung

$$\frac{d\varphi(x)}{dx} \in L_2[0, 1]$$

besitzen.

Wir haben somit alle möglichen selbstadjungierten Erweiterungen von A konstruiert, für welche $R(\tilde{A}) = H$ ist. Jede dieser Erweiterungen ist durch einen beliebigen reellen Parameter β oder, was das gleiche ist, durch eine reelle Zahl θ ($0 < \theta < 2\pi$) bestimmt.

Wir betrachten nun solche selbstadjungierten Erweiterungen A' von A , für welche $R(A') = R(\tilde{A})$ ist.

Existiert eine solche Erweiterung, so wird $D(\tilde{A})$ nur durch die Nullstellen des Operators A^* , d. h. die Elemente $\varphi(x) = \text{const}$ ergänzt. Der Operator D auf der Menge $D(A) + \text{const}$ ist der symmetrische Operator A' , der eine Erweiterung von $\tilde{A}$ ist. Die Menge $D(A')$ läßt sich dadurch charakterisieren, daß sie aus den Elementen $\varphi(x) \in D(A^*)$ besteht, für welche $\varphi(0) = \varphi(1)$ ist.

Wir beweisen nun, daß $D(A'^*)$ mit $D(A')$ übereinstimmt. Es sei $\varphi(x) \in D(A'^*)$, d. h., für jedes $\omega(x) \in D(A')$ gelte

$$0 = (A'\omega, \varphi) - (\omega, \varphi). \quad (26)$$

Wegen $D(A'^*) \subseteq D(A^*)$ ist

$$(A'\omega, \varphi) = \int_0^1 i \frac{d\omega(x)}{dx} \overline{\varphi(x)} dx = \int_0^1 \omega(x) i \frac{d\overline{\varphi(x)}}{dx} dx + i\omega(x) \overline{\varphi(x)} \Big|_{x=0}^{x=1},$$

woraus sich wegen $\omega(x) \in D(A')$ und (26)

$$\varphi(x) = i \frac{d\varphi(x)}{dx} \quad \text{und} \quad \varphi(0) = \varphi(1),$$

also $\varphi(x) \in D(A')$ ergibt. Damit ist gezeigt, daß A' eine selbstadjungierte Erweiterung von A ist. Vergleichen wir die Randbedingung $\varphi(0) = \varphi(1)$, der die Funktionen aus $D(A')$ unterworfen sind, mit der Bedingung (25) für die oben erhaltenen Erweiterungen, so sehen wir, daß sie dem Wert $\theta = 0$ (bzw. $\beta = \infty$) entspricht.

Somit haben wir alle möglichen selbstadjungierten Erweiterungen des Operators A untersucht. Außer ihnen hat A noch verschiedene nicht selbstadjungierte Erweiterungen, mit denen wir uns aber nicht beschäftigen wollen.

Die selbstadjungierten Erweiterungen $\tilde{A}$, die einem von 0 verschiedenen Wert von θ entsprechen, haben beschränkte Inverse $(\tilde{A})^{-1}$; denn ist $\tilde{A}\varphi = i \frac{d\varphi(x)}{dx} = 0$, so ist $\varphi(x) = c$, und wegen (25) muß $c = e^{i\theta}c$, also $c = 0$ sein. Daher existiert der Operator $(\tilde{A})^{-1}$, und da er auf ganz H definiert und selbstadjungiert ist, ist er auch beschränkt [187]. Der Operator A' besitzt kein Inverses auf $R(A') = R(\tilde{A})$.

2. Der Operator $D = i \frac{d}{dx}$ im Raum $L_2(-\infty, \infty)$. Den auf stetig differenzierbaren,

finiten Funktionen $\varphi(x)$ definierten Differentialoperator D bezeichnen wir mit A . Er ist symmetrisch, und sein Definitionsbereich $D(A)$ ist in H dicht.

Wir untersuchen zunächst seinen adjungierten Operator A^* . Eine Funktion $\psi(x)$ gehört zu $D(A^*)$, wenn für jedes $\varphi(x) \in D(A)$ die Beziehung

$$\int_{-\infty}^{\infty} i \frac{d\varphi(x)}{dx} \overline{\psi(x)} dx = \int_{-\infty}^{\infty} \varphi(x) \overline{\psi^*(x)} dx \quad (27)$$

mit $\psi^*(x) \in L_2(-\infty, \infty)$ erfüllt ist. Dann ist $\psi^*(x) = A^*\psi(x)$. Aus der ersten Definition der verallgemeinerten Ableitung folgt sofort, daß $D(A^*)$ gleich der Menge $W_2^{(1)}(-\infty, \infty)$ und daß $\psi^*(x) = D\psi(x)$ ist; dabei ist $W_2^{(1)}(-\infty, \infty)$ die Menge derjenigen Funktionen aus $L_2(-\infty, \infty)$, die auf jedem endlichen Intervall stetig sind und eine Ableitung aus $L_2(-\infty, \infty)$ besitzen. Wir zeigen zunächst, daß $\varphi(x)$ für $x \rightarrow \pm \infty$ gegen 0 strebt, wenn $\varphi(x)$ zu $W_2^{(1)}(-\infty, \infty)$ gehört. Aus der offensichtlich richtigen Formel

$$|\varphi(x)|^2 = |\varphi(a)|^2 + \int_a^x \frac{d\varphi(x)}{dx} \overline{\varphi(x)} dx + \int_a^x \varphi(x) \frac{d\overline{\varphi(x)}}{dx} dx$$

und $\varphi(x) \frac{d\varphi(x)}{dx} \in L_2(-\infty, \infty)$ ergibt sich nämlich, daß $|\varphi(x)|$ für $x \rightarrow \pm \infty$ einen endlichen Grenzwert hat und dieser gleich 0 ist.

Wir untersuchen nun A^{**} . Die Funktion $\psi(x)$ gehört zu $D(A^{**})$, wenn für jedes $\varphi(x) \in D(A^*)$ die Beziehung (27) mit $\psi^*(x) \in L_2(-\infty, \infty)$ erfüllt ist. Dann haben wir $\psi^*(x) = A^{**}\psi(x)$.

Mit Hilfe von $A \subset A^*$ und $A^{**} \subseteq A^*$ können wir behaupten, daß jede Funktion $\psi(x)$ aus $D(A^{**})$ auch zu $D(A^*)$, d. h. zu $W_2^{(1)}(-\infty, \infty)$ gehören muß und daß $A^{**}\psi(x) = D\psi(x)$ gilt. Andererseits können wir leicht die Beziehung (27) nachweisen, wenn wir $\varphi(x) \in D(A^*)$, $\psi(x) \in W_2^{(1)}(-\infty, \infty)$ und $\psi^*(x) = i \frac{d\varphi(x)}{dx}$ annehmen; denn partielle Integration über ein beliebiges endliches Intervall ergibt

$$\int_a^b i \frac{d\varphi(x)}{dx} \overline{\psi(x)} dx = \int_a^b \varphi(x) i \frac{d\overline{\psi(x)}}{dx} dx + i\varphi(x) \overline{\psi(x)} \Big|_{x=a}^{x=b}.$$

Hieraus erhalten wir die Beziehung (27), wenn $\varphi(x)$ und $\psi(x)$ für $x \rightarrow \pm \infty$ (für $a \rightarrow \infty$ und $b \rightarrow -\infty$) gegen 0 streben. Also ist $A^{**} = A^*$, d. h., A^* ist selbstadjungiert. Da wir schon wissen, daß $A^{**} = \bar{A}$ ist, führt also die Abschließung des Operators A auf den selbstadjungierten Operator A^* .

3. Der Operator $D = i \frac{d}{dx}$ im Raum $L_2(0, \infty)$. Es sei A der Operator D auf allen stetig differenzierbaren Funktionen, die im Unendlichen und in der Umgebung des Punktes $x=0$ *finit* sind. Mit Hilfe analoger Überlegungen wie in den vorhergehenden Beispielen läßt sich zeigen, daß A^* ebenfalls gleich dem Operator D ist, $D(A^*)$ mit der Menge derjenigen Funktionen $\varphi(x) \in H$ übereinstimmt, die auf einem beliebigen endlichen Intervall $[0, a]$ absolut stetig sind und deren Ableitungen in $L_2(0, \infty)$ liegen, und daß ferner $D(\bar{A}) = D(A^{**})$ die Menge derjenigen Funktionen $\varphi(x) \in D(A^*)$ ist, die der Forderung $\varphi(0) = 0$ genügen, wobei $\bar{A}\varphi(x) = i \frac{d\varphi(x)}{dx}$ ist. Somit ist $D(A^*)$ umfassender als $D(\bar{A})$, und $\bar{A}$ ist wegen $(\bar{A}^*) = A^*$ nicht selbstadjungiert. Der Operator $\bar{A}$ besitzt auch keine selbstadjungierten Erweiterungen. Zum Beweis nehmen wir das Gegenteil an, $\bar{A}$ sei eine solche Erweiterung von $\bar{A}$. Wegen $\bar{A} \subseteq A^*$ ist $\bar{A}\varphi = i \frac{d\varphi(x)}{dx}$, und $D(\bar{A})$ muß umfassen-

der als $D(\tilde{A})$ sein. Wir zeigen aber nun, daß die Funktion $\varphi(x)$ aus $D(\tilde{A})$ auch in $D(\tilde{A})$ liegt, und gelangen so zu einem Widerspruch.

Es sei $\varphi(x) \in D(\tilde{A})$ und somit $\varphi(x) \in D(A^*)$. Die Formel $(\tilde{A}\varphi, \varphi) = (\varphi, \tilde{A}\varphi)$ führt auf die Beziehung

$$\int_0^\infty i \frac{d\varphi}{dx} \overline{\varphi(x)} dx = \int_0^\infty \varphi(x) i \frac{d\overline{\varphi(x)}}{dx} dx,$$

aus der $\varphi(0) = 0$, also $\varphi(x) \in D(\tilde{A})$ wegen $\varphi(x) \rightarrow 0$ für $x \rightarrow \infty$ folgt. Der Operator $\tilde{A}$ hat somit keine selbstadjungierten Erweiterungen.

4. Der Operator $-\frac{d^2\varphi(x)}{dx^2} + q(x)\varphi(x)$ im Raum $L_2(0, 1)$. Es sei $q(x)$ eine reelle, auf dem Intervall $[0, 1]$ stetige Funktion und A der Operator $-\frac{d^2}{dx^2} + q(x)$ auf der Menge $D(A)$ der Funktionen $\varphi(x)$, die die folgenden Eigenschaften besitzen: $\varphi(x)$ und $\frac{d\varphi(x)}{dx}$ sind absolut stetig für $x \in [0, 1]$, $\varphi(0) = \varphi(1) = 0$ und $\frac{d^2\varphi(x)}{dx^2} \in L_2(0, 1)$. Man sieht sofort, daß A symmetrisch und $D(A)$ in H dicht ist.

Wir nehmen an, $q(x)$ sei so beschaffen, daß die Gleichung $-y'' + q(x)y = 0$ in den Punkten $x=0$ und $x=1$ außer der trivialen Lösung $y \equiv 0$ keine andere verschwindende Lösung hat, und zeigen, daß dann $R(A) = H$, also A ein selbstadjungierter Operator ist. Es sei $f(x) \in L_2(0, 1)$. Wir müssen beweisen, daß eine Funktion $\varphi(x) \in D(A)$ existiert, für welche $A\varphi(x) = f(x)$ ist. Dazu führen wir die Funktion

$$\psi(x) = -\int_0^x \left[\int_0^t f(\tau) d\tau \right] dt + x \int_0^1 \left[\int_0^t f(\tau) d\tau \right] dt$$

ein, die offenbar zu $D(A)$ gehört und für die Gleichung $-\psi''(x) = f(x)$ gilt; $\omega(x)$ sei diejenige Lösung von

$$-\omega''(x) + q(x)\omega(x) = -q(x)\psi(x),$$

die den Bedingungen $\omega(0) = \omega(1) = 0$ genügt. Eine solche Lösung (mit stetigen Ableitungen bis zur zweiten Ordnung) existiert [IV/2, 75]. Wir können unmittelbar nachprüfen, daß die Funktion $\varphi(x) = \psi(x) + \omega(x)$ zu $D(A)$ gehört und $A\varphi(x) = f(x)$ ist, was zu beweisen war. Somit ist A ein selbstadjungierter Operator. Er besitzt ein beschränktes Inverses [IV/2, 75],

$$A^{-1}f(x) = -\int_0^1 G(x, t) f(t) dt,$$

wobei $G(x, t)$ die Greensche Funktion von A bei den Randbedingungen $\omega(0) = \omega(1) = 0$ ist. Die oben erwähnte Forderung für die in $x=0$ und $x=1$ verschwindenden Lösungen der Gleichung $-y'' + q(x)y = 0$ ist erfüllt, wenn etwa $q(x) \geq 0$ ist.

5. Der Operator $D^k = i^k \frac{\partial^k}{\partial x_{i_1} \dots \partial x_{i_k}}$ im Raum $L_2(D)$, wobei D ein beschränkter Bereich des R_n ist. Den Operator D^k auf allen k -mal stetig differenzierbaren, in D finiten Funktionen bezeichnen wir mit A . Es ist bekannt, daß $D(A)$ in H dicht ist, und mit Hilfe von (123) aus [109] läßt sich leicht die Symmetrie von A nachweisen. Der Definitionsbereich des selbstadjungierten Operators A^* enthält alle Funktionen $\varphi(x) \in L_2(D)$, die innerhalb von D eine in $L_2(D)$ liegende verallgemeinerte Ableitung der Form D^k haben (dies folgt aus der Definition der verallgemeinerten Ableitung). Offenbar ist $D(A^*)$ umfassender als $D(A)$. Auf $R(A)$ existiert ein beschränktes Inverses von A ; denn gehört die Funktion $\varphi(x)$ zu $D(A)$, setzen wir sie außerhalb D gleich 0 und schließen wir D in den Würfel $-a \leq x_i \leq a$ ein, so läßt sich $\varphi(x)$ in der Form

$$\varphi(x) = \int_{-a}^{x_{i_1}} \dots \int_{-a}^{x_{i_k}} (-i)^k A\varphi(x_1, \dots, x_n) dx_{i_1} \dots dx_{i_k}$$

darstellen, woraus sich mit Hilfe der Schwarzschen Ungleichung leicht

$$\|\varphi\|_{L_4(D)} \leq c \|A\varphi\|_{L_4(D)}$$

ergibt, so daß also A^{-1} auf $R(A)$ existiert, $\|A^{-1}\| \leq c$ ist und $R(\bar{A})$ einen Unterraum von H bildet. Wie wir im folgenden in der Theorie der Erweiterungen von Operatoren beweisen werden, gibt es zu diesen symmetrischen Operatoren wenigstens eine selbstadjungierte Erweiterung.

6. Der Operator $-\Delta = -\sum_{k=1}^n \frac{\partial^2}{\partial x_k^2}$ im Raum $L_2(D)$, wobei D ein beschränkter Bereich des R_n ist. Mit A bezeichnen wir den auf allen zweimal stetig differenzierbaren, in D finiten Funktionen definierten Operator $-\Delta$, d. h., es sei $D(A) = \dot{C}^{(2)}(D)$. Die Menge $\dot{C}^{(2)}(D)$ ist in H dicht. Gehört $\varphi(x)$ zu $D(A)$, so ist

$$\int_D -\Delta\varphi(x) \overline{\varphi(x)} dx = \int_D \sum_{k=1}^n \left| \frac{\partial\varphi(x)}{\partial x_k} \right|^2 dx; \quad (28)$$

A ist also ein positiver und infolgedessen symmetrischer Operator. Außerdem ist bekannt [114], daß für alle $\varphi(x) \in D(A)$

$$\|\varphi\|_{L_4(D)} \leq c \left[\int_D \sum_{k=1}^n \left| \frac{\partial\varphi(x)}{\partial x_k} \right|^2 dx \right]^{1/2} \quad (29)$$

mit ein und derselben Konstanten c gilt, die allein von der Größe von D abhängt. Aus (28) und (29) ergibt sich

$$\|\varphi\|_{L_4(D)} \leq c^2 \|A\varphi\|_{L_4(D)}; \quad (30)$$

A ist somit positiv definit, auf $R(A)$ existiert ein beschränktes Inverses A^{-1} mit $\|A^{-1}\| \leq c^2$, und es ist $R(\bar{A}) = \bar{R}(\bar{A})$.

Diese Operatoren A gestatten, wie wir später sehen werden, selbstadjungierte Erweiterungen, wobei jeder Erweiterung ein Randwertproblem für den Laplaceschen Operator entspricht. Wir wollen nun die Struktur von $D(\bar{A})$ und $D(A^*)$ untersuchen und wählen dazu eine Funktion $\varphi(x) \in D(A)$. Mit Hilfe partieller Integration läßt sich die Gültigkeit der Beziehung

$$\int_D |A\varphi(x)|^2 dx = \int_D \sum_{i,k=1}^n \left| \frac{\partial^2\varphi(x)}{\partial x_i \partial x_k} \right|^2 dx \quad (31)$$

leicht nachprüfen, aus der wir, wenn wir (28) und (29) berücksichtigen, die Ungleichung

$$\|\varphi\|_{W_2^{(2)}(D)} \leq c_1 \|A\varphi(x)\|_{L_4(D)} \quad (32)$$

erhalten; hier ist c_1 eine gewisse, nur von D abhängende Konstante. Es sei nun $\varphi_m(x) \in D(A)$, $\varphi_m(x) \Rightarrow \varphi(x)$ und $A\varphi_m(x) \Rightarrow A\varphi(x)$. Aus (32) folgt dann, daß $\varphi_m(x)$ gegen $\varphi(x)$ der Norm von $W_2^{(2)}(D)$ nach konvergiert, so daß $\varphi(x)$ zu $W_2^{(2)}(D)$ gehört und $\bar{A}\varphi = -\Delta\varphi(x)$ gilt. Diese Vervollständigung von $D(A)$ bezeichnen wir in [113] mit $\dot{W}_2^{(2)}(D)$, so daß $D(\bar{A}) = \dot{W}_2^{(2)}(D)$ ist.

Die Zugehörigkeit von $\varphi(x)$ zu $D(A^*)$ bedeutet, daß es eine Funktion $\psi(x) \in H$ gibt derart, daß die Identität

$$\int_D -\Delta\omega(x) \overline{\varphi(x)} dx = \int_D \omega(x) \overline{\psi(x)} dx \quad (33)$$

für alle $\omega(x) \in D(A)$ gültig ist. Wir konstruieren nun das Newtonsche Potential

$$u(x) = \frac{1}{k_n} \int_D \frac{\psi(y)}{|x-y|^{n-2}} dy, \quad (34)$$

wobei k_n den Flächeninhalt der Einheitssphäre im R_n bezeichnet. Es ist bekannt

[II, 212], daß $u(x)$ zweimal stetig differenzierbar in D ist und

$$-\Delta u(x) = \psi(x) \quad (35)$$

gilt, wenn $\psi(y)$ eine stetig differenzierbare Funktion ist.

Wir beweisen nun, daß $u(x) = \hat{W}_2^{(2)}(D)$ für $\psi(x) \in L_2(D)$ ist. Dazu setzen wir voraus, $\psi(y)$ sei außerhalb D gleich 0, konstruieren die Mittelungen $\varphi_\varrho(y)$ und betrachten die Funktionen

$$u_\varrho(x) = \frac{1}{k_n} \int_{R_n} \frac{\varphi_\varrho(y)}{|x-y|^{n-2}} dy.$$

Sie sind zweimal stetig differenzierbar und genügen der Gleichung

$$-\Delta u_\varrho(x) = \varphi_\varrho(x). \quad (36)$$

Für $\varrho \rightarrow 0$ konvergieren $u_\varrho(x)$ und $\frac{\partial u_\varrho(x)}{\partial x_k}$ gegen $u(x)$ bzw. $\frac{\partial u(x)}{\partial x_k}$ der Norm von $L_2(D_1)$ nach [115], wobei D_1 ein beliebiger beschränkter Bereich ist. Wir setzen voraus, daß D echt in D_1 enthalten ist. Wegen (36) gilt für $v(x) = u_\varrho(x) - u_{\varrho'}(x)$ die Gleichung

$$\int_{R_n} |\Delta v(x)|^2 \zeta(x) dx = \int_{R_n} |\varphi_\varrho(x) - \varphi_{\varrho'}(x)|^2 \zeta(x) dx; \quad (37)$$

dabei ist $\zeta(x)$ eine feste, nichtnegative, zweimal stetig differenzierbare Funktion, die innerhalb von D gleich 1 und außerhalb von D_1 gleich 0 ist und ferner überall der Forderung

$$\zeta_{x_k}^2 \leq c_2 \zeta(x) \quad (38)$$

genügt (die Existenz solcher Funktionen läßt sich leicht nachweisen). Mit Hilfe partieller Integration bringen wir (37) auf die Form

$$\begin{aligned} \int_D |\varphi_\varrho(x) - \varphi_{\varrho'}(x)|^2 \zeta(x) dx &= \int_{D_1} \sum_{i,k=1}^n \left| \frac{\partial^2 v(x)}{\partial x_i \partial x_k} \right|^2 \zeta(x) dx \\ &+ \int_{D_1} \left[\sum_{i,k=1}^n \frac{\partial v(x)}{\partial x_i} \frac{\partial^2 v(x)}{\partial x_i \partial x_k} \frac{\partial \zeta(x)}{\partial x_k} - \sum_{i=1}^n \frac{\partial v(x)}{\partial x_i} \frac{\partial v(x)}{\partial x_i} \frac{\partial \zeta(x)}{\partial x_i} \right] dx. \end{aligned}$$

Hieraus ergibt sich, wenn wir (38) und die für jedes $\varepsilon > 0$ geltende Ungleichung

$$|2ab| \leq \varepsilon |a|^2 + \frac{1}{\varepsilon} |b|^2$$

benutzen,

$$\begin{aligned} \int_{D_1} \sum_{i,k=1}^n \left| \frac{\partial^2 v(x)}{\partial x_i \partial x_k} \right|^2 \zeta(x) dx &\leq \int_{D_1} |\varphi_\varrho(x) - \varphi_{\varrho'}(x)|^2 \zeta(x) dx \\ &+ c_3 \int_{D_1} \left[\varepsilon \sum_{i,k=1}^n \left| \frac{\partial^2 v(x)}{\partial x_i \partial x_k} \right|^2 \zeta(x) + \frac{1}{\varepsilon} \sum_{k=1}^n \left| \frac{\partial v(x)}{\partial x_k} \right|^2 \right] dx \end{aligned}$$

oder mit $\varepsilon = \frac{1}{2c_3}$

$$\frac{1}{2} \int_{D_1} \sum_{i,k=1}^n \left| \frac{\partial^2 v(x)}{\partial x_i \partial x_k} \right|^2 \zeta(x) dx \leq \int_{D_1} |\varphi_\varrho(x) - \varphi_{\varrho'}(x)|^2 \zeta(x) dx + 2c_3^2 \int_{D_1} \sum_{k=1}^n \left| \frac{\partial v(x)}{\partial x_k} \right|^2 dx.$$

Wir können also schließen, daß die Funktion $v(x) = u_\varrho(x) - u_{\varrho'}(x)$ für $\varrho, \varrho' \rightarrow 0$ nebst ihren Ableitungen erster und zweiter Ordnung gegen 0 konvergiert, und zwar der Norm von $L_2(D)$ nach. Die durch das Integral (34) definierte Grenzfunktion $u(x)$ von $u_\varrho(x)$ gehört somit zu $\hat{W}_2^{(2)}(D)$ und genügt der Gleichung (35). Also gilt für sie

$$\int_D -\Delta \omega(x) \overline{u(x)} dx = \int_D \omega(x) \overline{\psi(x)} dx$$

für jedes $\omega(x) \in D(A)$. Subtrahieren wir dies von (33), so erhalten wir für $\varphi(x)$ die Identität

$$\int_D -\Delta \omega(x) [\overline{\varphi(x)} - \overline{u(x)}] dx = 0.$$

In [119] hatten wir bewiesen, daß aus dieser Identität folgt, daß die Funktion $\varphi(x) - u(x)$ harmonisch ist, wobei für jede in D harmonische Funktion eine entsprechende Identität gilt.

Somit läßt sich eine Funktion $\varphi(x)$ aus $D(A^*)$ in der Form

$$\varphi(x) = \frac{1}{k_n} \int_D \frac{\psi(y)}{|x-y|^{n-2}} dy + v(x)$$

darstellen, wobei $\psi(x)$ zu $L_2(D)$ gehört und $v(x)$ eine in D harmonische Funktion ist. Da $\varphi(x)$ und das Integral in $L_2(D)$ liegen, ist auch $v(x) \in L_2(D)$. Daß jede Funktion $\varphi(x)$ dieser Art in $D(A^*)$ auftritt und

$$A^* \varphi = -\Delta \varphi(x) = \varphi(x)$$

ist, läßt sich leicht nachprüfen.

7. Der Operator $-\Delta = -\sum_{k=1}^n \frac{\partial^2}{\partial x_k^2}$ im Raum $L_2(R_n)$. Wir betrachten zunächst nicht den Operator $-\Delta$ selbst, sondern einen der Operatoren $-\Delta + \lambda E$ mit beliebigem reellem λ . Wir setzen λ als positiv, etwa $\lambda = 1$ voraus. Das garantiert, wie wir im folgenden sehen werden, die Existenz eines beschränkten inversen Operators zu $-\Delta + E$ und erleichtert somit die Lösung des Problems der selbstadjungierten Erweiterungen des Operators.

Es sei also A der Operator $-\Delta + E$, der auf der zweimal stetig differenzierbaren und finiten Funktion $\varphi(x)$ definiert ist. Der Definitionsbereich $D(A)$ ist in H dicht, und wir sehen leicht, daß $-\Delta + E$ symmetrisch ist. Aus Beispiel 6 folgt, daß die Elemente aus $D(\bar{A})$ und $D(A^*)$ verallgemeinerte Ableitungen bis zur zweiten Ordnung haben, die bezüglich eines endlichen Bereichs D aus R_n quadratisch summierbar sind. Der Operator A wird auf $D(A^*)$ durch den Differentialoperator

$$-\sum_{k=1}^n \frac{\partial^2 \varphi(x)}{\partial x_k^2} + \varphi(x)$$

dargestellt.

Wir zeigen, daß $D(\bar{A}) = W_2^{(2)}(R_n)$ ist. Einerseits ist $D(\bar{A}) \subset W_2^{(2)}(R_n)$, denn gilt $\varphi_m(x) \in D(A)$ und $\varphi_m(x) \rightarrow \varphi(x)$, $A\varphi_m(x) \rightarrow \bar{A}\varphi(x)$, so folgt aus (28) und (31)

$$\begin{aligned} & \int_{R_n} \left[\sum_{k=1}^n \left| \frac{\partial(\varphi_l(x) - \varphi_m(x))}{\partial x_k} \right|^2 + \sum_{i,k=1}^n \left| \frac{\partial^2(\varphi_l(x) - \varphi_m(x))}{\partial x_i \partial x_k} \right|^2 \right] dx \\ &= \int_{R_n} [-\Delta(\varphi_l(x) - \varphi_m(x)) (\overline{\varphi_l(x) - \varphi_m(x)}) + |\Delta(\varphi_l(x) - \varphi_m(x))|^2] dx \\ &\leq \|A(\varphi_l - \varphi_m)\|_{L_2(R_n)} \cdot \|\varphi_l - \varphi_m\|_{L_2(R_n)} + \|A(\varphi_l - \varphi_m)\|_{L_2(R_n)}^2 \rightarrow 0 \quad \text{für } l, m \rightarrow \infty. \end{aligned}$$

Andererseits ist $W_2^{(2)}(R_n) \subseteq D(\bar{A})$. Um dies zu zeigen, nehmen wir ein $\varphi(x) \in W_2^{(2)}(R_n)$ und konstruieren eine Folge von Funktionen $\psi_m(x) = \varphi_{1/m}(x) \zeta_m(x)$, wobei $\varphi_{1/m}(x)$ die Mittelung von $\varphi(x)$ mit dem Radius $1/m$ ist und die ζ_m zweimal stetig differenzierbare und finite Funktionen sind, die für $r \geq 0$ durch die Eigenschaften

$$\zeta_m(r) = \begin{cases} 1 & \text{für } 0 \leq r \leq m, \\ 0 & \text{für } r \geq m+1, \\ \zeta_{m-1}(r-1) & \text{für } m \leq r \leq m+1 \end{cases}$$

charakterisiert werden; $\zeta_0(r) \leq 1$ ist eine stetig differenzierbare, nichtnegative Funktion, die für $r \geq 1$ gleich 0 ist. Jede Funktion $\psi_m(x)$ gehört zu $D(A)$. Wir zeigen nun, daß $\psi_m(x)$ für $m \rightarrow \infty$ der Norm von $W_2^{(2)}(R_n)$ nach gegen $\varphi(x)$ konvergiert. Nehmen wir in dem Ausdruck für die Norm in $W_2^{(2)}(R_n)$ die Integration über R_n in zwei Schritten vor, und

zwar einmal über den Teil, in dem $|x| \leq l$, und dann über den Teil, in dem $|x| \geq l$ ist, so erhalten wir

$$\|\varphi - \psi_m\|_{W_2^{(2)}(R_n)} = \|\varphi - \psi_m\|_{W_2^{(2)}(|x| \leq l)} + \|\varphi - \psi_m\|_{W_2^{(2)}(|x| \geq l)}.$$

Es sei nun ein $\varepsilon > 0$ gegeben. Der zweite Summand kann für alle m kleiner oder gleich $\varepsilon/2$ gemacht werden, wenn l hinreichend groß gewählt wird. Das folgt aus $\|\varphi\|_{W_2^{(2)}(R_n)} < \infty$,

aus der Konstruktion der Funktionen $\zeta_m(x)$ und den Eigenschaften der Mittelung der Funktion $\varphi(x)$ und ihrer verallgemeinerten Ableitungen. Wählen wir auf diese Weise ein festes l , so gilt $\psi_m(x) = \varphi_{1/m}(x)$ für $m > l$ und $|x| \leq l$ sowie $\varphi_{1/m}(x) \rightarrow \varphi(x)$ der Norm von $W_2^{(2)}(|x| \leq l)$ nach. Somit gilt für alle hinreichend großen m die Ungleichung

$$\|\varphi - \psi_m\|_{W_2^{(2)}(|x| \leq l)} \leq \frac{\varepsilon}{2} \quad \text{bzw.} \quad \|\varphi - \psi_m\|_{W_2^{(2)}(R_n)} \leq \varepsilon.$$

Wir können also auf $W_2^{(2)}(R_n) \subseteq D(\bar{A})$ und zusammen mit dem oben Bewiesenen auf $D(\bar{A}) = W_2^{(2)}(R_n)$ schließen.

Wir zeigen nun, daß $\bar{A} = A^*$, also $\bar{A}$ selbstadjungiert ist. Es genügt zu zeigen, daß der Wertebereich des Operators $\bar{A}$ mit $L_2(R_n)$ übereinstimmt. Der Einfachheit wegen setzen wir $n = 3$; ferner wählen wir aus $L_2(R_n)$ eine beliebige finite, stetig differenzierbare Funktion $\varphi(x)$. Die Funktion

$$u(x) = \frac{1}{4\pi} \int_{R_n} \frac{e^{-|x-y|} \varphi(y)}{|x-y|} dy$$

ist bekanntlich [IV/2, 133] zweimal stetig differenzierbar, genügt der Gleichung $-\Delta u + u = \varphi$ und fällt nebst ihren Ableitungen exponentiell für $|x| \rightarrow \infty$, so daß $u(x)$ zu $W_2^{(2)}(R_n)$ gehört. Also ist $(u(x) \in D(\bar{A})) \quad \Delta u = \varphi$, und der Wertebereich $R(\bar{A})$ von $\bar{A}$ ist in H dicht.

Nun beweisen wir, daß auf $R(\bar{A})$ ein beschränktes Inverses $\bar{A}^{-1}$ existiert. Daraus folgt dann, daß $R(\bar{A})$ ein Unterraum von H und somit also $R(\bar{A}) = H$ ist.

Es sei $v(x) \in D(\bar{A})$ und $-\Delta v(x) + v(x) = \varphi(x)$. Multiplizieren wir diese Gleichung mit $v(x)$ und integrieren wir über R_n , so erhalten wir, wenn wir ihr mit Hilfe partieller Integration eine andere Form geben,

$$\int_{R_n} \varphi(x) \overline{v(x)} dx = \int_{R_n} \left[\sum_{k=1}^n \left| \frac{\partial v(x)}{\partial x_k} \right|^2 + |v(x)|^2 \right] dx.$$

Mit Hilfe der Cauchyschen Ungleichung ergibt sich hieraus

$$\|v\|_{L_2(R_n)} \leq \|\varphi\|_{L_2(R_n)} = \|Av\|_{L_2(R_n)},$$

d. h., A^{-1} existiert tatsächlich auf $R(A)$, und es ist $\|A^{-1}\| \leq 1$.

Damit ist endlich bewiesen, daß $\bar{A} = A^*$ gilt oder, was das gleiche ist, die Abschließung des Operators A auf seine (eindeutige) selbstadjungierte Erweiterung führt, wobei $D(\bar{A}) = W_2^{(2)}(R_n)$ ist. Dies gilt auch für $-\Delta = A - E$, d. h., der Operator $-\Delta$ ist auf $W_2^{(2)}(R_n)$ selbstadjungiert. Jedoch besitzt $-\Delta$ zum Unterschied von $-\Delta + E$ kein beschränktes Inverses ($-\Delta$ hat zwar ein Inverses, es ist jedoch nicht beschränkt).

189. Das Spektrum eines selbstadjungierten Operators. Wie in [128] läßt sich auch hier zeigen, daß die Eigenwerte eines selbstadjungierten Operators reell und die zu verschiedenen Eigenwerten gehörigen Eigenelemente zueinander orthogonal sind und ein selbstadjungierter Operator ein normiertes Orthogonalsystem der Eigenelemente erzeugt. Da ein selbstadjungierter Operator abgeschlossen ist, bilden die einem festen Eigenwert zugeordneten Eigenelemente (zuzüglich des Nullelements) einen Unterraum.

Auch die Definitionen des regulären Punktes des Spektrums sind die gleichen wie

in [129]. Wir beweisen nun für einen selbstadjungierten Operator Sätze, die denen aus [129] analog sind.

Satz 1. *Ist λ kein Eigenwert des selbstadjungierten Operators A , so ist die lineare Mannigfaltigkeit $R(A - \lambda E)$ in H dicht.*

Ist λ eine reelle Zahl, so ist $A - \lambda E$ ein selbstadjungierter Operator, und die Behauptung folgt auf Grund von Satz 1 aus [187]. Ist λ komplex, so würde, wenn $R(A - \lambda E) \neq H$ wäre, ein vom Nullelement verschiedenes und zu $R(A - \lambda E)$ orthogonales Element z existieren, d. h., es wäre

$$((A - \lambda E)x, z) = 0 \quad \text{oder} \quad ((A - \lambda E)x, z) = (x, 0).$$

Hieraus folgt $z \in D(A)$ und $(A - \lambda E)^* z = (A - \bar{\lambda} E)z = 0$, also $Az = \bar{\lambda}z$. Das ist aber nicht möglich, denn A kann nur reelle Eigenwerte haben.

Satz 2. *Der Punkt λ ist genau dann regulärer Punkt des Spektrums des selbstadjungierten Operators A , wenn eine positive Zahl p existiert, für die die Ungleichung*

$$\|(A - \lambda E)x\| \geq p\|x\| \quad (x \in D(A)) \quad (39)$$

erfüllt ist. Die nicht reellen Werte λ sind reguläre Punkte des selbstadjungierten Operators.

Der Beweis verläuft wie in [129], nur mit dem Unterschied, daß statt der Stetigkeit die Abgeschlossenheit von A benutzt wird.

Wie in [129] ergibt sich auch hier die Folgerung: *Die Punkte des Spektrums bilden eine abgeschlossene Menge.*

Im folgenden werden wir sehen, daß das Spektrum jedes selbstadjungierten Operators wenigstens einen Punkt enthält. Die nicht reellen Werte von λ sind reguläre Punkte des selbstadjungierten Operators A ; daher werden wir nur von reellen λ sprechen. In diesem Fall ist der Operator $A - \lambda E$ selbstadjungiert.

Es sei λ kein Eigenwert. Ist $R(A - \lambda E) = H$, so ist der abgeschlossene, auf ganz H definierte Operator $(A - \lambda E)^{-1}$ beschränkt, also λ ein regulärer Punkt. Stimmt jedoch $R(A - \lambda E)$ nicht mit H überein, so folgt aus Satz 3 von [187], daß $(A - \lambda E)^{-1}$ ein auf $R(A - \lambda E)$ nicht beschränkter Operator ist. So gelangen wir zu dem

Satz 3. *Es sei die reelle Zahl λ kein Eigenwert. Ist $R(A - \lambda E) = H$, so ist λ ein regulärer Punkt. Stimmt die lineare Mannigfaltigkeit $R(A - \lambda E)$ nicht mit ganz H überein (sie ist in H dicht), so ist λ ein Punkt des Spektrums.*

Wir untersuchen nun den Fall, daß λ ein Eigenwert ist, und bezeichnen mit P_λ den Unterraum der entsprechenden Eigenelemente (zuzüglich des Nullelements). Es gibt einen einzigen Operator, für den $P_\lambda = H$ ist; das ist der Operator, der die Multiplikation mit der Zahl λ vermittelt: $Ax = \lambda x$ für alle $x \in H$. In allen übrigen Fällen ist P_λ ein echter Unterraum von H .

Wir wollen nun zeigen, daß P_λ die Menge der Elemente x ist, die zu der linearen Mannigfaltigkeit des Elemente $(A - \lambda E)z$ mit $z \in D(A)$ orthogonal sind. Aus $((A - \lambda E)z, x) = 0$ folgt wegen der Selbstadjungiertheit von $A - \lambda E$, daß $x \in D(A)$ und $(A - \lambda E)x = 0$ ist, und umgekehrt folgt aus $x \in D(A)$ und $(A - \lambda E)x = 0$ die Beziehung $((A - \lambda E)z, x) = 0$. Also ist der zu P_λ komplementäre Unterraum Q_λ [122] die abgeschlossene Hülle der linearen Mannigfaltigkeit der durch

$$y = (A - \lambda E)z \quad (z \in D(A))$$

definierten Elemente y , d. h. $Q_\lambda = \overline{R(A - \lambda E)}$. Wir schreiben ein Element $z \in D(A)$ in der Form $z = z_1 + z_2$ mit $z_1 \in P_\lambda$ und $z_2 \in Q_\lambda$. Aus der Definition von P_λ folgt $z_1 \in D(A)$

und $Az_1 = \lambda z_1$ und infolgedessen auch $z_2 = z - z_1 \in D(A)$, d. h., die Projektion der linearen Mannigfaltigkeit $D(A)$ auf Q_λ ist eine lineare Mannigfaltigkeit aus $D(A)$, die wir mit $D_\lambda(A)$ bezeichnen wollen. $D_\lambda(A)$ ist offenbar die lineare Mannigfaltigkeit derjenigen Elemente von Q_λ , die zu $D(A)$ gehören. Auf $D_\lambda(A)$ ist ein Operator A definiert, und wir erkennen leicht, daß Ay für $y \in D_\lambda(A)$ in Q_λ liegt. Nach Voraussetzung ist nämlich $(y, x) = 0$, wenn $x \in D(A)$ ist und der Gleichung $Ax = \lambda x$ genügt; daraus folgt $(Ay, x) = (y, Ax) = (y, \lambda x) = 0$. Wegen des oben Gesagten können wir A als Operator im Unterraum Q_λ auffassen, den wir als neuen Raum H ansehen. Wir bezeichnen diesen Operator mit A_λ , so daß A_λ auf $D_\lambda(A)$ definiert und $A_\lambda y = Ay$ für $y \in D_\lambda(A)$ ist. Statt $D_\lambda(A)$ können wir auch $D(A_\lambda)$ schreiben, wenn wir die üblichen Bezeichnungen benutzen.

Aus einem allgemeinen Satz, den wir in [191] beweisen werden, folgt, daß $D(A_\lambda)$ in Q_λ dicht und A_λ ein in Q_λ selbstadjungierter Operator ist. Der Wert λ kann nach Konstruktion von Q_λ kein Eigenwert von A_λ , aber sowohl regulärer Punkt von A_λ als auch ein Punkt des Spektrums von A_λ sein. Im ersten Fall bildet $(A_\lambda - \lambda E)z$ mit $z \in D(A_\lambda)$ den Bereich $D(A_\lambda)$ in den vollständigen Raum Q_λ und im zweiten Fall in eine in Q_λ dichte lineare Mannigfaltigkeit ab. Statt $(A_\lambda - \lambda E)z$ können wir auch $(A - \lambda E)z$ schreiben. Für alle $z \in P_\lambda$ gilt $(A - \lambda E)z = 0$.

Dies führt auf die folgende Klassifikation der Werte von λ :

I. Die regulären Werte von λ , für welche die Tatsache $R(A - \lambda E) = H$ charakteristisch ist und der beschränkte inverse Operator $(A - \lambda E)^{-1}$ existiert.

II. Die Werte von λ , für welche $R(A - \lambda E)$ eine von H verschiedene lineare Mannigfaltigkeit und $\overline{R(A - \lambda E)} = H$ ist. Auf $R(A - \lambda E)$ existiert ein inverser, nicht beschränkter Operator $(A - \lambda E)^{-1}$. Gewöhnlich sagt man, diese Werte von λ gehören zu einem kontinuierlichen Spektrum.

III. Die Werte von λ , welche Eigenwerte von A sind und für welche A_λ den Punkt λ als regulären Punkt besitzt. Für diese λ ist $R(A - \lambda E)$ ein Unterraum, der nicht mit H übereinstimmt. Gewöhnlich sagt man, diese Werte von λ gehören nur zu einem Punktspektrum.

IV. Die Werte von λ , welche Eigenwerte von A sind und für welche A_λ den Punkt λ als Punkt des Spektrums hat. Für diese λ ist $R(A - \lambda E)$ kein Unterraum, sondern eine lineare Mannigfaltigkeit, deren abgeschlossene Hülle $\overline{R(A - \lambda E)}$ ein nicht mit H übereinstimmender Unterraum ist. Von diesem λ sagt man, sie gehören gleichzeitig zum Punktspektrum und zum kontinuierlichen Spektrum.

In dem Spektrum des selbstadjungierten Operators A können auch einige Typen von λ fehlen. Für beschränkte selbstadjungierte Operatoren haben wir bewiesen, daß das Spektrum wenigstens einen Punkt enthält. Wir können leicht zeigen, daß dies auch für einen nicht beschränkten selbstadjungierten Operator A gilt. Dazu setzen wir voraus, daß für jedes reelle λ der Operator $A - \lambda E$ ein beschränktes Inverses besitzt. Offenbar gilt die Gleichung

$$\frac{1}{\lambda} (A - \lambda E) A^{-1} = \frac{1}{\lambda} - A^{-1} \quad (\lambda \neq 0),$$

wobei beide Seiten einen selbstadjungierten beschränkten Operator darstellen. Aus dieser Gleichung und aus der Voraussetzung folgt, daß $R(\mu E - A^{-1})$ für jedes reelle μ mit H übereinstimmt. Das widerspricht jedoch der Tatsache, daß der beschränkte Operator A^{-1} Punkte des Spektrums hat.

Im folgenden werden wir beweisen, daß ein nicht beschränkter selbstadjungierter

Operator unendlich viele Punkte des Spektrums hat, die außerhalb eines beliebigen festen Intervalls der λ -Achse liegen. Ist λ kein Eigenwert von A , so ist der Operator

$$R_\lambda = (A - \lambda E)^{-1}, \quad (40)$$

wie wir wissen, die Resolvente des Operators A . Er ist auf $R(A - \lambda E)$ definiert und bildet diese lineare Mannigfaltigkeit eineindeutig in $D(A)$ ab. Aus der Definition des inversen Operators folgt, daß sich aus $x \in D(A)$ und $R_\lambda x = 0$ die Beziehung $x = 0$ ergibt (vgl. [144]). Wie für beschränkte Operatoren gilt auch hier die Formel

$$R_\lambda^* = R_{\bar{\lambda}}. \quad (41)$$

Ist λ reell, so folgt (41) aus der Selbstadjungiertheit von R_λ . Für komplexe λ ergibt sich (41) aus den Beziehungen

$$(A - \lambda E)^* = A - \bar{\lambda} E \quad \text{und} \quad [(A - \lambda E)^{-1}]^* = [(A - \bar{\lambda} E)^{-1}].$$

Sind λ und μ reguläre Werte, so läßt sich ebenso wie für beschränkte Operatoren die Formel

$$R_\mu - R_\lambda = (\mu - \lambda) R_\mu R_\lambda \quad (42)$$

beweisen [144].

190. Das Punktspektrum. Man sagt, der selbstadjungierte Operator A habe ein *Punktspektrum* oder ein *reines Punktspektrum*, wenn das normierte Orthogonalsystem seiner Eigenelemente in H abgeschlossen ist. Es sei x_k ($k=1, 2, \dots$) dieses beliebig numerierte System, und λ_k seien die entsprechenden Eigenwerte: $Ax_k = \lambda_k x_k$. Nach Voraussetzung läßt sich jedes Element $x \in H$ als Fourierreihe

$$x = \sum_{k=1}^{\infty} a_k x_k \quad (43)$$

darstellen.

Satz 1. *Das Element x gehört genau dann zu $D(A)$, wenn die Reihe*

$$\sum_{k=1}^{\infty} \lambda_k^2 |a_k|^2 \quad (44)$$

konvergiert. Ist diese Bedingung erfüllt, so gilt

$$Ax = \sum_{k=1}^{\infty} \lambda_k a_k x_k. \quad (45)$$

Beweis. Gehört x zu $D(A)$, so ist $(Ax, x_k) = (x, Ax_k) = (x, \lambda_k x_k) = \lambda_k a_k$ der Fourierkoeffizient des Elements Ax ; daraus ergibt sich die Konvergenz von (44). Wir nehmen nun an, daß (44) konvergiert, und beweisen $x \in D(A)$. Aus der Konvergenz von (44) folgt, daß wir ein Element

$$x' = \sum_{k=1}^{\infty} \lambda_k a_k x_k \quad (46)$$

konstruieren können. Bezeichnen wir die n -te Partialsumme von (43) mit y_n , so ist $y_n \in D(A)$, $y_n \rightarrow x$ und $Ay_n \rightarrow x'$; wegen der Abgeschlossenheit von A ist also $x \in D(A)$ und $Ax = x'$. Damit ist der Satz bewiesen.

Schon früher sahen wir, daß ein vollstetiger selbstadjungierter Operator ein reines Punktspektrum hat, wobei bei einer beliebigen Numerierung der Eigenwerte $\lambda_k \rightarrow 0$ für $k \rightarrow \infty$ gilt. Wir weisen jetzt auf den wichtigen Fall hin, daß ein selbstad-

jungierter Operator ein Punktspektrum hat. Es sei A ein selbstadjungierter Operator mit einem vollstetigen Inversen A^{-1} . Auf Grund der Definition des inversen Operators hat die Gleichung $A^{-1}x = 0$ außer der trivialen Lösung $x = 0$ keine andere Lösung mehr. Der vollstetige selbstadjungierte Operator A^{-1} hat ein Punktspektrum, und seine Eigenwerte μ_k ($k = 1, 2, \dots$) lassen sich nach ihren nicht wachsenden Absolutbeträgen anordnen: $|\mu_1| \geq |\mu_2| \geq \dots$, wobei wegen des oben Gesagten $\mu_k \neq 0$ für alle k ist. Sind x_k die entsprechenden Eigenelemente, die ein (in H abgeschlossenes) normiertes Orthogonalsystem bilden, so können wir $A^{-1}x_k = \mu_k x_k$ schreiben; daraus folgt sofort $Ax_k = \lambda_k x_k$ mit $\lambda_k = 1/\mu_k$. Unter Berücksichtigung des Inhalts von [136] gelangen wir zu dem folgenden Satz:

Satz 2. *Besitzt der selbstadjungierte Operator A ein vollstetiges Inverses A^{-1} , so hat A ein Punktspektrum, alle seine Eigenwerte sind von endlichem Rang, und in jedem endlichen Intervall liegen nur endlich viele Eigenwerte von A .*

Die Eigenwerte λ_k eines solchen Operators A lassen sich also nach nichtfallenden Absolutwerten anordnen: $|\lambda_1| \leq |\lambda_2| \leq \dots$ mit $|\lambda_k| \rightarrow \infty$ für $k \rightarrow \infty$.

Zu dem Spektrum eines Operators mit Punktspektrum brauchen aber nicht nur Eigenwerte zu gehören. Ist λ ein Verdichtungspunkt der λ_k , so liegt er unbedingt im Spektrum, denn die regulären Punkte bilden eine offene Menge. Wir zeigen nun, daß die übrigen Werte von λ nicht zum Spektrum gehören.

Satz 3. *Hat der selbstadjungierte Operator A ein Punktspektrum, so ist jedes reelle λ , das von den Eigenwerten verschieden und kein Verdichtungspunkt dieser Eigenwerte ist, ein regulärer Punkt von A .*

Nach Voraussetzung existiert eine solche positive Zahl m , daß $|\lambda_k - \lambda| \geq m$ für alle k ist. Ferner nehmen wir $x \in D(A)$. Aus (43) und (45) folgt

$$\|(A - \lambda E)x\|^2 = \sum_{k=1}^{\infty} |\lambda_k - \lambda|^2 |a_k|^2 \geq m^2 \sum_{k=1}^{\infty} |a_k|^2 = m^2 \|x\|^2;$$

der Punkt λ ist also regulär. Man sagt auch, der Operator A habe ein *reines Punktspektrum*.

Hat der selbstadjungierte Operator A keine Eigenwerte, so sagt man, A habe ein rein kontinuierliches Spektrum. In diesem Fall läßt sich jedes Element $x \in H$ nicht als Fourierreihe (43), sondern als Summe von Integralen darstellen [149]. Im folgenden werden wir zeigen, daß bei jedem selbstadjungierten Operator das Punktspektrum von dem rein kontinuierlichen Spektrum abgetrennt werden kann (analog zu [189], wo wir einen Eigenwert von dem übrigen Teil des Spektrums abtrennten). Dazu benötigen wir jedoch noch einige neue Begriffe, die auch in der Theorie der Operatoren von Interesse sind.

191. Invariante Unterräume und Reduzibilität eines Operators. Ehe wir den Begriff des invarianten Unterraums einführen, erklären wir zunächst die Vertauschbarkeit eines beschränkten, auf ganz H definierten Operators B mit einem nicht in ganz H definierten Operator A .

Definition. Ein beschränkter, überall definierter Operator B ist mit einem Operator A *vertauschbar*, wenn die folgenden Bedingungen erfüllt sind; a) Mit $x \in D(A)$ ist auch $Bx \in D(A)$; b) für $x \in D(A)$ ist $BAx = ABx$.

Ist auch A ein überall definierter beschränkter Operator, so entfällt die erste Bedingung, und wir erhalten die frühere Definition für die Vertauschbarkeit zweier Operatoren.

Satz 1. *Für die Vertauschbarkeit des Operators B mit einem selbstadjungierten Operator A ist notwendig, daß B mit der Resolvente R_λ für jedes reguläre λ vertauschbar ist, und hinreichend, daß B mit R_λ für mindestens ein reguläres λ vertauscht werden kann.*

Beweis. Die Vertauschbarkeit von B und R_λ (für reguläres λ) ist die früher definierte Vertauschbarkeit ($BR_\lambda x = R_\lambda Bx$ für $x \in H$). Es sei nun B mit A vertauschbar, λ ein regulärer Punkt und y ein beliebiges Element aus H . Dabei gehören $R_\lambda y$ und $BR_\lambda y$ zu $D(A)$, und außerdem gilt

$$ABR_\lambda y = BAR_\lambda y. \quad (47)$$

Nun ist $(A - \lambda E) R_\lambda y = y$, woraus $AR_\lambda y = (\lambda R_\lambda + E) y$ folgt. Die Gleichung (47) läßt sich dann in der Form

$$ABR_\lambda y = \lambda BR_\lambda y + By, \quad \text{also} \quad (A - \lambda E) BR_\lambda y = By \quad (48)$$

darstellen. Wenden wir auf beide Seiten von (48) den Operator R_λ an, so erhalten wir $BR_\lambda y = R_\lambda By$. Damit ist der erste Teil des Satzes bewiesen.

Wir nehmen zum weiteren Beweis an, daß ein reguläres λ existiert, so daß also $BR_\lambda y = R_\lambda By$ ist. Die Form der rechten Seite besagt, daß beide Seiten für jedes $y \in H$ zu $D(A)$ gehören. Durchläuft y ganz H , so durchläuft $x = R_\lambda y$ ganz $D(A)$. Aus der obigen Beziehung folgt also, daß neben x auch By zu $D(A)$ gehört. Wenden wir auf beide Seiten der Beziehung den Operator $A - \lambda E$ an, so erhalten wir (48) und (47), und (47) läßt sich in der Form $ABx = BAx$ für $x \in D(A)$ darstellen. Damit ist der Satz vollständig bewiesen.

Folgerung. *Ist B für jedes reguläre λ mit R_λ vertauschbar, so ist B auch für alle regulären λ mit R_λ vertauschbar.*

Wir kommen nun zur Definition des invarianten Unterraums.

Definition. Ein Unterraum L heißt *invariant* bezüglich eines Operators A , wenn mit $x \in D(A)$ und $x \in L$ auch Ax zu L gehört.

Ist L ein bezüglich A invarianter Unterraum und $D_L(A)$ die lineare Mannigfaltigkeit der Elemente x , die gleichzeitig zu $D(A)$ und zu L gehören, so induziert A in L einen Operator A_1 , der auf $D_L(A)$ definiert und dort gleich A ist. Den Unterraum L können wir als Hilbertschen Raum betrachten (der auch endlichdimensional sein kann). Wie wir im folgenden sehen werden, ist wesentlich, daß nicht nur L , sondern auch der komplementäre Unterraum $H \ominus L$ bezüglich A invariant ist und die Projektion eines beliebigen Elements $x \in D(A)$ auf L ebenfalls zu $D(A)$ gehört. Dies führt uns auf die folgende

Definition. Der Raum L *reduziert* den Operator A , wenn die beiden folgenden Bedingungen erfüllt sind:

- a) L und $M = H \ominus L$ sind bezüglich A invariante Unterräume;
- b) ist $x \in D(A)$, so gehört auch die Projektion von x in den Unterraum L zu $D(A)$.

Im folgenden bezeichnen wir mit P_K den Projektionsoperator in den Unterraum K . Es gilt

$$x = P_L x + P_M x. \quad (49)$$

Für $x \in D(A)$ folgt aus der Definition, daß $P_L x \in D(A)$ und somit $P_M x \in D(A)$ ist, d. h., reduziert L den Operator A , so wird A auch von M reduziert. Es seien A_1 und

A_2 die von A in L bzw. M induzierten Operatoren. Dann gilt für jedes $x \in D(A)$

$$Ax = A_1(P_L x) + A_2(P_M x); \quad (50)$$

wir haben auf diese Weise den Operator A in die beiden in L bzw. M wirkenden Operatoren A_1 und A_2 aufgeteilt.

Satz 2. *Der Unterraum L reduziert genau dann den Operator A , wenn P_L und A vertauschbar sind.*

Beweis. Wir zeigen zunächst die Notwendigkeit der Bedingung. Der Operator A werde von L reduziert. Dann folgt aus der Definition, daß mit x auch $P_L x$ zu $D(A)$ gehört. Wir brauchen also nur nachzuweisen, daß $P_L A x = A P_L x$ für $x \in D(A)$ gilt. Dies ist tatsächlich der Fall; denn wenden wir auf beide Seiten der Gleichung (50), in der $A_1(P_L x) \in L$ und $A_2(P_M x) \in M$ ist, den Operator P_L an, so ergibt sich

$$P_L A x = A_1(P_L x) = A P_L x.$$

Daß die Bedingung auch hinreichend ist, läßt sich folgendermaßen beweisen. Aus der Bedingung für die Vertauschbarkeit von P_L und A folgt, daß mit x auch $P_L x$ in $D(A)$ liegt. Es muß also noch festgestellt werden, daß $A x \in L$ für $x \in D(A)$, $x \in L$ gilt (und das Analoge für M). Das erste ergibt sich unmittelbar aus der Formel $P_L A x = A P_L x$, deren linke Seite offenbar zu L gehört; ihre rechte Seite ist wegen $x \in L$ gleich $A x$. Das gleiche gilt auch für M , denn wenn A mit P_L vertauschbar ist, so ist A auch mit P_M vertauschbar.

Wir gehen jetzt zu dem Fall über, daß A selbstadjungiert ist.

Satz 3. *Dafür, daß der bezüglich des selbstadjungierten Operators A invariante Unterraum L diesen Operator A reduziert, ist*

$$P_L x \in D(A) \quad \text{für} \quad x \in D(A)$$

eine hinreichende Bedingung.

Beweis. Wir müssen zeigen, daß aus den Bedingungen des Satzes $A y \in M$ für $y \in D(A)$, $y \in M$ folgt. Für ein solches y und für jedes $x \in D(A)$ gilt

$$(P_L A y, x) = (y, A P_L x).$$

Nun ist nach Voraussetzung $A P_L x \in L$, woraus sich $(y, A P_L x) = 0$ und somit $(P_L A y, x) = 0$ für jedes $x \in D(A)$ ergibt. Die lineare Mannigfaltigkeit $D(A)$ ist in H dicht; also gilt $P_L A y = 0$, d. h. $A y \in M$, was zu beweisen war.

Wir bezeichnen nun wieder mit $D_L(A)$ die Projektion von $D(A)$ in den Raum L , d. h. die lineare Mannigfaltigkeit derjenigen Elemente von L , auf denen der Operator A definiert ist, und mit A_L den durch A in L induzierten Operator.

Satz 4. *Reduziert der Unterraum L den selbstadjungierten Operator A , so ist $D_L(A)$ in L dicht, und A_L ist ein in L selbstadjungierter Operator.*

Beweis. Es sei y ein gegebenes Element aus L , und $\varepsilon > 0$ eine gegebene Zahl. Wir müssen zeigen, daß es ein Element $x \in D_L(A)$ gibt, für welches $\|y - x\| \leq \varepsilon$ gilt. Da $D(A)$ in H dicht ist, gibt es ein Element $z \in D(A)$ mit $\|y - z\| \leq \varepsilon$. Es ist sogar $\|P_L y - P_L z\| \leq \varepsilon$. Wegen $P_L y = y$ und $P_L z \in D_L(A)$ ist die erste Behauptung bewiesen.

Wir müssen noch zeigen: Gilt für jedes $x \in D_L(A)$ die Beziehung

$$(A_L x, y) = (x, y^*), \quad (51)$$

wobei y und y^* zu L gehören, so ist $y \in D_L(A)$ und $y^* = A_L y$. Um dies zu beweisen, setzen wir $x = P_L z$ mit z beliebig aus $D(A)$ und erhalten so $(AP_L z, y) = (P_L z, y^*)$ oder wegen Satz 2

$$(P_L A z, y) = (P_L z, y^*) .$$

Daraus folgt $(Az, P_L y) = (z, P_L y^*)$ oder $(Az, y) = (z, y^*)$, denn es ist $y, y^* \in L$. Aus der letzten Gleichung ergibt sich wegen der Selbstadjungiertheit des Operators A

$$y \in D_L(A) \quad \text{und} \quad y^* = Ay = A_L y .$$

Damit ist der Satz vollständig bewiesen.

Dieser Satz wurde schon in [189] benutzt.

Es seien nun L_k ($k=1, 2, \dots$) paarweise orthogonale Unterräume, und L sei ihre orthogonale Summe [139]:

$$L = L_1 \oplus L_2 \oplus \dots .$$

Satz 5. Reduzieren die paarweise orthogonalen Unterräume L_k ($k=1, 2, \dots$) den Operator A , so reduziert ihre orthogonale Summe ebenfalls den Operator A .

Beweis. Wir betrachten den Fall unendlich vieler Summanden und müssen zeigen, daß die Operatoren P_L und A vertauschbar sind. Es sei Q_n derjenige Projektionsoperator, der gleich der Summe der ersten n der P_{L_k} ist. Liegt x in $D(A)$, so gehört $Q_n x$ zu L_k den Operator A reduziert, zu $D(A)$, und es ist $A Q_n x = Q_n A x$. Wegen $Q_n x \Rightarrow P_L x$ ist $A Q_n x = Q_n A x \Rightarrow P_L A x$; daraus folgt wegen der Abgeschlossenheit von A

$$P_L x \in D(A) \quad \text{und} \quad A P_L x = P_L A x ,$$

was zu beweisen war.

Es seien nun A ein selbstadjungierter Operator, λ_k seine verschiedenen Eigenwerte und L_k die entsprechenden Unterräume der Eigenelemente (zuzüglich des Nullelements). Die Anzahl dieser Unterräume kann auch endlich sein. Der Operator A wird offenbar von jedem L_k reduziert. Stimmt die orthogonale Summe L der L_k mit H überein, so hat A ein Punktspektrum. Ist L von H verschieden, so gilt die orthogonale Zerlegung von H :

$$H = L \oplus M \quad \text{und} \quad x = P_L x + P_M x \quad (x \in H) , \quad (52)$$

wobei L und M den Operator A reduzieren. Dieser Operator induziert in L bzw. M zwei Operatoren A_1 und A_2 derart, daß $Ax = A_1(P_L x) + A_2(P_M x)$ ist; dabei ist $A_1(P_L x) = A(P_L x)$ und $A_2(P_M x) = A(P_M x)$ für $x \in D(A)$. Der Operator A_1 hat ein Punktspektrum in L , der Operator A_2 ein rein kontinuierliches Spektrum in M .

192. Zerlegung der Einheit. Das Stieltjessche Integral. Wir gehen jetzt zum Studium der Theorie der Spektralschar (Zerlegung der Einheit) für selbstadjungierte Operatoren über. Es ist in vielem dem Fall beschränkter selbstadjungierter Operatoren analog. Wir werden die Stellen, an denen die Nichtbeschränktheit eines Operators herangezogen werden muß, besonders hervorheben.

Unter der *Zerlegung der Einheit* verstehen wir die Schar $\mathcal{E}_\lambda$ von Projektionsoperatoren, die auf dem Intervall $(-\infty, \infty)$ von dem reellen Parameter λ abhängt und die folgenden Bedingungen erfüllt: a) Für $\mu > \lambda$ gilt $\mathcal{E}_\mu \supseteq \mathcal{E}_\lambda$; b) $\mathcal{E}_\lambda$ strebt für $\lambda \rightarrow -\infty$ gegen den Nulloperator und für $\lambda \rightarrow \infty$ gegen E ; c) $\mathcal{E}_\lambda$ ist von rechts stetig, d. h., es ist $\mathcal{E}_\lambda \rightarrow \mathcal{E}_{\lambda'}$ für $\lambda = \lambda' + 0$.

Ferner ist $\mathbb{E}_\lambda \mathbb{E}_\mu = \mathbb{E}_\mu \mathbb{E}_\lambda = \mathbb{E}_\lambda$ für $\lambda < \mu$, und ist Δ das Intervall $[\alpha, \beta]$, so gilt, wenn wir $\Delta \mathbb{E}_\lambda = \mathbb{E}_\beta - \mathbb{E}_\alpha$ setzen, wie früher [141]

$$\Delta' \mathbb{E}_\lambda x \perp \Delta'' \mathbb{E}_\lambda x \quad (53)$$

(Δ' und Δ'' haben keine gemeinsamen inneren Punkte) und

$$\Delta' \mathbb{E}_\lambda \cdot \Delta'' \mathbb{E}_\lambda = \Delta_0 \mathbb{E}_\lambda \quad (54)$$

(Δ_0 ist der gemeinsame Teil von Δ' und Δ'').

Es sei δ eine Zerlegung des Intervalls $(-\infty, \infty)$:

$$\dots < \lambda_{-2} < \lambda_{-1} < \lambda_0 < \lambda_1 < \lambda_2 < \dots,$$

wobei die obere Grenze ω_δ der Differenzen $\lambda_k - \lambda_{k-1}$ ($k=0, \pm 1, \pm 2, \dots$) endlich sei. Wir bilden die unendliche Summe

$$\sum_{k=-\infty}^{+\infty} \lambda'_k \Delta_k \mathbb{E}_\lambda x = \sum_{k=-\infty}^{+\infty} \lambda'_k (\mathbb{E}_{\lambda_k} - \mathbb{E}_{\lambda_{k-1}}) x; \quad (55)$$

dabei ist $\lambda_{k-1} \leq \lambda'_k \leq \lambda_k$ und x ein Element aus H . Wegen (53) besteht diese Summe aus paarweise orthogonalen Elementen, und sie konvergiert dann und nur dann, wenn die Reihe

$$\sum_{k=-\infty}^{\infty} \lambda_k'^2 \|\Delta_k \mathbb{E}_\lambda x\|^2 = \sum_{k=-\infty}^{\infty} \lambda_k'^2 \Delta_k \|\mathbb{E}_\lambda x\|^2 \quad (56)$$

konvergiert [121].

Diese Reihe stellt die Summe σ_δ [3] für das Integral

$$\int_{-\infty}^{\infty} \lambda^2 d(\mathbb{E}_\lambda x, x) = \int_{-\infty}^{\infty} \lambda^2 d\|\mathbb{E}_\lambda x\|^2 \quad (57)$$

dar, und bekanntlich [5] konvergiert die Reihe (56) für jede Zerlegung des Intervalls und jedes λ'_k , wenn sie für eine gewisse Zerlegung δ und ein gewisses λ'_k konvergiert. Dabei ist der Grenzwert der Summen (56) für $\omega_\delta \rightarrow 0$ gleich dem Integral (57), und die Existenz dieses uneigentlichen Integrals ist der Konvergenz von (56) äquivalent. Somit dürfen wir die Summen (55) für solche x betrachten, für die die Reihen (56) konvergieren oder, was dasselbe ist, das Integral (57) einen endlichen Wert hat. Die Menge dieser x bezeichnen wir mit l . Wegen

$$\|\Delta_k \mathbb{E}_\lambda (x+y)\|^2 \leq (\|\Delta_k \mathbb{E}_\lambda x\| + \|\Delta_k \mathbb{E}_\lambda y\|)^2 \leq 2\|\Delta_k \mathbb{E}_\lambda x\|^2 + 2\|\Delta_k \mathbb{E}_\lambda y\|^2$$

können wir behaupten, daß mit $x \in l$ und $y \in l$ auch $x+y \in l$ ist. Außerdem ist offenbar $ax \in l$, wenn $x \in l$ und a eine komplexe Zahl ist, d. h., l ist eine lineare Mannigfaltigkeit. Gehört x zu dem Unterraum, in den der Operator $\mathbb{E}_\beta - \mathbb{E}_\alpha$ projiziert, so sind die Summanden von (56), für welche $\lambda_{k-1} > \beta$ oder $\lambda_k < \alpha$ ist, gleich 0; diese x liegen also in l . Wegen $\mathbb{E}_\beta - \mathbb{E}_\alpha \rightarrow E$ für $\alpha \rightarrow -\infty$ und $\beta \rightarrow \infty$ ist die lineare Mannigfaltigkeit l in H dicht. Ist ferner $x \in l$, so können wir wie in [141] zeigen, daß die Summen (55) für $\omega_\delta \rightarrow 0$ einen wohlbestimmten Grenzwert im Sinne der Konvergenz in H haben. Dieser Grenzwert läßt sich auf natürliche Weise als Stieltjessches Integral darstellen, und er definiert auf der linearen Mannigfaltigkeit l einen gewissen distributiven Operator A :

$$Ax = \int_{-\infty}^{\infty} \lambda d\mathbb{E}_\lambda x. \quad (58)$$

Die lineare Mannigfaltigkeit l wollen wir wie stets mit $D(A)$ bezeichnen. Sie besteht aus denjenigen Elementen x , für welche (57) einen endlichen Wert hat. Multi-

plizieren wir die Summe (55) skalar mit sich selbst, beachten wir (54) und gehen wir zur Grenze über, so erhalten wir

$$\int_{-\infty}^{\infty} \lambda^2 d\|\mathcal{E}_\lambda x\|^2 = \|Ax\|^2 \quad (x \in D(A)). \quad (59)$$

Dieses Integral ist der Grenzwert der Summen (56) für $\omega_\delta \rightarrow 0$, oder wir können es als uneigentliches Integral mit unendlichen Grenzen auffassen. Multiplizieren wir (56) skalar mit einem beliebigen Element y und gehen wir zur Grenze über, so finden wir das bilineare Funktional

$$(Ax, y) = \int_{-\infty}^{\infty} \lambda d(\mathcal{E}_\lambda x, y) \quad (x \in D(A), y \in H). \quad (60)$$

Es ist der Grenzwert der entsprechenden Summen σ_δ für $\omega_\delta \rightarrow 0$. Ersetzen wir y durch $(\mathcal{E}_\beta - \mathcal{E}_\alpha)y$, so ergibt sich wegen (54)

$$(Ax, (\mathcal{E}_\beta - \mathcal{E}_\alpha)y) = \int_{\alpha}^{\beta} \lambda d(\mathcal{E}_\lambda x, y)$$

oder, wenn wir für $\alpha \rightarrow -\infty$, $\beta \rightarrow \infty$ zur Grenze übergehen,

$$(Ax, y) = \lim_{\substack{\alpha \rightarrow -\infty \\ \beta \rightarrow \infty}} \int_{\alpha}^{\beta} \lambda d(\mathcal{E}_\lambda x, y). \quad (61)$$

Das Integral (60) läßt sich also als gewöhnliches uneigentliches Stieltjessches Integral auffassen, wobei $(\mathcal{E}_\lambda x, y)$ eine Funktion von beschränkter Variation ist.

Das unendliche Intervall $(-\infty, +\infty)$ hat bezüglich der nichtfallenden Funktion $\|\mathcal{E}_\lambda x\|^2$ ein endliches Maß, und das Integral (57) können wir auch als Lebesgue-Stieltjessches Integral einer nicht beschränkten, nichtnegativen Funktion λ^2 über eine Menge von endlichem Maß erklären [50].

Es sei nun x ein beliebiges Element aus H . Dann gehört $(\mathcal{E}_\beta - \mathcal{E}_\alpha)x$ zu demjenigen Unterraum, in den der Operator $\mathcal{E}_\beta - \mathcal{E}_\alpha$ projiziert. Daraus folgt, wie wir oben sahen, $(\mathcal{E}_\beta - \mathcal{E}_\alpha)x \in D(A)$ für jedes Element x . Dies läßt sich aber noch nicht von dem Element $\mathcal{E}_\mu x$ behaupten. Ist $x \in D(A)$, d. h. konvergiert die Reihe (56), so konvergiert sie auch, wenn wir x durch $\mathcal{E}_\mu x$ ersetzen, denn es ist

$$\|A_k \mathcal{E}_\lambda (\mathcal{E}_\mu x)\| = \|\mathcal{E}_\mu A_k \mathcal{E}_\lambda x\| \leq \|A_k \mathcal{E}_\lambda x\|.$$

Mit x gehört also auch $\mathcal{E}_\mu x$ für jedes μ zu $D(A)$. Nehmen wir $x \in D(A)$ an, ersetzen wir in (55) x durch $\mathcal{E}_\mu x$ und wählen wir μ als einen der Teilpunkte, etwa $\mu = \lambda_p$, so gehen alle Summanden mit $k > p$ in das Nullelement über und die Summanden mit $k \leq p$ bleiben unverändert. Nach Grenzübergang erhalten wir also

$$A\mathcal{E}_\mu x = \int_{-\infty}^{\mu} \lambda d\mathcal{E}_\lambda x \quad (x \in D(A)). \quad (62)$$

Dieses Integral ist der Grenzwert von Summen der Form (55) bei der Zerlegung von $(-\infty, \mu)$ in Teilintervalle. Wenden wir andererseits auf die Summe (55) den beschränkten und daher stetigen Operator $\mathcal{E}_\mu$ an, so bleibt das oben über die Summanden Gesagte gültig, und für $\omega_\delta \rightarrow 0$ ergibt sich unter Berücksichtigung der Stetigkeit von $\mathcal{E}_\mu$

$$\mathcal{E}_\mu Ax = \int_{-\infty}^{\mu} \lambda d\mathcal{E}_\lambda x \quad (x \in D(A)). \quad (62_1)$$

Ein Vergleich mit (62) zeigt

$$\mathfrak{E}_\mu Ax = A\mathfrak{E}_\mu x \quad (x \in D(A)). \quad (63)$$

Genau so erhalten wir für jedes x

$$A(\mathfrak{E}_\beta - \mathfrak{E}_\alpha)x = \int_\alpha^\beta \lambda d\mathfrak{E}_\lambda x \quad (x \in H). \quad (64)$$

Außerdem folgt aus (62) für $x \in D(A)$

$$(\mathfrak{E}_\beta - \mathfrak{E}_\alpha)Ax = \int_\alpha^\beta \lambda d\mathfrak{E}_\lambda x \quad (x \in D(A)) \quad (64_1)$$

oder, wenn α gegen $-\infty$ und β gegen ∞ strebt,

$$\int_\alpha^\beta \lambda d\mathfrak{E}_\lambda x \Rightarrow Ax; \quad (65)$$

d. h., auch das Integral (58) können wir wie (57) als uneigentliches Integral erklären. Aus den vorhergehenden Formeln finden wir

$$(A\mathfrak{E}_\mu x, y) = (\mathfrak{E}_\mu Ax, y) = \int_{-\infty}^\mu \lambda d(\mathfrak{E}_\lambda x, y) \quad (x \in D(A), y \in H), \quad (66_1)$$

$$(A(\mathfrak{E}_\beta - \mathfrak{E}_\alpha)x, y) = \int_\alpha^\beta \lambda d(\mathfrak{E}_\lambda x, y) \quad (x \in H, y \in H). \quad (66_2)$$

Liegt nicht nur x , sondern auch y in $D(A)$, so gilt, wenn wir in (58) y statt x setzen und von links skalar mit x multiplizieren,

$$(x, Ay) = \int_{-\infty}^\infty \lambda d(x, \mathfrak{E}_\lambda y).$$

Hieraus und aus (60) finden wir wegen $(x, \mathfrak{E}_\lambda y) = (\mathfrak{E}_\lambda x, y)$ die Beziehung

$$(Ax, y) = (x, Ay),$$

d. h., A ist ein symmetrischer Operator.

Dafür, daß A selbstadjungiert ist, brauchen wir nur zu zeigen, daß $z \in D(A)$ aus $z \in D(A^*)$ folgt. Es sei δ eine Zerlegung des Intervalls $(-\infty, \infty)$ und P_j der durch den Projektionsoperator $\mathfrak{E}_{\lambda_j} - \mathfrak{E}_{\lambda_{j-1}}$ definierte Unterraum. Ist $x \in P_j$, so sind in (55) und (56) alle Summanden mit $k \geq j+1$ und $k \leq -j$ gleich 0, x gehört zu $D(A)$, und für $j+1 > k > -j$ sind die durch die Projektionsoperatoren $A_k \mathfrak{E}_\lambda$ definierten Unterräume in P_j enthalten, so daß die Summe (55) und ihr Grenzwert Ax in P_j liegen. Daraus folgt $Ax \in D(A)$ und $A^2x \in P_j$. Wir setzen nun $z \in D(A^*)$ und beweisen $z \in D(A)$. Ist z_j die Projektion von z in den Unterraum P_j , so gehört z_j zu $D(A)$ und somit auch zu $D(A^*)$. Dann ist auch das Element $z - z_j$ in $D(A^*)$ enthalten und zu P_j orthogonal. Nach Definition von A^* haben wir

$$(A^2z_j, z - z_j) = (Az_j, A^*(z - z_j)).$$

Da A^2z_j in P_j liegt und $z - z_j$ zu P_j orthogonal ist, folgt

$$(Az_j, A^*(z - z_j)) = 0.$$

Mit Hilfe der offenbar gültigen Gleichung

$$\begin{aligned} (A^*z, A^*z) &= (A^*(z - z_j), A^*(z - z_j)) + (A^*z_j, A^*z_j) \\ &\quad + (A^*(z - z_j), A^*z_j) + (A^*z_j, A^*(z - z_j)), \end{aligned}$$

der vorhergehenden Formel und der Gleichung $A^*z_j = Az_j$ erhalten wir

$$\|A^*z\|^2 = \|A^*(z - z_j)\|^2 + \|Az_j\|^2, \quad (67)$$

also

$$\|Az_j\|^2 \leq \|A^*z\|^2.$$

Wir untersuchen nun die Summe (56) für $x = z_j$. In ihr sind alle Summanden mit $k \geq j+1$ und $k \leq -j$ gleich 0, und für die übrigen Summanden gilt wegen (54)

$$\Delta_k \mathfrak{E}_{\lambda} z_j = (\mathfrak{E}_{\lambda_k} - \mathfrak{E}_{\lambda_{k-1}}) (\mathfrak{E}_{\lambda_j} - \mathfrak{E}_{\lambda_{-j}}) z = (\mathfrak{E}_{\lambda_k} - \mathfrak{E}_{\lambda_{k-1}}) z = \Delta_k \mathfrak{E}_{\lambda} z.$$

Somit liefert die Summe (56) nach Grenzübergang wegen (59) die Beziehung

$$\|Az_j\|^2 = \int_{\lambda_{-j}}^{\lambda_j} \lambda^2 d\|\mathfrak{E}_{\lambda} z\|^2$$

oder wegen (67)

$$\int_{\lambda_{-j}}^{\lambda_j} \lambda^2 d\|\mathfrak{E}_{\lambda} z\|^2 \leq \|A^*z\|^2.$$

Lassen wir j gegen ∞ streben, so hat das Integral (57) für $x = z$ einen endlichen Wert, d. h., z gehört zu $D(A)$. Der Operator A ist also selbstadjungiert.

Diese Überlegungen führen uns auf den folgenden Satz.

Satz 1. *Jeder Zerlegung der Einheit $\mathfrak{E}_{\lambda}$ entspricht ein selbstadjungierter Operator A , der für diejenigen x definiert ist, für die das Integral (57) endliche Werte annimmt. Der Operator A läßt sich als Grenzwert der Summen (55) oder, was das gleiche ist, als Integral (58) definieren. Das ihm entsprechende bilineare Funktional wird durch (60) gegeben.*

Wir können auch die Umkehrung dieses Satzes beweisen.

Satz 2. *Zu jedem gegebenen selbstadjungierten Operator A existiert eine solche Zerlegung der Einheit $\mathfrak{E}_{\lambda}$, daß A durch die Formel (58) ausgedrückt wird.*

Den Beweis dieses Satzes führen wir später. Wir werden später auch eine Formel angeben, die $\mathfrak{E}_{\lambda}$ durch A definiert (vgl. [144]); aus ihr folgt, daß verschiedenen $\mathfrak{E}_{\lambda}$ verschiedene A entsprechen. Der Operator $\mathfrak{E}_{\lambda}$ heißt die Spektralschar des selbstadjungierten Operators A . Genau wie in [144] läßt sich zeigen, daß die regulären Punkte $\lambda = \mu$ des Spektrums durch die Existenz eines Intervalls charakterisiert werden, in dem $\mathfrak{E}_{\lambda}$ konstant ist und welches μ im Innern enthält. Dabei muß beachtet werden, daß $(\mathfrak{E}_{\beta} - \mathfrak{E}_{\alpha})x$ für jedes $x \in H$ und beliebige endliche α und β zu $D(A)$ gehört. Die Eigenwerte $\lambda = \nu$ sind dadurch charakterisiert, daß $\mathfrak{E}_{\lambda}$ für $\lambda = \nu$ Sprungstellen hat, wobei der Operator $\mathfrak{E}_{\nu} - \mathfrak{E}_{\nu-0}$ in den Unterraum der entsprechenden Eigenelemente (zuzüglich des Nullelements) projiziert [145].

Wir nehmen an, der selbstadjungierte Operator sei halbbeschränkt, und m_A sei seine untere Grenze:

$$m_A = \inf (Ax, x) \quad \text{für } x \in D(A) \quad \text{und} \quad \|x\| = 1.$$

Für jedes reelle λ gilt

$$((A - \lambda E)x, x) \geq (m_A - \lambda)(x, x) \quad (x \in D(A));$$

daraus folgt

$$\|(A - \lambda E)x\| \cdot \|x\| \geq (m_A - \lambda) \|x\|^2$$

oder

$$\|(A - \lambda E)x\| \equiv (m_A - \lambda) \|x\|.$$

Somit ergibt sich, daß *alle Werte von λ , die der Bedingung $\lambda < m_A$ genügen, reguläre Punkte von A sind*. Wir können nun zeigen, daß $\lambda = m_A$ ein Punkt des Spektrums von A ist. Ist dies nicht der Fall, so existiert ein $m_1 > m_A$ derart, daß alle Werte $\lambda \leq m_1$ reguläre Punkte von A sind, und $\mathcal{E}_\lambda$ ist für $\lambda \leq m_1$ der Nulloperator, so daß

$$(Ax, x) = \int_{m_1}^{\lambda} \lambda d(\mathcal{E}_\lambda x, x) \quad (x \in D(A))$$

gilt. Hieraus folgt $(Ax, x) \equiv m_1(x, x)$ für $x \in D(A)$; dies widerspricht jedoch der Definition von m_A . Der Wert $\lambda = m_A$ wird auch *untere Schranke des Spektrums von A* genannt.

Der Operator $\mathcal{E}_\lambda$ heißt *Spektralschar* des selbstadjungierten Operators A .

Wir gehen jetzt zu einer kurzen Wiederholung der Eigenschaften allgemeiner selbstadjungierten Operatoren über; sie sind denen beschränkter selbstadjungierter Operatoren vollkommen analog.

193. Stetige Funktionen eines selbstadjungierten Operators. Es sei $f(\lambda)$ eine beschränkte und auf dem Intervall $(-\infty, \infty)$ gleichmäßig stetige Funktion (z. B. ist $f(\lambda)$ stetig in dem abgeschlossenen Intervall $[-\infty, \infty]$). Wir bilden die der Summe (55) analoge Summe

$$\sum_{k=-\infty}^{\infty} f(\lambda'_k) \Delta_k \mathcal{E}_\lambda x \quad (68)$$

mit einem beliebigen Element $x \in H$. Diese aus paarweise orthogonalen Elementen bestehende Reihe konvergiert für jedes x ; denn nach Voraussetzung ist $|f(\lambda)| \leq k$, wobei k eine bestimmte Zahl ist, und

$$\left\| \sum_{k=-\infty}^{\infty} f(\lambda'_k) \Delta_k \mathcal{E}_\lambda x \right\|^2 = \sum_{k=-\infty}^{\infty} |f(\lambda'_k)|^2 \|\Delta_k \mathcal{E}_\lambda x\|^2 \leq k^2 \sum_{k=-\infty}^{\infty} \|\Delta_k \mathcal{E}_\lambda x\|^2 = k^2 \|x\|^2, \quad (69)$$

so daß die zu (56) analoge Reihe konvergiert. Genau wie in [141] können wir zeigen, daß die Summe (68) für jedes $x \in H$ einen bestimmten Grenzwert für $\omega_\delta \rightarrow 0$ hat. Dieser Grenzwert ergibt einen auf ganz H definierten distributiven Operator $f(A)$. Aus (69) folgt unmittelbar $\|f(A)x\| \leq k\|x\|$, d. h., der Operator $f(A)$ ist beschränkt. Die Summe (68) läßt sich natürlich in Form eines Stieltjesschen Integrals schreiben:

$$f(A)x = \int_{-\infty}^{\infty} f(\lambda) d\mathcal{E}_\lambda x \quad (x \in H), \quad (70)$$

$$(f(A)x, y) = \int_{-\infty}^{\infty} f(\lambda) d(\mathcal{E}_\lambda x, y) \quad (x \in H, y \in H). \quad (71)$$

Das letzte Integral können wir als gewöhnliches Stieltjessches Integral auffassen, so wie wir es in [4] definiert haben.

Analog den Formeln (62) und (66) gilt

$$\mathcal{E}_\mu f(A)x = f(A)\mathcal{E}_\mu x = \int_{-\infty}^{\mu} f(\lambda) d\mathcal{E}_\lambda x \quad (72)$$

und

$$(\mathcal{E}_\mu f(A)x, y) = (f(A)\mathcal{E}_\mu x, y) = \int_{-\infty}^{\mu} f(\lambda) d(\mathcal{E}_\lambda x, y) \quad (x \in H, y \in H). \quad (73)$$

Wir könnten die obige Definition von $f(A)$ auch für eine beliebige beschränkte und auf dem Intervall $(-\infty, \infty)$ stetige Funktion benutzen, wenn wir nicht gleichmäßige Stetigkeit fordern. Im folgenden werden wir beweisen, daß es möglich ist, den Begriff der Funktion eines selbstadjungierten Operators auf eine umfassendere Klasse von Funktionen $f(\lambda)$ zu übertragen.

Für den Operator $A - lE$, wobei l eine beliebige Zahl ist, gelten offenbar die Formeln

$$\left. \begin{aligned} (A - lE)x &= \int_{-\infty}^{\infty} (\lambda - l) d\mathcal{E}_{\lambda} x, \\ ((A - lE)x, y) &= \int_{-\infty}^{\infty} (\lambda - l) d(\mathcal{E}_{\lambda} x, y) \end{aligned} \right\} (x \in D(A), y \in H). \quad (74)$$

194. Die Resolvente. Wir wollen nun die Resolvente durch die Spektralschar ausdrücken.

Ist l nicht reell, so ist $1/(\lambda - l)$ eine Funktion von λ , die im abgeschlossenen Intervall $[-\infty, \infty]$ stetig ist, und wir können den beschränkten Operator

$$R_l x = \int_{-\infty}^{\infty} \frac{1}{\lambda - l} d\mathcal{E}_{\lambda} x \quad (75)$$

bilden und zeigen, daß er alle Eigenschaften einer Resolvente für $\lambda = l$ besitzt, was die Bezeichnung R_l rechtfertigt. Für jedes x gehört das Element $(\mathcal{E}_{\beta} - \mathcal{E}_{\alpha}) R_l x$ zu $D(A)$, und es gilt wegen (66₂)

$$((A - lE)(\mathcal{E}_{\beta} - \mathcal{E}_{\alpha}) R_l x, y) = \int_{\alpha}^{\beta} (\lambda - l) d(\mathcal{E}_{\lambda} R_l x, y).$$

Andererseits ist wegen (73) und (75)

$$(\mathcal{E}_{\lambda} R_l x, y) = \int_{-\infty}^{\lambda} \frac{1}{\mu - l} d(\mathcal{E}_{\mu} x, y).$$

Setzen wir dies in die vorhergehende Formel ein und benutzen wir die Eigenschaft des Stieltjesschen Integrals [9], so erhalten wir

$$((A - lE)(\mathcal{E}_{\beta} - \mathcal{E}_{\alpha}) R_l x, y) = \int_{\alpha}^{\beta} d(\mathcal{E}_{\lambda} x, y) = ((\mathcal{E}_{\beta} - \mathcal{E}_{\alpha}) x, y)$$

oder, da y beliebig ist,

$$(A - lE)(\mathcal{E}_{\beta} - \mathcal{E}_{\alpha}) R_l x = (\mathcal{E}_{\beta} - \mathcal{E}_{\alpha}) x. \quad (76)$$

Wir konstruieren zwei Folgen von Zahlen α_n und β_n mit $\alpha_n \rightarrow -\infty$ und $\beta_n \rightarrow \infty$ und eine Folge von Elementen $y_n = (\mathcal{E}_{\beta_n} - \mathcal{E}_{\alpha_n}) R_l x$. Dabei ist $y_n \in D(A)$, $y_n \Rightarrow R_l x$ und wegen (76)

$$(A - lE) y_n \Rightarrow x.$$

Daraus folgt wegen der Abgeschlossenheit des Operators A , daß für jedes x

$$R_l x \in D(A)$$

und $(A - lE) R_l x = x$ ist.

Es bleibt noch zu zeigen, daß $R_l(A - lE)x = x$ für $x \in D(A)$ ist. Dies folgt sofort

aus den Formeln

$$(R_l(A - lE)x, y) = \int_{-\infty}^{\infty} \frac{1}{\lambda - l} d(\mathfrak{E}_\lambda(A - l\mathfrak{E})x, y)$$

und

$$(\mathfrak{E}_\lambda(A - lE)x, y) = \int_{-\infty}^{\lambda} (\mu - l) d(\mathfrak{E}_\mu x, y),$$

die eine Folgerung aus (71) und (66₁) sind.

Ferner bleibt die Formel

$$\frac{1}{2} [(\mathfrak{E}_{\lambda-0}x, y) + (\mathfrak{E}_\lambda x, y)] = \lim_{\tau \rightarrow 0} \int_{-\infty}^{\lambda} ((R_{\sigma+i\tau} - R_{\sigma-i\tau})x, y) d\sigma \quad (77)$$

gültig, mit deren Hilfe die Spektralschar durch die Resolvente ausgedrückt wird.

Einem selbstadjungierten Operator entspricht eine bestimmte Spektralschar $\mathfrak{E}_\lambda$, und der Operator ist genau dann beschränkt, wenn $\mathfrak{E}_\lambda$ nur auf einem endlichen Intervall variabel ist.

In [191] bewiesen wir, daß ein beschränkter Operator B genau dann mit dem selbstadjungierten Operator A vertauschbar ist, wenn für jedes l , für das die Resolvente existiert, die Bedingung

$$BR_l = R_l B \quad (78)$$

erfüllt ist. Nun beweisen wir den folgenden

Satz. Der Operator B ist genau dann mit A vertauschbar, wenn für jedes reelle λ die Bedingung

$$B\mathfrak{E}_\lambda = \mathfrak{E}_\lambda B \quad (79)$$

gilt.

Beweis. Es genügt zu zeigen, daß die Bedingungen (78) und (79) äquivalent sind. Auf Grund von (75) gilt für alle Elemente x und y

$$(BR_l x, y) = (R_l x, B^* y) = \int_{-\infty}^{\infty} \frac{1}{\lambda - l} d(\mathfrak{E}_\lambda x, B^* y),$$

d. h.

$$(BR_l x, y) = \int_{-\infty}^{\infty} \frac{1}{\lambda - l} d(B\mathfrak{E}_\lambda x, y), \quad (R_l Bx, y) = \int_{-\infty}^{\infty} \frac{1}{\lambda - l} d(\mathfrak{E}_\lambda Bx, y). \quad (80)$$

Ist (79) erfüllt, so stimmen die rechten und demnach auch die linken Seiten der Gleichungen (80) überein, und die Bedingung (78) ist erfüllt, denn x und y sind beliebig. Ist dagegen (78) erfüllt, so können sich $(B\mathfrak{E}_\lambda x, y)$ und $(\mathfrak{E}_\lambda Bx, y)$ wegen der Eindeutigkeit der Umkehrung des Cauchy-Stieltjesschen Integrals [29] nur um eine additive Konstante (auch in den Unstetigkeitspunkten) unterscheiden. Diese beiden Funktionen streben aber für $\lambda \rightarrow -\infty$ gegen 0 und sind in den Unstetigkeitsstellen von rechts stetig; für beliebige x und y gilt dann

$$(B\mathfrak{E}_\lambda x, y) = (\mathfrak{E}_\lambda Bx, y),$$

d. h., (79) ist erfüllt. Damit ist der Satz bewiesen.

Auf Grund der Resultate aus [156] ist $\mathcal{E}_\mu$ mit A für jedes μ vertauschbar. Dies folgt auch aus dem eben bewiesenen Satz. Aus ihm und aus (70) ergibt sich, daß sich $f(A)$ mit A vertauschen läßt.

195. Eigenwerte. Wie wir uns erinnern, ist $\lambda = \lambda'$ genau dann ein Eigenwert von A , wenn λ' eine Unstetigkeitsstelle von $\mathcal{E}_\lambda$ ist. Dann ist $\mathcal{E}_\lambda - \mathcal{E}_{\lambda-0}$ ein Projektionsoperator in den Unterraum der entsprechenden Eigenelemente (zuzüglich des Nullelements). Wir nehmen H wie üblich als separabel an und behaupten, daß die Anzahl der Eigenwerte, falls sie existieren, endlich oder abzählbar unendlich ist. Der Rang der Eigenwerte ist wie früher definiert, und wir können annehmen, daß die Gesamtheit aller Eigenelemente $x_1, x_2, \dots$ ein normiertes Orthogonalsystem ist. Es seien λ_k die Unstetigkeitsstellen von $\mathcal{E}_\lambda$, L_k die Unterräume der entsprechenden Eigenelemente und $P_{L_k} = \mathcal{E}_{\lambda_k} - \mathcal{E}_{\lambda_k-0}$ die in diese Unterräume projizierenden Operatoren. Wir bilden die orthogonale Summe

$$H' = L_1 \oplus L_2 \oplus L_3 \oplus \dots \quad (81)$$

Der Unterraum H' reduziert A , und der durch A in H' induzierte Operator A' ist selbstadjungiert und besitzt ein reines Punktspektrum. Hat schon in H der Operator ein reines Punktspektrum, so ist

$$\mathcal{E}_\lambda = \sum_{\lambda_k \leq \lambda} (\mathcal{E}_{\lambda_k} - \mathcal{E}_{\lambda_k-0}). \quad (82)$$

Es sei nun $x_1, x_2, \dots$ ein abgeschlossenes normiertes Orthogonalsystem in H und $\mu_1, \mu_2, \dots$ eine Folge reeller Zahlen (unter denen auch gleiche auftreten können), wobei jeweils μ_k und x_k zugeordnet sein sollen. Ferner seien $\lambda_1, \lambda_2, \dots$ voneinander verschiedene Zahlen aus der Folge der μ_s ; L_k sei der von denjenigen x_s gebildete Unterraum, denen diejenigen μ_s entsprechen, die gleich λ_k sind, und P_{L_k} sei Projektionsoperator in den Unterraum L_k .

Wir definieren den Projektionsoperator

$$\mathcal{E}_\lambda = \sum_{\lambda_k \leq \lambda} P_{L_k}.$$

Dies ist die Zerlegung der Einheit, der ein selbstadjungierter Operator C mit reinem Punktspektrum zugeordnet ist. Seine Eigenwerte sind die λ_k , und $x_1, x_2, \dots$ ist das abgeschlossene System seiner Eigenfunktionen. Liegen alle λ_k in einem endlichen Intervall, so ist C ein beschränkter Operator.

196. Das gemischte Spektrum. Wir wollen zunächst das in [191] über die Zerlegung eines selbstadjungierten Operators A in Operatoren mit reinem Punktspektrum und Operatoren mit rein kontinuierlichem Spektrum Gesagte noch etwas ergänzen. Wir nehmen an, ein Unterraum H' von H reduziert den Operator A . Dann reduziert er auch $\mathcal{E}_\lambda$, und bezeichnen wir die durch A und $\mathcal{E}_\lambda$ in H' induzierten Operatoren mit A' bzw. $\mathcal{E}'_\lambda$, so sehen wir, daß $\mathcal{E}'_\lambda$ die Zerlegung der Einheit in H' und ferner

$$A'x = \int_{-\infty}^{\infty} \lambda d\mathcal{E}'_\lambda x \quad (x \in D(A')) \quad (83)$$

ist, d. h., $\mathcal{E}'_\lambda$ ist die Spektralschar von A' . Nun seien A'' und $\mathcal{E}''_\lambda$ zwei von A bzw. $\mathcal{E}_\lambda$ im Unterraum $H'' = H \ominus H'$ induzierte Operatoren. Ist $x = x' + x''$ ($x \in D(A)$) bzw. $y = y' + y''$ die Zerlegung von x bzw. y in H' bzw. H'' , so gilt [191]

$$Ax = A'x' - A''x'', \quad \mathcal{E}_\lambda y = \mathcal{E}'_\lambda y' + \mathcal{E}''_\lambda y'', \quad (Ax, y) = (Ax', y') + (Ax'', y''). \quad (84)$$

Analoge Formeln gelten auch im Fall endlich oder abzählbar unendlich vieler paarweise orthogonaler Unterräume, die A reduzieren. Wir kehren zu den Bezeichnungen von [191] zurück und nehmen an, daß H' nicht gleich dem ganzen Raum H ist. In H' hat der Operator A' ein reines Punktspektrum, in H'' der Operator A'' ein rein kontinuierliches Spektrum. Dabei ist

$$\mathcal{E}'_\lambda = \sum_{\lambda_k \leq \lambda} (\mathcal{E}_{\lambda_k} - \mathcal{E}_{\lambda_k - 0}),$$

und die durch $\mathcal{E}'_\lambda = \mathcal{E}_\lambda - \mathcal{E}'_\lambda$ definierte Spektralschar $\mathcal{E}'_\lambda$ des Operators A'' hat keine Unstetigkeiten. Ist A ein nicht beschränkter Operator, so kann doch einer der Operatoren A' oder A'' beschränkt sein. So ist z. B. der Operator A' beschränkt, wenn alle Unstetigkeitsstellen von $\mathcal{E}_\lambda$ in einem endlichen Intervall liegen. Wir nehmen an, A habe ein rein kontinuierliches Spektrum, und bezeichnen wie in [147] mit C_x die lineare abgeschlossene Hülle der Elemente $\mathcal{E}_\lambda x$. Dann hat A ein einfaches kontinuierliches Spektrum, wenn es ein Element x gibt, für welches C_x mit H übereinstimmt. Dabei gelten die Formeln (254) und (256) aus [147] mit Hellingerschen Integralen über ein unendliches Intervall. Diese Integrale sind die Grenzwerte der entsprechenden Summen bei der Zerlegung des unendlichen Intervalls in endlich viele Teilintervalle. Ist $y \in D(A)$, so gelten auch die Formeln (259) und (261) aus [147]. Die entsprechenden Integrale lassen sich als uneigentliche Integrale mit unendlichem Integrationsintervall auffassen. Mit Hilfe der Ungleichung

$$|\Delta \varphi_y(\lambda)|^2 \leq \Delta \varrho(\lambda) \cdot \Delta \|\mathcal{E}_\lambda y\|^2$$

(vgl. [147]) können wir wie in [192] in bezug auf die Summen (55) zeigen, daß die dem Integral (261) aus [147] entsprechenden unendlichen Summen wegen $y \in D(A)$ konvergente Reihen ergeben. Im allgemeinen Fall des kontinuierlichen Spektrums folgt aus dem Beweis von Satz 2 aus [147], daß C_x für jedes λ den Operator $\mathcal{E}_\lambda$ und infolgedessen auch den Operator A reduziert. Der in C_x durch A induzierte Operator hat ein rein kontinuierliches Spektrum, und wir können genau wie in [147] einen Operator mit rein kontinuierlichem Spektrum in zueinander orthogonalen Unterräumen zerlegen, deren orthogonale Summe den Raum H ergibt. Dementsprechend muß dann in allen Formeln statt eines Hellingerschen Integrals die Summe solcher Integrale stehen. Der Zusammenhang zwischen C_x und $L_2^{(x)}$ läßt sich wie in [152] herstellen.

Es sei nun A ein selbstadjungierter und U ein unitärer Operator. Der Operator $A' = UAU^{-1}$ ist auf der linearen Mannigfaltigkeit $D(A')$ definiert, die sich durch Anwendung von U auf $D(A)$ ergibt. Der Operator A' ist selbstadjungiert. Zum Beweis setzen wir

$$(UAU^{-1}x, y) = (x, y^*)$$

für alle $x \in D(A')$. Wir müssen also die Beziehungen $y \in D(A')$ und $y^* = UAU^{-1}y$ nachweisen. Die obige Gleichung können wir auch in der Form $(Ax', U^{-1}y) = (Ux', y^*)$, wobei $x' = U^{-1}x$ ein beliebiges Element aus $D(A)$ ist, oder in der Form $(Ax', U^{-1}y) = (x', U^{-1}y^*)$ schreiben. Hieraus ergibt sich wegen der Selbstadjungiertheit von A

$$U^{-1}y \in D(A) \quad \text{und} \quad U^{-1}y^* = AU^{-1}y,$$

also $y \in D(A')$ und $y^* = UAU^{-1}y$, was zu beweisen war.

Ist $\mathcal{E}_\lambda$ die Spektralschar des Operators A , dann besitzt $\mathcal{E}'_\lambda = U\mathcal{E}_\lambda U^{-1}$ alle Eigenschaften der Zerlegung der Einheit, und es gilt $\|\mathcal{E}'_\lambda x\| = \|\mathcal{E}_\lambda U^{-1}x\|$. Liegt x in

$D(A')$, so gehört $U^{-1}x$ zu $D(A)$; folglich konvergiert die Reihe

$$\sum_{k=-\infty}^{\infty} \lambda_k'^2 \Delta_k \|\mathcal{E}_\lambda' x\|^2,$$

und die Summe

$$\sum_{k=-\infty}^{\infty} \lambda_k' \Delta_k \mathcal{E}_\lambda' x$$

hat wie in [192] den Grenzwert $A'x = UA U^{-1}x$. Daraus ist ersichtlich, daß $\mathcal{E}_\lambda' = U\mathcal{E}_\lambda U^{-1}$ die Spektralschar von A' ist.

Auch das in [153] gezeigte Kriterium für die unitäre Äquivalenz von Operatoren bleibt gültig. Ferner bleiben die Begriffe der Differentiallösung und des abgeschlossenen Systems von Differentiallösungen unverändert. Jede (im Sinne von H) stetige Differentiallösung $x(\lambda)$ hat die Form $\mathcal{E}_\lambda x$ für $x \in D(A)$; dabei muß $x(\lambda) \Rightarrow 0$ für $\lambda \rightarrow -\infty$ angenommen werden.

197. Funktionen eines selbstadjungierten Operators. Ist $\mathcal{E}_\lambda$ die Spektralschar des selbstadjungierten Operators A und $f(\lambda)$ eine im Intervall $-\infty < \lambda < \infty$ beschränkte Funktion, die bezüglich aller nichtfallenden Funktionen $\|\mathcal{E}_\lambda z\|^2$ meßbar ist, so definiert wie in [155] die Formel

$$(f(A)x, y) = \int_{-\infty}^{\infty} f(\lambda) d\mathcal{E}_\lambda x, y \quad (x \in H, y \in H) \quad (85)$$

einen beschränkten, in ganz H definierten Operator $f(A)$, der alle in [155] erwähnten Eigenschaften besitzt. Die Werte von $f(\lambda)$ auf einer Menge vom Maß 0 bezüglich aller $\|\mathcal{E}_\lambda z\|^2$ haben auf das Integral (85) und damit auch auf $f(A)$ keinen Einfluß.

Wir verallgemeinern nun den Begriff der Funktionen eines Operators $f(A)$ auf reelle Funktionen $f(\lambda)$, welche endliche Werte besitzen und wie früher bezüglich aller $\|\mathcal{E}_\lambda z\|^2$ meßbar, aber nicht beschränkt sind. Mit $f_N(\lambda)$ sei diejenige Abschnittsfunktion bezeichnet, die durch die Beziehungen

$$f_N(\lambda) = \begin{cases} f(\lambda) & \text{für } |f(\lambda)| \leq N, \\ N & \text{für } f_N(\lambda) > N, \\ -N & \text{für } f_N(\lambda) < -N \end{cases}$$

definiert wird. Zu der beschränkten Funktion $f_N(\lambda)$ können wir den beschränkten Operator $f_N(A)$ konstruieren. Dann folgt auf Grund der Formel (302) aus [155] für $y = x$ und $f_1(A) = f_2(A)$

$$\|f_N(A)x - f_M(A)x\|^2 = \int_{-\infty}^{\infty} |f_N(\lambda) - f_M(\lambda)|^2 d\|\mathcal{E}_\lambda x\|^2. \quad (86)$$

Gehört $f(\lambda)$ bezüglich $\|\mathcal{E}_\lambda x\|^2$ zu L_2 , so strebt $f_N(\lambda)$ wegen $|f_N(\lambda)| \leq |f(\lambda)|$ fast überall bezüglich $\|\mathcal{E}_\lambda x\|^2$ gegen $f(\lambda)$. Die rechte Seite von (86) strebt für N und $M \rightarrow \infty$ gegen 0, d. h., die Folge der $f_N(A)x$ konvergiert stark gegen ein Grenzelement, das wir mit $f(A)x$ bezeichnen. Es gilt also $f_N(A)x \Rightarrow f(A)x$ für $N \rightarrow \infty$, wenn

$$\int_{-\infty}^{\infty} f^2(\lambda) d\|\mathcal{E}_\lambda x\|^2 < \infty \quad (87)$$

ist. Die Menge der Elemente x , die dieser Bedingung genügen, bezeichnen wir mit $D[f(A)]$. Im folgenden geben wir ohne Beweis einige diesbezügliche Resultate an.

Die lineare Mannigfaltigkeit $D[f(A)]$ ist in H dicht, $f(A)$ ist ein selbstadjungierter Operator, und es gilt

$$(f(A)x, y) = \int_{-\infty}^{\infty} f(\lambda) d\mathcal{E}_\lambda x, y \quad (x \in D[f(A)], y \in H). \quad (88)$$

Für komplexe Funktionen $f(\lambda) = f_1(\lambda) + f_2(\lambda)i$ nehmen wir $f(A) = f_1(A) + f_2(A)i$ an, und die lineare Mannigfaltigkeit $D[f(A)]$ wird wie oben durch die Bedingung (87) definiert, nur muß $f^2(\lambda)$ durch $|f(\lambda)|^2$ ersetzt werden. Wie in [156] können wir behaupten: Ein Operator ist genau dann eine Funktion des selbstadjungierten Operators A , wenn er abgeschlossen und mit jedem beschränkten, mit A vertauschbaren Operator ebenfalls vertauschbar ist.

Wir untersuchen nun die Vertauschbarkeit allgemeiner selbstadjungierter Operatoren. Erinnern wir uns an den Satz 1 aus [156], so gelangen wir zu der folgenden Definition: Zwei selbstadjungierte Operatoren A und B sind *vertauschbar*, wenn ihre Spektralscharen $\mathfrak{S}_\lambda$ und F_μ (als beschränkte Operatoren) für alle λ und μ vertauschbar sind. Auf Grund des erwähnten Satzes ist diese Definition mit der gewöhnlichen äquivalent, wenn A und B beschränkte Operatoren sind. Ist A nicht beschränkt, B dagegen beschränkt, so erhalten wir die Vertauschbarkeitsdefinition aus [191]. Wir sehen sofort, daß auch sie mit der eben gegebenen Definition übereinstimmt, wenn A und B selbstadjungiert sind. Auf Grund von Satz 1 aus [191] ist nämlich die Vertauschbarkeit im früheren Sinne mit derjenigen Definition äquivalent, die besagt, daß sich B mit $\mathfrak{S}_\lambda$ für jedes λ vertauschen läßt; diese Tatsache ist wiederum damit gleichwertig [143], daß F_μ für jedes μ mit allen $\mathfrak{S}_\lambda$ vertauschbar ist, d. h., wir gelangen zu einer neuen Vertauschbarkeitsdefinition.

Ausgehend von dieser neuen Definition für selbstadjungierte Operatoren kann man zeigen, daß reelle Funktionen eines selbstadjungierten Operators A vertauschbar sind und daß die selbstadjungierten Operatoren $A_1, A_2, \dots$ Funktionen von A sind, wenn sie paarweise vertauschbar sind (vgl. [156]).

Wir betrachten nun die Summe und das Produkt nicht beschränkter Operatoren. Der Operator $(A+B)x = Ax + Bx$ ist für diejenigen Elemente x definiert, die gleichzeitig zu $D(A)$ und zu $D(B)$ gehören. Der Operator $(AB)x = A(Bx)$ ist für alle x mit $x \in D(B)$ und $Bx \in D(A)$ definiert. Ist a eine beliebige komplexe Zahl, so ist der Operator $(aA)x = a(Ax)$ auf $D(A)$ definiert. Sind A und B vertauschbare selbstadjungierte Operatoren, wobei B außerdem beschränkt und auf ganz H definiert sei, so ist der Operator ABx auf der linearen Mannigfaltigkeit l' derjenigen x definiert, für die $Bx \in D(A)$ gilt. Auf Grund der Vertauschbarkeitsdefinition folgt $Bx \in D(A)$ für $x \in D(A)$, d. h., $D(A)$ ist in l' enthalten, und l' kann dabei umfassender als $D(A)$ sein. Wir zeigen nun, daß AB ein auf l' selbstadjungierter Operator ist. Es ist $(ABx, y) = (x, y^*)$ für $x \in l'$ und somit erst recht für $x \in D(A)$. Wir müssen zeigen, daß y ebenfalls zu l' gehört und $y^* = AB y$ ist. Unter der Voraussetzung $x \in D(A)$ läßt sich die obige Beziehung in der Form

$$(BAx, y) = (x, y^*) \quad \text{für } x \in D(A)$$

oder, da B beschränkt und selbstadjungiert ist,

$$(Ax, By) = (x, y^*) \quad \text{für } x \in D(A)$$

schreiben. Daraus ergibt sich wegen der Selbstadjungiertheit von A

$$By \in D(A) \quad \text{und} \quad y^* = AB y,$$

was zu beweisen war. Sind A und B vertauschbar und selbstadjungiert, aber nicht beschränkt, so braucht der Operator AB nicht selbstadjungiert zu sein; jedoch ist der zu ihm adjungierte Operator $(AB)^*$ stets selbstadjungiert.

Wir wenden nun die Summen- und Produktdefinition auf Potenzen des Operators A an. Die lineare Mannigfaltigkeit $D(A^2)$ besteht aus denjenigen x , für die $x \in D(A)$ und $Ax \in D(A)$ ist, d. h., $D(A^2)$ ist in $D(A)$ enthalten und kann eventuell

mit $D(A)$ übereinstimmen. Analog besteht $D(A^3)$ aus den Elementen x mit $x \in D(A^2)$ und $A^2x \in D(A)$, so daß folglich $D(A^3)$ in $D(A^2)$ enthalten ist. Ein Polynom der Form $a_1A + a_2A^2 + \dots + a_nA^n$ ist offenbar auf der linearen Mannigfaltigkeit $D(A^n)$ definiert. Es stimmt mit der oben definierten Funktion des Operators A überein, wenn man $f(\lambda) = a_0 + a_1\lambda + \dots + a_n\lambda^n$ setzt. Die Menge der Elemente, für die diese Polynome definiert sind, ist eine in H dichte lineare Mannigfaltigkeit.

Ist der selbstadjungierte Operator A positiv, d. h., ist die untere Grenze m_A des Spektrums nichtnegativ, so läßt sich wie in [143] der positive selbstadjungierte Operator

$$A^{1/2} = \int_{m_A}^{\infty} \sqrt{\lambda} \, d\mathcal{E}_\lambda$$

konstruieren, dessen Quadrat gleich A ist.

Es existiert nur ein einziger positiver selbstadjungierter Operator B , dessen Quadrat gleich A ist. Um dies zu beweisen, nehmen wir an, $\mathcal{E}'_\mu$ sei die Zerlegung der Einheit für B . Es muß also

$$B = \int_0^{\infty} \mu \, d\mathcal{E}'_\mu \quad \text{und} \quad A = B^2 = \int_0^{\infty} \mu^2 \, d\mathcal{E}'_\mu$$

oder

$$A = B^2 = \int_0^{\infty} \lambda \, d\mathcal{E}'_{\sqrt{\lambda}}$$

gelten. Die von dem Parameter λ abhängige Schar $\mathcal{E}'_{\sqrt{\lambda}}$ ist eine Zerlegung der Einheit. Wegen der Eindeutigkeit der Spektralschar gilt $\mathcal{E}'_{\sqrt{\lambda}} = \mathcal{E}_\lambda$. Daraus folgt $B = A^{1/2}$.

198. Kleine Störungen des Spektrums. Wir untersuchen nun, wie sich das Spektrum eines selbstadjungierten Operators bei Addition eines anderen selbstadjungierten Operators ändert. Zunächst weisen wir darauf hin, daß für nicht beschränkte Operatoren der Satz 1 aus [157] gilt. Ihm wollen wir noch zwei weitere Sätze zur Seite stellen.

Es sei L ein Unterraum von H . Seine Dimension r ist gleich der Anzahl der Elemente eines in L abgeschlossenen normierten Orthogonalsystems. Die Zahl r kann sowohl endlich als auch unendlich sein. Man kann aber leicht zeigen, daß sie von der Wahl dieses Systems unabhängig ist.

Hilfssatz 1. *Es seien L_1 und L_2 zwei Unterräume der Dimension r_1 bzw. r_2 . Ist $r_1 < r_2$, so gibt es in L_2 ein vom Nullelement verschiedenes Element, das zu allen Elementen aus L_1 orthogonal ist.*

Beweis. Da H separabel ist, ist r_1 endlich und r_2 entweder endlich oder unendlich. Den Beweis des Hilfssatzes führen wir indirekt und nehmen an, L_2 besitze keine vom Nullelement verschiedenen Elemente, die zu L_1 orthogonal sind; $x_1, x_2, \dots, x_{r_1}$ sei ein abgeschlossenes normiertes Orthogonalsystem in L_1 (eine Basis in L_1) und P der Projektionsoperator in den Unterraum L_2 . Für jedes Element $v \in L_2$ gilt

$$(v, Px_k) = (Pv, x_k) = (v, x_k) .$$

Die Elemente $Px_k \in L_2$ definieren einen gewissen Unterraum L_3 der Dimension $r_3 \leq r_1$ ($r_3 < r_1$, wenn die Elemente Px_k linear abhängig sind). Wir zeigen, daß L_3 mit L_2 übereinstimmen muß. Wäre es nicht der Fall, so würde ein vom Nullelement verschiedenes und zu L_3 orthogonales Element $y \in L_2$ existieren, es wäre $(y, Px_k) = 0$ und somit auf Grund der obigen Gleichung

$$(y, x_k) = 0 \quad (k = 1, 2, \dots, r_1) ;$$

d. h., y wäre zu L_1 orthogonal. Nach Voraussetzung gibt es aber kein solches Element y . Also stimmt L_3 mit L_2 überein, und es ist $r_2 = r_3$. Wegen $r_3 \leq r_1$ ist dann auch $r_2 \leq r_1$. Dies widerspricht jedoch der Bedingung $r_1 < r_2$.

Hilfssatz 2. Es sei A ein (beschränkter oder nicht beschränkter) selbstadjungierter Operator, B ein beschränkter selbstadjungierter Operator und $\mathcal{S}_\lambda$ bzw. $\mathcal{S}'_\lambda$ die Spektralschar von A bzw. $A' = A + B$. Ferner sei Δ ein endliches Intervall $[a, b]$ und $\Delta_{B,\varepsilon}$ das Intervall $[a - \|B\| - \varepsilon, b + \|B\| + \varepsilon]$, wobei ε eine beliebige positive Zahl sei. Dann ist die Dimension des Unterraums $L_1 = (\Delta_{B,\varepsilon} \mathcal{S}'_\lambda) x, x \in H$, nicht kleiner als die Dimension des Unterraums $L_2 = (\Delta \mathcal{S}_\lambda) x, x \in H$.

Beweis. Wir benutzen die Bezeichnungen aus [141] (etwa $\Delta \mathcal{S}_\lambda = \mathcal{S}_b - \mathcal{S}_a$). Auch diesen Beweis führen wir indirekt, indem wir annehmen, die Dimension von L_1 sei kleiner als die von L_2 . Auf Grund des Hilfssatzes 1 existiert ein Element $y \in L_2$, das vom Nullelement verschieden und zu L_1 orthogonal ist. Wir können $\|y\| = 1$ annehmen. Setzen wir $\alpha = (a+b)/2$, so ist für $y \in L_2$

$$\|Ay - \alpha y\|^2 = \int_a^b (\lambda - \alpha)^2 d(\mathcal{S}_\lambda y, y) \leq \left(\frac{b-a}{2}\right)^2$$

und

$$\|A'y - \alpha y\| \leq \|Ay - \alpha y\| + \|By\| \leq \frac{b-a}{2} + \|B\|. \quad (89)$$

Andererseits gilt wegen $y \perp L_1$

$$\|A'y - \alpha y\|^2 = \int_{-\infty}^{\infty} (\lambda - \alpha)^2 d(\mathcal{S}'_\lambda y, y) = \int_{-\infty}^{a-\|B\|-\varepsilon} + \int_{b+\|B\|+\varepsilon}^{\infty},$$

woraus sich, wieder wegen $y \perp L_1$,

$$\|A'y - \alpha y\| \geq \left(\frac{b-a}{2} + \|B\| + \varepsilon\right) \int_{-\infty}^{\infty} d(\mathcal{S}'_\lambda y, y),$$

d. h.

$$\|A'y - \alpha y\| \geq \frac{b-a}{2} + \|B\| + \varepsilon \quad (90)$$

ergibt. Durch den Widerspruch zwischen (89) und (90) wird der Hilfssatz bewiesen.

Satz 1. Im Innern des Intervalls Δ bestehe das Spektrum des Operators A aus endlich vielen Eigenwerten, deren Vielfachheiten, addiert, die endliche Zahl k ergeben, und der Abstand zwischen Δ und dem übrigen Teil des Spektrums von A sei größer als $2\|B\|$. Dann besteht das Spektrum von A' im Intervall $\Delta_{B,0} = [a - \|B\|, b + \|B\|]$ aus den Eigenwerten, deren Vielfachheiten, addiert, die Zahl k ergeben.

Beweis. Auf Grund von Hilfssatz 2 ist für jedes $\varepsilon > 0$ die Dimension von L_1 mindestens gleich k . Ist sie größer als k , so können wir, wenn wir den Hilfssatz 2 und die Beziehung $A = A' - B$ benutzen, behaupten, daß die Dimension des Unterraums $(\Delta_{B,2\varepsilon} \mathcal{S}_\lambda) x$ mit $x \in H$ größer als k ist. Nach Voraussetzung kann dies jedoch nicht für alle hinreichend nahe bei 0 liegenden ε gelten, d. h., für diese ε ist die Dimension von L_1 gleich k . Damit ist der Satz bewiesen.

Bemerkung. Benutzen wir den Satz für $k=1$, so haben wir bei kleinen Störungen die Möglichkeit, die Änderung eines isolierten einfachen Eigenwerts zu beobachten.

Satz 2. Liegt im Innern des Intervalls Δ wenigstens ein Verdichtungspunkt des Spektrums von A , so befindet sich im Intervall $\Delta_{B,0}$ wenigstens ein Verdichtungspunkt des Spektrums von A' .

Hier ist $k = \infty$, so daß dieser Satz aus Hilfssatz 2 folgt.

Bemerkung. Ist λ_0 ein Verdichtungspunkt des Spektrums von A , so gibt es im Intervall $[\lambda_0 - ||B||, \lambda_0 + ||B||]$ wenigstens einen Verdichtungspunkt des Spektrums von A' .

199. Der Operator der Multiplikation. Wir betrachten nun den Raum $L_2(-\infty, \infty)$ und den Operator der Multiplikation mit der unabhängigen Veränderlichen:

$$Af(x) = xf(x). \quad (91)$$

Die lineare Mannigfaltigkeit $D(A)$ enthält diejenigen Funktionen $f(x) \in L_2$, für die auch $xf(x) \in L_2$ ist. Insbesondere gehören zu $D(A)$ alle Funktionen, die nur auf endlichen Intervallen ungleich 0 sind. Daraus folgt, daß $D(A)$ in H dicht ist. Wir wollen zeigen, daß der Operator A selbstadjungiert ist; es muß also $y \in D(A)$ und $y^* = Ay$ sein, wenn $(Ax, y) = (x, y^*)$ für $x \in D(A)$ ist. In diesem Fall gilt also

$$\int_{-\infty}^{\infty} xf(x) \overline{\varphi(x)} dx = \int_{-\infty}^{\infty} f(x) \overline{\varphi^*(x)} dx \quad (92)$$

für alle $f(x) \in D(A)$ und gewisse $\varphi(x)$ und $\varphi^*(x)$ aus L_2 . Wir brauchen somit nur nachzuweisen, daß $\varphi(x) \in D(A)$ und $\varphi^*(x) = x\varphi(x)$ ist. Dazu wenden wir (92) auf die nur auf einem gewissen endlichen Intervall $(-a, a)$ von 0 verschiedene Funktion $f(x) \in L_2$ an; diese Funktion gehört auch zu $D(A)$, und es ist

$$\int_{-a}^a f(x) \overline{[\varphi^*(x) - x\varphi(x)]} dx = 0.$$

Da $f(x)$ willkürlich ist, folgt hieraus unmittelbar (vgl. [51]), daß $\varphi^*(x) - x\varphi(x)$ auf dem Intervall $(-a, a)$ und wegen der Willkürlichkeit von a auch auf dem Intervall $(-\infty, \infty)$ äquivalent 0 ist. Wir können also $\varphi^*(x) - x\varphi(x) = 0$ annehmen. Nun ist $\varphi^*(x) \in L_2$ und somit $x\varphi(x) \in L_2$ sowie $\varphi^*(x) = x\varphi(x)$, was zu beweisen war. Die Spektralschar des mit Hilfe von (92) definierten Operators wird wie in [152] durch

$$\mathcal{E}_\lambda f(x) = \begin{cases} f(x) & \text{für } x \leq \lambda, \\ 0 & \text{für } x > \lambda \end{cases} \quad (93)$$

gegeben. Der Operator hat ein rein kontinuierliches Spektrum auf dem Intervall $-\infty < \lambda < \infty$.

Der durch (91) definierte Operator ist offenbar nicht beschränkt. Wir erwähnen noch, daß jede Funktion $f(x) \in D(A)$ auf dem Intervall $(-\infty, \infty)$ summierbar ist. Setzen wir nämlich $xf(x) = \omega(x)$ mit $\omega(x) \in L_2$, so können wir $f(x) = \omega(x)/x$ schreiben. Da $1/x$ und $\omega(x)$ auf jedem Intervall $(-\infty, -a)$ und (a, ∞) mit $a > 0$ zu L_2 gehören, folgt hieraus die Summierbarkeit von $f(x)$ im Intervall $(-\infty, \infty)$.

Analog wie in [152] können wir den Operator der Multiplikation mit einer reellen und meßbaren Funktion untersuchen:

$$A'f(x) = \omega(x) f(x). \quad (94)$$

Ist $\omega(x)$ nicht beschränkt, so erhalten wir auch einen nicht beschränkten Operator. Wir nehmen nun an, $\omega(x)$ sei auf dem ganzen Intervall $(-\infty, \infty)$ mit Ausnahme hinreichend kleiner Umgebungen endlich vieler Punkte beschränkt, und wählen ein beliebiges abgeschlossenes Intervall, das keinen dieser Punkte enthält. Dann können wir uns wie oben davon überzeugen, daß der durch (94) für solche $f(x) \in L_2$, für die auch $\omega(x) f(x)$ zu L_2 gehört, definierte Operator selbstadjungiert ist. Seine

Spektralschar läßt sich wie in [152] durch

$$\mathfrak{S}_\lambda f(x) = \begin{cases} f(x) & \text{für } \omega(x) \leq \lambda, \\ 0 & \text{für } \omega(x) > \lambda \end{cases} \quad (95)$$

ausdrücken.

Ist beispielsweise $\omega(x) \in L_2$, so enthält die lineare Mannigfaltigkeit $D(A)$ alle beschränkten Funktionen aus L_2 , und A' ist selbstadjungiert. Der Operator A' läßt sich auch als Funktion $\omega(A)$ auffassen, wobei A der Operator der Multiplikation mit der unabhängigen Veränderlichen für die Klasse der in [197] betrachteten Funktionen $\omega(x)$ ist.

Mit B bezeichnen wir nun den folgenden selbstadjungierten Operator:

$$B\Phi(x) = i \frac{d\Phi(x)}{dx} = i\varphi(x), \quad (96)$$

der in $H = L_2(-\infty, \infty)$ auf der Menge der Funktionen $\Phi(x)$, die auf jedem endlichen Intervall absolut stetig sind und Ableitungen in $L_2(-\infty, \infty)$ besitzen, definiert ist [188].

Mit Hilfe einer Fouriertransformation wollen wir nun die linearen Mannigfaltigkeiten $D(A)$ und $D(B)$ vergleichen. Wir setzen

$$\Psi_N(x) = \frac{1}{\sqrt{2\pi}} \int_{-N}^N \Phi(t) e^{ixt} dt, \quad \psi_N(x) = \frac{1}{\sqrt{2\pi}} \int_{-N}^N i\varphi(t) e^{ixt} dt, \quad (97)$$

wobei $\Phi(t)$ eine beliebige Funktion aus $D(B)$ ist, und erhalten nach partieller Integration

$$x\Psi_N(x) - \psi_N(x) = \frac{1}{\sqrt{2\pi i}} [\Phi(N) e^{ixN} - \Phi(-N) e^{-ixN}]; \quad (98)$$

hier strebt die rechte Seite für $N \rightarrow \infty$ auf dem ganzen unendlichen Intervall gleichmäßig bezüglich x gegen 0 [188]. Die Funktionen $\psi_N(x)$ und $\Psi_N(x)$ konvergieren im gleichen Intervall im Mittel gegen Funktionen $\psi(x)$ bzw. $\Psi(x)$ aus L_2 [178], also ist dies in jedem endlichen Intervall $[-a, a]$ erst recht der Fall. Außerdem strebt $x\Psi_N(x)$ in einem solchen Intervall im Mittel gegen $x\Psi(x)$. Aus (98) folgt, daß für jedes gegebene positive ε ein hinreichend großes N existiert, so daß die Ungleichung

$$\int_{-a}^a |x\Psi_N(x) - \psi_N(x)|^2 dx \leq \varepsilon$$

oder, wenn wir zur Grenze übergehen,

$$\int_{-a}^a |x\Psi(x) - \psi(x)|^2 dx \leq \varepsilon$$

erfüllt ist. Hieraus erhalten wir, da ε und a beliebig sind,

$$\int_{-\infty}^{\infty} |x\Psi(x) - \psi(x)|^2 dx = 0,$$

also $x\Psi(x) = \psi(x)$, oder

$$x \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \Phi(t) e^{ixt} dt = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} i\varphi(t) e^{ixt} dt. \quad (99)$$

Dieser Gleichung können wir auch die Gestalt

$$T^*(\Phi) = \frac{1}{x} T^*(i\varphi) = f \quad (100)$$

geben. Aus ihr können wir ablesen, daß wegen $T^*(i\varphi) \in L_2$ sowohl $f(x)$ als auch $xf(x)$ zu L_2 gehören, d. h., es ist $f(x) \in D(A)$.

Wir zeigen nun, daß $T(f)$ in $D(B)$ liegt, wenn $f(x)$ eine beliebige Funktion aus $D(A)$ ist. Wegen $f(x) \in L_2$ und $xf(x) \in L_2$ können wir $T(f)$ und $T(xf)$ bilden. Unter Berücksichtigung von (100) setzen wir

$$\Phi(x) = T(f) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t) e^{-ixt} dt, \quad i\varphi(x) = T(xf) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} tf(t) e^{-ixt} dt. \quad (101)$$

Nun gilt die Beziehung (vgl. [178])

$$i \int_0^N \varphi(t) dt = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} tf(t) \frac{1}{i} (1 - e^{-iNt}) dt$$

oder, wenn wir die erste Gleichung (101) benutzen, aus der wegen der Summierbarkeit von $f(t)$ auf dem Intervall $(-\infty, \infty)$

$$\Phi(0) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t) dt$$

folgt,

$$i \int_0^N \varphi(t) dt = \frac{1}{i} \Phi(0) - \frac{1}{i} \Phi(N).$$

Also gehört $\Phi(x)$ zu $D(B)$, und es ist $i\varphi(x) = B\Phi(x)$. Oben sahen wir, daß T^* jedes Element $\Phi \in D(B)$ in ein Element $f \in D(A)$ abbildet, und eben haben wir gezeigt, daß T jedes Element $f \in D(A)$ in ein Element $\Phi \in D(B)$ transformiert. Damit haben wir also eine eindeutige Zuordnung zwischen $D(A)$ und $D(B)$ hergestellt, wobei T die lineare Mannigfaltigkeit $D(A)$ in $D(B)$ abbildet und der Übergang von $f(x)$ zu $xf(x)$ vor der Transformation dem Übergang von $\Phi(x)$ zu $B\Phi(x) = i\varphi(x)$ nach der Transformation entspricht, d. h., es ist

$$B = TAT^*. \quad (102)$$

Diese Überlegungen führen uns auf den folgenden

Satz. *Zwei selbstadjungierte Operatoren A und B sind unitär äquivalent, und für sie gilt die Formel (102), wobei T eine Fouriertransformation ist.*

Für die Spektralschar $\mathcal{E}'_\lambda$ des Operators B gilt $\mathcal{E}'_\lambda = T\mathcal{E}_\lambda T^*$, wobei $\mathcal{E}_\lambda$ durch (93) definiert ist, d. h.

$$\mathcal{E}'_\lambda f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\lambda} \left[\int_{-\infty}^{\infty} f(t) e^{iyt} dt \right] e^{-ixy} dy \quad (103)$$

oder

$$\mathcal{E}'_\lambda f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\lambda} T^*(f) e^{-ixy} dy. \quad (104)$$

Wir können $\Phi(x)$ in der Form

$$\Phi(x) = \int_{-\infty}^{\infty} \varphi(t) dt = - \int_x^{\infty} \varphi(t) dt \quad (105)$$

angeben, aber nichts über die absolute Integrierbarkeit von $\varphi(t)$ im Intervall $(-\infty, \infty)$ aussagen; die Integrale fassen wir als Grenzwerte von Integralen über ein endliches Intervall auf. Der durch $B^{-1}\varphi = -i\Phi$ definierte selbstadjungierte Operator B^{-1} ist nicht auf ganz L_2 , sondern nur für solche $\varphi(x)$ gegeben, für welche das durch (105) definierte $\Phi(x)$ ebenfalls in L_2 liegt. Dies ist eine Folge davon, daß $\lambda' = 0$ zum kontinuierlichen Spektrum von B gehört.

Betrachten wir den Operator (94), so können wir den Operator $TA'T^*$ als Funktion $\omega(B)$ des Operators B darstellen, so daß

$$\omega(B)f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \omega(y) \left[\int_{-\infty}^{\infty} f(t) e^{iyt} dt \right] e^{-ixy} dy \quad (106)$$

gilt. Ist $\omega(x)$ eine nicht beschränkte Funktion, so läßt sich die lineare Mannigfaltigkeit, auf der dieser Operator definiert ist, aus $D(A')$ mit Hilfe des Operators T gewinnen. Wir erwähnten, daß die obigen Integrale mit unendlichen Grenzen im Sinne der Konvergenz im Mittel aufgefaßt werden können. Sowohl B als auch A haben ein rein kontinuierliches Spektrum. Die Funktion $\omega(x)$ aus (106) muß also allen in [197] für $f(\lambda)$ formulierten Bedingungen genügen.

200. Integraloperatoren. Wir betrachten nun auf dem Intervall $[a, b]$ Integraloperatoren mit dem Kern $K(x, y)$, der der Bedingung

$$K(y, x) = \overline{K(x, y)} \quad (107)$$

genügt und für den die Beziehung

$$K^2(x) = \int_a^b |K(x, y)|^2 dy < \infty \quad (K(x) \geq 0) \quad (108)$$

für fast alle x aus $[a, b]$ gilt. Der entsprechende, durch

$$\varphi(x) = Kf(x) = \int_a^b K(x, y) f(y) dy \quad (109)$$

definierte Operator K ist auf der linearen Mannigfaltigkeit derjenigen $f(x) \in L_2[a, b]$ definiert, für welche (109) ebenfalls zu L_2 gehört. Wir betrachten noch wie in [173] die lineare Mannigfaltigkeit l . Sie enthält diejenigen $f(x) \in L_2$, für die

$$\int_a^b K(x) |f(x)| dx < \infty \quad (110)$$

gilt. Die lineare Mannigfaltigkeit l ist in L_2 dicht [173]. Wir wollen nun zeigen: *Liegt die Funktion $f(x)$ in l , so liegt sie erst recht in $D(K)$.* Hieraus ergibt sich unter anderem, daß $D(K)$ in L_2 dicht ist. Es sei also $f(x) \in l$. Nun können wir schreiben:

$$|\varphi(x)|^2 = \int_a^b K(x, y) f(y) dy \cdot \int_a^b \overline{K(x, t) f(t)} dt,$$

also

$$\int_a^b |\varphi(x)|^2 dx \leq \int_a^b \int_a^b |K(x, y)| \cdot |K(x, t)| \cdot |f(y)| \cdot |f(t)| dy dt dx. \quad (111)$$

Es genügt nachzuweisen, daß das rechts stehende Integral bei gewisser Reihenfolge der Integration einen endlichen Wert hat. Auf Grund der Schwarzschen Ungleichung ist

$$\int_a^b |K(x, y)| \cdot |K(x, t)| \, dx \leq \left[\int_a^b |K(x, y)|^2 \, dx \cdot \int_a^b |K(x, t)|^2 \, dx \right]^{1/2} = K(y) K(t),$$

und die rechte Seite von (111) ist nicht größer als

$$\int_a^b K(y) |f(y)| \, dy \cdot \int_a^b K(t) |f(t)| \, dt.$$

Dieses Produkt hat wegen $f(x) \in l$ einen endlichen Wert. Wir bezeichnen den durch (109) auf der linearen Mannigfaltigkeit l definierten Operator mit $Af(x)$. Dieser Operator ist symmetrisch, d. h., es ist

$$\int_a^b \left[\int_a^b K(x, y) f(y) \, dy \right] \overline{\omega(x)} \, dx = \int_a^b \left[\int_a^b K(x, y) \overline{\omega(x)} \, dx \right] f(y) \, dy, \quad (112)$$

wenn $f(x)$ und $\omega(x)$ zu l gehören; dabei ist wegen (107)

$$\int_a^b K(x, y) \overline{\omega(x)} \, dx = \overline{\int_a^b K(y, x) \omega(x) \, dx},$$

und diese Funktion liegt als Funktion von y in L_2 . Zum Beweis von (112) brauchen wir uns nur davon zu überzeugen, daß das Integral

$$\int_a^b \int_a^b |K(x, y)| \cdot |f(y)| \cdot |\omega(x)| \, dy \, dx \quad (113)$$

bei gewisser Reihenfolge der Integration endlich ist. Integrieren wir zuerst nach x , so ist das Ergebnis auf Grund der Schwarzschen Ungleichung nicht größer als

$$||\omega|| \cdot \int_a^b K(y) |f(y)| \, dy.$$

Die Größe ist wegen $f(y) \in l$ endlich.

Der auf l durch (109) definierte symmetrische Operator A ist aber nicht immer selbstadjungiert; er hat jedoch einen adjungierten Operator A^* . Wir können beweisen, daß A^* mit dem Operator K übereinstimmt; K ist ebenfalls durch (109) definiert, aber nicht auf l , sondern auf der linearen Mannigfaltigkeit $D(K)$ derjenigen $f(x) \in L_2$, für die auch $\varphi(x) \in L_2$ gilt. Es gelte für alle $f(x) \in l$ die Formel

$$\int_a^b \left[\int_a^b K(x, y) f(y) \, dy \right] \overline{\omega(x)} \, dx = \int_a^b f(x) \overline{\omega^*(x)} \, dx \quad (114)$$

mit $\omega(x) \in L_2$ und $\omega^*(x) \in L_2$. Wir müssen also zeigen, daß

$$\omega^*(x) = \int_a^b K(x, y) \omega(y) \, dy \quad (115)$$

gilt, woraus sich auch $\omega(x) \in D(K)$ ergibt. Beim Beweis der Endlichkeit von (113) benutzten wir nur die Zugehörigkeit der Funktionen $f(x)$ zu l . Daher können wir in dem links stehenden Integral von (114) die Reihenfolge der Integration ändern

und (114) auf die Form

$$\int_a^b f(x) \left[\omega^*(x) - \int_a^b K(x, y) \omega(y) dy \right] dx = 0 \quad (116)$$

bringen. Der Beweis von Satz 2 aus [173] führt uns unmittelbar darauf, daß die in eckigen Klammern stehende Differenz verschwinden muß, so daß wir (115) erhalten.

Ist umgekehrt $\omega(x) \in D(K)$ und die Funktion $\omega^*(x)$ durch (115) gegeben, so folgt aus den vorhergehenden Überlegungen sofort, daß (114) gilt. Damit erhalten wir den folgenden

Satz. *Es sei $K(y, x) = \overline{K(x, y)}$, und die Forderung (108) sei erfüllt. Ferner sei l die lineare Mannigfaltigkeit der Funktionen $f(x) \in L_2[a, b]$, die der Bedingung (110) genügen, und $D(K)$ die lineare Mannigfaltigkeit der Funktionen $f(x) \in L_2$, für die das durch (109) gegebene $\varphi(x)$ ebenfalls zu L_2 gehört. Dann ist l in L_2 dicht und in der linearen Mannigfaltigkeit $D(K)$ enthalten. Ist A ein auf l durch (109) und K ein auf $D(K)$ ebenfalls durch (109) definierter Operator, so ist A symmetrisch, und es ist $A^* = K$.*

Notwendig für die Selbstadjungiertheit des Operators K ist seine Symmetrie; wir gelangen also wegen (107) zu der Beziehung

$$\int_a^b \left[\int_a^b K(y, x) f(x) dx \right] \overline{\omega(y)} dy = \int_a^b \left[\int_a^b K(y, x) \overline{\omega(y)} dy \right] f(x) dx, \quad (117)$$

die für alle $f(x)$ und $\omega(x)$ aus $D(K)$ erfüllt sein muß. Diese Bedingung ist für die Selbstadjungiertheit von K auch hinreichend. Zum Beweis dieser Tatsache nehmen wir an, daß für alle $f(x) \in D(K)$

$$\int_a^b \left[\int_a^b K(x, y) f(y) dy \right] \overline{\omega(x)} dx = \int_a^b f(x) \overline{\omega^*(x)} dx$$

mit $\omega(x) \in L_2$ und $\omega^*(x) \in L_2$ gilt. Wir müssen dann zeigen, daß $\omega(x)$ zu $D(K)$ gehört und (115) gilt. Auf Grund der Definition von $D(K)$ brauchen wir also nur die Gültigkeit von (115) nachzuweisen. Subtrahieren wir (117) von der letzten Gleichung und wenden wir (107) an, so erhalten wir die Beziehung (116). Aus ihr ergibt sich (115), da $f(x)$ beliebig aus $D(K)$ ist und l in $D(K)$ enthalten ist. *Der Operator K ist also genau dann selbstadjungiert, wenn für alle $f(x) \in D(K)$ und $\omega(x) \in D(K)$ die Relation (117) erfüllt ist.*

Wir geben nun einfache Beispiele für selbstadjungierte Operatoren mit Differenzkernen auf dem Intervall $(-\infty, \infty)$ an. Es sei $g(t)$ eine reelle gerade Funktion aus L_2 auf diesem Intervall und $f(x)$ eine beliebige Funktion aus L_2 . Ferner setzen wir $G(t) = T^*(g)$ und $F(t) = T^*(f)$. Die Funktion $G(t)$ ist reell, da $g(t)$ reell und gerade ist. Wir können dann schreiben (vgl. [181]):

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} g(x-y) f(y) dy = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} G(t) F(t) e^{-ixt} dt. \quad (118)$$

Dabei ist das Produkt $G(t) F(t)$ auf dem Intervall $(-\infty, \infty)$ summierbar, da beide Faktoren aus L_2 stammen. Es sei l' die lineare Mannigfaltigkeit derjenigen Funktionen $F(t) \in L_2$, für die auch $G(t) F(t) \in L_2$ ist. Auf l' läßt sich die rechte Seite von (118) durch

$T(GF)$ darstellen, so daß $\varphi(x)$ zu L_2 gehört. Wir wenden nun unsere Aufmerksamkeit dem mittleren Teil von (118) zu und behaupten, daß auf der durch die Transformation T aus U hervorgehenden linearen Mannigfaltigkeit U'_1 der Integraloperator K mit dem Kern $g(x-y)$ unitär äquivalent der Multiplikation mit der Funktion $G(t) \in L_2$ und infolgedessen selbstadjungiert ist. Schon früher haben wir die lineare Mannigfaltigkeit derjenigen Funktionen $f(x) \in L_2$, für die das durch (118) definierte $\varphi(x)$ ebenfalls in L_2 liegt, mit $D(K)$ bezeichnet. Aus den vorhergehenden Überlegungen folgt nur, daß U'_1 in $D(K)$ enthalten ist. Wir können aber zeigen, daß U'_1 mit $D(K)$ sogar übereinstimmt. Diese Behauptung ist der folgenden äquivalent: Ist in Formel (118) $\varphi(x) \in L_2$, so ist auch $G(t) F(t)$ in L_2 enthalten. Die Übereinstimmung mit $D(K)$ folgt unmittelbar aus der Unmöglichkeit, einen selbstadjungierten Operator so zu erweitern, daß wieder ein selbstadjungierter Operator entsteht. Die Unmöglichkeit einer solchen Erweiterung haben wir schon in [187] bewiesen. Ist also $g(t)$ eine reelle gerade Funktion aus L_2 , so ist der Integraloperator mit dem Kern $g(x-y)$ in $D(K)$ selbstadjungiert.

201. Erweiterung eines abgeschlossenen symmetrischen Operators. Wir nehmen im folgenden den Operator A als abgeschlossen und symmetrisch an und beweisen zwei für das Weitere grundlegende Sätze.

Satz 1. *Die lineare Mannigfaltigkeit $D(A)$ der Elemente x wird vermöge der Formeln*

$$y = (A + iE)x, \quad (119)$$

$$z = (A - iE)x \quad (120)$$

eindeutig auf gewisse Unterräume $L_1(A)$ und $L_{-1}(A)$ abgebildet. Sind y und z Elemente dieser Unterräume, die ein und demselben Element x aus $D(A)$ entsprechen, so transformiert der distributive Operator U , der y in z abbildet,

$$z = Uy, \quad (121)$$

den Unterraum $L_1(A)$ eineindeutig in $L_{-1}(A)$, wobei er Norm und Skalarprodukt invariant läßt:

$$\|Uy\| = \|y\|, \quad (Uy_1, Uy_2) = (y_1, y_2). \quad (122)$$

Beweis. Es gilt

$$\begin{aligned} \|(A + iE)x\|^2 &= ((A + iE)x, (A + iE)x) \\ &= (Ax, Ax) + i(x, Ax) - i(Ax, x) + (x, x) \end{aligned}$$

oder, da A symmetrisch ist,

$$\|(A + iE)x\|^2 = \|Ax\|^2 + \|x\|^2 \quad (x \in D(A)). \quad (123)$$

Wir zeigen zunächst, daß die Formel (119) für verschiedene x auch verschiedene y liefert. Ergeben nämlich zwei verschiedene Elemente x_1 und x_2 aus $D(A)$ ein einziges Element y , so ergibt ihre Differenz $x = x_1 - x_2$ das Element $y = 0$. Es ist also zu zeigen, daß $x = 0$ aus $(A + iE)x = 0$ folgt. Das erkennt man aber sofort aus (123). Somit führt (119) die lineare Mannigfaltigkeit $D(A)$ eineindeutig in eine lineare Mannigfaltigkeit $L_1(A)$ über. Diese Mannigfaltigkeit ist abgeschlossen, d. h., $L_1(A)$ ist ein Unterraum. In der Tat, mit Hilfe von (123) erhalten wir $\|(A + iE)x\| \leq \|x\|$; also ist $(A + iE)^{-1}$ beschränkt. In [184] zeigten wir, daß $R(B)$ ein Unterraum ist, wenn B ein abgeschlossener Operator ist und auf $R(B)$ ein beschränktes Inverses B^{-1} existiert. Damit ist bewiesen, daß $L_1(A)$ ein Unterraum ist. Aus der zu (123)

analogen Formel

$$|(A - iE)x|^2 = |Ax|^2 + |x|^2 \quad (123_1)$$

ergibt sich, daß vermöge (120) die lineare Mannigfaltigkeit $D(A)$ eineindeutig in einen Unterraum $L_{-1}(A)$ übergeführt wird. Nehmen wir zwei Elemente y und z , die beide ein und demselben Element x entsprechen, so existiert die wohlbestimmte distributive Abbildung (121), die $L_1(A)$ eineindeutig in $L_{-1}(A)$ transformiert, wobei (123) und (123₁) in der Form

$$|y|^2 = |Ax|^2 + |x|^2 \quad \text{bzw.} \quad |z|^2 = |Ax|^2 + |x|^2$$

geschrieben werden können, d. h., es ist $|z| = |y|$ oder $|Uy| = |y|$. Die zweite Formel (122) läßt sich analog nachweisen wie für unitäre Operatoren [137]. Damit ist der Satz bewiesen.

Ist der Operator A selbstadjungiert, so stimmen, da $\pm i$ reguläre Punkte von A sind, $L_1(A)$ und $L_{-1}(A)$ mit H überein [189], und U ist eine unitäre Transformation. Stimmen beide Unterräume oder wenigstens einer von ihnen nicht mit H überein, so nennt man U im allgemeinen einen isometrischen Operator, d. h., ein distributiver Operator U , der auf einem Unterraum L' definiert ist und diesen eineindeutig auf einen anderen Unterraum L'' unter Beibehaltung der Norm (und somit auch des Skalarprodukts) abbildet, heißt *isometrisch*. Der inverse Operator U^{-1} , der L'' in L' überführt, ist offenbar ebenfalls isometrisch. Stimmen L' und L'' mit H überein, so ist U unitär und auf ganz H definiert. Vermöge der Formeln $y = Ax + ix$ und $Uy = Ax - ix$ wird $D(A)$ eineindeutig in $L_1(A)$ bzw. $L_{-1}(A)$ transformiert, und aus ihnen ergibt sich

$$x = \frac{1}{2i} (y - Uy), \quad Ax = \frac{1}{2} (y + Uy).$$

Die erste dieser Beziehungen bildet $L_1(A)$ eineindeutig in $D(A)$ ab. Ersetzen wir y durch $2iy$, was auf dieselbe lineare Mannigfaltigkeit $L_1(A)$ der Elemente y führt, so erhalten wir die einfacheren Formeln

$$x = y - Uy, \quad (124)$$

$$Ax = i(y + Uy). \quad (125)$$

Die erste dieser Formeln transformiert wie früher $L_1(A)$ eineindeutig in $D(A)$, die zweite gibt das entsprechende Element Ax an. Wir weisen noch darauf hin, daß die durch (124) definierte lineare Mannigfaltigkeit $D(A)$ in H dicht ist.

Den isometrischen Operator U nennen wir die *Cayleytransformierte* des abgeschlossenen symmetrischen Operators A .

Wir beweisen nun einen Satz, der in gewissem Sinne die Umkehrung des vorhergehenden ist.

Satz 2. *Ist U ein isometrischer Operator, der einen Unterraum L' in einen Unterraum L'' transformiert, und definiert die Formel (124) für $y \in L'$ eine in H dichte lineare Mannigfaltigkeit l , so definiert die Formel (125) auf l einen abgeschlossenen symmetrischen Operator A , wobei U die Cayleytransformierte für A ist und L' bzw. L'' mit $L_1(A)$ bzw. $L_{-1}(A)$ übereinstimmen.*

Beweis. Wir müssen zunächst zeigen, daß (124) für verschiedene $y \in L'$ auch verschiedene x ergibt, d. h., daß wie oben aus $y_0 - Uy_0 = 0$ die Beziehung $y_0 = 0$ folgt.

Dazu bilden wir das Skalarprodukt (y_0, x) . Wenn wir nachweisen, daß es für jedes $x \in l$ verschwindet, so muß, da die lineare Mannigfaltigkeit l in H dicht ist, $y_0 = 0$ sein. Tatsächlich ist

$$(y_0, x) = (y_0, y - Uy) = (y_0, y) - (y_0, Uy)$$

oder, da U isometrisch ist,

$$(y_0, x) = (Uy_0, Uy) - (y_0, Uy) = (Uy_0 - y_0, Uy) = (0, Uy) = 0,$$

was zu beweisen war.

Für ein beliebiges $x \in l$ erhalten wir aus (124) ein bestimmtes $y \in L'$ und aus (125) ein bestimmtes Ax . Damit ist der distributive Operator A gebildet. Sind x' und x'' zwei Elemente aus l und y' bzw. y'' die ihnen entsprechenden Elemente aus L' , so ergibt sich, da U isometrisch ist,

$$\begin{aligned} (Ax', x'') &= (i(y' + Uy'), y'' - Uy'') \\ &= i(y', y'') + i(Uy'y'') - i(y', Uy'') - i(Uy', Uy'') \\ &= i(Uy', y'') - i(y', Uy''). \end{aligned}$$

Dasselbe Resultat finden wir für

$$(x', Ax'') = (y' - Uy', i(y'' + Uy'')),$$

d. h., es ist $(Ax', x'') = (x', Ax'')$, also A symmetrisch. Um die Abgeschlossenheit von A zu zeigen, wählen wir aus l eine Folge von Elementen x_n , für die

$$x_n \Rightarrow x \quad \text{und} \quad Ax_n \Rightarrow w \tag{126}$$

ist. Nachzuweisen ist $x \in l$ und $w = Ax$. Das dem Element x_n entsprechende Element aus L' sei y_n , also

$$x_n = y_n - Uy_n, \quad Ax_n = i(y_n + Uy_n); \tag{127}$$

hieraus folgt

$$y_n = \frac{1}{2i} (Ax_n + ix_n) \Rightarrow \frac{1}{2i} (w + ix).$$

Bezeichnen wir dieses Grenzelement zur Abkürzung mit y , so können wir behaupten, daß y zu L' gehört (denn L' ist ein Unterraum) und daß $Uy_n \Rightarrow Uy$ gilt (denn U ist ein isometrischer Operator und folglich $\|U(y - y_n)\| = \|y - y_n\|$). Gehen wir in (127) zur Grenze über, so erhalten wir $x = y - Uy$ und $w = i(y + Uy)$ mit $y \in L'$. Also liegt x in l , und es ist $w = Ax$. Schließlich ergibt sich aus (124) und (125), wenn wir x durch $2ix$ ersetzen,

$$y = (A + iE)x, \quad Uy = (A - iE)x \quad (x \in l).$$

Hieraus folgt sofort, daß U die Cayleytransformierte für A ist und L' bzw. L'' mit $L_1(A)$ bzw. $L_{-1}(A)$ übereinstimmen. Damit ist der Satz vollständig bewiesen.

Die Sätze 1 und 2 lassen erkennen, ob eine Erweiterung eines abgeschlossenen symmetrischen Operators A möglich ist. Es sei B eine abgeschlossene symmetrische Erweiterung von A (die von A selbst verschieden sein soll). Dann sind die rechten Seiten von

$$y = (B + iE)x, \quad z = (B - iE)x$$

auf der $D(A)$ enthaltenden linearen Mannigfaltigkeit $D(B)$ definiert. Für Elemente $x \in D(A)$ ergeben diese Formeln dasselbe Resultat wie die rechten Seiten von (119) und (120). Somit sind die Unterräume $L_1(A)$ und $L_{-1}(A)$ in den Unterräumen $L_1(B)$ bzw. $L_{-1}(B)$ echt enthalten, und bezeichnen wir mit V die Cayleytransformierte des Operators B , so können wir behaupten, daß V die lineare Mannigfaltigkeit $L_1(B)$ in $L_{-1}(B)$ überführt und auf $L_1(A)$ mit U übereinstimmt; d. h., *der isometrische Operator V ist eine Erweiterung des isometrischen Operators U* . Auf Grund von Satz 2 läßt sich auch die Umkehrung behaupten: *Jede Erweiterung des isometrischen Operators U , die auf einen isometrischen Operator V führt, ergibt vermöge der Formeln*

$$x = y - Vy, \quad Bx = i(y + Vy) \quad (y \in D(V)) \quad (128)$$

einen abgeschlossenen symmetrischen Operator B , der eine Erweiterung von A ist. Nach dem oben Gesagten sieht man, daß man nur auf diese Art alle Erweiterungen von A erhalten kann, welche abgeschlossene symmetrische Operatoren sind.

Ist A selbstadjungiert, so stimmen $L_1(A)$ und $L_{-1}(A)$ mit dem ganzen Raum H überein; eine Erweiterung ist dann unmöglich.

202. Defektindizes. Wir bezeichnen mit $M_1(A)$ und $M_{-1}(A)$ die zu $L_1(A)$ bzw. $L_{-1}(A)$ komplementären Unterräume und mit p und q die Dimensionen von $M_1(A)$ bzw. $M_{-1}(A)$. Stimmt zum Beispiel $L_1(A)$ mit H überein, so existiert $M_1(A)$ nicht, und wir nehmen $p = 0$ an.

Ist die lineare Mannigfaltigkeit $M_1(A)$ endlichdimensional, so gibt p ihre Dimension an; ist $M_1(A)$ unendlichdimensional, so ist $p = \infty$. Das Zahlenpaar (p, q) definiert die *Defektindizes*¹⁾ des Operators A . Für sie beweisen wir jetzt einige einfache Sätze.

Satz 1. *Ein symmetrischer abgeschlossener Operator A ist genau dann selbstadjungiert, wenn seine beiden Defektindizes gleich 0 sind.*

Beweis. Ist A ein selbstadjungierter Operator, so bilden, wie wir aus [201] wissen, die Formeln (119) und (120) die lineare Mannigfaltigkeit $D(A)$ in den Raum H ab. Folglich ist $p = q = 0$. Wir nehmen nun an, es sei $p = q = 0$. Dann stimmen $L_1(A)$ und $L_{-1}(A)$ mit H überein, und U ist ein unitärer Operator, der wie U^{-1} auf ganz H definiert ist. Es gelte $(Ax, v) = (x, v^*)$ für alle $x \in D(A)$. Wir müssen zeigen, daß auch v zu $D(A)$ gehört und $v^* = Av$ ist. Die Gleichung $(Ax, v) = (x, v^*)$ läßt sich wegen (124) und (125) in die Form

$$i(y, v) + i(Uy, v) = (y, v^*) - (Uy, v^*)$$

bringen, woraus sich, da U unitär ist,

$$(y, v^* - U^{-1}v^* + iv + iU^{-1}v) = 0$$

ergibt. Da y beliebig ist, gilt also $v^* + iv = U^{-1}(v^* - iv)$. Setzen wir $v^* + iv = 2iy'$, so folgt $v^* - iv = 2iUy'$, also $v = y' - Uy'$ und $v^* = i(y' + Uy')$. Damit liegt v in $D(A)$, und es ist $v^* = Av$. Den Beweis, daß die Bedingung hinreichend ist, können wir auch mit Hilfe der Folgerung von Satz 2 aus [187] erbringen.

Vor dem Beweis der folgenden Sätze klären wir erst die Struktur isometrischer Operatoren. Wie für unitäre Operatoren [137] läßt sich zeigen, daß sich isometrische Transformationen U des Unterraumes L' in den Unterraum L'' auf die Transfor-

¹⁾ Für (p, q) ist ebenfalls die Bezeichnung *Defektindex* gebräuchlich; dabei werden p und q die *Defektzahlen* genannt (Anm. d. Red.).

mation eines abgeschlossenen normierten Orthogonalsystems $x_1, x_2, \dots$ aus L' in ein ebensolches System $y_1, y_2, \dots$ aus L'' zurückführen lassen, so daß $Ux_k = y_k$ (H soll separabel sein) und

$$U \sum_k a_k x_k = \sum_k a_k y_k$$

ist. Dabei müssen beide Unterräume entweder unendliche Dimension oder die gleiche endliche Dimensionszahl haben. Ist diese Bedingung erfüllt, so können wir, da die Wahl der Einheitsvektoren willkürlich ist, unendlich viele isometrische Transformationen von L' in L'' bilden. Gibt es einen isometrischen Operator U von $L_1(A)$ in die lineare Mannigfaltigkeit $L_{-1}(A)$, so läßt er sich nur erweitern, wenn die gleiche Anzahl neuer Einheitsvektoren aus $M_1(A)$ bzw. $M_{-1}(A)$ hinzugefügt und zwischen ihnen paarweise Zuordnungen hergestellt werden. Damit können wir den folgenden Satz formulieren:

Satz 2. Eine symmetrische Erweiterung des abgeschlossenen symmetrischen Operators A ist dann und nur dann möglich, wenn beide Defektindizes des Operators A von 0 verschieden sind. Ist dies der Fall, so gibt es unendlich viele Erweiterungen. Der Operator A läßt sich genau dann zu einem selbstadjungierten Operator erweitern, wenn beide von 0 verschiedenen Defektindizes von A gleich sind; dann gibt es unendlich viele Erweiterungen.

Wir beschreiben nun das allgemeine Schema von Erweiterungen des Operators A . Von $M_1(A)$ und $M_{-1}(A)$ trennen wir zwei Unterräume N_1 bzw. N_{-1} mit gleichen Dimensionen ab und konstruieren einen isometrischen Operator V , der N_1 in N_{-1} überführt. Zu dem erweiterten Operator B definieren wir den Unterraum $L_1(B)$ als orthogonale Summe $L_1(A) \oplus N_1 = L_1(B)$, so daß jedes Element $y \in L_1(B)$ eindeutig als Summe $y = y' + y''$ mit $y' \in L_1(A)$ und $y'' \in N_1$ dargestellt werden kann. Der erweiterte isometrische Operator V wird durch die Formel $Vy = Uy' + Vy''$ gegeben, deren rechte Seite die Zerlegung von $Vy \in L_{-1}(A) \oplus N_{-1}$ in die orthogonalen Unterräume $L_{-1}(A)$ und N_{-1} darstellt. Auf Grund von (124) ist die lineare Mannigfaltigkeit $D(B)$ durch die Beziehung

$$x = y' + y'' - Uy' - Vy'' = (y' - Uy') + (y'' - Vy'')$$

definiert, wobei $y' - Uy'$ ein beliebiges Element aus $D(A)$ und y'' ein beliebiges Element aus N_1 ist. Statt dessen können wir auch

$$x_B = x_A + x_{N_1} - Vx_{N_1} \quad (129)$$

schreiben. Analog ergibt (125)

$$Bx = i(y' + Uy') + iy'' + iVy'',$$

d. h.

$$Bx_B = Ax_A + ix_{N_1} + iVx_{N_1}. \quad (130)$$

Stimmen N_1 und N_{-1} mit $M_1(A)$ bzw. $M_{-1}(A)$ überein, so definieren (129) und (130) eine selbstadjungierte Erweiterung von A . Es läßt sich leicht zeigen, daß die Darstellung von x_B als Summe (129) eindeutig ist. Dazu weisen wir nach, daß jeder der Summanden von (129) gleich dem Nullelement ist, wenn die ganze Summe das Nullelement ergibt. Ist $x_B = 0$, so gilt auch $Bx_B = 0$, und aus (129) und (130) folgt $x_A + x_{N_1} - Vx_{N_1} = 0$ bzw. $Ax_A + ix_{N_1} + iVx_{N_1} = 0$. Multiplikation der ersten Gleichung

mit i und Addition zur zweiten Gleichung liefert

$$(A + iE)x_A + 2ix_{N_1} = 0.$$

Hierin gehört der erste Summand zu $L_1(A)$, der zweite ist orthogonal zu $L_1(A)$; d. h., beide sind gleich 0, also $x_{N_1} = 0$ und daher $Vx_{N_1} = 0$ und $x_A = 0$, was zu beweisen war.

Ist einer der Defektindizes gleich 0, der andere aber nicht, so gestattet A keine abgeschlossene symmetrische Erweiterung. In diesem Fall heißt der Operator A *maximal*. Im Zusammenhang mit dieser Bezeichnung wird manchmal ein selbstadjungierter Operator, bei dem also beide Defektindizes gleich 0 sind, *hypermaximal* genannt.

Wir nehmen nun an, A habe die Defektindizes $(1, 1)$, d. h., die Unterräume $M_1(A)$ und $M_{-1}(A)$ seien eindimensional, und v_0 und w_0 seien beliebige Elemente aus $M_1(A)$ bzw. $M_{-1}(A)$ mit $\|v_0\| = \|w_0\| \neq 0$. Dann sind alle Elemente dieser linearen Mannigfaltigkeit in der Form $v = av_0$, $w = aw_0$ (a beliebig komplex) darstellbar. Die Formel $V(av_0) = e^{-i\theta} aw_0$, wobei θ eine beliebige gegebene Zahl aus dem Intervall $0 \leq \theta < 2\pi$ ist, vermittelt eine isometrische Transformation von $M_1(A)$ in $M_{-1}(A)$. Fügen wir diese Transformation zur Transformation U , die $L_1(A)$ in $L_{-1}(A)$ überführt, hinzu, so erhalten wir einen unitären Operator V . Die Formel (128) definiert einen selbstadjungierten Operator B , der von der Wahl von θ abhängt.

Sind die Defektindizes von A gleich $(2, 2)$, so setzen wir

$$Vv_k = w_k \quad (k = 1, 2),$$

wobei v_1, v_2 zwei zueinander orthogonale Elemente aus $M_1(A)$ und w_1, w_2 zwei solche Elemente aus $M_{-1}(A)$ sind. Halten wir v_k fest und wählen wir alle möglichen w_k , so erhalten wir verschiedene V . Vervollständigen wir die Transformation U zu dem unitären Operator V , so ergibt sich wieder vermöge der Formeln (128) der selbstadjungierte Operator A .

Wir beweisen noch einen Satz, der die Unterräume $M_1(A)$ und $M_{-1}(A)$ neu charakterisiert.

Satz 3. $M_1(A)$ ist der Unterraum der Eigenelemente des Operators A^* , die dem Eigenwert $\lambda = i$ entsprechen, d. h. der Unterraum der Lösungen der Gleichung $A^*x = ix$; ferner ist $M_{-1}(A)$ der Unterraum der Lösungen von $A^*x = -ix$.

Beweis. Da der Operator A^* abgeschlossen ist, ist die lineare Mannigfaltigkeit seiner Eigenelemente, die einem beliebigen Eigenwert entsprechen, stets abgeschlossen, also ein Unterraum. Die Elemente v aus dem Unterraum $M_1(A)$ sind dadurch charakterisiert, daß sie für jedes $x \in D(A)$ zu $(A + iE)x$ orthogonal sind. Für sie gilt infolgedessen $((A + iE)x, v) = 0$ oder, in anderer Form, $(Ax, v) = -(x, iv)$, wobei x ein beliebiges Element aus $D(A)$ ist. Der letzten Gleichung können wir auf Grund der Definition von A^* entnehmen, daß $v \in D(A^*)$ und $A^*v = iv$ ist. Der erste Teil des Satzes ist damit bewiesen. Der zweite Teil läßt sich analog nachweisen.

Aus diesem Satz folgt, daß die Existenz von 0 verschiedener Defektindizes damit zusammenhängt, daß der Operator A^* nicht schon auf $D(A^*)$ symmetrisch ist und i oder $-i$ oder beide Zahlen gleichzeitig als Eigenwerte hat.

An Stelle der Zahlen $\pm i$ könnten wir auch eine beliebige komplexe Zahl λ aus der oberen Halbebene und die zu ihr konjugiert komplexe Zahl $\bar{\lambda}$ nehmen. Dann

vermitteln die Formeln

$$y = (A + \lambda E) x, \quad z = (A + \bar{\lambda} E) x \quad (x \in D(A))$$

eine eindeutige Transformation von $D(A)$ in Unterräume $L_\lambda(A)$ bzw. $L_{\bar{\lambda}}(A)$, und es existiert eine isometrische Transformation $z = Uy$ von $L_\lambda(A)$ in $L_{\bar{\lambda}}(A)$.

Die komplementären Unterräume $M_\lambda(A)$ und $M_{\bar{\lambda}}(A)$ sind die Unterräume der Lösungen von $Ax^* = -\bar{\lambda}x$ bzw. $A^*x = -\lambda x$. Die Dimensionen dieser Unterräume bezeichnen wir mit p_λ bzw. q_λ . Die Zahlen (p_λ, q_λ) sind die Defektindizes von A . Im folgenden wollen wir zeigen, daß sie nicht von der Wahl von λ aus der oberen Halbebene abhängen. Die Formeln (129) und (130) haben die Gestalt

$$x_B = x_A + x_{N_\lambda} - Vx_{N_\lambda}, \quad Bx_B = Ax_A - \bar{\lambda}x_{N_\lambda} + \lambda Vx_{N_\lambda}.$$

Die erste von ihnen definiert $D(B)$, und wir können offenbar

$$D(B) = D(A) + (E - V) N_\lambda \quad (129_1)$$

schreiben, wobei N_λ ein Unterraum von $M_\lambda(A)$ und V ein isometrischer Operator ist, der N_λ in einen in $M_{\bar{\lambda}}(A)$ liegenden Unterraum $N_{\bar{\lambda}}$ abbildet.

Sind L_k ($k = 1, 2, \dots, n$) lineare Mannigfaltigkeiten, so nennen wir die Menge L der Elemente der Form $x = x_1 + x_2 + \dots + x_n$ ($x_k \in L_k$) die *direkte Summe* der L_k , wenn die Darstellung von x eindeutig ist (L ist ebenfalls eine lineare Mannigfaltigkeit). Die Formel (129₁) gibt eine direkte Summe an. Für die direkte Summe wird oft auch das Zeichen $+$ verwendet.

203. Der adjungierte Operator. In Satz 3 aus [202] stellten wir den Zusammenhang zwischen den von uns eingeführten Unterräumen und dem adjungierten Operator her. Wir untersuchen nun die Zusammensetzung der linearen Mannigfaltigkeit $D(A^*)$ und den Zusammenhang zwischen A^* und A . Mit x_A , x_1 und x_{-1} bezeichnen wir beliebige Elemente aus $D(A)$, $M_1(A)$ bzw. $M_{-1}(A)$, und es gilt der folgende

Satz. *Die durch die Formel*

$$v = x_A + x_1 + x_{-1} \quad (131)$$

definierten Elemente bilden die ganze lineare Mannigfaltigkeit $D(A^*)$, und es gilt

$$A^*v = Ax_A + ix_1 - ix_{-1}. \quad (132)$$

Jedes Element aus $D(A^*)$ ist durch (131) eindeutig darstellbar.

Beweis. Da A^* eine Erweiterung von A ist, folgt aus Satz 3 aus [202] unmittelbar, daß die durch (131) gegebenen Elemente v zu $D(A^*)$ gehören und daß (132) gilt. Wir werden zunächst zeigen, daß v durch (131) eindeutig dargestellt ist, d. h., aus

$$x_A + x_1 + x_{-1} = 0. \quad (133)$$

folgt, daß alle drei Summanden gleich dem Nullelement sind.

Wenden wir auf (133) den Operator A^* an, so ergibt sich $Ax_A + ix_1 - ix_{-1} = 0$ oder, wenn wir hierzu die mit i multiplizierte Gleichung (133) addieren,

$$(A + iE)x_A + 2ix_1 = 0.$$

Beide Summanden sind zueinander orthogonal, also müssen beide gleich 0 sein, d. h., es ist $x_1 = 0$. Multiplizieren wir (133) mit $-i$, so ergibt sich $x_{-1} = 0$ und wegen

(133) schließlich auch $x_A = 0$. Nun bleibt nur noch zu zeigen, daß jedes Element $v \in D(A^*)$ durch (131) dargestellt wird. Ein solches Element ist dadurch charakterisiert, daß für alle $x \in D(A)$ die Beziehung $(Ax, v) = (x, v^*)$ oder wegen (124) und (125)

$$(i(y + Uy), v) = (y - Uy, v^*)$$

gilt, woraus sich

$$(Uy, v^* - iv) = (y, v^* + iv) \quad (y \in L_1(A)) \quad (134)$$

ergibt. Projizieren wir in zueinander orthogonale Unterräume, so können wir

$$v^* - iv = x_{L_1} + x_{-1}, \quad v^* + iv = x_{L_1} + x_1 \quad (135)$$

schreiben; dabei ist $x_{L_1} \in L_1(A)$ und $x_{L_1} \in L_{-1}(A)$. Setzen wir (135) in (134) ein, so erhalten wir, wenn wir $x_{-1} \perp Uy$ und $x_1 \perp y$ berücksichtigen, $(Uy, x_{L_1}) = (y, x_{L_1})$ oder mit $x_{L_1} = y'$ und $x_{L_1} = Uy''$ ($y', y'' \in L_1(A)$)

$$(Uy, Uy'') = (y, y').$$

Ferner ist $(y, y'' - y') = 0$ für $y \in L_1(A)$, da U isometrisch ist. Dies gilt insbesondere für $y = y'' - y'$. Damit erhalten wir $y'' = y'$, also $x_{L_1} = y'$ und $x_{L_1} = Uy'$. Setzen wir dies in (135) ein, so finden wir

$$v = \frac{1}{2i} (y' - Uy') + \frac{1}{2i} x_1 - \frac{1}{2i} x_{-1}.$$

Da $\frac{1}{2i} x_1$ zu $M_1(A)$ und $-\frac{1}{2i} x_{-1}$ zu $M_{-1}(A)$ gehört, ist also v durch eine Beziehung der Gestalt (131) darstellbar.

Wir wollen noch auf eine Folgerung aus diesem Satz eingehen. Aus ihm ergibt sich unmittelbar, daß die lineare Mannigfaltigkeit $D(A^*)$ der Elemente v mit Hilfe der Formel

$$w = (A^* - iE)v \quad (136)$$

in den Raum H abgebildet wird. Wegen (131) und (132) erhalten wir tatsächlich $w = (A - iE)x_A - 2ix_{-1}$; der erste Summand ist ein beliebiges Element aus $L_{-1}(A)$, der zweite ein beliebiges Element aus dem komplementären Unterraum $M_{-1}(A)$. Fassen wir (136) als Gleichung in v auf, so hat sie für jedes $w \in H$ Lösungen. Dann besitzt die homogene Gleichung $(A^* - iE)v = 0$ als Lösungen die Elemente des Unterraums $M_1(A)$. Ist er nicht leer, d. h., ist der erste Defektindex p von 0 verschieden, so hat (136) für jedes w unendlich viele Lösungen $v = v_0 + x_1$, wobei v_0 eine spezielle Lösung von (136) und x_1 ein beliebiges Element aus $M_1(A)$ ist. Die spezielle Lösung v_0 hat die Gestalt (131), und da sich zur Lösung ein beliebiges Element aus $M_1(A)$ hinzufügen läßt, können wir annehmen, daß v_0 nicht x_1 enthält, d. h., die Lösung von (136) kann man in der Form

$$v = (x_A + x_{-1}) + x_1$$

schreiben, wobei x_A und x_{-1} bestimmte Elemente sind und x_1 ein beliebiges Element ist. Für $p = 0$ tritt x_1 nicht auf, und es ergibt sich eine wohlbestimmte Lösung. Analog hat die Gleichung

$$(A^* + iE)v = w'$$

für jedes $w' \in H$ die allgemeine Lösung $v' = (x'_A + x'_1) + x'_{-1}$, wobei x'_A und x'_1 bestimmt und x'_{-1} beliebig sind. Ist der zweite Defektindex q gleich 0, so tritt x'_{-1} nicht auf.

Wir weisen noch darauf hin, daß die zu (131) und (132) analogen Formeln

$$v = x_A + x_\lambda + x_{\bar{\lambda}}, \quad A^*v = Ax_A - \lambda x_\lambda - \bar{\lambda} x_{\bar{\lambda}} \quad (\text{Im } \lambda > 0)$$

gelten.

204. Maximale Operatoren. Wir geben ein einfaches Verfahren zur Konstruktion maximaler Operatoren an. Dazu benutzen wir ein abgeschlossenes normiertes Orthogonalsystem

$$x_1, x_2, \dots \quad (137)$$

aus H und definieren eine isometrische Transformation U vermöge $Ux_k = x_{k+1}$ ($k = 1, 2, \dots$); für jedes Element

$$y = \sum_{k=1}^{\infty} a_k x_k \quad \left(\sum_{k=1}^{\infty} |a_k|^2 < \infty \right)$$

aus H gilt also

$$Uy = \sum_{k=1}^{\infty} a_k x_{k+1}.$$

Bei Verwendung der Bezeichnungen aus [201] können wir sagen, L' stimmt mit H überein, L'' wird von allen Einheitsvektoren (137) mit Ausnahme von x_1 gebildet, und die Formeln (124) und (125), in denen y beliebig aus H ist, führen uns auf einen abgeschlossenen symmetrischen Operator A mit den Defektindizes $(0, 1)$. Wir brauchen nur nachzuprüfen, ob die von den Elementen $y - Uy$ gebildete lineare Mannigfaltigkeit $D(A)$ in H dicht ist, d. h., ob es in $D(A)$ solche Elemente x gibt, daß für einen gegebenen Einheitsvektor x_k die Norm $\|x_k - x\|$ hinreichend klein wird. Wir bilden das Element

$$y = \sum_{s=0}^{m-1} \frac{m-s}{m} x_{k+s}$$

(m eine gewisse ganze positive Zahl). Dann ist

$$x = y - Uy = \sum_{s=0}^{m-1} \frac{m-s}{m} x_{k+s} - \sum_{s=0}^{m-1} \frac{m-s}{m} x_{k+s+1} = x_k - \frac{1}{m} \sum_{s=1}^m x_{k+s};$$

hieraus folgt auf Grund des Satzes des PYTHAGORAS und der Normiertheit der x_{k+s}

$$\|x_k - x\|^2 = \frac{1}{m^2} \sum_{s=1}^m \|x_{k+s}\|^2 = \frac{1}{m}.$$

Lassen wir m unbeschränkt wachsen, so erhalten wir beliebig kleine Werte für $\|x_k - x\|$, was zu beweisen war.

Einen maximalen Operator A dieser Art nennen wir einen *elementaren symmetrischen* Operator. Wird $B = -A$ gesetzt, so stimmt $D(B)$ mit $D(A)$ überein, und statt (119) und (120) gilt jetzt

$$y = (-B + iE)x, \quad z = (-B - iE)x.$$

Hieran sehen wir, wenn wir x durch $-x$ ersetzen, daß $(B + iE)x$ die lineare Mannigfaltigkeit $D(B)$ in den Unterraum $L_{-1}(A)$ und $(B - iE)x$ die lineare Mannigfaltigkeit $D(B)$ in $L_1(A)$ abbildet, d. h., bei Ersetzen von A durch $-A$ werden die Unterräume L_1 und L_{-1} vertauscht. Ist also A der oben erwähnte Operator mit den Defektindizes $(0, 1)$, so hat $-A$ die Defektindizes $(1, 0)$.

Es sei U_0 ein unitärer Operator, der die x_k in x'_k überführt. Wenden wir das vorhin benutzte Verfahren auf die x'_k an, so erhalten wir einen isometrischen Operator U' mit $U'x'_k = x'_{k+1}$ und einen elementaren symmetrischen Operator A' , wobei offenbar $U' = U_0 U U_0^{-1}$ und $A' = U_0 A U_0^{-1}$ ist; $D(A')$ ergibt sich aus $D(A)$ vermöge des Operators U_0 . Wir können zeigen, daß sich H als orthogonale Summe von Unterräumen darstellen läßt,

$$H = L_0 \oplus L_1 \oplus \dots \oplus L_q,$$

wenn A ein beliebiger abgeschlossener symmetrischer Operator mit den Defektindizes $(0, q)$ ist ($q > 0$ und endlich). Diese Unterräume reduzieren den Operator A und sind so beschaffen, daß jedes L_k für $k \geq 1$ unendlichdimensional und der durch A in L_k induzierte Operator A_k ein elementarer symmetrischer Operator ist. Der Unterraum L_0 , der nicht zu existieren braucht, kann sowohl unendlich- als auch endlichdimensional sein, und der in ihm induzierte Operator A_0 ist selbstadjungiert. Ein analoges Resultat findet man auch für $q = \infty$.

Sind $(p, 0)$ mit $p > 0$ die Defektindizes, so sind die Operatoren A_k zu elementaren symmetrischen Operatoren in bezug auf das Vorzeichen invers.

205. Erweiterung symmetrischer halbbeschränkter Operatoren. Es sei A ein symmetrischer halbbeschränkter Operator mit der unteren Schranke m_A :

$$(Ax, x) \geq m_A(x, x). \quad (138)$$

Wir werden vorläufig $m_A > 0$, d. h. den Operator A positiv definit annehmen und ihn so zu erweitern suchen, daß er symmetrisch bleibt und sein Wertebereich mit ganz H übereinstimmt. Eine solche Erweiterung führt, wie wir aus [187] wissen, auf einen selbstadjungierten Operator.

Dem Operator A ordnen wir ein quadratisches Funktional mit reellen Werten vermöge der Formel

$$J_y(x) = (Ax, x) - (y, x) - (x, y) \quad (x \in D(A)) \quad (139)$$

zu, wobei y ein festes Element aus H ist, und betrachten für dieses Funktional das Minimumproblem.

Satz 1. *Hat die Gleichung*

$$Ax = y \quad (140)$$

eine Lösung x aus $D(A)$, so ist $J_y(x) \leq J_y(z)$; dabei ist z ein beliebiges Element aus $D(A)$, und das Gleichheitszeichen gilt nur für $z = x$. Gilt umgekehrt für ein $x \in D(A)$ die Ungleichung $J_y(x) \leq J_y(z)$, wobei z ein beliebiges Element aus $D(A)$ ist, so genügt x der Gleichung (140).

Beweis. Es genüge $x \in D(A)$ der Gleichung (140). Für jedes $z \in D(A)$ gilt, da A symmetrisch ist,

$$\begin{aligned} J_y(z) &= (Az, z) - (y, z) - (z, y) = (Az, z) - (Ax, z) - (z, Ax) \\ &= (A(x-z), x-z) - (Ax, x) = (A(x-z), x-z) + J_y(x), \end{aligned}$$

und aus

$$(A(x-z), x-z) \geq m_A \|x-z\|^2$$

und $m_A > 0$ folgt der erste Teil des Satzes.

Ist umgekehrt $x \in D(A)$ und $J_y(x) \leq J_y(z)$ für alle $z \in D(A)$, so hat das quadratische Funktional $J_y(x+tz)$, wenn der reelle Parameter t gleich 0 ist, ein Minimum

für jedes feste $z \in D(A)$. Hieraus folgt

$$(Az, x) + (Ax, z) - (y, z) - (z, y) = 0$$

oder

$$(Ax, z) + (z, Ax) - (y, z) - (z, y) = 0,$$

d. h., es ist $\operatorname{Re} (Ax - y, z) = 0$. Wenn wir nun z durch iz ersetzen, dann erhalten wir $\operatorname{Im} (Ax - y, z) = 0$, also $(Ax - y, z) = 0$ oder, da $D(A)$ in H dicht ist, $Ax - y = 0$. Damit ist auch der zweite Teil des Satzes bewiesen.

Aus diesem Satz folgt nicht, daß das Funktional $J_y(x)$ für jedes $y \in H$ und $x \in D(A)$ seinen kleinsten Wert erreicht. Daher werden wir im folgenden den Definitionsbereich des Funktional vergrößern. Wegen (139) ist es auf $D(A)$ definiert. Wir führen in $D(A)$ ein neues Skalarprodukt ein, indem wir

$$[x, y]_A = (Ax, y) \quad \text{oder einfach} \quad [x, y] = (Ax, y) \quad (141)$$

setzen, und eine neue Norm

$$\|x\|_A^2 = [x, x] = (Ax, x). \quad (142)$$

Auf Grund von (138) gilt die Ungleichung

$$\|x\| \leq \frac{1}{\sqrt{m_A}} \|x\|_A. \quad (143)$$

Es läßt sich leicht nachweisen, daß die Elemente der linearen Mannigfaltigkeit $D(A)$ bei der früheren Definition der Multiplikation mit einer Zahl und Addition und bei dem Skalarprodukt (141) alle Axiome des Hilbertschen Raumes erfüllen, außer eventuell das Vollständigkeitsaxiom. Ist dieses Axiom nicht erfüllt, so können wir $D(A)$ durch neue ideale Elemente so vervollständigen, daß sich ein vollständiger Hilbertscher Raum ergibt, den wir mit H_A bezeichnen. Wir wollen nun diese Vervollständigung näher betrachten [85]. In $D(A)$ existiere eine Fundamentalfolge mit dem Skalarprodukt (141). Wegen (143) ist es auch eine Fundamentalfolge in H , die in H (auf Grund der Vollständigkeit von H) ein Grenzelement x' besitzt. Die Fundamentalfolgen aus $D(A)$ mit dem Skalarprodukt (141), die in ein und derselben Klasse liegen, konvergieren gegen das gleiche Element $x' \in H$, d. h., gehören x_n, y_n ($n = 1, 2, \dots$) zu $D(A)$ und ist $\|y_n - x_n\|_A \rightarrow 0$ für $n \rightarrow \infty$, so geht auch $\|y_n - x_n\|$ gegen 0. Dies folgt aus (143). Wir weisen nun noch nach, daß den zu verschiedenen Klassen gehörenden Fundamentalfolgen von Elementen x_n und y_n aus $D(A)$ mit dem Skalarprodukt (141) verschiedene Grenzwerte x' bzw. y' aus H entsprechen.

Für jedes $z \in D(A)$ gilt

$$(Az, x_n - y_n) = [z, x_n - y_n];$$

durch Grenzübergang erhalten wir

$$(Az, x' - y') = [z, v],$$

wobei rechts das Element v aus H_A und von 0 verschieden ist, da die Folgen der Elemente x_n bzw. y_n zu verschiedenen Klassen in H_A gehören. Wäre nämlich $x' = y'$, so hätten wir $[z, v] = 0$. Dies ist aber nicht möglich, denn $D(A)$ ist in H_A dicht, und das Element $v \in H_A$ ist vom Nullelement verschieden.

Wir erweitern die Definition des Funktional $J_y(x)$ auf ganz H_A , indem wir

$$J_A(x) = [x, x] - (y, x) - (x, y) \quad (144)$$

setzen. Für dieses Funktional untersuchen wir das Minimumproblem. Die Größe (x, y) ist für jedes feste $y \in H$ ein lineares beschränktes Funktional über x in H_A ,

denn es ist

$$|(x, y)| \leq \|y\| \cdot \|x\| \leq \frac{1}{\sqrt{m_A}} \|y\| \cdot \|x\|_A,$$

und nach einem bekannten Satz [123] existiert ein einziges Element $x_0 \in H_A$, für das

$$(x, y) = [x, x_0] \quad (x \in H_A, y \in H) \quad (145)$$

und infolgedessen $(y, x) = [x_0, x]$ ist. Den Ausdruck (144) können wir auf die Gestalt

$$J_y(x) = [x, x] - [x_0, x] - [x, x_0] = [x - x_0, x - x_0] - [x_0, x_0]$$

bringen, und wir sehen, daß $J_y(x_0) \equiv J_y(x)$ für alle $x \in H_A$ gilt, wobei Gleichheit nur für $x = x_0$ eintritt. Außerdem folgt aus (145) und der Tatsache, daß H_A in H dicht liegt, die Zuordnung verschiedener x_0 zu verschiedenen $y \in H$.

Ferner ergibt sich aus (145) unmittelbar, daß die Menge aller Lösungen x_0 der Variationsprobleme, die allen möglichen $y \in H$ entsprechen, eine lineare Mannigfaltigkeit l ist. Das oben Gesagte ermöglicht es, auf l einen distributiven Operator $\tilde{A}$ vermöge

$$\tilde{A}x_0 = y \quad (146)$$

zu definieren, der auf ganz H ein Inverses besitzt. Auf Grund von Satz 1 ist $\tilde{A}$ eine Erweiterung von A . Statt l können wir natürlich auch $D(\tilde{A})$ schreiben. Es gilt $D(\tilde{A}) \subseteq H_A$. Der Operator $\tilde{A}$ ist in H symmetrisch, denn aus (145) folgt

$$(\tilde{A}x_0, x) = [x_0, x] \quad (x_0 \in D(\tilde{A}), x \in H_A), \quad (147)$$

woraus sich für $x = x_0$

$$(\tilde{A}x_0, x_0) \geq 0$$

ergibt, d. h., $(\tilde{A}x_0, x_0)$ ist für $x_0 \in D(\tilde{A})$ reell; damit ist $\tilde{A}$ symmetrisch [187]. Das Inverse $\tilde{A}^{-1}$ ist in ganz H definiert, ebenfalls symmetrisch und infolgedessen ein beschränkter selbstadjungierter Operator; also ist auch $\tilde{A}$ in H selbstadjungiert. Damit gelangen wir zu dem folgenden Satz.

Satz 2. *Ein symmetrischer Operator A , der der Bedingung (138) für $m_A > 0$ genügt, gestattet eine selbstadjungierte Erweiterung $\tilde{A}$ derart, daß $\tilde{A}^{-1}$ in ganz H definiert und beschränkt ist.*

Die oben durchgeführte Konstruktion der Erweiterung eines halbbeschränkten symmetrischen Operators stammt von FRIEDRICHS¹⁾, der Beweis von MICHLIN²⁾.

Wir nehmen nun an, daß der symmetrische Operator A der Forderung (138) für $m_A \leq 0$ genügt. Dann befriedigt der symmetrische Operator $B = A + (\varepsilon - m_A) E$ ($\varepsilon > 0$), für den $D(B) = D(A)$ ist, die Bedingung $(Bx, x) \geq \varepsilon(x, x)$ für $x \in D(B)$, und der oben bewiesene Satz führt uns auf den folgenden

Satz 3. *Jeder symmetrische halbbeschränkte Operator A erlaubt eine selbstadjungierte Erweiterung $\tilde{A}$ derart, daß für jedes $\varepsilon > 0$ der Operator $[\tilde{A} + (\varepsilon - m_A) E]^{-1}$ auf ganz H definiert und beschränkt ist.*

Der Operator A läßt, wenn er nicht selbstadjungiert ist, unendlich viele selbst-

¹⁾ K. FRIEDRICHS, Spektraltheorie halbbeschränkter Operatoren, Math. Ann. 109 (1934), 465–487, 685–713.

²⁾ S. G. MICHLIN, Das Minimumproblem eines quadratischen Funktionals (russ.), Gostechisdat, Moskau-Leningrad 1952.

adjungierte Erweiterungen zu. Die obige Erweiterung $\tilde{A}$ heißt im allgemeinen *Erweiterung nach FRIEDRICHS*.

Liegt x in $D(A)$, so gilt (143) für $m_A > 0$. Nehmen wir nun $x \in H_A$, aber $x \notin D(A)$ an, so gibt es eine Folge von Elementen $x_n \in D(A)$ ($n = 1, 2, \dots$) derart, daß $x_n \rightarrow x$ der Norm von H_A nach und damit erst recht der Norm von H nach ist.

Schreiben wir die Ungleichung (143) für x_n und benutzen wir die Stetigkeit der Norm, so können wir behaupten, daß für $m_A > 0$ die Ungleichung (143) in ganz H_A gilt.

Ist $x \in D(\tilde{A})$ und damit auch $x \in H_A$, so ergibt sich aus (145)

$$(\tilde{A}x, x) = [x, x] = \|x\|_A^2$$

und aus (143)

$$(\tilde{A}x, x) \geq m_A(x, x). \quad (148)$$

Zum Beweis setzten wir $m_A > 0$ voraus. Ist $m_A \leq 0$, so können wir den Operator $B = A + (\varepsilon - m_A) E$ ($\varepsilon > 0$) konstruieren. Auf Grund des oben Gesagten gilt $(\tilde{B}x, x) \geq \varepsilon(x, x)$ mit $\tilde{B} = \tilde{A} + (\varepsilon - m_A) E$, woraus sich (148) ergibt. Damit erhalten wir das folgende Resultat.

Satz 4. Für $\tilde{A}$ gilt die Ungleichung (148).

Wir untersuchen nun den Fall $m_A > 0$ und wollen den folgenden Satz beweisen.

Satz 5. Für einen positiv definiten symmetrischen Operator A gilt $H_A = D(\tilde{A}^{1/2})$ und

$$[x, x] = \|\tilde{A}^{1/2}x\|^2. \quad (149)$$

Beweis. Wir zeigen zunächst $H_A \subset D(\tilde{A}^{1/2})$. Dazu wählen wir ein x aus H_A . Es gibt eine Folge von Elementen $x_n \in D(\tilde{A})$, $n = 1, 2, \dots$, derart, daß $\|x - x_n\|_A \rightarrow 0$ ist; für x_n gilt

$$[x_n, x_n] = (\tilde{A}x_n, x_n) = \|\tilde{A}^{1/2}x_n\|^2 = \int_{m_A=0}^{\infty} \lambda d(\mathcal{E}_\lambda x_n, x_n), \quad (150)$$

wobei $\mathcal{E}_\lambda$ die Spektralschar von $\tilde{A}$ ist. Wegen

$$\|x - x_n\| \rightarrow 0, \quad \|x_n - x_m\|_A^2 = \|\tilde{A}^{1/2}(x_n - x_m)\|^2 \rightarrow 0$$

und der Abgeschlossenheit von $\tilde{A}^{1/2}$ gilt $x \in D(\tilde{A}^{1/2})$, $\tilde{A}^{1/2}x_n \rightarrow \tilde{A}^{1/2}x$ und $\|\tilde{A}^{1/2}x\| = \|x\|_A$ (Stetigkeit der Norm), d. h., x genügt der Beziehung (149). Damit der Satz vollständig bewiesen ist, müssen wir noch nachweisen, daß $D(\tilde{A}^{1/2}) \subset H_A$ ist. Wir nehmen $x \in D(\tilde{A}^{1/2})$ an und zeigen, daß dann auch x zu H_A gehört. Dazu betrachten wir die Folge der Elemente $x_n = \mathcal{E}_n x$ ($n = 1, 2, \dots$). Daß $x_n \in D(\tilde{A})$ und $\|x - x_n\| \rightarrow 0$ für $n \rightarrow \infty$ gilt, sehen wir sofort. Wegen der Konvergenz des Integrals

$$\int_{m_A=0}^{\infty} \lambda d(\mathcal{E}_\lambda x, x)$$

konvergieren die Elemente

$$\tilde{A}^{1/2}x_n = \int_{m_A=0}^n \sqrt{\lambda} d\mathcal{E}_\lambda x$$

für $n \rightarrow \infty$ gegen $\tilde{A}^{1/2}x$ in H , und daher ist

$$\|\tilde{A}^{1/2}(x_n - x_m)\|^2 = (\tilde{A}(x_n - x_m), x_n - x_m) = \|x_n - x_m\|_A^2 \rightarrow 0$$

für $m, n \rightarrow \infty$. Das bedeutet, die Elemente x_n bilden eine Fundamentalfolge in H_A . Ist $\tilde{x}$ das ihr entsprechende Element in H_A , so konvergiert in H_A die Folge der x_n stark gegen $\tilde{x}$. Wie wir aber oben sahen, ist $x_n \Rightarrow x$ in H , so daß also $x = \tilde{x} \in H_A$ gilt, was zu beweisen war.

Für $m_A > 0$ ist der Operator $\tilde{A}^{-1}$ auf ganz H definiert und beschränkt. Wir wollen nun untersuchen, wann er vollstetig ist, und führen dazu den Operator W ein, der jedem Element $x \in H_A$ das gleiche Element x , aber als Element aus H zuordnet.

Satz 6. *Der Operator $\tilde{A}^{-1}$ ist genau dann vollstetig, wenn der Operator W der Einbettung von H_A in H vollstetig ist.*

Beweis. Wir zeigen zuerst die Notwendigkeit der Bedingung. Es sei $\tilde{A}^{-1}$ ein vollstetiger Operator. Er hat ein reines Punktspektrum, das zum Intervall $[0, 1/m_A]$ gehört, alle seine von 0 verschiedenen Eigenwerte haben endlichen Rang, und nur der Punkt $\lambda = 0$ ist Verdichtungspunkt des Spektrums [136]. Das Spektrum des selbstadjungierten Operators $\tilde{A}^{-1/2}$ hat denselben Charakter: Jeder Eigenwert λ wird durch $\sqrt{\lambda}$ ersetzt, und die Eigenelemente bleiben unverändert. Somit ist $\tilde{A}^{-1/2}$ ein vollstetiger positiver Operator. Wir wählen nun eine beliebige, in H_A beschränkte Menge U derart, daß für $x \in U$ die Ungleichung $\|x\|_A \leq c$ gilt, wobei c eine bestimmte Konstante ist. Wir können auf x den Operator $\tilde{A}^{1/2}$ anwenden. Es sei $y = \tilde{A}^{1/2}x$, so daß $x = \tilde{A}^{-1/2}y$ gilt. Dann ist $\|y\|^2 = (\tilde{A}^{1/2}x, \tilde{A}^{1/2}x) = \|x\|_A^2 \leq c^2$, d. h., die Menge der Elemente $\tilde{A}^{1/2}x$ ist in H beschränkt, und der vollstetige Operator $\tilde{A}^{-1/2}$ bildet sie in eine in H kompakte Menge ab. Tatsächlich ist einer der Norm von H_A nach beschränkte Menge aus H_A in H kompakt. Also ist der Operator W vollstetig.

Nun zeigen wir, daß die Bedingung hinreichend ist. Es sei W ein vollstetiger Operator und U eine in H beschränkte Menge von Elementen; d. h., für $x \in U$ sei $\|x\| \leq c$. Wir müssen also zeigen, daß die Menge der Elemente $y = \tilde{A}^{-1}x$ in H kompakt ist. Die Elemente y gehören offenbar zu $D(\tilde{A})$, und es gilt

$$\|y\|_A^2 = (\tilde{A}y, y) = (x, y) \leq c\|y\| \leq \frac{c}{\sqrt{m_A}} \|y\|_A,$$

woraus

$$\|y\|_A \leq \frac{c}{\sqrt{m_A}}$$

folgt, d. h., die Menge der Elemente y ist der Norm von H_A nach beschränkt und, da W vollstetig ist, in H kompakt. Damit ist der Satz vollständig bewiesen.

206. Vergleich halbbeschränkter Operatoren. Es seien A und B halbbeschränkte selbstadjungierte Operatoren. Wir sagen, A sei nicht kleiner als B , und schreiben $A \geq B$, wenn $D(A) \subseteq D(B)$ und

$$(Ax, x) \geq (Bx, x) \quad \text{für } x \in D(A) \quad (151)$$

gilt. Haben A und B ein reines Punktspektrum und kann man die Eigenwerte von A und B in nichtfallender Reihenfolge unter Berücksichtigung der Vielfachheit numerieren, so können wir zeigen, wenn wir das Minimum-Maximum-Prinzip [136] auf nichtbeschränkte Operatoren übertragen, daß $\lambda_n(A) \geq \lambda_n(B)$ ist, wobei $\lambda_n(A)$ und $\lambda_n(B)$ der n -te Eigenwert von A bzw. B ist.

Wir beweisen nun einen etwas allgemeineren Satz.

Satz. Es sei $A \cong B$, und das Spektrum von B , das sich auf der Halbgeraden $\lambda < \beta$ (β beliebig) befindet, bestehe aus Eigenwerten endlicher Vielfachheit, die keine Verdichtungspunkte links von β haben mögen. Dann besitzt das Spektrum von A dieselbe Eigenschaft, und auf der Halbgeraden $\lambda < \beta$ gilt $\lambda_n(A) \cong \lambda_n(B)$.

Beweis. Wir brauchen nur zu zeigen, daß für ein beliebiges, aber von den Eigenwerten von A und B verschiedenes $\delta > \beta$ die Dimension des dem Projektionsoperator $\mathcal{E}_\delta$ entsprechenden Unterraums nicht größer ist als die Dimension des dem Projektionsoperator F_δ entsprechenden Unterraums, wobei $\mathcal{E}_\lambda$ und F_λ die Spektralscharen von A bzw. B sind, d. h., wir brauchen nur die Beziehung

$$\dim \mathcal{E}_\delta H \leq \dim F_\delta H \quad (152)$$

nachzuweisen. Den Beweis führen wir indirekt; wir nehmen also an, es gelte die umgekehrte Ungleichung. Dann existiert in $\mathcal{E}_\delta H$ ein normiertes Element x_0 , das zu ganz $F_\delta H$ orthogonal ist. Nun ist $x_0 \in D(A)$ und somit $x_0 \in D(B)$, denn x_0 gehört zu

$$\mathcal{E}_\delta H = (\mathcal{E}_\delta - \mathcal{E}_{m_A - 0}) H.$$

Ferner gilt

$$(Ax_0, x_0) = \int_{m_A - 0}^{\infty} \lambda d(\mathcal{E}_\lambda x_0, x_0) = \int_{m_A - 0}^{\delta} \lambda d(\mathcal{E}_\lambda x_0, x_0) \leq \delta \|x_0\| = \delta. \quad (153)$$

Andererseits ist wegen $x_0 \perp F_\delta H$

$$(Bx_0, x_0) = \int_{m_B - 0}^{\infty} \lambda d(F_\lambda x_0, x_0) = \int_{\delta}^{\infty} \lambda d(F_\lambda x_0, x_0).$$

Da F_λ in einer Umgebung des Punktes $\lambda = \delta$ konstant ist, gibt es ein $\varepsilon > 0$ derart, daß

$$(Bx_0, x_0) = \int_{\delta + \varepsilon}^{\infty} \lambda d(F_\lambda x_0, x_0) \geq \delta + \varepsilon$$

gilt. Diese Ungleichung steht aber in Widerspruch zu (151) und (153). Damit ist (152) bewiesen.

Bemerkung. Wir kehren nun wieder zu symmetrischen halbbeschränkten Operatoren zurück. Es sei A ein solcher (nicht unbedingt positiv definiter) Operator. Für ihn definieren wir den Raum H_A . Es sei α eine beliebige Zahl, die die Ungleichung $\alpha > -m_A$ erfüllt, so daß der Operator $A + \alpha E$ positiv definit ist. Wir werden annehmen, daß H_A aus allen Elementen von $H_{A+\alpha E}$ besteht, und in H_A ein bilineares Funktional

$$[x, y]_A = [x, y]_{A+\alpha E} - \alpha(x, y) \quad (154)$$

eingeführen, das eine Erweiterung von (Ax, y) auf ganz H_A ist. Das Funktional (154) ist auf H_A stetig. Ferner können wir leicht die Beziehung

$$[x, y]_A = \int_{m_A - 0}^{\infty} \lambda d(\mathcal{E}_\lambda x, y) \quad (155)$$

nachweisen, wobei $\mathcal{E}_\lambda$ die Spektralschar des selbstadjungierten Operators $\tilde{A}$ ist. Wir weisen noch darauf hin, daß $H_{A+\alpha E}$ für alle $\alpha > -m_A$ aus ein und denselben Elementen besteht. Das folgt daraus, daß $D(A + \alpha E)$ nicht von α abhängt und die Normen von $H_{A+\alpha E}$ für alle $\alpha > -m_A$ äquivalent sind. Der Operator A kann auch selbstadjungiert sein.

Die Bedingung (151) ist mit

$$[x, x]_A \cong [x, x]_B$$

äquivalent, und das Spektrum des selbstadjungierten Operators $\tilde{A}$ auf der Halbgeraden $\lambda < \beta$, wobei β in dem obigen Satz definiert ist, kann man als aufeinanderfolgende untere Schranken von (Ax, x) für $x \in D(A)$ und $\|x\| = 1$ bestimmen, wenn wir berücksichtigen, daß x zu den schon gefundenen Eigenelementen orthogonal ist (vgl. [136]). Dieses Problem läßt sich auch durch das der sukzessiven Minima von $[x, x]_A$ für $x \in H_A$, $\|x\| = 1$ ersetzen.

207. Beispiele aus der Theorie der Erweiterungen. 1. Wir zeigten früher [188], daß der Operator $D = i \frac{d}{dx}$ im Raum $H = L_2(0, \infty)$ keine selbstadjungierten Erweiterungen hat.

Wir übernehmen nun die Bezeichnungen aus [188] und beweisen den oben zitierten Satz mit Hilfe der Defektindizes. Wie wir wissen, stimmt der abgeschlossene symmetrische Operator $\tilde{A}$ mit dem Operator D überein, der auf der Menge der Funktionen $\varphi(x)$ definiert ist, die auf jedem endlichen Intervall $[0, a]$ absolut stetig sind, in $L_2(0, \infty)$ eine Ableitung besitzen und der Bedingung $\varphi(0) = 0$ genügen. Der Operator $\tilde{A}^*$ ist auf der Menge der Funktionen $\varphi(x)$ gegeben, die allen obigen Forderungen, außer der Bedingung $\varphi(0) = 0$, genügen.

Wir konstruieren die Räume $M_1(\tilde{A})$ und $M_{-1}(\tilde{A})$ der Eigenelemente von $\tilde{A}^*$, die den Eigenwerten $\pm i$ entsprechen, d. h. die Räume der Lösungen von $\tilde{A}^* \psi(x) = \pm i \psi(x)$ oder der von $i \psi'(x) = \pm \psi(x)$. Daraus ergibt sich $\psi(x) = ce^x$ und $\psi(x) = ce^{-x}$. Jedoch gehört e^x nicht zu $L_2(0, \infty)$; somit sind $(0, 1)$ die Defektindizes von $\tilde{A}$. Der Operator $\tilde{A}$ ist maximal, $L_1(\tilde{A})$ ist der ganze Raum $L_2(0, \infty)$, und $L_{-1}(\tilde{A})$ besteht aus den Funktionen, die zu $L_2(0, \infty)$ gehören und auf dem Intervall $(0, \infty)$ zu e^{-x} orthogonal sind.

Führen wir das normierte Orthogonalsystem der Laguerreschen Funktionen $\varphi_k(x) = e^{-x} p_k(x)$, $k = 0, 1, 2, \dots$, ein, wobei $p_k(x)$ ein Polynom k -ten Grades ist, so können wir leicht zeigen, daß $U \varphi_k(x) = \varphi_{k+1}(x)$ gilt; dabei ist U ein isometrischer Operator, der $L_1(\tilde{A})$ in $L_{-1}(\tilde{A})$ abbildet, d. h., $\tilde{A}$ ist ein elementarer symmetrischer Operator.

2. Wir untersuchen den Operator $L(y) = -y''$ im Raum $H = L_2(-\infty, \infty)$. Mit A bezeichnen wir diesen Operator auf der linearen Mannigfaltigkeit $D(A)$ der finiten Funktionen, die stetige Ableitungen bis zur zweiten Ordnung haben. Der Operator A ist symmetrisch. Der zu ihm adjungierte Operator A^* ist, wie man leicht zeigen kann, der gleiche Operator $L(y)$, der auf der linearen Mannigfaltigkeit der Funktionen $y(x)$ definiert ist, welche folgende Eigenschaften besitzen: $y(x)$ und $y'(x)$ sind auf jedem endlichen Intervall absolut stetig, und $y(x)$ und $y''(x)$ gehören zu $L_2(-\infty, \infty)$. Man kann zeigen, daß $y'(x)$ in $L_2(-\infty, \infty)$ liegt. Der Operator $A^{**} = \tilde{A}$ stimmt mit A^* überein, d. h., $\tilde{A}$ ist selbstadjungiert (vgl. [188]). Die Gleichungen $-y'' = \pm i y$ können in $L_2(-\infty, \infty)$ nicht gelöst werden.

Wir betrachten nun den gleichen Operator $L(y)$ auf dem Intervall $[0, \infty]$. Es sei l' die lineare Mannigfaltigkeit der Funktionen $y(x)$ mit den folgenden Eigenschaften: $y(x)$ und $y'(x)$ sind auf jedem endlichen Intervall $[0, a]$ absolut stetig; $y(x)$ und $y''(x)$ gehören zu $L_2(0, \infty)$. Wir definieren noch die lineare Mannigfaltigkeit l der $y(x) \in l'$, die den Bedingungen $y(0) = y'(0) = 0$ und

$$\lim_{x \rightarrow \infty} (-y'z + yz') = 0$$

für jedes $z(x) \in l'$ genügen.

Bezeichnen wir den Operator $L(y)$ auf l mit A , so ist A^* der gleiche Operator, aber auf l' . Der Operator A ist abgeschlossen und symmetrisch. Die Gleichungen $-y'' = \pm i y$ haben in $L_2(0, \infty)$ (bis auf einen konstanten Faktor) genau eine Lösung

$$y = e^{-\frac{\sqrt{2}}{2}(1 \mp i)x},$$

so daß die Defektindizes von A gleich $(1, 1)$ sind. Um eine selbstadjungierte Erweiterung zu erreichen, muß bei $x=0$ eine Randbedingung aufgestellt werden. Bei der Forderung $y(0)=0$ hat der Operator kein Punktspektrum; das kontinuierliche Spektrum füllt die Halbgerade $\lambda \geq 0$ aus. Es existiert die einzige Differentiallösung

$$y(x) = \frac{1}{x} \left(\sqrt{\lambda} \cos \lambda x - \frac{1}{x} \sin \sqrt{\lambda} x \right).$$

Konstruieren wir die Resolvente, d. h. die Lösung der Gleichung $-y'' + (\sigma + \lambda i)y = f(x)$ unter den Bedingungen $y(0)=0$, $\tau > 0$, und gehen wir zur Grenze über, so erhalten wir die Spektralschar

$$\mathcal{E}_\lambda f(x) = \frac{1}{\pi} \int_0^\infty \frac{\sin \sqrt{2\lambda}(x-t)}{x-t} f(t) dt - \frac{1}{\pi} \int_0^\infty \frac{\sin \sqrt{\lambda}(x+t)}{x+t} f(t) dt.$$

Alle diese Resultate ergeben sich mit Hilfe einfacher Rechnungen. Auch können wir leicht nachweisen, daß $y'(x)$ in $L_2(0, \infty)$ liegt, wenn $y(x) \in D(A^*)$ ist.

3. In [188] betrachteten wir den Laplaceschen Operator im Raum $L_2(D)$. Für den dort erwähnten Operator A sind alle Bedingungen von Satz 3 aus [187] erfüllt, und wir untersuchen nun für A eine selbstadjungierte Erweiterung nach FRIEDRICHS. Der Raum H_A ergibt sich durch Vervollständigung von $D(A)$ bezüglich der Metrik:

$$\|u\|_A = \sqrt{(Au, u)} = \sqrt{\int_D -\Delta u \cdot \bar{u} dx} = \sqrt{\int_D \sum_i |u_{x_i}|^2 dx}.$$

Diese Norm ist derjenigen von $\dot{W}_2^{(1)}(D)$ äquivalent [114]; folglich stimmen H_A und $\dot{W}_2^{(1)}(D)$ überein ($\dot{W}_2^{(1)}(D)$ ist die Vervollständigung der Menge aller einmal stetig differenzierbaren finiten Funktionen bezüglich der Norm von $\dot{W}_2^{(1)}(D)$). Man kann leicht sehen (mit Hilfe des Mittelungsprozesses), daß man als Ausgangsmenge die unendlich oft stetig differenzierbaren finiten Funktionen wählen kann. Das Funktional $J_f(u)$ hat für $u \in D(A)$ die Form

$$J_f(u) = \int_D [-\Delta u \cdot \bar{u} - 2\operatorname{Re}(u, \bar{f})] dx, \quad (156)$$

wobei $f(x) \in L_2(D)$ ist, oder

$$J_f(u) = \int_D \left[\sum_i |u_{x_i}|^2 - 2\operatorname{Re}(u, \bar{f}) \right] dx. \quad (157)$$

In dieser Gestalt ist das Funktional für jede Funktion $u \in \dot{W}_2^{(1)}(D)$ sinnvoll; das Variationsproblem aus [205] besteht darin, eine solche Funktion $u \in \dot{W}_2^{(1)}(D)$ zu bestimmen, für die das Funktional (157) sein Minimum erreicht. Wir wissen, daß dieses Problem für jedes $f(x) \in L_2(D)$ eindeutig lösbar ist. Fügen wir alle diese Lösungen, die wir bei verschiedener Wahl von $f(x) \in L_2(D)$ erhalten, zu $D(A)$ hinzu, so finden wir eine selbstadjungierte Erweiterung $\bar{A}$ von A mit

$$\bar{A}u = f. \quad (158)$$

Da $\bar{A}$ eine selbstadjungierte Erweiterung von A ist, gilt $D(\bar{A}) \in D(A^*)$. Nun haben die Funktionen aus $D(A^*)$ innerhalb von D verallgemeinerte Ableitungen bis zur zweiten Ordnung, die in jedem in D echt enthaltenen Gebiet D' quadratisch summierbar sind, und der Operator A^* auf ihnen wirkt wie der Laplacesche Operator [188]. Daher ist auch $\bar{A}$ der Laplacesche Operator, d. h., die Gleichung (158) hat die Gestalt

$$-\sum_i u_{x_i x_i} = f(x). \quad (159)$$

Ferner zeigten wir, daß die Lösung des oben erwähnten Variationsproblems nicht nur in $\dot{W}_2^{(1)}(D)$, sondern auch in $\dot{W}_2^{(2)}(D')$ liegt, wobei D' ein beliebiges, aber in D echt enthaltenes Gebiet ist, und daß diese Lösung der Poissonschen Gleichung genügt.

Andererseits hat (159) unendlich viele Lösungen aus $L_2(D)$. Es genügt, zu der oben erwähnten Lösung eine harmonische Funktion aus $L_2(D)$ hinzuzufügen. Die Bedingung für eine Lösung, zu $\hat{W}_2^{(1)}(D)$ zu gehören, spaltet von dieser Klasse eine Lösung ab, die wir auch aus dem Variationsproblem erhielten. Diese Lösung muß auf dem Rand S von D in bestimmten Sinne gegen 0 gehen [113]. Daraus erklärt sich der Zusammenhang zwischen der konstruierten Erweiterung von A und dem Dirichletschen Problem für die Poissonsche Gleichung:

$$-\Delta u = f(x), \quad u|_S = 0. \quad (160)$$

Alles oben über den Laplaceschen Operator Gesagte gilt auch für beliebige lineare elliptische selbstadjungierte Operatoren zweiter Ordnung [IV/2, 48]. Aus der Friedrichschen Erweiterungstheorie folgt für sie die Lösbarkeit des Dirichletschen Problems in verallgemeinerten Sinne (im Sinne der Zugehörigkeit der Lösung zu $\hat{W}_2^{(1)}(D)$). In Wirklichkeit gehört diese verallgemeinerte Lösung des Dirichletschen Problems, die einer Erweiterung eines elliptischen Operators nach FRIEDRICHS entspricht, zu $W_2^{(2)}(D')$ oder sogar zu $W_2^{(2)}(D)$, sobald S eine hinreichend glatte Fläche ist. Darüber findet sich Näheres in Arbeiten von LADYSHENSKAJA.¹⁾

208. Das Spektrum eines symmetrischen Operators. Wir führten schon früher den Begriff des Spektrums eines selbstadjungierten Operators ein und klassifizierten seine Punkte. Im folgenden tun wir dies für einen abgeschlossenen symmetrischen Operator und untersuchen, wie sich das Spektrum bei symmetrischen Erweiterungen eines solchen Operators ändert.

Es sei A ein abgeschlossener symmetrischer Operator. Die Zahl λ nennen wir einen *Punkt regulären Typs* des Operators A , wenn es ein solches $k > 0$ gibt, daß für alle $x \in D(A)$

$$\|(A - \lambda E)x\| \geq k\|x\| \quad (x \in D(A)) \quad (161)$$

gilt.

Aus der Abgeschlossenheit von A und aus (161) folgt, daß $R(A - \lambda E)$ ein Unterraum von H und $(A - \lambda E)^{-1}$ ein auf $R(A - \lambda E)$ linearer beschränkter Operator ist. Existiert umgekehrt auf $R(A - \lambda E)$ der Operator $(A - \lambda E)^{-1}$ und ist er beschränkt, so folgt daraus (161), d. h., λ ist ein Punkt regulären Typs. Wie in [129] läßt sich leicht zeigen, daß *die Menge der Punkte regulären Typs offen ist*. Die Zahl λ heißt *regulärer Punkt* von A , wenn (161) erfüllt ist und $R(A - \lambda E)$ mit H übereinstimmt. Ist $R(A - \lambda E) = H$, so ist λ kein Eigenwert von A (denn $R(A - \lambda E)$ muß zu den Eigenelementen orthogonal sein), und $(A - \lambda E)^{-1}$ ist ein in H beschränkter Operator [186], d. h., es ist (161) erfüllt.

Ist λ ein reeller regulärer Punkt von A , so ist $(A - \lambda E)^{-1}$ ein beschränkter selbstadjungierter Operator und infolgedessen A selbstadjungiert. Um zu zeigen, daß die regulären Punkte eine offene Menge bilden, genügt der Nachweis, daß die Gleichung

$$(A - \lambda E)x = y \quad (162)$$

für jedes $y \in H$ eindeutig lösbar ist, wenn λ hinreichend nahe bei einem regulären

¹⁾ O. A. LADYSHENSKAJA, Über die Abschließung eines elliptischen Operators (russ.), Doklady Akad. Nauk SSSR 79 (1951), 723–725; O. A. LADYSHENSKAJA, Ein einfacher Beweis für die Lösbarkeit von Randwertproblemen und des Eigenwertproblems für lineare elliptische Operatoren (russ.), Vestnik Univ. Leningrad (Ser. mat. fis. chim. 4). 10: 11 (1955), 23–29.

Punkt λ_0 liegt. Wir setzen

$$|\lambda - \lambda_0| < \frac{1}{\|(A - \lambda_0 E)^{-1}\|}$$

und schreiben die Gleichung (162) in der Form

$$(A - \lambda_0 E) x + (\lambda_0 - \lambda) x = y;$$

sie ist äquivalent mit der Gleichung

$$x = (\lambda - \lambda_0) (A - \lambda_0 E)^{-1} x + (A - \lambda_0 E)^{-1} y,$$

die sich für jedes $y \in H$ eindeutig lösen läßt [88], denn es ist

$$\|(\lambda - \lambda_0) (A - \lambda_0 E)^{-1}\| < 1.$$

Wie in [129] können wir zeigen, daß für $\lambda = \sigma + \tau i$ und $\tau \neq 0$ die Ungleichung

$$\|(A - \lambda E) x\| \geq |\tau| \cdot \|x\| \quad (x \in D(A)) \quad (163)$$

gilt, d. h., alle komplexen λ sind Punkte regulären Typs.

Wir nehmen an, der Punkt $\lambda_0 = \sigma + \tau i$ ($\tau \neq 0$) sei ein regulärer Punkt. Dann sind alle λ , die der Bedingung $|\lambda - \lambda_0| < |\tau|$ genügen, wegen (163) und des oben über die Lösbarkeit von (162) Gesagten ebenfalls reguläre Punkte. Gehen wir von einem regulären Wert λ_0 aus und wenden wir die eben durchgeführten Überlegungen hinreichend oft an, so sehen wir, daß jede komplexe Zahl $\lambda = \sigma' + \tau' i$, bei der das Vorzeichen von τ' mit dem von τ übereinstimmt, ein regulärer Punkt ist. Diese Behauptung können wir folgendermaßen formulieren:

Hilfssatz. *Ist einer der Defektindizes p_λ oder q_λ des Operators A für $\lambda = \lambda_0$ (Im $\lambda_0 > 0$) gleich 0, so verschwindet er für alle λ aus der Halbebene Im $\lambda > 0$.*

Bei einem selbstadjungierten Operator sind alle komplexen λ reguläre Punkte. Wir geben nun ein Beispiel für einen abgeschlossenen symmetrischen Operator A an, der keine regulären Punkte besitzt. Es sei H der Raum $L_2(0, 1)$ und A der Operator $i \frac{d}{dx}$, der auf der Menge der Funktionen $\varphi(x)$ betrachtet wird; dabei ist $\varphi(x)$ absolut stetig auf dem Intervall $[0, 1]$, es gilt $\varphi(0) = \varphi(1) = 0$, und $\varphi'(x)$ liegt in $L_2(0, 1)$. Der Operator A ist abgeschlossen und symmetrisch [188]. Für jedes λ gehört die Funktion $e^{-i\lambda x}$ zu $L_2(0, 1)$ und ist zu allen Funktionen $\psi(x)$, die sich in der Gestalt

$$\psi(x) = i \frac{d\varphi(x)}{dx} - \lambda \varphi(x)$$

mit $\varphi(x) \in D(A)$ darstellen lassen, orthogonal, d. h., $e^{-i\lambda x}$ ist orthogonal zu $R(A - \lambda E)$; der Punkt λ ist also nicht regulär.

Wir nennen die Menge der Punkte der λ -Ebene, die zur Menge der regulären Punkte komplementär ist, das *Spektrum* des Operators A . Das ist die Menge derjenigen λ , in denen der Operator $A - \lambda E$ kein auf ganz H definiertes beschränktes Inverses besitzt. Der *Kern des Spektrums* von A ist dabei die Menge derjenigen Punkte, die zu der Menge der Punkte regulären Typs komplementär ist. Das Spektrum und der Kern des Spektrums sind abgeschlossene Mengen und die erste Menge (das Spektrum) enthält die zweite (den Kern des Spektrums). Der Kern muß sich auf der reellen Achse befinden; das Spektrum kann die ganze Ebene bedecken, wie aus dem obigen Beispiel ersichtlich ist.

Ist A ein selbstadjungierter Operator, so stimmt der Kern des Spektrums mit dem Spektrum selbst überein [189].

Es ist leicht zu erkennen, daß der Kern des Spektrums von A zum Kern des Spektrums jeder abgeschlossenen symmetrischen Erweiterung von A gehört. Die Zugehörigkeit von λ zum Kern des Spektrums von A ist nämlich damit äquivalent, daß eine Folge von normierten Elementen $x_n \in D(A)$ derart existiert, daß $(A - \lambda E)x_n \rightarrow 0$ für $n \rightarrow \infty$ ist. Diese Eigenschaft wird offenbar bei den oben erwähnten Erweiterungen von A bewahrt.

Wir wollen nun die Punkte des Kerns im Spektrum von A klassifizieren. Zunächst untersuchen wir den Fall, daß λ ein Eigenwert von A ist (λ sei eine reelle Zahl). P_λ sei der Unterraum der entsprechenden Eigenelemente (zuzüglich des Nullelements). Die lineare Mannigfaltigkeit $D(A)$ läßt sich als orthogonale Summe

$$D(A) = P_\lambda \oplus D_\lambda(A) \quad (164)$$

darstellen, wobei $D_\lambda(A)$ die lineare Mannigfaltigkeit derjenigen Elemente ist, die gleichzeitig in $H \ominus P_\lambda$ und $D(A)$ liegen. Mit A_λ bezeichnen wir den Operator A , der auf der linearen Mannigfaltigkeit $D_\lambda(A)$ definiert ist und dort mit A übereinstimmt. Ist λ kein Eigenwert, so ist P_λ nicht vorhanden und $D_\lambda(A) = D(A)$. In diesem Fall sagen wir, daß A_λ und A übereinstimmen. Der Operator $A_\lambda - \lambda E$, betrachtet in $H \ominus P_\lambda$, hat für jedes λ ein auf $R(A - \lambda E)$ definiertes Inverses $(A_\lambda - \lambda E)^{-1}$. Die Werte von λ , für die $(A_\lambda - \lambda E)^{-1}$ ein nicht beschränkter Operator ist, liegen im Kern des Spektrums. Diesen Teil nennen wir den *kontinuierlichen Teil* des Kerns. Die Eigenwerte gehören ebenfalls zum Kern und bilden den *Punkteil*. Jeder Punkt des Kerns liegt in einem der beiden Teile, kann aber auch beiden angehören. Wir werden sagen, die Eigenwerte gehören zum *reinen Punkteil*, wenn $(A_\lambda - \lambda E)^{-1}$ ein auf $R(A - \lambda E)$ beschränkter Operator ist. Jeder Punkt des Kerns liegt entweder im kontinuierlichen Teil oder im reinen Punkteil des Kerns und nur in einem von ihnen. Bei einer abgeschlossenen symmetrischen Erweiterung von A können die beiden Teile des Kerns nur erweitert werden. Wegen der Abgeschlossenheit von A und A_λ ist der kontinuierliche Teil des Kerns im Spektrum von A dadurch charakterisiert, daß $R(A - \lambda E)$ eine nicht abgeschlossene lineare Mannigfaltigkeit ist.

209. Einige Sätze über Erweiterungen und ihre Spektren. Wir beginnen mit dem Beweis des folgenden Satzes:

Satz 1. *Ist λ ein reeller Punkt regulären Typs eines abgeschlossenen symmetrischen Operators A , so gibt es eine selbstadjungierte Erweiterung $\tilde{A}$ von A , für welche λ ein regulärer Punkt ist.*

Beweis. Ohne Beschränkung der Allgemeinheit können wir $\lambda = 0$ voraussetzen. Aus den Bedingungen des Satzes folgt, daß $R(A)$ ein Unterraum ist; auf ihm ist ein beschränkter inverser Operator A^{-1} definiert [208]. Wir müssen zeigen, daß es eine selbstadjungierte Erweiterung $\tilde{A}$ von A gibt, für welche $R(\tilde{A}) = H$ ist [189]. Unter den erwähnten Bedingungen gilt

$$H = R(A) \oplus U;$$

U ist der Unterraum aller Lösungen der Gleichung $A^*u = 0$ [185]. Bekanntlich ist in unserem Fall $R(A^*) = H$ [187]; folglich gibt es für jedes $u \in U$ wenigstens eine Lösung der Gleichung

$$A^*v = u. \quad (165)$$

Mit V bezeichnen wir die lineare Mannigfaltigkeit aller Lösungen von (165), wenn u ganz U durchläuft. Offenbar ist $U \subseteq V \subseteq D(A^*)$. Ferner sei $\tilde{U}$ die lineare Mannigfaltigkeit derjenigen Elemente von V , die zu U orthogonal sind, und wir konstruieren die lineare Mannigfaltigkeit l der Elemente x , die sich in der Form

$$x = y + \tilde{u} \quad (166)$$

mit $y \in D(A)$ und $\tilde{u} \in \tilde{U}$ darstellen lassen. Diese Darstellung ist eindeutig. Wäre sie es nicht, so würde ein vom Nullelement verschiedenes Element z existieren, das gleichzeitig zu $D(A)$ und $\tilde{U}$ gehört. Dann wäre $Az \in R(A)$ und $Az = A^*z \in U$ wegen $z \in \tilde{U}$, also erst recht $z \in V$. Somit ist $R(A) \perp U$ und infolgedessen $Az = 0$, also $z \in U$, woraus zusammen mit $z \in \tilde{U}$ die Beziehung $z = 0$ folgt. Damit ist die Eindeutigkeit der Darstellung (166) nachgewiesen, d. h., l ist die direkte Summe $D(A) \dot{+} \tilde{U}$.

Wir definieren nun auf der linearen Mannigfaltigkeit l einen Operator $\tilde{A}$ mit Hilfe der Formel $\tilde{A}x = Ay + A^*\tilde{u}$ und wählen die Bezeichnung $D(\tilde{A})$ statt l . Offenbar ist $A \subseteq \tilde{A} \subseteq A^*$. Der Operator $\tilde{A}$ erfüllt alle Forderungen des Satzes. Auf der linearen Mannigfaltigkeit $D(A)$ stimmt $\tilde{A}$ mit A überein, und auf $\tilde{U}$ liefert der Operator A^* den ganzen Unterraum U , was aus der Definition von $\tilde{U}$ und der Tatsache folgt, daß $A^*u = 0$ für $u \in U$ gilt. Damit ist $R(\tilde{A}) = H$.

Wir müssen noch die Symmetrie von $\tilde{A}$ auf $D(\tilde{A})$ nachweisen [187]. Es gilt

$$(A^*\tilde{u}_1, \tilde{u}_2) = (\tilde{u}_1, A^*\tilde{u}_2) = 0 \quad (\tilde{u}_1, \tilde{u}_2 \in \tilde{U}),$$

und x_1 und x_2 seien Elemente aus $D(\tilde{A})$. Dann ist auf Grund von (166) $x_1 = y_1 + \tilde{u}_1$, $x_2 = y_2 + \tilde{u}_2$, und wir erhalten

$$\begin{aligned} (\tilde{A}x_1, x_2) &= (Ay_1 + A^*\tilde{u}_1, x_2) = (y_1, A^*x_2) + (A^*\tilde{u}_1, y_2) \\ &= (y_1, A^*x_2) + (\tilde{u}_1, Ay_2) = (y_1, A^*x_2) + (u_1, Ay_2) + (\tilde{u}_1, A^*\tilde{u}_2) \\ &= (y_1, A^*x_2) + (\tilde{u}_1, \tilde{A}x_2) = (x_1, \tilde{A}x_2). \end{aligned}$$

Damit ist der Satz bewiesen.

Man kann zeigen, daß mit A^{-1} auch $\tilde{A}^{-1}$ vollstetig ist. Dem eingehenden Studium solcher Erweiterungen von Operatoren (sowohl abstrakten Operatoren als auch Differentialoperatoren) sind Arbeiten von WISCHIK¹⁾ und HÖRMANDER²⁾ gewidmet.

Folgerung 1. *Gehört eine reelle Zahl λ zum reinen Punktteil des Kerns im Spektrum, so gibt es eine selbstadjungierte Erweiterung $\tilde{A}$ von A , für welche λ auch zum reinen Punktspektrum mit dem gleichen Unterraum P_λ der Eigenelemente wie A gehört.*

Der Operator A , betrachtet auf dem Unterraum $H_1 = H \ominus P_\lambda$, ist, wie man leicht sieht, abgeschlossen und symmetrisch und genügt den Bedingungen von Satz 1. Deshalb läßt er in H_1 eine selbstadjungierte Erweiterung $\tilde{A}_1$ zu, für welche λ ein regulärer Punkt ist. Der Operator $\tilde{A}$ mit dem Definitionsbereich $D(\tilde{A}) = D(\tilde{A}_1) \oplus P_\lambda$ stimmt auf $D(\tilde{A}_1)$ mit $\tilde{A}_1$ und auf P_λ mit A überein und ist offenbar eine Erweiterung aus Folgerung 1.

Folgerung 2. *Ist A maximal, aber nicht selbstadjungiert, so belegt der kontinuierliche Teil des Kerns im Spektrum von A die gesamte reelle Achse.*

Anderenfalls hätte der Operator A eine selbstadjungierte Erweiterung.

¹⁾ M. I. WISCHIK, Über allgemeine Randwertprobleme für elliptische Differentialgleichungen (russ.), Arb. Moskauer Math. Ges. 1 (1952), 187–246.

²⁾ L. HÖRMANDER, On the theory of general partial differential operators, Acta Math. 94 (1955), 161–248.

Folgerung 3. *Gibt es ein reelles λ , das nicht zum kontinuierlichen Teil des Kerns im Spektrum von A gehört, so sind für jedes λ ($\text{Im } \lambda > 0$) die Defektindizes p_λ und q_λ gleich.*

Das folgt daraus, daß der Operator A unter dieser Bedingung eine selbstadjungierte Erweiterung besitzt.

Satz 2. *Es sei ein reelles λ ein Punkt regulären Typs für den abgeschlossenen symmetrischen Operator A , und es sei $U = H \ominus R(A - \lambda E)$. Dann existiert eine selbstadjungierte Erweiterung $\tilde{A}$ von A , für welche λ zum reinen Punktspektrum gehört und U der Unterraum der Eigenelemente ist, die λ entsprechen.*

Beweis. Ohne Beschränkung der Allgemeinheit nehmen wir $\lambda = 0$ an. $R(A)$ und U sind Unterräume, wobei U die Menge der Lösungen von $A^*z = 0$ ist. Mit $D(A) + U$ bezeichnen wir die Menge der Elemente $x = y + z$, $y \in D(A)$, $z \in U$.

Diese Darstellung von x ist eindeutig. Anderenfalls gäbe es ein vom Nullelement verschiedenes Element $x_0 \in D(A)$, das auch in U liegt. Daraus folgt $(x_0, Ax) = 0$ ($x \in D(A)$) und $(Ax_0, x) = 0$. Da $D(A)$ in H dicht ist, erhalten wir $Ax_0 = 0$; dies widerspricht aber der Tatsache, daß $\lambda = 0$ ein Punkt regulären Typs ist.

Somit können wir auf der direkten Summe $D(\tilde{A}) = D(A) + U$ den Operator $\tilde{A}$ definieren, indem wir $\tilde{A}x = Ay$ setzen, wenn $x = y + z$ mit $y \in D(A)$, $z \in U$ ist.

Die Symmetrie von $\tilde{A}$ läßt sich sofort nachweisen, denn es ist $(\tilde{A}z, y) = 0$ und

$$(z, \tilde{A}y) = (z, Ay) = (A^*z, y) = 0.$$

Wir zeigen nun die Selbstadjungiertheit von $\tilde{A}$. Es sei

$$(\tilde{A}x, u) = (x, u^*) \quad (x \in D(\tilde{A})). \quad (167)$$

Wegen $\tilde{A}^* \subseteq A^*$ ist $u \in D(A^*)$ und $u^* = A^*u$. Setzen wir x in der Form $x = y + z$ an, so erhalten wir

$$(\tilde{A}x, u) = (Ay, u) = (y, u^*) + (z, u^*) = (y, A^*u) + (z, u^*)$$

oder

$$(Ay, u) = (Ay, u) + (z, u^*),$$

d. h.

$$(z, u^*) = (z, A^*u) = 0.$$

Diese Gleichung gilt für alle $z \in U$. Daher ist $u^* = A^*u \in R(A)$; folglich existiert ein Element $y_0 \in D(A)$, für welches $Ay_0 = u^*$ gilt. Daraus ergibt sich

$$A^*u - Ay_0 = A^*(u - y_0) = 0,$$

also $u - y_0 = z_0 \in U$ und $u = y_0 + z_0 \in D(\tilde{A})$. Damit ist wegen (167) die Selbstadjungiertheit von $\tilde{A}$ bewiesen. Die übrigen Behauptungen über die Eigenschaften des Operators $\tilde{A}$ folgen unmittelbar aus seiner Konstruktion.

Folgerung. *Gehört ein reelles λ zum reinen Punktspektrum von A , so gestattet A eine selbstadjungierte Erweiterung $\tilde{A}$, für welche λ ebenfalls nur zum Punktteil des Kerns des Spektrums von $\tilde{A}$ gehört; dabei stimmt der Unterraum der Eigenelemente von $\tilde{A}$, die λ entsprechen, mit dem Unterraum der Eigenelemente von A^* überein, die ebenfalls λ entsprechen.*

Der Beweis verläuft analog dem der Folgerung 1 aus Satz 1.

210. Die Unabhängigkeit der Defektindizes von λ . Wir erwähnten schon früher, daß die Defektindizes p_λ und q_λ nicht von der Wahl der komplexen Zahl λ aus der oberen Halbebene abhängen. Diese Behauptung wollen wir jetzt beweisen. Zuvor

bemerken wir folgendes: Bildet ein beliebiger linearer (nicht notwendig beschränkter) Operator B einen Unterraum V in einen Unterraum W eindeutig ab, so stimmen die Dimensionen von V und W überein: $\dim V = \dim W$; die linear unabhängigen Elemente $\{x_1, x_2, \dots\}$ des Unterraumes V werden nämlich von B in linear unabhängige Elemente $\{Bx_1, Bx_2, \dots\}$ aus W übergeführt und umgekehrt.

Wir geben noch eine Definition. Eine Zahl m heißt *Dimension der linearen Mannigfaltigkeit l modulo der linearen Mannigfaltigkeit l'* , wenn m und nicht mehr als m linear unabhängige Elemente in l existieren, von denen keine Linearkombination zu l' gehört (außer im Fall, daß alle Koeffizienten gleich 0 sind). Im allgemeinen schreibt man $m = \dim l \pmod{l'}$. Es kann auch $m = \infty$ sein.

Es sei A ein abgeschlossener symmetrischer Operator, und p_λ und q_λ seien seine dem Wert λ aus der oberen Halbebene zugeordneten Defektindizes. Dann gelten, wie wir sahen [202], die Formeln

$$H = L_\lambda(A) \oplus M_\lambda(A), \quad H = L_{\bar{\lambda}}(A) \oplus M_{\bar{\lambda}}(A); \quad (168)$$

$M_\lambda(A)$ ist die Gesamtheit aller Nullstellen des Operators $A^* + \lambda E$, $L_\lambda(A)$ die Gesamtheit aller Elemente der Form $y = (A - \lambda E)x$ mit $x \in D(A)$. Analoges gilt für $M_{\bar{\lambda}}(A)$ und $L_{\bar{\lambda}}(A)$. Ferner ist [203]

$$D(A^*) = D(A) \dot{+} M_\lambda(A) \dot{+} M_{\bar{\lambda}}(A). \quad (169)$$

Jede symmetrische Erweiterung von A läßt sich mit Hilfe eines isometrischen Operators konstruieren, der die gleichdimensionalen Unterräume $N_\lambda(A)$ und $N_{\bar{\lambda}}(A)$ der Räume $M_\lambda(A)$ bzw. $M_{\bar{\lambda}}(A)$ einander eindeutig zuordnet [202].

Aus (169) folgt

$$p_\lambda + q_\lambda = \dim D(A^*) \pmod{D(A)}; \quad (170)$$

die Summe $p_\lambda + q_\lambda$ hängt also nicht von λ ab. Zuerst setzen wir voraus, $p_\lambda + q_\lambda$ sei endlich, konstruieren eine maximale Erweiterung A_0 von A und betrachten einen Wert $\lambda = \lambda'$ aus der oberen Halbebene. Ohne Beschränkung der Allgemeinheit nehmen wir $p_{\lambda'} \leq q_{\lambda'}$ an. Dann sind die Defektindizes von A_0 für $\lambda = \lambda'$ gleich $(0, r_{\lambda'})$ mit $r_{\lambda'} = q_{\lambda'} - p_{\lambda'}$. Auf Grund des Hilfssatzes aus [208] sind dann die Defektindizes von A für alle λ aus der oberen Halbebene gleich $(0, r_\lambda)$ mit $r_\lambda = q_\lambda - p_\lambda$ [202]. Ferner ergibt sich aus (169)

$$D(A_0^*) = D(A_0) \dot{+} M_{\bar{\lambda}}(A_0),$$

so daß

$$r_\lambda = \dim M_{\bar{\lambda}}(A_0) = \dim D(A_0^*) \pmod{D(A_0)}$$

ist; also ist $r_\lambda = q_\lambda - p_\lambda$ und demzufolge auch p_λ und q_λ von λ unabhängig. Wir setzen nun voraus, die Summe $p_\lambda + q_\lambda$ sei unendlich. Gilt für ein $\lambda = \lambda'$ dann $p_{\lambda'} = \infty$, $q_{\lambda'} = \infty$, so sind selbstadjungierte Erweiterungen von A möglich. Also ist $p_\lambda = \infty$, $q_\lambda = \infty$ für jedes λ .

Wir müssen nun noch den Fall betrachten, daß für $\lambda = \lambda'$ der eine Defektindex endlich, der andere unendlich ist.

Es sei $p_{\lambda'}$ endlich und $q_{\lambda'} = \infty$. Dann folgt unmittelbar aus dem Vorhergehenden, daß für jedes λ auch p_λ endlich und $q_\lambda = \infty$ ist. Daß p_λ nicht von λ abhängt, folgt leicht aus der Formel

$$D(A_0) = D(A) \dot{+} (E - V_0) M_\lambda(A) \quad (171)$$

(vgl. [202]), wobei A_0 eine feste maximale Erweiterung und V_0 ein von A_0 und λ abhängiger isometrischer Operator ist. Da $(E - V_0)^{-1}$ existiert, ergibt sich aus (171)

$$p_\lambda = \dim D(A_0) \pmod{D(A)}. \quad (172)$$

Also ist p_λ von λ unabhängig.

Aus Satz 1 von [129] folgt: Existiert ein reelles $\lambda = \lambda_0$ als Punkt regulären Typs des abgeschlossenen symmetrischen Operators A , so gestattet A eine selbstadjungierte Erweiterung; somit stimmen die (von λ unabhängigen) Defektindizes (p, q) bei Existenz eines solchen Punktes überein. Sie sind gleich der Dimension des Unterraumes U der Eigenelemente von A^* , die dem Eigenwert $\lambda = \lambda_0$ entsprechen. Um das zu zeigen, setzen wir $\lambda_0 = 0$ und bezeichnen mit $\tilde{A}$ diejenige selbstadjungierte Erweiterung von A , für welche $\lambda = 0$ ein regulärer Punkt ist [209]. Dann ist im betrachteten Fall $H = R(\tilde{A}) \oplus D$ und $D(A) = D(\tilde{A}) \dot{+} \tilde{A}^{-1}U$. In diesen Formeln ist die Darstellung eines Elements der linken Seite als Summe eindeutig, und folglich ist

$$p = \dim D(\tilde{A}) \pmod{D(A)} = \dim \tilde{A}^{-1}U = \dim U.$$

211. Über die Invarianz des kontinuierlichen Teils des Kerns bei symmetrischen Erweiterungen. In diesem Abschnitt werden wir voraussetzen, daß ein abgeschlossener symmetrischer Operator A endliche Defektindizes (p, q) besitzt, und seine abgeschlossenen symmetrischen Erweiterungen untersuchen.

Zunächst beweisen wir einen einfachen Hilfssatz.

Hilfssatz. Sind U und W zwei Unterräume, von denen W endlichdimensional sein soll, so gilt

$$V = U + W, \quad (173)$$

d. h., die Menge der Elemente $x = y + z$ mit $y \in U$ und $z \in W$ ist ebenfalls ein Unterraum.

Beweis. Offenbar können wir annehmen, daß U und W außer dem Nullelement keine gemeinsamen Elemente haben. Es sei $(w_1, w_2, \dots, w_n)$ eine Basis von W . Jedes ihrer Elemente läßt sich in der Form $w_k = w'_k + w''_k$ mit $w'_k \in U$ und $w''_k \perp U$ darstellen; die lineare Hülle der w''_k ($k = 1, 2, \dots, n$) bezeichnen wir mit W'' . Die Menge V können wir als orthogonale Summe zweier Unterräume darstellen:

$$V = U \oplus W''.$$

Damit ist der Hilfssatz bewiesen.

Satz. Der kontinuierliche Teil des Kerns im Spektrum einer beliebigen abgeschlossenen symmetrischen Erweiterung $\tilde{A}$ eines Operators A stimmt mit dem von A überein.

Beweis. Bei einer Erweiterung von A kann sich der kontinuierliche Teil des Kerns nicht verkleinern [208]. Zum Beweis nehmen wir an, er vergrößere sich, d. h., es gäbe eine reelle Zahl λ_0 , die nicht im kontinuierlichen Teil des Kerns von A , aber im kontinuierlichen Teil des Kerns von $\tilde{A}$ auftritt. Für diese Zahl ist $R(A - \lambda_0 E)$ ein Unterraum und $R(\tilde{A} - \lambda_0 E)$ eine nicht abgeschlossene lineare Mannigfaltigkeit. Mit Hilfe der Formel für $D(\tilde{A})$ aus [203] können wir, da die Defektindizes für A endlich sind,

$$R(\tilde{A} - \lambda_0 E) = R(A - \lambda_0 E) + W$$

schreiben, wobei W endlichdimensional ist. Diese Formel und die Nichtabgeschlossenheit von $R(\tilde{A} - \lambda_0 E)$ widersprechen dem oben bewiesenen Hilfssatz.

212. Über die Spektren selbstadjungierter Erweiterungen. Ist λ ein reeller Punkt regulären Typs für A , so existieren bekanntlich selbstadjungierte Erweiterungen $\tilde{A}$ der beiden folgenden Arten: Bei den einen Erweiterungen ist λ ein regulärer Punkt von $\tilde{A}$, bei den anderen ein Eigenwert, dessen Vielfachheit gleich den miteinander übereinstimmenden Defektindizes von A ist.

Diese Resultate wollen wir nun noch ergänzen.

Satz 1. Sind die Defektindizes (p, p) eines abgeschlossenen symmetrischen Operators A endlich, so läßt sich bei jeder selbstadjungierten Erweiterung $\tilde{A}$ von A die Vielfachheit eines beliebigen Eigenwertes nicht um mehr als p erhöhen, und dasjenige reelle λ , das kein Eigenwert von $\tilde{A}$ ist, kann nicht ein mehr als p -facher Eigenwert von A sein.

Beweis. Es sei λ kein Eigenwert von A , aber ein Eigenwert von $\tilde{A}$ mit der Vielfachheit $k > p$. Aus der Formel

$$p = \dim D(\tilde{A}) \pmod{D(A)}$$

(vgl. (172)) und $k > p$ folgt, daß es ein Eigelement von $\tilde{A}$ gibt, das zu $D(A)$ gehört, d. h., λ ist Eigenwert von A . Das widerspricht aber der Voraussetzung. Somit ist $k \leq p$. Der Fall, daß λ ein Eigenwert von A ist, läßt sich bei Verwendung des Operators A_λ analog behandeln [208].

Satz 2. Ist A ein halbbeschränkter abgeschlossener symmetrischer Operator und sind seine Defektindizes (p, p) endlich, so besteht bei einer selbstadjungierten Erweiterung von A das Spektrum, das sich links von der unteren Schranke von A befindet, nur aus endlich vielen Eigenwerten, wobei die Summe der Vielfachheiten dieser Eigenwerte höchstens gleich p ist.

Beweis. Ohne Beschränkung der Allgemeinheit setzen wir A als positiven Operator voraus, und $\mathfrak{E}_\lambda$ sei die Spektralschar einer selbstadjungierten Erweiterung $\tilde{A}$. Wir zeigen, daß für alle α und β ($0 > \beta > \alpha$) die Ungleichung

$$\dim \mathcal{A}\mathfrak{E}_\lambda H = \dim (\mathfrak{E}_\beta - \mathfrak{E}_\alpha) H \leq p \quad (174)$$

gilt. Dazu nehmen wir das Gegenteil an:

$$\dim \mathcal{A}\mathfrak{E}_\lambda H > p. \quad (175)$$

Da $\mathcal{A}\mathfrak{E}_\lambda x$ für $x \in H$ zu $D(\tilde{A})$ gehört und $p = \dim D(\tilde{A}) \pmod{D(A)}$ gilt, folgt auf Grund von (175), daß der Unterraum $\mathcal{A}\mathfrak{E}_\lambda H$ ein normiertes Element $x \in D(A)$ enthält. Dann ist

$$(Ax, x) = (\tilde{A}x, x) = \int_{-\infty}^{\infty} \lambda d(\mathfrak{E}_\lambda x, x) = \int_{\alpha}^{\beta} \lambda d(\mathfrak{E}_\lambda x, x) \leq \beta < 0;$$

dies gilt aber nicht, wenn A positiv ist. Also ist die Ungleichung (174) richtig. Damit ist der Satz bewiesen.

213. Beispiele. 1. In [188] untersuchten wir im Raum $H = L_2(D)$ den Operator A , der auf allen stetig differenzierbaren und in D finiten Funktionen definiert und auf diesen Funktionen gleich dem Differentialoperator

$$D^k = i^k \frac{\partial^k}{\partial x_{l_1} \dots \partial x_{l_k}} \quad (176)$$

ist. Ferner zeigten wir, daß $\tilde{A}$ ein symmetrischer Operator mit einem beschränkten Inversen auf $R(\tilde{A})$ ist. Daraus folgt [209], daß $\tilde{A}$ eine solche selbstadjungierte Erweiterung $\tilde{A}$ gestattet, für die die Gleichung

$$\tilde{A}\varphi = \psi \quad (\varphi \in D(\tilde{A})) \quad (177)$$

für jedes $\psi(x) \in L_2(D)$ eindeutig lösbar ist. Der Definitionsbereich von $\tilde{A}$ wird bei dieser Erweiterung durch Funktionen $v(x) \in L_2(D)$ ergänzt, für die $A^*A^*v = 0$ gilt. Diese Funktionen haben verallgemeinerte Ableitungen D^k und $D^k D^k$ und werden durch

$$D^k D^k v = \frac{\partial^{2k} v(x)}{\partial x_{l_1}^2 \dots \partial x_{l_k}^2} = 0$$

definiert. Für $D(\bar{A})$ wählen wir hiervon diejenigen aus, die zu den Lösungen von $D^k u = 0$ orthogonal sind. Der Operator $\bar{A}$ hat auf $D(\bar{A})$ die Form D^k , wobei D^k die verallgemeinerte Ableitung (176) ist.

2. Wir betrachten im Raum $H = L_2(0, \infty)$ den Operator

$$A = -\frac{d^2}{dx^2};$$

er ist auf allen stetig differenzierbaren und in der Umgebung von $x = 0$ und $x = \infty$ finiten Funktionen definiert. $D(A^*)$ bezeichnet die Menge aller Funktionen $\varphi(x)$ mit folgenden Eigenschaften: $\varphi(x)$ und $\varphi'(x)$ sind absolut stetig auf jedem endlichen Intervall $[0, a]$; $\varphi(x)$ und $\varphi''(x)$ gehören zu $L_2(0, \infty)$. Dann ist, wie wir erwähnten, auch $\varphi'(x) \in L_2(0, \infty)$. Für eine Funktion $\varphi(x) \in D(A^*)$ gilt $A^*\varphi(x) = -\varphi''(x)$ [188]. Der Bereich $D(\bar{A})$ besteht aus denjenigen Elementen von $D(A^*)$, die den Bedingungen $\varphi(0) = \varphi'(0) = 0$ genügen.

Wir können leicht nachprüfen, daß die folgenden Behauptungen gelten: a) $\bar{A}$ ist positiv; b) die Defektindizes von $\bar{A}$ sind $(1, 1)$; c) der kontinuierliche Teil des Kerns im Spektrum liegt auf der Halbachse $0 \leq \lambda < \infty$ (dasselbe gilt auch für jede selbstadjungierte Erweiterung von A); d) jede symmetrische Erweiterung von $\bar{A}$ ist selbstadjungiert; sie ist definiert auf den Elementen von $D(A^*)$, die einer der beiden Bedingungen $\varphi'(0) - h\varphi(0) = 0$ (h eine feste reelle Zahl) oder $\varphi(0) = 0$ genügen. Der letzten Bedingung entspricht eine Erweiterung nach FRIEDRICHS; dabei bleibt der Operator positiv und hat ein rein kontinuierliches Spektrum, das aus der Halbachse $0 \leq \lambda < \infty$ besteht. Ist $h \leq 0$, so hat der selbstadjungierte Operator, der der Bedingung $\varphi'(0) - h\varphi(0) = 0$ entspricht, ein rein kontinuierliches Spektrum. Für $h > 0$ hat er einen einfachen Eigenwert $\lambda = -h^2$.

214. Unendliche Matrizen. In [200] betrachteten wir Integraloperatoren, deren Kerne den Bedingungen (107) und (108) genügen. Analog können wir Operatoren in l_2 untersuchen:

$$y_i = \sum_{k=1}^{\infty} a_{ik} x_k \quad (i = 1, 2, \dots); \quad (178)$$

sie werden durch Matrizen realisiert, für welche $a_{ki} = \overline{a_{ik}}$ gilt und die Bedingungen

$$d_k^2 = \sum_{i=1}^{\infty} |a_{ik}|^2 = \sum_{i=1}^{\infty} |a_{ki}|^2 < \infty \quad (k = 1, 2, \dots; d_k \geq 0) \quad (179)$$

erfüllt sind. Dabei konvergiert die Reihe (178) absolut für jedes Element x aus l_2 , während die aus den $|y_i|^2$ gebildete Reihe nicht unbedingt zu konvergieren braucht, d. h., $(y_1, y_2, \dots)$ braucht kein Element aus l_2 zu sein. Mit $D(A_0)$ bezeichnen wir die lineare Mannigfaltigkeit derjenigen Elemente $x \in l_2$, für die

$$\sum_{k=1}^{\infty} d_k |x_k| < \infty$$

gilt, und mit $D(B)$ die lineare Mannigfaltigkeit der x mit $(y_1, y_2, \dots) \in l_2$. Wie in [200] läßt sich zeigen, daß $D(A_0)$ in l_2 überall dicht ist und $D(A_0) \subseteq D(B)$ gilt. Ferner seien A_0 und B die in $D(A_0)$ bzw. $D(B)$ durch (178) definierten Operatoren. Dann können wir behaupten, daß A_0 symmetrisch und $B = A^*$ ist (vgl. [200]). Wegen (179) gehören alle Einheitsvektoren zu $D(B)$ und sogar zu $D(A_0)$. Der Operator B ist genau dann selbstadjungiert, wenn für alle x und y aus $D(B)$ die Gleichung

$$\sum_{i=1}^{\infty} \left(\sum_{k=1}^{\infty} a_{ik} x_k \right) \bar{y}_i = \sum_{k=1}^{\infty} x_k \left(\sum_{i=1}^{\infty} a_{ik} \bar{y}_i \right) \quad (180)$$

(vgl. [200]) erfüllt ist; dabei gilt wegen $a_{ik} = \overline{a_{ki}}$

$$\sum_{i=1}^{\infty} a_{ik} \bar{y}_i = \overline{\sum_{i=1}^{\infty} a_{ki} y_i}.$$

Der symmetrische Operator A_0 braucht nicht abgeschlossen zu sein. Wir führen noch einen neuen Operator A ein, der sich, wie wir jetzt zeigen wollen, als Abschließung von A_0 erweisen wird. Es sei $D(A)$ die lineare Mannigfaltigkeit derjenigen x aus $D(B)$, für welche

$$(Bx, y) = (x, By) \quad (y \in D(B)) \quad (181)$$

gilt, und A der auf $D(A)$ durch (178) definierte Operator. Andererseits ist A_0^{**} auf der linearen Mannigfaltigkeit $D(A_0^{**})$ derjenigen Elemente x definiert, mit denen $(By, x) = (y, x^*)$ für alle $y \in D(B)$ gilt. Wegen $A_0^{**} \subset A_0^*$ wird x^* durch x vermöge der Formeln (178) ausgedrückt, d. h., es ist $x^* = Bx$. Vergleichen wir diese Beziehung mit der Definition von A , so erkennen wir, daß A_0^{**} mit A übereinstimmt. A_0^{**} und demzufolge auch A ist die Abschließung von A_0 , und es gilt $A^* = A_0^* = B$.

Wir erklären nun eine Eigenschaft der linearen Mannigfaltigkeit $D(A)$ und des Operators A . Wir nennen jedes Element $(x_1, x_2, \dots)$ aus l_2 , welches endlich viele von 0 verschiedene Komponenten x_k besitzt, ein *endliches Element*. Es sei $D(A')$ eine lineare Mannigfaltigkeit endlicher Elemente und A' der auf $D(A')$ definierte Operator (178). Auf Grund von $\overline{a_{ki}} = a_{ik}$ ist dieser Operator symmetrisch, und jedes Element von $D(A')$ ist offenbar auch in $D(A_0)$ enthalten, d. h., es ist $A' \subseteq A_0$. Also erhalten wir, wenn wir mit $\bar{A}'$ die Abschließung von A' bezeichnen, $\bar{A}' \subseteq A$ und infolgedessen $A^* \subseteq (\bar{A}')^*$. Jedes Element $y \in (\bar{A}')^*$ muß der Beziehung $(Be_k, y) = (e_k, y^*)$ genügen, wobei $y^* = (\bar{A}')^* y$ ist und mit e_k der k -te Einheitsvektor bezeichnet wird. Sind y_i und y_i^* die Komponenten von y bzw. y^* , so erhält die erwähnte Gleichung die Form

$$\sum_{i=1}^{\infty} a_{ik} \bar{y}_i = \bar{y}_k^* \quad \text{bzw.} \quad y_k^* = \sum_{i=1}^{\infty} a_{ki} y_i.$$

Daraus erkennen wir, daß y zu $D(A^*)$ gehört und $y^* = A^* y$ ist. Durch Vergleich dieses Resultats mit $A^* \subseteq (\bar{A}')^*$ finden wir $(\bar{A}')^* = A^*$, also $A^{**} = \bar{A}'$. Da aber $A^{**} = A$ gilt, muß $\bar{A}' = A$ sein. Dieses Ergebnis läßt sich folgendermaßen formulieren:

Satz. Ist $x \in D(A)$, so existiert eine Folge von endlichen Elementen ξ_n , für die $\xi_n \rightarrow x$, $A\xi_n \rightarrow Ax$ und $A = \bar{A}'$ gilt.

Die Selbstadjungiertheit von A führt auf die Beziehung $A^* = A$. Ist dies nicht der Fall, so werden die Defektindizes des Operators A durch die Dimensionen der Unterräume bestimmt, die von den Lösungen der Gleichungen $A^*x = ix$ und $A^*x = -ix$ gebildet werden. Wir beweisen noch: Ist die Matrix der a_{ik} reell und gehört das komplexe Element $x' + ix''$ zu $D(A)$, so liegen auch x' , x'' sowie $x' - ix''$ in $D(A)$. Auf Grund des bewiesenen Satzes gibt es eine Folge von endlichen Elementen $\xi_n = \xi'_n + i\xi''_n$ derart, daß

$$\|x - \xi_n\|^2 = \|x' - \xi'_n\|^2 + \|x'' - \xi''_n\|^2 \rightarrow 0$$

und

$$\|Ax - A\xi_n\|^2 = \|Ax' - A\xi'_n\|^2 + \|Ax'' - A\xi''_n\|^2 \rightarrow 0$$

gilt. Hieraus folgt

$$\|x' - \xi'_n\| \rightarrow 0, \quad \|Ax' - A\xi'_n\| \rightarrow 0, \quad \|x'' - \xi''_n\| \rightarrow 0, \quad \|Ax'' - A\xi''_n\| \rightarrow 0.$$

Unter Berücksichtigung von $A = \overline{A'}$ erhalten wir die obige Behauptung.

215. Jacobische Matrizen. Wir wenden nun die vorhergehenden Resultate auf Jacobische Matrizen

$$\begin{pmatrix} a_0 & b_0 & 0 & 0 & 0 & \dots \\ b_0 & a_1 & b_1 & 0 & 0 & \dots \\ 0 & b_1 & a_2 & b_2 & 0 & \dots \\ 0 & 0 & b_2 & a_3 & b_3 & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \end{pmatrix} \quad (182)$$

an, wobei die a_i reell und die b_i positiv sind. Die Bedingung (179) ist offenbar erfüllt. Die Numerierung der Einheitsvektoren e_k beginnen wir bei $k=0$.

Vermöge der Formeln

$$\begin{aligned} \lambda P_k(\lambda) &= b_k P_{k+1}(\lambda) + a_k P_k(\lambda) + b_{k-1} P_{k-1}(\lambda), \\ P_{-1}(\lambda) &= 0, \quad P_0(\lambda) = 1 \end{aligned} \quad (183)$$

(vgl. [167]) konstruieren wir reelle Polynome $P_k(\lambda)$. Aus (183) folgt

$$e_k = P_k(A) e_0. \quad (184)$$

Satz 1. *Konvergiert die Reihe*

$$\sum_{k=0}^{\infty} |P_k(i)|^2 \quad (i = \sqrt{-1}), \quad (185)$$

so ist der Operator A nicht selbstadjungiert.

Beweis. Wegen der Konvergenz von (185) läßt sich ein Element $x \in l_2$ mit den Komponenten $x_k = P_k(i)$ konstruieren, für welches $(x, e_k) = P_k(i)$ ist. Berücksichtigen wir die Beziehung

$$Ae_k = b_{k-1}e_{k-1} + a_k e_k + b_k e_{k+1}$$

und die Formel (183), so erhalten wir

$$(Ae_k, x) = b_{k-1} \overline{P_{k-1}(i)} + a_k \overline{P_k(i)} + b_k \overline{P_{k+1}(i)} = i \overline{P_k(i)}$$

oder wegen $(e_k, x) = \overline{P_k(i)}$

$$(Ae_k, x) = (e_k, ix)$$

Da der Operator A und das Skalarprodukt distributiv sind, gilt $(Ay, x) = (y, ix)$ für jedes endliche Element y ; diese Gleichung behält wegen des Satzes aus [214] auch für jedes $y \in D(A)$ ihre Gültigkeit. Also ist $x \in D(A^*)$ und $A^*x = ix$; daraus folgt, daß A nicht selbstadjungiert ist.

Satz 2. *Divergiert die Reihe (185), so ist der Operator A selbstadjungiert.*

Beweis. Es genügt zu zeigen, daß A^* nicht die Eigenwerte $\pm i$ besitzt. Dazu nehmen wir das Gegenteil an; es sei also $A^*x = ix$, wobei das Element $x(x_1, x_2, \dots)$ von 0 verschieden sei. Auf Grund der Definition von A^* und der Tatsache, daß e_k zu $D(A)$ gehört, gilt $(Ae_k, x) = (e_k, ix)$ oder $(x, Ae_k) = i(x, e_k) = ix_k$, also

$$(x, b_{k-1}e_{k-1} + a_k e_k + b_k e_{k+1}) = ix_k.$$

Rechnen wir das Skalarprodukt aus, so erhalten wir

$$b_{k-1}x_{k-1} + a_k x_k + b_k x_{k+1} = i x_k.$$

Mit Hilfe von (183) und der Methode der vollständigen Induktion ergibt sich hieraus $x_k = P_k(i) x_0$ und $x_0 \neq 0$. Dies widerspricht jedoch der Divergenz von (185). Ersetzen wir in (185) i durch $-i$, so gelangen wir wegen $P_k(-i) = \overline{P_k(i)}$ wieder zu einer divergenten Reihe; A^* hat also auch nicht den Eigenwert $-i$. Damit ist der Satz bewiesen.

Die Divergenz der Reihe (185) ist also eine notwendige und hinreichende Bedingung für die Selbstadjungiertheit von A . Wiederholen wir wörtlich den Beweis der letzten beiden Sätze, so finden wir: Konvergiert die Reihe

$$\sum_{k=0}^{\infty} |P_k(\alpha)|^2 \quad (186)$$

für ein gewisses nicht reelles α , so sind α und $\bar{\alpha}$ Eigenwerte von A^* ; divergiert (186) für ein gewisses nicht reelles α , so sind α und $\bar{\alpha}$ keine Eigenwerte von A^* . Daraus folgt die Selbstadjungiertheit von A , d. h., auch die Reihe (185) divergiert. Konvergiert aber (186) für ein nicht reelles α , so hat A^* nicht reelle Eigenwerte, und A ist nicht selbstadjungiert; die Reihe (185) konvergiert dann ebenfalls. Diese Überlegungen führen auf den folgenden Satz:

Satz 3. *Es sind nur die beiden folgenden Fälle möglich: Die Reihe (186) divergiert oder konvergiert für jedes nicht reelle α ; im ersten Fall ist A selbstadjungiert, im zweiten Fall nicht.*

Aus dem Beweis von Satz 2 folgt ferner: Konvergiert (185), so genügen die Komponenten desjenigen Eigenelements von A^* , das dem Eigenwert i entspricht, den Gleichungen $x_k = P_k(i) x_0$ ($k = 1, 2, \dots$), wobei x_0 beliebig und von 0 verschieden ist. Der Unterraum $M_1(A)$ ist also eindimensional. Das gleiche gilt von dem Unterraum $M_{-1}(A)$, der sich offenbar aus $M_1(A)$ durch Ersetzen der Elemente x_k durch ihre Konjugierten ergibt. Somit sind im zweiten Fall die Defektindizes gleich $(1, 1)$. Die Elemente x_1 und x_{-1} der Unterräume $M_1(A)$ bzw. $M_{-1}(A)$ können wir bis auf einen beliebigen komplexen Faktor durch die Formeln

$$x_1 = \sum_{k=0}^{\infty} P_k(i) e_k, \quad x_{-1} = \sum_{k=0}^{\infty} P_k(-i) e_k$$

beschreiben. Die Elemente des Unterraumes $D(A_\theta)$ einer selbstadjungierten Erweiterung A_θ von A sind eindeutig durch $v = x_A + \alpha x_\theta$ definiert, wobei $x_A \in D(A)$, α eine beliebige komplexe Zahl, $x_\theta = i(e^{-\theta/2} x_{-1} + e^{\theta/2} x_1)$ und $0 \leq \theta < 2\pi$ ist. Es sei nun $\mathcal{E}_\lambda$ die Spektralschar von A im ersten Fall oder von A_θ im zweiten Fall sowie $\varrho(\lambda) = (\mathcal{E}_\lambda e_0, e_0)$. Genau wie in [167] gilt

$$\int_{-\infty}^{\infty} P_k(\lambda) P_l(\lambda) d\varrho(\lambda) = \begin{cases} 0 & \text{für } k \neq l, \\ 1 & \text{für } k = l, \end{cases}$$

$$(A e_k, e_l) = \int_{-\infty}^{\infty} \lambda P_k(\lambda) P_l(\lambda) d\varrho(\lambda), \quad e_k = \int_{-\infty}^{\infty} P_k(\lambda) d\mathcal{E}_\lambda e_0,$$

wobei für den zweiten Fall nur A durch A_θ zu ersetzen ist. Die Elemente der Matrix (182) lassen sich offenbar durch

$$a_{ik} = \int_{-\infty}^{\infty} \lambda P_k(\lambda) P_i(\lambda) d\varrho(\lambda)$$

ausdrücken.

Satz 4. Die Polynome $P_k(\lambda)$ bilden ein bezüglich $\varrho(\lambda)$ vollständiges System.

Beweis. Es sei $\varphi_\mu(\lambda)$ eine Funktion, die für $-\infty < \lambda \leq \mu$ gleich 1 und für $\lambda > \mu$ gleich 0 ist. Jede Funktion $\pi(\lambda)$, die endlich viele Werte $a_1, a_2, \dots, a_m$ (jeden Wert a_k auf einem endlichen Intervall) annimmt, läßt sich als endliche Linearkombination der Funktionen $\varphi_\mu(\lambda)$ für verschiedene μ darstellen. Wenn wir für jedes μ die Parsevalsche Gleichung für $\varphi_\mu(\lambda)$ beweisen können, dann gilt sie auch für jede Linearkombination der $\varphi_\mu(\lambda)$ und somit für alle Funktionen $\pi(\lambda)$ des oben beschriebenen Typs. Die lineare Mannigfaltigkeit dieser Funktionen ist bezüglich $\varrho(\lambda)$ in L_2 überall dicht [60]; also bilden die $P_k(\lambda)$ ein vollständiges System [60]. Wir müssen nun noch die Parsevalsche Gleichung für die $\varphi_\mu(\lambda)$ beweisen. Dazu berechnen wir das Integral von $\varphi_\mu^2(\lambda)$ und die Fourierkoeffizienten dieser Funktion:

$$\int_{-\infty}^{\infty} \varphi_\mu^2(\lambda) d\varrho(\lambda) = \int_{(-\infty, \mu]} d\varrho(\lambda) = \varrho(\mu),$$

$$a_k = \int_{-\infty}^{\infty} \varphi_\mu(\lambda) P_k(\lambda) d\varrho(\lambda) = \int_{(-\infty, \mu]} P_k(\lambda) d\varrho(\lambda).$$

Es muß also für jedes μ

$$\varrho(\mu) = \sum_{k=0}^{\infty} \int_{(-\infty, \mu]} P_k(\lambda) d\varrho(\lambda) \cdot \int_{(-\infty, \mu]} P_k(\lambda) d\varrho(\lambda)$$

gelten. Auf Grund der Parsevalschen Gleichung haben wir

$$\varrho(\mu) = \|\mathcal{E}_\mu e_0\|^2 = \sum_{k=0}^{\infty} (\mathcal{E}_\mu e_0, e_k) (e_k, \mathcal{E}_\mu e_0),$$

so daß es genügt, die Gültigkeit von

$$(\mathcal{E}_\mu e_0, e_k) = (e_k, \mathcal{E}_\mu e_0) = \int_{(-\infty, \mu]} P_k(\lambda) d\varrho(\lambda)$$

nachzuprüfen (die rechte Seite dieser Beziehung ist reell). Diese Beziehung folgt schon aus der früher angegebenen Integraldarstellung der e_k (vgl. [192]). Damit ist der Satz bewiesen.

Wir geben nun ein einfaches hinreichendes Kriterium für die Selbstdjungiertheit der Matrix (182). Aus (183) folgt unmittelbar

$$b_k \frac{P_{k+1}(\alpha) \overline{P_k(\alpha)} - \overline{P_{k+1}(\alpha)} P_k(\alpha)}{\alpha - \bar{\alpha}} = |P_k(\alpha)|^2 + b_{k-1} \frac{P_k(\alpha) \overline{P_{k-1}(\alpha)} - \overline{P_k(\alpha)} P_{k-1}(\alpha)}{\alpha - \bar{\alpha}};$$

summieren wir hier über k von $k=0$ bis $k=n-1$, so ergibt sich

$$\sum_{k=0}^{n-1} |P_k(\alpha)|^2 = b_{n-1} \frac{P_n(\alpha) \overline{P_{n-1}(\alpha)} - \overline{P_n(\alpha)} P_{n-1}(\alpha)}{\alpha - \bar{\alpha}}$$

und insbesondere für $\alpha = i$

$$\sum_{k=0}^{n-1} |P_k(i)|^2 = b_{n-1} \frac{P_n(i) \overline{P_{n-1}(i)} - \overline{P_n(i)} P_{n-1}(i)}{2i} = b_{n-1} \operatorname{Im} [P_n(i) \overline{P_{n-1}(i)}].$$

Die linke Seite ist wegen $P_0(\lambda) \equiv 1$ mindestens gleich 1, und daraus folgt

$$\frac{1}{b_{n-1}} \leq \operatorname{Im} [P_n(i) \overline{P_{n-1}(i)}] \leq |P_n(i)| \cdot |P_{n-1}(i)| \leq \frac{1}{2} [|P_n(i)|^2 + |P_{n-1}(i)|^2]$$

oder, wenn wir hier von $n=1$ bis $n=m+1$ summieren,

$$\sum_{n=0}^m \frac{1}{b_n} \leq \sum_{n=0}^{m+1} |P_n(i)|^2.$$

Auf Grund von Satz 2 erhalten wir hieraus den folgenden Satz:

Satz 5. *Divergiert die aus den Elementen $1/b_n$ gebildete Reihe, so ist A ein selbstadjungierter Operator.*

Wir erwähnen ohne Beweis noch zwei Tatsachen, die mit den vorhergehenden Darlegungen eng zusammenhängen. Hat A die Defektindizes $(1, 1)$, so konvergiert (186) für jedes α . Ist A ein selbstadjungierter Operator, so divergiert (186) auch für jedes reelle α , außer denen, die dem Punktspektrum von A entsprechen, falls es ein solches Spektrum gibt. Außerdem hat jede selbstadjungierte Erweiterung von A auf Grund der Defektindizes $(1, 1)$ ein reines Punktspektrum.¹⁾

Als Beispiel für eine Jacobische Matrix betrachten wir die Hermiteschen Polynome

$$H_k(\lambda) = (-1)^k e^{\lambda^2} \frac{d^k}{d\lambda^k} (e^{-\lambda^2}); \quad (187)$$

für sie gilt die Rekursionsformel

$$\lambda H_k(\lambda) = \frac{1}{2} H_{k+1}(\lambda) + k H_{k-1}(\lambda) \quad (188)$$

und die Integralgleichung [III/2, 158]

$$\int_{-\infty}^{\infty} e^{-\lambda^2} H_k^2(\lambda) d\lambda = 2^k k! \sqrt{\pi}. \quad (189)$$

Da wir mit normierten Polynomen arbeiten wollen, führen wir statt (187) die Polynome

$$P_k(\lambda) = \frac{(-1)^k}{\sqrt{2^k k!}} e^{\lambda^2} \frac{d^k}{d\lambda^k} (e^{-\lambda^2}) \quad (190)$$

ein. Dann erhält (188) die Gestalt

$$\lambda P_k(\lambda) = \sqrt{\frac{k+1}{2}} P_{k+1}(\lambda) + \sqrt{\frac{k}{2}} P_{k-1}(\lambda) \quad \text{mit} \quad P_0(\lambda) \equiv 1.$$

Wählen wir eine Jacobische Matrix, indem wir

$$a_k = 0, \quad b_k = \sqrt{\frac{k+1}{2}} \quad (k=0, 1, 2, \dots)$$

setzen, so gelangen wir zu den Polynomen (190); dabei gelten die Formeln (vgl. [III/2, 158])

$$\begin{aligned} \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} e^{-\lambda^2} P_k(\lambda) P_l(\lambda) d\lambda &= \begin{cases} 0 & \text{für } k \neq l, \\ 1 & \text{für } k = l, \end{cases} \\ \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} e^{-\lambda^2} \lambda P_k(\lambda) P_l(\lambda) d\lambda &= \begin{cases} 0 & \text{für } |k-l| \neq 1, \\ \sqrt{\frac{k}{2}} & \text{für } l = k+1. \end{cases} \end{aligned}$$

Aus Satz 5 folgt, daß A in diesem Fall ein selbstadjungierter Operator ist. Berücksichtigen wir die oben angegebenen Integralformeln, so gilt für den Operator A

$$\varrho(\lambda) = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\lambda} e^{-\mu^2} d\mu.$$

¹⁾ N. I. ACHESER, Unendliche Jacobische Matrizen und das Momentenproblem (russ.), Uspechi mat. nauk 9 (1941), 126–156.

Der Operator A hat ein einfaches kontinuierliches Spektrum, das das ganze Intervall $(-\infty, \infty)$ bedeckt.

216. Matrizen und Operatoren. Wir untersuchen nun den Zusammenhang zwischen Matrizen und symmetrischen Operatoren des Hilbertschen Raumes H . Zunächst nehmen wir an, daß in diesem Raum ein beschränkter selbstadjungierter Operator A gegeben und $\varphi_1, \varphi_2, \dots$ ein vollständiges normiertes Orthogonalsystem von Elementen aus H ist. Mit Hilfe der Formeln

$$a_{nk} = (A\varphi_k, \varphi_n) = (\varphi_k, A\varphi_n) \quad (a_{kn} = \overline{a_{nk}}) \quad (191)$$

definieren wir Elemente einer gewissen Matrix und erhalten

$$x'_n = \sum_{k=1}^{\infty} a_{nk} x_k, \quad (192)$$

wobei die x_k die Komponenten eines Elements x sind, also $x_k = (x, \varphi_k)$, und die x'_k die der transformierten Elemente sind, also $x'_k = (Ax, \varphi_k)$. Somit wird bei einer bestimmten Wahl der Einheitsvektoren der Operator A vermöge der Formeln (192) als Matrix ausgedrückt. Greifen wir ein anderes System von Einheitsvektoren $\psi_1, \psi_2, \dots$ heraus und ist U eine unitäre Transformation derart, daß $U\varphi_k = \psi_k$ ($k=1, 2, \dots$) und $u_{pq} = (U\varphi_q, \varphi_p) = (\psi_q, \varphi_p)$ ist, so entspricht dem Operator A eine Matrix mit den Elementen

$$\begin{aligned} b_{nk} &= (A\psi_k, \psi_n) = \sum_{s=1}^{\infty} (A\psi_k, \varphi_s) \overline{(\psi_n, \varphi_s)} = \sum_{s=1}^{\infty} (\psi_k, A\varphi_s) \overline{(\psi_n, \varphi_s)} \\ &= \sum_{s=1}^{\infty} \overline{(\psi_n, \varphi_s)} \sum_{t=1}^{\infty} (\psi_k, \varphi_t) \overline{(A\varphi_s, \varphi_t)}; \end{aligned} \quad (193)$$

dabei benutzen wir die verallgemeinerte Parsevalsche Gleichung (18₁) aus [121]. Mit Hilfe der oben eingeführten Bezeichnungen ist

$$b_{nk} = \sum_{s=1}^{\infty} \overline{u_{sn}} \sum_{t=1}^{\infty} a_{st} u_{tk}. \quad (194)$$

Wenden wir diese Formel auf $b_{kn} = \overline{b_{nk}}$ an und gehen wir zu den konjugierten Größen über, wobei wir noch auf der rechten Seite von (194) s durch t und t durch s ersetzen, so erhalten wir

$$b_{nk} = \sum_{t=1}^{\infty} u_{tk} \sum_{s=1}^{\infty} \overline{u_{sn}} a_{st}. \quad (195)$$

Analog ist

$$a_{nk} = \sum_{s=1}^{\infty} u_{ns} \sum_{t=1}^{\infty} b_{st} \overline{u_{kt}} = \sum_{t=1}^{\infty} \overline{u_{kt}} \sum_{s=1}^{\infty} u_{ns} b_{st}. \quad (196)$$

Ist umgekehrt eine Matrix von Elementen a_{nk} gegeben, die der Bedingung $a_{kn} = \overline{a_{nk}}$ und der Beschränktheitsbedingung [163] genügen, und wählen wir ein System von Einheitsvektoren φ_k ($k=1, 2, \dots$) in H , so definieren die Formeln (192) einen beschränkten selbstadjungierten Operator A in H . Dabei gibt die k -te Spalte der Matrix der a_{nk} die Komponenten des aus den φ_k erhaltenen transformierten Elements an, so daß wir

$$A\varphi_k = \sum_{n=1}^{\infty} a_{nk} \varphi_n$$

erhalten.

Komplizierter ist der Zusammenhang zwischen Matrizen und nichtbeschränkten Operatoren. Im folgenden werden wir die Matrizen, deren Elemente den Bedingungen $a_{kn} = \overline{a_{nk}}$ und (179) genügen, *C-Matrizen* nennen. Gegeben sei ein abgeschlossener symmetrischer Operator F mit einem in H dichten Definitionsbereich $D(F)$. Wir wählen

dann ein vollständiges normiertes Orthogonalsystem von Elementen φ_k derart, daß alle φ_k zu $D(F)$ gehören, und definieren die Elemente a_{nk} mit Hilfe von (191), wobei wir A durch F ersetzen. Wegen der Symmetrie von F und der Parsevalschen Gleichung können wir behaupten, daß die Matrix der a_{nk} eine C -Matrix ist. Wenden wir auf das Skalarprodukt $(Fx, \varphi_n) = (x, F\varphi_n)$ mit $x \in D(F)$ die verallgemeinerte Parsevalsche Gleichung an, so erhalten wir

$$x'_n = (Fx, \varphi_n) = \sum_{k=1}^{\infty} (x, \varphi_k) \overline{(F\varphi_n, \varphi_k)} = \sum_{k=1}^{\infty} a_{nk} x_k, \quad (197)$$

d. h., der Operator F wird durch (192) in Einheitsvektoren φ_k ausgedrückt. Die lineare Mannigfaltigkeit $D(F)$ enthält offenbar alle endlichen Elemente, d. h. alle endlichen Linearkombinationen der Einheitsvektoren. Da der in [214] mit Hilfe der Matrix der a_{nk} definierte Operator A die Abschließung des Operators A' ist, der mit Hilfe von (192) auf der linearen Mannigfaltigkeit der endlichen Elemente definiert ist, können wir behaupten, daß $A \subseteq F$ ist. Folglich gilt $F^* \subseteq A^*$, d. h., F^* ist ebenfalls durch (192) auf der entsprechenden linearen Mannigfaltigkeit $D(F^*)$ gegeben. Ist F eine Erweiterung von A , so ist $D(F^*)$ wegen Satz 1 aus [186] nur ein Teil der in [214] definierten linearen Mannigfaltigkeit $D(B)$. In diesem Fall ergibt ein und dieselbe Matrix verschiedene Operatoren.

Nehmen wir statt der φ_k ein anderes System von Einheitsvektoren ψ_k , die ebenfalls zu $D(F)$ gehören, so läßt sich der Operator F durch eine Matrix von Elementen b_{nk} ausdrücken, wobei wie oben die Formeln (192) und (197) gelten, wenn a_{nk} durch b_{nk} ersetzt wird. Wir setzen jetzt voraus, daß nicht ein Operator, sondern die C -Matrix der Elemente a_{nk} gegeben ist. Ferner wählen wir ein beliebiges System von Einheitsvektoren φ_k aus H und definieren den Operator A' vermöge (192) oder (197) auf der linearen Mannigfaltigkeit $D(A')$ der endlichen Elemente. Die Abschließung A' führt auf einen abgeschlossenen symmetrischen Operator A . Wir wollen sagen, der Operator A wird von der Matrix der a_{nk} und dem System der φ_k erzeugt, und schreiben $A \sim a_{nk}\{\varphi_k\}$. Auf diese Art erhalten wir jeden symmetrischen abgeschlossenen Operator in H .

Satz. Jeder symmetrische abgeschlossene Operator A mit einem in H dichten Definitionsbereich $D(A)$ kann durch eine C -Matrix und ein System von Einheitsvektoren φ_k erzeugt werden.

Beweis. Es genügt, die entsprechenden Einheitsvektoren φ_k aus $D(A)$ zu konstruieren. Die Matrix ist durch (191) bestimmt. Die φ_k müssen der folgenden Forderung genügen: Für jedes Element $x \in D(A)$ existiert eine solche Folge von endlichen Elementen ω_n , daß $\omega_n \rightarrow x$ und $A\omega_n \rightarrow Ax$ gilt. Um solche φ_k zu erhalten, brauchen wir nur eine Folge von ω'_n aus $D(A)$ zu bilden, so daß es für jedes $x \in D(A)$ eine Teilfolge $\omega'_{n_1}, \omega'_{n_2}, \dots$ gibt, für welche $\omega'_{n_k} \rightarrow x$ und $A\omega'_{n_k} \rightarrow Ax$ gilt. Die Orthogonalisierung der ω'_n führt offenbar auf die φ_k . Dabei ist wegen $\omega'_{n_k} \rightarrow x$ und der Tatsache, daß $D(A)$ in H dicht ist, auch die Folge der ω'_n in H dicht; das System der φ_k ist also abgeschlossen. Wir gehen jetzt zur Konstruktion der ω'_n über und wählen eine beliebige Folge von Elementen $\chi_1, \chi_2, \dots$, die in H dicht ist. Ferner sei p, q, r ein Tripel von natürlichen Zahlen. Existiert ein Element $x \in D(A)$, für welches die Ungleichungen $\|\chi_p - x\| \leq 1/r$ und $\|\chi_q - Ax\| \leq 1/r$ erfüllt sind, so ordnen wir diesem Element x das Tripel p, q, r zu und schreiben $x_{p,q,r}$. Solche Elemente können offenbar numeriert werden [1], und wir werden zeigen, daß sie allen an ω'_n gestellten Forderungen genügen. Es sei $x \in D(A)$ und ε eine gegebene positive Zahl. Wir wählen ein r , das der Ungleichung $1/r \leq \varepsilon/2$ genügt, und Elemente χ_p, χ_q derart, daß $\|\chi_p - x\| \leq 1/r$ und $\|\chi_q - Ax\| \leq 1/r$ erfüllt ist. Das ist möglich, da die χ_n in H dicht sind. Dann existiert ein Element $x_{p,q,r}$, für welches $\|\chi_p - x_{p,q,r}\| \leq 1/r$ und $\|\chi_q - Ax_{p,q,r}\| \leq 1/r$ ist; hieraus erhalten wir

$$\|x - x_{p,q,r}\| \leq \varepsilon \quad \text{und} \quad \|Ax - Ax_{p,q,r}\| \leq \varepsilon,$$

wenn wir die Ungleichungen

$$\begin{aligned} \|x - x_{p,q,r}\| &\leq \|\chi_p - x\| + \|\chi_p - x_{p,q,r}\|, \\ \|Ax - Ax_{p,q,r}\| &\leq \|\chi_q - Ax\| + \|\chi_q - Ax_{p,q,r}\| \end{aligned}$$

und $1/r \equiv \varepsilon/2$ berücksichtigen. Das Element $x_{p,q,r}$ hat also, da ε beliebig ist, alle Eigenschaften, mit deren Hilfe wir oben die Folge der ω'_n charakterisierten. Damit ist der Satz bewiesen.

Ein und derselbe abgeschlossene symmetrische Operator kann auch von verschiedenen Matrizen und Einheitsvektoren erzeugt werden. Ist $A \sim a_{nk}\{\varphi_k\}$ und $A \sim b_{nk}\{\psi_k\}$, so erhalten wir die Formeln (194) bis (196), wenn wir den oben erwähnten unitären Operator U einführen und $u_{pq} = (\psi_q, \varphi_p)$ setzen.

Bemerkung. Ist F ein symmetrischer abgeschlossener Operator, gehört das System der Einheitsvektoren φ_k zu $D(F)$, ist die durch (191) definierte Matrix der a_{nk} gegeben (wobei A durch F ersetzt werden muß) und ist $A \sim a_{nk}\{\varphi_k\}$, so stimmt F entweder mit A überein, oder F ist eine Erweiterung von A , wie wir schon oben gesehen haben.

217. Unitäre Äquivalenz von C -Matrizen. In [216] betrachteten wir zwei Systeme $a_{nk}\{\varphi_k\}$ und $b_{nk}\{\psi_k\}$, die ein und denselben Operator A erzeugten. Die Zahlen u_{pq} waren dabei Elemente der Matrix, die der unitären Transformation $U\varphi_k = \psi_k$ entspricht. Die innere Summe in (194) stellt das Skalarprodukt $(A\psi_k, \varphi_s)$ dar, und wegen der Parsevalschen Gleichung können wir behaupten, daß das Quadrat des Absolutbetrages dieser Summe eine bei Summation über s konvergente Reihe bildet. Die innere Summe in (195) ergibt sich aus der inneren Summe von (194) durch Übergang zu den konjugierten Größen, durch Vertauschen von s und t und von k und n ; da auch für sie das oben über das Quadrat des Absolutbetrages Gesagte (jetzt bei Summation über t) gilt, haben wir

$$\sum_{s=1}^{\infty} \left| \sum_{t=1}^{\infty} a_{st} u_{tk} \right|^2 < \infty, \quad \sum_{t=1}^{\infty} \left| \sum_{s=1}^{\infty} \overline{u_{sn}} a_{st} \right|^2 < \infty; \quad (198)$$

dies führt auf die folgende Definition:

Definition 1. Die unitäre Matrix der u_{pq} ist auf die C -Matrix der a_{nk} anwendbar, wenn die Bedingungen (198) erfüllt sind und die Doppelsummen (194) und (195) auf das gleiche Resultat führen. Die erhaltene Matrix von Elementen b_{nk} nennen wir die *transformierte Matrix*.

Aus der Tatsache, daß die a_{nk} eine C -Matrix und die u_{pq} eine unitäre Matrix bilden, folgt auf Grund der Cauchyschen Ungleichung die absolute Konvergenz der inneren Reihen in (194) und (195); aus (198) ergibt sich dann die Konvergenz der äußeren Reihen. Auf Grund des oben über die inneren Summen Gesagten zieht eine der Bedingungen (198) die andere Bedingung nach sich. Aus (194) und (195) folgt unmittelbar $b_{kn} = b_{nk}$, und aus (198) und der Tatsache, daß die Matrix der u_{pq} unitär ist, sehen wir, daß die Glieder $|b_{nk}|^2$, über k summiert, eine endliche Summe ergeben, d. h., die b_{nk} bilden eine C -Matrix.

Satz 1. Ist eine unitäre Matrix U auf die Matrix der a_{nk} anwendbar, so ist die inverse Matrix U^{-1} auf die Matrix der b_{nk} anwendbar; die a_{nk} bilden die transformierte Matrix.

Beweis. Wir müssen die Gültigkeit der Beziehungen

$$\sum_{n=1}^{\infty} u_{pn} b_{nk} = \sum_{t=1}^{\infty} a_{pt} u_{tk}, \quad \sum_{k=1}^{\infty} b_{nk} \overline{u_{qk}} = \sum_{s=1}^{\infty} \overline{u_{sn}} a_{sq} \quad (199)$$

nachweisen, brauchen jedoch nur die erste Formel zu betrachten, da der Beweis für die zweite analog geführt wird. Es gilt

$$\sum_{n=1}^{\infty} u_{pn} b_{nk} = \sum_{n=1}^{\infty} \left[\sum_{s=1}^{\infty} \left(\sum_{t=1}^{\infty} a_{st} u_{tk} \right) \overline{u_{st}} \right] u_{pn}.$$

Wegen (198) können wir die Summe in eckigen Klammern als Komponente ξ_s eines Elements $\xi \in l_2$ auffassen. Bezeichnen wir mit γ den durch die Matrix der u_{pq} realisierbaren unitären Operator aus l_2 , so können wir der rechten Seite der letzten Formel die

Gestalt

$$\sum_{n=1}^{\infty} (\gamma^{-1}\xi)_n u_{pn} = [\gamma(\gamma^{-1}\xi)]_p = \xi_p = \sum_{t=1}^{\infty} a_{pt} u_{tk} \quad (200)$$

geben; daraus folgt sofort die erste der Beziehungen (199). Aus (199) ergibt sich unmittelbar

$$\sum_{k=1}^{\infty} \left| \sum_{n=1}^{\infty} u_{pn} b_{nk} \right|^2 < \infty, \quad \sum_{n=1}^{\infty} \left| \sum_{k=1}^{\infty} b_{nk} \overline{u_{qk}} \right|^2 < \infty. \quad (201)$$

Setzen wir nämlich $\eta_t = \overline{a_{pt}}$, so gibt es ein Element η aus l_2 mit den Komponenten η_t ; die rechte Seite der ersten Beziehung (199) läßt sich dann in der Form $\overline{(\gamma^{-1}\eta)_k}$ schreiben, so daß sich sofort die erste Beziehung (201) ergibt. Die zweite Beziehung wird analog bewiesen. Führen wir noch ein Element η' aus l_2 mit den Komponenten $\eta'_n = b_{nk}$ ein, so können wir die erste Formel (199) in der Gestalt $(\gamma\eta')_p = \overline{(\gamma^{-1}\eta)_k}$ schreiben, woraus sich $(\gamma^{-1}\eta)_k = \overline{(\gamma\eta')_p}$ ergibt. Nach Multiplikation mit u_{qk} und Summation über k finden wir

$$\overline{a_{pq}} = \sum_{k=1}^{\infty} \left(\sum_{n=1}^{\infty} \overline{u_{pn} b_{nk}} \right) u_{qk},$$

d. h. die eine der Beziehungen (196). Die zweite läßt sich analog nachweisen. Damit ist der Satz bewiesen.

Wir geben nun eine Definition für die unitäre Äquivalenz zweier Systeme $a_{nk}\{\varphi_k\}$ und $b_{nk}\{\psi_k\}$, wobei jedes ein gewisses abgeschlossenes System von Einheitsvektoren enthält.

Definition 2. Zwei Systeme $a_{nk}\{\varphi_k\}$ und $b_{nk}\{\psi_k\}$ heißen *unitär äquivalent*, wenn die unitäre Matrix U der Elemente $u_{pq} = (\varphi_q, \varphi_p)$ auf die Matrix der a_{nk} anwendbar ist und dies auf die transformierte Matrix der b_{nk} führt.

Ist dies der Fall, so folgt aus Satz 1, daß die unitäre Matrix, die zur Matrix der Elemente u_{pq} invers ist, d. h. die Matrix der Elemente $u_{pq}^* = \overline{u_{qp}} = (\varphi_q, \psi_p)$, auf die Matrix der b_{nk} anwendbar ist und die a_{nk} die transformierte Matrix bilden, d. h., die *unitäre Äquivalenz zweier Systeme ist eine Wechselbeziehung*. Wir weisen noch darauf hin, daß $U\varphi_k = \psi_k$ und $U^{-1}\psi_k = \varphi_k$ ist.

Der folgende Satz ist für die Äquivalenz von Systemen grundlegend:

Satz 2. Zwei Systeme $a_{nk}\{\varphi_k\}$ und $b_{nk}\{\psi_k\}$ sind genau dann unitär äquivalent, wenn die von ihnen erzeugten abgeschlossenen symmetrischen Operatoren A und A_1 ein und denselben symmetrischen Operator F als Erweiterung haben.

Beweis. Wir beweisen zunächst die Notwendigkeit der Bedingung; wir nehmen also an, die Systeme seien unitär äquivalent. Es gilt

$$\begin{aligned} (A\varphi_n, \psi_q) &= \sum_{s=1}^{\infty} (A\varphi_p, \varphi_s) (\varphi_s, \psi_q) = \sum_{s=1}^{\infty} a_{sp} \overline{u_{sq}}, \\ (\varphi_p, A_1\psi_q) &= \sum_{s=1}^{\infty} (\psi_s, A_1\psi_q) (\varphi_p, \psi_s) = \sum_{s=1}^{\infty} b_{qs} \overline{u_{ps}}. \end{aligned}$$

Mit Hilfe der zweiten Gleichung von (199) erhalten wir $(A\varphi_p, \psi_q) = (\varphi_p, A_1\psi_q)$. Diese Beziehung gilt offenbar auch für endliche Linearkombinationen der φ_p und ψ_q . Berücksichtigen wir die Definition von $D(A)$ und $D(A_1)$ und gehen wir in den Skalarprodukten zur Grenze über, so ergibt sich

$$(Ax, y) = (x, A_1y) \quad \text{für } x \in D(A) \quad \text{und} \quad y \in D(A_1). \quad (202)$$

Gehört x gleichzeitig zu $D(A)$ und $D(A_1)$, so gilt außer (202) noch $(A_1x, y) = (x, A_1y)$ und auf Grund von (202) die Beziehung $(Ax - A_1x, y) = 0$ für alle $y \in D(A_1)$. Da $D(A_1)$ in

H dicht ist, folgt hieraus $Ax = A_1x$. Es sei nun $D(F)$ die lineare Mannigfaltigkeit derjenigen Elemente x , die sich in der Form $x = x' + y'$ mit $x' \in D(A)$ und $y' \in D(A_1)$ darstellen lassen, und wir setzen $Fx = Ax' + A_1y'$. Gibt es von x zwei Darstellungen $x = x' + y'$ und $x = x'' + y''$ mit $x', x'' \in D(A)$ und $y', y'' \in D(A_1)$, so folgt aus $x' - x'' = y'' - y'$, daß $x' - x''$ und $y'' - y'$ gleichzeitig zu $D(A)$ und $D(A_1)$ gehören. Auf Grund des oben Gesagten ist dann

$$Ax' - Ax'' = A_1y'' - A_1y',$$

also $Ax' + A_1y' = Ax'' + A_1y''$. Damit ist Fx eindeutig bestimmt. Wegen (202) und der Symmetrie von A und A_1 ist der Operator F symmetrisch. Er ist offenbar die Erweiterung von A und A_1 . Damit ist die Notwendigkeit bewiesen.

Wir nehmen nun an, der symmetrische Operator F sei die Erweiterung von A und A_1 . Es gilt

$$\begin{aligned} \sum_{t=1}^{\infty} a_{pt} u_t k &= \sum_{t=1}^{\infty} (A\varphi_t, \varphi_p) (\psi_k, \varphi_t) = \sum_{t=1}^{\infty} (\psi_k, \varphi_t) \overline{(A\varphi_p, \varphi_t)} \\ &= (\psi_k, A\varphi_p) = (\psi_k, F\varphi_p) = (F\psi_k, \varphi_p) = (A_1\psi_k, \varphi_p) \\ &= \sum_{n=1}^{\infty} (A_1\psi_k, \psi_n) (\psi_n, \varphi_p) = \sum_{n=1}^{\infty} u_{pn} b_{nk}, \end{aligned}$$

d. h., wir erhalten die erste der Formeln (199); analog ergibt sich auch die zweite. Wiederholen wir den Beweis von Satz 1, so sehen wir, daß die Matrix der u_{pq} auf die der a_{nk} anwendbar ist und die Formeln (194) bis (196) gelten. Damit ist die unitäre Äquivalenz der Systeme bewiesen.

Aus Satz 2 folgt, daß die unitäre Äquivalenz von Systemen durch die von ihnen erzeugten abgeschlossenen symmetrischen Operatoren vollständig bestimmt ist. In diesem Fall sagen wir, daß diese Operatoren *matrix-äquivalent* sind. Hieraus folgt unmittelbar, daß ein beschränkter Operator nur sich selbst und ein maximaler Operator nur seinem Teil matrix-äquivalent ist. Diese Äquivalenz ist eine Wechselbeziehung, aber nicht transitiv, d. h., ist ein Operator C_1 matrix-äquivalent einem Operator C_2 und ist C_2 matrix-äquivalent einem Operator C_3 , so folgt daraus noch nicht, daß auch C_1 matrix-äquivalent C_3 ist. Das ist offenbar nur der Fall, wenn C_2 maximal ist, denn C_1 und C_3 haben in diesem Fall die gemeinsame Erweiterung C_2 .

Die allgemeine Theorie der C -Matrizen unterscheidet sich wesentlich von der Theorie derjenigen Matrizen, denen beschränkte selbstadjungierte Operatoren entsprechen. Diese Theorie der C -Matrizen wurde von VON NEUMANN¹⁾ und WINTNER²⁾ dargelegt.

218. Die Existenz einer Spektralschar. Wir kommen nun zum Beweis des Hauptsatzes aus [192], nach dem es zu jedem selbstadjungierten Operator A eine Zerlegung der Einheit $\mathcal{E}_\lambda$ gibt, mit deren Hilfe der Operator durch das Stieltjesche Integral (58) aus [192] ausgedrückt werden kann. Wir benutzen dazu die Eigenschaften von A , die wir ohne die Formel (58) hergeleitet haben. In [189] bewiesen wir ohne diese Formel, daß für jedes komplexe λ ein beschränkter, in ganz H definierter Operator $(A - \lambda E)^{-1}$ existiert, so daß die Beziehung $x = (A - \lambda E)u$ eine eindeutige Abbildung von $D(A)$ in H vermittelt. Wir betrachten den Fall $\lambda = \pm i$ und setzen

$$x = (A - iE)u, \quad y = (A + iE)v \quad (u, v \in D(A))$$

und

$$u = (A - iE)^{-1}x, \quad v = (A + iE)^{-1}y \quad (x, y \in H).$$

Da A selbstadjungiert ist, gilt

$$[(A - iE)^{-1}]^* = (A + iE)^{-1}.$$

¹⁾ J. v. NEUMANN, Zur Theorie der unbeschränkten Matrizen, J. reine angew. Math. 161 (1929), 208–236.

²⁾ A. WINTNER, Spektraltheorie unendlicher Matrizen, Leipzig 1929.

Wir führen noch die selbstadjungierten Operatoren

$$\begin{aligned} C &= \frac{1}{2} [(A - iE)^{-1} + (A + iE)^{-1}], \\ B &= \frac{1}{2i} [(A - iE)^{-1} - (A + iE)^{-1}] \end{aligned} \quad (203)$$

ein und gelangen so zu

$$(A - iE)^{-1} = C + iB, \quad (A + iE)^{-1} = C - iB; \quad (204)$$

bei beliebiger Wahl von $x \in H$ gehören die Elemente Cx und Bx zu $D(A)$. Aus (204) folgt

$$\begin{aligned} (A - iE)(C + iB) &= (AC + B) + i(AB - C) = E, \\ (A + iE)(C - iB) &= (AC + B) - i(AB - C) = E, \end{aligned}$$

also

$$AC = E - B, \quad (205)$$

$$AB = C. \quad (206)$$

Für $x \in D(A)$ können wir auch

$$(C + iB)(A - iE)x = x \quad \text{und} \quad (C - iB)(A + iE)x = x$$

schreiben. Daraus ergibt sich, wenn wir die Klammern ausrechnen und das Ergebnis mit den obigen Formeln vergleichen,

$$ACx = CAx, \quad ABx = BAx \quad (x \in D(A)), \quad (207)$$

d. h., die beschränkten Operatoren B und C sind mit A im Sinne der Definition aus [191] vertauschbar.

Aus $BC = BAB = ABB = CB$ folgt

$$CB = BC; \quad (208)$$

auch B und C sind also miteinander vertauschbar. Mit Hilfe von (205) und (206) erhalten wir

$$B = BE = B(B + AC) = B^2 + BAC = B^2 + ABC = B^2 + C^2;$$

B ist also ein positiver Operator. Ferner folgt aus den Formeln

$$\begin{aligned} \|u\|^2 &= \|(A - iE)u\|^2 = ((A - iE)u, (A - iE)u) = \|Au\|^2 + \|u\|^2, \\ \|v\|^2 &= \|Av\|^2 + \|v\|^2, \end{aligned}$$

daß $\|u\| \leq \|x\|$ und $\|v\| \leq \|y\|$ ist, d. h., die Normen der Operatoren $(A - iE)^{-1}$ und $(A + iE)^{-1}$ sind höchstens gleich 1; dasselbe gilt wegen (203) auch für B und C . Wir zeigen noch, daß $x = 0$ aus $Bx = 0$ folgt. Ist nämlich $Bx = 0$, so ist auch

$$(Bx, x) = (B^2x, x) + (C^2x, x) = 0,$$

also $(Bx, Bx) + (Cx, Cx) = 0$ oder $\|Bx\|^2 + \|Cx\|^2 = 0$; daraus ergibt sich $Cx = 0$. Aus (205) folgt demnach $x = Bx + ACx = 0$.

Außer den oben genannten Eigenschaften der Operatoren B und C benötigen wir noch einen Hilfssatz, den wir aber erst später (am Schluß dieses Abschnitts) beweisen werden.

Hilfssatz. *Es seien M_n ($n = 1, 2, \dots$) paarweise orthogonale Unterräume, deren orthogonale Summe gleich H ist, und in jedem M_n sei ein selbstadjungierter beschränkter Operator A_n definiert. Dann existiert ein einziger selbstadjungierter Operator A in H , der in M_n mit A_n übereinstimmt. Die lineare Mannigfaltigkeit $D(A)$ besteht aus denjenigen Elementen x , für welche die Reihe aus den Elementen $\|A_n x_n\|^2$ konvergiert, wobei mit x_n die Projektion von x in den Unterraum M_n bezeichnet ist. Für diese x gilt*

$$Ax = \sum_n A_n x_n. \quad (209)$$

Wir kehren zu den Operatoren B und C zurück. Das Spektrum von B liegt auf dem Segment $[0, 1]$, und da $x=0$ aus $Bx=0$ folgt, gehört der Punkt $\lambda=0$ nicht zum Punktspektrum, so daß, wenn wir mit $\mathcal{E}'_\lambda$ die Spektralschar von B bezeichnen, $\mathcal{E}'_0=0$ ist. Bezeichnen wir mit M_n den Unterraum, in den der Operator $(\mathcal{E}'_{1/n} - \mathcal{E}'_{1/(n+1)})$ projiziert, so sind die M_n paarweise orthogonal, und ihre orthogonale Summe ist gleich H . Ist der beschränkte selbstadjungierte Operator F mit $\mathcal{E}'_\lambda$ vertauschbar, so reduziert M_n auf Grund eines Satzes aus [148] den Operator F . Dies gilt auch für C und jede reelle stetige Funktion von B [193]. Wir setzen $\varphi_n(\lambda)$ gleich $1/\lambda$ für $1/(n+1) \leq \lambda \leq 1/n$ und gleich einer Konstanten außerhalb dieses Intervalls (unter Beibehaltung der Stetigkeit in den Endpunkten) und führen den selbstadjungierten beschränkten Operator $\varphi_n(B)$ ein. Ist $z \in M_n$, so ist $\mathcal{E}'_\lambda z = z$ für $\lambda \geq 1/n$ und $\mathcal{E}'_\lambda z = 0$ für $\lambda \leq 1/(n+1)$. Dies folgt unmittelbar aus der Formel

$$\mathcal{E}'_\lambda z = \mathcal{E}'_\lambda (\mathcal{E}'_{1/n} - \mathcal{E}'_{1/(n+1)}) z$$

und der Beziehung $\mathcal{E}'_\mu \mathcal{E}'_\lambda = \mathcal{E}'_\lambda \mathcal{E}'_\mu = \mathcal{E}'_\mu$ für $\mu \leq \lambda$. Drücken wir $\varphi_n(B) z$ und Bz mit Hilfe von Stieltjeschen Integralen durch $\mathcal{E}'_\lambda$ aus und benutzen wir die Definition von $\varphi_n(\lambda)$ sowie das eben über $\mathcal{E}'_\lambda z$ Gesagte, so finden wir

$$\varphi_n(B) B = B \varphi_n(B) = E \quad \text{in } M_n,$$

also sind B und $\varphi_n(B)$ in M_n zueinander invers. Gehört z zu M_n , so können wir $z = B \varphi_n(B) z$ schreiben; daraus ist ersichtlich, daß mit $z \in M_n$ auch $z \in D(A)$ ist. Mit Hilfe von (206) erhalten wir $Az = A B \varphi_n(B) z = C \varphi_n(B) z$; also ist A ein in M_n beschränkter Operator. Berücksichtigen wir noch, daß C und $\varphi_n(B)$ vertauschbar sind und M_n die Operatoren C und $\varphi_n(B)$ reduziert, so sehen wir, daß A beschränkt und selbstadjungiert in M_n ist. Mit $\mathcal{E}_\lambda^{(n)}$ bezeichnen wir nun die diesem Operator entsprechende Zerlegung der Einheit in M_n . Auf Grund des Hilfssatzes können wir den durch die Beziehung

$$\mathcal{E}_\lambda x = \sum_{n=1}^{\infty} \mathcal{E}_\lambda^{(n)} x, \quad (210)$$

definierten, in H selbstadjungierten Operator $\mathcal{E}_\lambda$ bilden und leicht nachprüfen, daß $\mathcal{E}_\lambda$ die Zerlegung der Einheit in H ist. Bilden wir das Stieltjesche Integral

$$\int_{-\infty}^{\infty} \lambda d\mathcal{E}_\lambda x,$$

das einen selbstadjungierten Operator definiert, so sehen wir, wenn wir $\mathcal{E}_\lambda x = \mathcal{E}_\lambda^{(n)} x$ für $x \in M_n$ berücksichtigen, daß A durch dieses Integral für $x \in M_n$ gegeben wird; infolgedessen definiert es auf Grund des Hilfssatzes den Operator A .

Wir müssen nun noch den Hilfssatz beweisen. Es gehöre x zu der linearen Mannigfaltigkeit $D(A)$ derjenigen Elemente, für welche

$$\sum_n \|A_n x_n\|^2 < \infty$$

gilt. Den Operator A definieren wir durch $Ax = A_1 x_1 + A_2 x_2 + \dots$ und zeigen, daß A selbstadjungiert ist. Zu $D(A)$ gehören offenbar die endlichen Summen der Elemente x_n ; daher ist $D(A)$ in H dicht. Die Symmetrie von A folgt aus der Selbstadjungiertheit der A_n und der Formel

$$(Ax, y) = \left(\sum_n A_n x_n, \sum_n y_n \right) = \sum_n (A_n x_n, y_n) = \sum_n (x_n, A_n y_n) = \left(\sum_n x_n, \sum_n A_n y_n \right) = (x, Ay),$$

wobei x und y zu $D(A)$ gehören und die Stetigkeit und die Distributivität des Skalarprodukts sowie die Orthogonalität der Unterräume M_n benutzt wurde. Somit ist $A^* \supseteq A$. Wir müssen noch zeigen, daß mit $x \in D(A^*)$ auch $x \in D(A)$ ist. Da $A^2 x_n$ zu M_n gehört und $x - x_n \perp M_n$ ist, können wir

$$(A^*(x - x_n), Ax_n) = (x - x_n, A^2 x_n) = 0$$

schreiben. Hieraus folgt für $n=1$ mit Hilfe des Satzes des PYTHAGORAS

$$||A^*x||^2 = ||A^*(x-x_1)||^2 + ||Ax_1||^2.$$

Analog können wir

$$||A^*(x-x_1)||^2 = ||A^*(x-x_1-x_2)||^2 + ||Ax_2||^2,$$

d. h.

$$||A^*x||^2 = ||A^*(x-x_1-x_2)||^2 + ||Ax_1||^2 + ||Ax_2||^2$$

und allgemein

$$||A^*x||^2 = ||A^*(x-x_1-x_2-\dots-x_n)||^2 + \sum_{k=1}^n ||Ax_k||^2$$

schreiben; hieraus folgt

$$\sum_{k=1}^{\infty} ||Ax_k||^2 \leq ||A^*x||^2$$

und somit $x \in D(A)$. Damit ist die Selbstadjungiertheit von A bewiesen. Wir müssen nun noch zeigen, daß nur ein einziger selbstadjungierter Operator existiert, der mit A_n auf M_n übereinstimmt. Wir nehmen dazu an, daß es außer A noch einen Operator A' gibt. Aus der Selbstadjungiertheit des Operators A' folgt seine Abgeschlossenheit. Für endliche Summen gilt

$$A' \sum_{k=1}^m x_k = A \sum_{k=1}^m x_k = \sum_{k=1}^m Ax_k,$$

denn A und A' stimmen auf M_n überein. Wegen der Abgeschlossenheit von A' können wir behaupten, daß A' auf $D(A)$ definiert ist und dort mit A übereinstimmt, d. h., daß $A' \supseteq A$ ist. Andererseits gelangen wir, wenn wir in dem obigen Beweis A^* durch A' ersetzen, zu dem Schluß, daß mit $x \in D(A')$ auch $x \in D(A)$ ist; folglich stimmt A' mit A überein. Damit ist der Hilfssatz bewiesen.

Literaturhinweise

Die folgende Zusammenstellung erhebt keinen Anspruch auf Vollständigkeit; es sind vor allem neuere Werke aufgeführt und solche, die sich an den üblichen Leserkreis dieses Buches wenden. Einige der angegebenen Werke enthalten selbst ein ausführliches Literaturverzeichnis. Angaben über speziellere Einzelergebnisse lassen sich leicht aus den mathematischen Referatezeitschriften ermitteln.

Zur Ergänzung des behandelten Stoffes sei auf folgende Lehrbücher verwiesen:

- ALEXANDROFF, P. S.: Einführung in die Mengenlehre und in die allgemeine Topologie, VEB Deutscher Verlag der Wissenschaften, Berlin 1984 (Übersetzung aus dem Russischen).
- BÖGEL, K.; TASCHKE, M.: Analysis in normierten Räumen, Akademie-Verlag, Berlin 1974.
- DIEUDONNÉ, J.: Grundzüge der modernen Analysis, Bd. 1; 3. ber. und erg. Aufl., Vieweg, Braunschweig – Wiesbaden / VEB Deutscher Verlag der Wissenschaften, Berlin 1984 (Übersetzung aus dem Englischen und Französischen).
- DIEUDONNÉ, J.: History of functional analysis, North-Holland, Amsterdam 1981; 1st reprint 1983.
- EDWARDS, R. E.: Functional analysis. Theory and applications, Holt, Rinehart and Winston, New York 1965.
- FENYÖ, ST.; STOLLE, H. W.: Theorie und Praxis der linearen Integralgleichungen, Bd. 1, 2, 3, 4; VEB Deutscher Verlag der Wissenschaften, Berlin/Birkhäuser, Basel – Boston 1982–1984.
- GARNIR, H.; DE WILDE, M.; SCHMETS, J.: Analyse fonctionnelle. Théorie constructive des espaces linéaires à seminormes, T. 1: Théorie générale, 1968; T. 2: Mesure et intégration dans l'espace euclidien, 1972; T. 3: Espaces fonctionnels usuels, Birkhäuser, Basel – Boston, 1968–1973.
- KANTOROWITSCH, L. V.; AKILOW, G. P.: Funktionalanalysis in normierten Räumen, 2. ber. Aufl., Akademie-Verlag, Berlin / Verlag Harri Deutsch, Thun-Frankfurt/Main 1978 (Übersetzung aus dem Russischen).
- KÖTHE, G.: Topological vector spaces, Vol. I, 2nd rev. print. 1983, Vol. II, 1979, Springer, Berlin – Heidelberg – New York.
- KOLMOGOROV, A. N.; FOMIN, S. V.: Reelle Funktionen und Funktionalanalysis, VEB Deutscher Verlag der Wissenschaften, Berlin 1976 (Übersetzung aus dem Russischen).
- LANGENBACH, A.: Vorlesungen zur höheren Analysis, VEB Deutscher Verlag der Wissenschaften, Berlin 1984.
- LJUSTERNIK, L. A.; SOBOLEW, W. I.: Elemente der Funktionalanalysis, 5. ber. Aufl., Akademie-Verlag, Berlin / Verlag Harri Deutsch, Thun – Frankfurt/Main 1976 (Übersetzung aus dem Russischen).
- NATANSON, I. P.: Theorie der Funktionen einer reellen Veränderlichen, 2. erg. und überarb. Aufl., Akademie-Verlag, Berlin 1961 (Übersetzung aus dem Russischen).
- REDDY, J. N.: Applied functional analysis and variational methods in engineering, McGraw-Hill, New York 1986.

- REED, M.; SIMON, B.: *Methods of modern mathematical physics*, Vol 1.: *Functional analysis*, rev. and enlarg. ed., 1980; Vol. 2.: *Fourier analysis, self-adjointness*, 1975; Vol. 3.: *Scattering theory*, 1979; Vol. 4.: *Analysis of operators*, 1978, Academic-Press, New York.
- RIESZ, F.; SZ.-NAGY, B.: *Vorlesungen über Funktionalanalysis*, 4. Aufl., VEB Deutscher Verlag der Wissenschaften, Berlin/Verlag Harri Deutsch, Thun – Frankfurt/Main 1982 (Übersetzung aus dem Französischen).
- TRIEBEL, H.: *Analysis und mathematische Physik*, 2. bearb. Aufl., BSB B. G. Teubner Verlagsgesellschaft, Leipzig 1984.
- WULICH, B. S.: *Einführung in die Funktionalanalysis*, Teil 1, 1961, Teil 2, 1962, BSB B. G. Teubner Verlagsgesellschaft, Leipzig (Übersetzung aus dem Russischen).
- YOSIDA, K.: *Functional analysis*, 6th ed., Springer, Berlin – Göttingen – Heidelberg – New York 1980.

Zu den Kapiteln I–III sind speziell folgende Bücher zu nennen:

- AUMANN, G.: *Reelle Funktionen*, 2. Aufl., Springer-Verlag, Berlin – Göttingen – Heidelberg 1969.
- BELKNER, H.; BREHMER, S.: *Lebesguesche Integrale*, VEB Deutscher Verlag der Wissenschaften, Berlin 1984.
- BELKNER, H.; BREHMER, S.: *Riemannsche Integrale*, VEB Deutscher Verlag der Wissenschaften, Berlin 1984.
- BELKNER, H.; BREHMER, S.: *Riemannsche und Lebesguesche Integrale*, Verlag Harri Deutsch, Thun – Frankfurt/Main 1985.
- BENEDETTO, J. J.: *Real variable and integration, with historical notes*, B. G. Teubner, Stuttgart 1976.
- BERBERIAN, S. K.: *Measure and integration*, Chelsea 1970.
- BERNAYS, P.; FRAENKEL, A. A.: *Axiomatic set theory*, 2nd ed., North-Holland Publ. Comp., Amsterdam 1963.
- BOURBAKI, N.: *Théorie des ensembles*, Chap. I, II, 2ème éd., Hermann et Cie., Paris 1960.
- BREHMER, S.: *Einführung in die Maßtheorie*, Akademie-Verlag, Berlin 1975.
- BURKILL, J. C.: *The Lebesgue integral*, Cambridge University Press, Cambridge 1951.
- CARATHÉODORY, C.: *Maß und Integral und ihre Algebraisierung*, Birkhäuser, Basel 1956.
- CARATHÉODORY, C.: *Vorlesungen über reelle Funktionen*, 2. Aufl., B. G. Teubner, Leipzig 1927. – In diesem Werk befindet sich ein systematisch geordnetes Literaturverzeichnis mit vielen Hinweisen auf Originalarbeiten. Eine Neubearbeitung dieses Werkes ist unter dem Titel: *Reelle Funktionen*, Bd. I: *Zahlen/Punktmengen*, B. G. Teubner, Leipzig – Berlin 1939, erschienen.
- CHAE, S. B.: *Lebesgue integration*, M. Dekker, Basel 1980.
- COHN, C. P.: *Measure theory*, Birkhäuser, Basel-Boston 1980.
- CONSTANTINESCU, C.: *Spaces of measures*, W. de Gruyter, Berlin 1984.
- CRAVEN, B. D.: *Lebesgue measure and integral*, Pitman, Oxford 1982.
- DE BARRA, G.: *Introduction to measure and integration*, Wiley, New York 1981.
- DINCULEANU, N.: *Vector measures*, VEB Deutscher Verlag der Wissenschaften, Berlin 1966.
- ELSTRODT, J.: *Maß- und Integrationstheorie*, Springer-Verlag, Berlin – Heidelberg – New York 1986.
- FEDERER, H.: *Geometric measure theory*, Springer-Verlag, Berlin – Göttingen – Heidelberg 1969.
- FLORET, K.: *Maß- und Integrationstheorie*, B. G. Teubner, Stuttgart 1981.
- FRAENKEL, A.: *Einleitung in die Mengenlehre*, 3. Aufl., Springer, Berlin 1928.
- FRAENKEL, A.: *Zehn Vorlesungen über die Grundlagen der Mengenlehre*, B. G. Teubner, Leipzig – Berlin 1927.

- FRAENKEL, A. A.; BAR-HILLEL, Y.; LEVY, A.: Foundations of set theory, 3rd rev. ed., North-Holland Publ. Comp., Amsterdam 1984.
- ГЛИВЕНКО, В. И.: Интеграл Стильбеса, Гостехиздат, Москва – Ленинград 1936.
- HADWIGER, H.: Vorlesungen über Inhalt, Oberfläche und Isoperimetrie, Springer-Verlag, Berlin – Göttingen – Heidelberg 1957.
- HAHN, H.; ROSENTHAL, A.: Set functions, University of New Mexico Press, Albuquerque (New Mexico) 1948.
- HALMOS, P. R.: Measure theory, 2nd print., D. van Nostrand Comp., Inc., Toronto – New York – London 1950.
- HAUSDORFF, F.: Grundzüge der Mengenlehre, Veit & Comp., Leipzig 1914.
- HAUSDORFF, F.: Mengenlehre, 2. Aufl., W. de Gruyter, Berlin 1927.
- HAWKINS, T.: Lebesgue's theory of integration, 2nd ed., Chelsea Publ., New York 1975.
- HENZE, E.: Einführung in die Maßtheorie, 2. überarb. Aufl., Bibliographisches Institut, Mannheim 1985.
- HILDEBRANDT, T. H.: Introduction to the theory of integration, Academic Press, New York 1963.
- JACOBS, K.: Measure and integral, Academic Press, New York 1978.
- JANSEN, A. J. E. M.: Integration theory, Springer-Verlag, Berlin – Heidelberg – New York 1984.
- KAMKE, E.: Das Lebesgue-Stieltjessche Integral, 2. Aufl., BSB B. G. Teubner Verlagsgesellschaft, Leipzig 1960.
- KAMKE, E.: Mengenlehre, 5. Aufl., Sammlung Götschen, W. de Gruyter, Berlin 1965.
- KLAUA, D.: Allgemeine Mengenlehre, Akademie-Verlag, Berlin 1964.
- LASS, H.: Elements of pure and applied mathematics, McGraw-Hill, London 1957.
- LINDE, W.: Infinitely divisible and stable measures on Banach spaces, BSB B. G. Teubner Verlagsgesellschaft, Leipzig 1983.
- VON MANGOLDT, H.; KNOPP, K.; LÖSCH, F.: Einführung in die höhere Mathematik, Bd. 4: Mengenlehre, Lebesguesches Maß und Integral, topologische Räume, Vektorräume, Funktionalanalysis, Integralgleichungen, 2. durchges. Aufl., Hirzel-Verlag, Leipzig/Hirzel-Verlag, Stuttgart 1975.
- MAYRHOFER, K.: Inhalt und Maß, Springer-Verlag, Wien 1952.
- MICHEL, H.: Maß- und Integrationstheorie I, VEB Deutscher Verlag der Wissenschaften, Berlin 1978.
- MIKUSIŃSKI, J.: The Bochner integral, Birkhäuser, Basel 1978.
- MUNROE, M. E.: Measure and integration, Addison – Wesley, Reading 1971.
- NIELSEN, O. A.: Direct integral theory, M. Dekker, Basel 1980.
- OXTOBY, J. C.: Maß und Kategorie, Springer-Verlag, Berlin – Heidelberg – New York 1971.
- PFEFFER, W. F.: Integrals and measures, 2nd print., M. Decker, Basel 1977.
- PHILIPS, E. R.: An introduction to analysis and integration theory, Dover Publ., New York 1984.
- SAKS, S.: Theorie de l'intégrale, Warszawa 1933.
- SAKS, S.: Theory of the integral, 2nd rev. ed., Hafner Publ. Comp., New York 1937.
- SCHLESINGER, L.; PLESSNER, A.: Lebesguesche Integrale und Fouriersche Reihen, W. de Gruyter, Berlin 1926.
- SEGAL, I. E.; KUNZE, R. A.: Integral and operators, 2nd rev. and enlarg. ed., Springer-Verlag, Berlin – Heidelberg – New York 1978.
- TAYLOR, S. J.: Introduction to measure and integration, Cambridge Univ. Press, Cambridge 1966.
- TEMPLE, G. F. J.: The structure of Lebesgue integration theory, Clarendon Press, Oxford 1971.
- Theory of sets and topology. A collection of papers in honour of Felix Hausdorff, VEB Deutscher Verlag der Wissenschaften, Berlin 1972.
- TITCHMARSH, E. C.: The theory of functions, 2nd ed., Oxford Univ. Press, Oxford 1939; reissued 1975.

- TOLSTOW, G. P.: Maß und Integral, Akademie-Verlag, Berlin 1981 / Hanser-Verlag, München 1982.
- DE LA VALLÉE-POUSSIN, CH.: Intégrales de Lebesgue, fonctions d'ensembles, classes de Baire, Gauthier-Villars, Paris 1916.
- WEIR, A. J.: General integration and measure, Cambridge Univ. Press, Cambridge 1972.
- WEIR, A. J.: Lebesgue integration and measure, Cambridge Univ. Press, Cambridge 1973.
- WHEEDEN, R.; ZYGMUND, A.: Measure and integral; an introduction to real analysis, 3rd print., M. Dekker, Basel 1977.
- YOUNG, W. H.; YOUNG, G. C.: The theory of sets of points, Cambridge University Press, Cambridge 1906. – Dieses Werk enthält ein umfassendes Literaturverzeichnis mit allen bis zu diesem Zeitpunkt erschienenen Originalarbeiten.
- ZAHN, P.: Ein konstruktiver Weg zur Maßtheorie und Funktionalanalysis, Wissenschaftliche Buchgesellschaft, Darmstadt 1978.

Für die Kapitel IV und V sei noch speziell auf folgende Literatur verwiesen:

- ACHESER, N. I.; GLASMANN, I. M.: Theorie der linearen Operatoren im Hilbertraum, 8. erw. Aufl., Akademie-Verlag, Berlin 1981 (Übersetzung aus dem Russischen).
- Ахизер, Н. И.: Бесконечные матрицы Якоби и проблема моментов, Успехи мат. наук 9 (1941), 126–156.
- ADAMS, R. A.: Sobolev spaces, Academic Press, New York 1975.
- ALT, H. W.: Funktionalanalysis. Eine anwendungsorientierte Einführung, Springer-Verlag, Berlin – Heidelberg – New York 1985.
- AUBIN, J. B.: Applied functional analysis, Wiley, New York 1980.
- BACHMAN, G.; NARICI, L.: Functional analysis, Academic Press, New York 1966.
- BALAKRISHNAN, A. V.: Applied functional analysis, 2nd ed., Springer-Verlag, Berlin – Heidelberg – New York 1981.
- BANACH, S.: Théorie des opérations linéaires, Warszawa 1932.
- BARTLE, G.: Studies in functional analysis, Wiley, New York 1981.
- BEAUZAMY, B.: Introduction to Banach spaces and their geometry, North-Holland Publ. Comp., Amsterdam 1982.
- BERBERIAN, S. K.: Lectures in functional analysis and operator theory, Springer-Verlag, Berlin – Heidelberg – New York 1974.
- BERG, L.: Einführung in die Operatorenrechnung, 2. Aufl., VEB Deutscher Verlag der Wissenschaften, Berlin 1965.
- BERG, L.: Operatorenrechnung I, II, VEB Deutscher Verlag der Wissenschaften, Berlin 1972 bzw. 1974.
- BONIC, R. A.: Linear functional analysis, Gordon and Breach, London 1969.
- BREHMER, S.: Hilbert-Räume und Spektralmaße, Akademie-Verlag, Berlin / Vieweg, Braunschweig – Wiesbaden 1979.
- BROWN, A.; PEARCY, C.: Introduction to operator theory, 1. Elements of functional analysis, Springer-Verlag, Berlin – Heidelberg – New York 1977.
- BROWN, A. L.; PAGE, A.: Elements of functional analysis, van Nostrand – Reinhold, London – New York 1970.
- CONWAY, J. B.: A course in functional analysis, Springer-Verlag, Berlin – Heidelberg – New York 1985.
- COOKE, R. G.: Linear operators, Macmillan, London 1953.
- COURANT, R.; HILBERT, D.: Methoden der mathematischen Physik, Bd. I (3. Aufl.), Bd. II (2. Aufl.), Springer-Verlag, Berlin – Heidelberg – New York 1968.
- CURTAIN, R. F.; PRITCHARD, A. J.: Functional analysis in modern applied mathematics, Academic Press, New York 1977.
- DAY, M. M.: Normed linear spaces, 3rd ed., Springer-Verlag, Berlin – Heidelberg – New York 1973.

- DEIMLING, K.: Nonlinear functional analysis, Springer-Verlag, Berlin – Heidelberg – New York 1985.
- DE VITO, C. L.: Functional analysis, Academic Press, New York 1978.
- DIESTEL, J.: Geometry of Banach spaces: selected topics, Springer-Verlag, Berlin – Heidelberg – New York 1975.
- DUNFORD, N.; SCHWARTZ, J. T.: Linear operators I, Interscience Publ., New York 1958.
- EGLE, K.: Elemente der Funktionalanalysis, Verlagsgruppe Athenäum, Königstein 1980.
- EPSTEIN, B. S.: Linear functional analysis: Introduction to Lebesgue integration and infinite-dimensional problems, Saunders, London – New York 1970.
- GÖPFERT, A.; RIEDRICH, T.: Funktionalanalysis, 2. bearb. Aufl., BSB B. G. Teubner Verlagsgesellschaft, Leipzig 1986.
- GOFFMAN, C.: First course in functional analysis, 2nd ed., Chelsea Publ., New York 1983.
- GOHBERG, I.; GOLDBERG, S.: Basic operator theory, Birkhäuser, Basel.– Boston 1981.
- GRIFFEL, D. H.: Applied functional analysis, Wiley, New York 1981.
- GROETSCH, C. W.: Elements of applicable functional analysis, 2nd print., M. Dekker, Basel 1980.
- HALMOS, P. R.: A Hilbert space problem book, Springer-Verlag, Berlin – Heidelberg – New York 1974 / Van Nostrand, Princeton 1967.
- HALMOS, P. R.: Introduction to Hilbert space and the theory of spectral multiplicity, New York 1951.
- HEUSER, H.: Funktionalanalysis, B. G. Teubner, Stuttgart 1975.
- HEUSER, H.: Functional analysis, Wiley, New York 1981.
- HILLE, E.; PHILLIPS, R. S.: Functional analysis and semigroups, American Math. Soc., Providence 1957.
- HILLE, E.: Methods in classical and functional analysis, Addison – Wesley, Reading 1972.
- HOFFMAN, K.: Banach spaces of analytic functions, Prentice-Hall, Englewood Cliffs 1962.
- HOLMES, R. B.: Geometric functional analysis and its applications, Springer-Verlag, Berlin – Heidelberg – New York 1975.
- HUBER, A.: Funktionalanalysis, Verlag des Fachvereins an den Schweizer Hochschulen und Techniken, Zürich 1968.
- HUTSON, V.; PYM, J. S.: Applications of functional analysis and operator theory, Academic Press, New York 1979.
- ISTRATESCU, V.: Introduction to linear operator theory, M. Dekker, Basel 1981.
- JOSHI, M. C.; BOSE, R. K.: Some topics in nonlinear functional analysis, Wiley, New York 1985.
- JUNEK, H.: Locally convex spaces and operator ideals, BSB B. G. Teubner Verlagsgesellschaft, Leipzig 1983.
- KIRILLOV, A. A.; GVISHIANI, A. D.: Theorems and problems in functional analysis, Springer-Verlag, Berlin – Heidelberg – New York 1982.
- KÖTHE, G.: Topologische lineare Räume I, 2. verb. Aufl., Springer-Verlag, Berlin – Göttingen – Heidelberg 1966.
- KOWALSKY, H.-J.: Topologische Räume, Birkhäuser-Verlag, Basel – Stuttgart 1961.
- KREYSZIG, E.: Introductory functional analysis with applications, Wiley, New York 1978.
- KÜHNER, E.; LESKY, P.: Grundlagen der Funktionalanalysis und Approximationstheorie, Vandenhoeck & Ruprecht, Göttingen 1977.
- LARSEN, R.: Functional analysis. An introduction, 2nd print., M. Dekker, Basel 1973.
- LEIGH, J. R.: Functional analysis and linear control theory, Academic Press, New York 1981.
- LIMAYE, B. V.: Functional analysis, Wiley, New York 1981.
- LINDENSTRAUSS, J.; TZAFRIRI, L.: Classical Banach spaces; I: Sequence spaces, II: Function spaces, Springer-Verlag, Berlin – Heidelberg – New York 1977 bzw. 1979.

- MADDOX, I. J.: Elements of functional analysis, Cambridge Univ. Press, Cambridge 1973.
- MAURIN, J.: Methods of Hilbert spaces, PWN, Warszawa 1972.
- MAURIN, K.: Analysis, I: Elements, PWN, Warszawa 1976.
- MIKUSIŃSKI, J.: Operatorenrechnung, VEB Deutscher Verlag der Wissenschaften, Berlin 1957 (Übersetzung aus dem Polnischen).
- MOORE, R.: An introduction to numerical functional analysis, Wiley, New York 1985.
- NACHBIN, L.: Introduction to functional analysis: Banach spaces and differential calculus, M. Dekker, Basel 1981.
- NAYLOR, A. W.; SELL, G. R.: Linear operator theory in engineering and science, Springer-Verlag, Berlin – Heidelberg – New York 1982.
- VON NEUMANN, J.: Functional operators, I: Measures and integrals; II: The geometry of orthogonal spaces, Princeton Univ. Press, Princeton 1950.
- NEUMARK, M. A.: Normierte Algebren, VEB Deutscher Verlag der Wissenschaften, Berlin 1959 (Übersetzung aus dem Russischen).
- NEUMARK, M. A.: Lineare Differentialoperatoren, 2. Aufl., Akademie-Verlag, Berlin 1963 (Übersetzung aus dem Russischen).
- ПЛЕШЕР, А. И.: Спектральная теория линейных операторов, I, II, Успехи мат. наук 9 (1941), 3–125; 1: 1 (11) (1946), 71–191.
- ПЛЕШЕР, А. И.: Спектральная теория линейных операторов, Наука, Москва 1965.
- REISEL, R.: Elementary theory of metric spaces, Springer-Verlag, New York – Berlin 1982.
- ROLEWICZ, S.: Metric linear spaces, Reidel Publ. Comp., Dordrecht 1985.
- RUDIN, W.: Functional analysis, McGraw-Hill, London – New York 1973.
- RUDIN, W.: Principles of mathematical analysis, 3rd ed., McGraw-Hill, London – New York 1976.
- SCHECHESTER, M.: Principles of functional analysis, Academic Press, New York 1975.
- SCHMEIDLER, W.: Lineare Operatoren im Hilbertschen Raum, B. G. Teubner, Stuttgart 1954.
- SCHWARTZ, L.: Analyse, Part 2: Topologie générale et analyse fonctionnelle, Hermann et Cie., Paris 1970.
- SIDDIQI, A. H.: Functional analysis, with applications, McGraw-Hill, New York 1986.
- SOBOLEW, S. L.: Einige Anwendungen der Funktionalanalysis auf Gleichungen der mathematischen Physik, Akademie-Verlag, Berlin 1964 (Übersetzung aus dem Russischen).
- STONE, M. H.: Linear transformations in Hilbert space and their applications to analysis, reprint, American Math. Soc., New York 1948.
- SUTHERLAND, W. A.: Introduction to metric and topological space, Oxford Univ. Press, Oxford 1975.
- SZ. NAGY, B.: Spektraldarstellung linearer Transformationen des Hilbertschen Raumes, Springer-Verlag, Berlin 1942; bericht. Nachdruck 1967.
- TAYLOR, A.; LAY, D.: Introduction to functional analysis, 2nd ed., Wiley, New York 1980.
- TRIEBEL, H.: Höhere Analysis, VEB Deutscher Verlag der Wissenschaften, Berlin 1972 / 2. Aufl., Verlag Harri Deutsch, Thun – Frankfurt/Main 1980.
- TRIEBEL, H.: Theory of function spaces, Akademische Verlagsgesellschaft Geest & Portig, Leipzig 1983 / Birkhäuser, Basel – Boston 1983.
- WLOKA, J.: Funktionalanalysis und Anwendungen, W. de Gruyter, Berlin – New York 1971.
- WOUK, A.: A course of applied functional analysis, Wiley, New York 1979.
- ZAANEN, A. C.: Linear analysis, 3rd ed., P. Noordhoff, Groningen 1960.

Namen- und Sachverzeichnis

- Ableitung einer Mengenfunktion 189
- , verallgemeinerte 277, 281
- Abschließung eines Operators 452
- Abschnittselement 411
- Abschnittsfunktion 128, 485
- Abstand zwischen Elementen 144, 151
- zwischen Mengen 86
- Abzählbarkeit einer Menge 15
- ACHESER, N. I. 436, 525
- Äquivalenz von Funktionen 110
- von Normen 285
- , unitäre, von Systemen 529

- Bairesche Klassen 119
- Banachscher Raum 246
- BERNSTEIN, S. N. 49
- Besselsche Ungleichung 146, 321
- B -Funktion 119
- BOCHNER, S. 438

- Cantorsches Diskontinuum 199
- CARLEMAN, T. 436
- Cauchyfolge 145
- Cayleytransformierte 496
- COURANT, R. 166, 291

- Darbouxsche Summe 21
- Darstellung, kanonische, einer Funktion [38]
- Defektindizes 498, 516f.
- Delta-Funktion 308
- Differentialgleichungen; Lösung in Form von Funktionalen 313ff.
- Differentiallösungen 389ff., 425ff.
- Differentialoperatoren; Beispiele 459ff.
- Differenzkerne 446ff.
- Dilatation eines Intervalls 90
- Dimension einer linearen Mannigfaltigkeit 247
- — — modulo einer anderen 517
- Distribution *siehe* Funktion, verallgemeinerte
- Dreiecksungleichung 144, 224

- EGOROW, D. F. 116
- EIDELMANN, S. D. 312

- Eigenelemente eines Operators 335, 377f.
- Eigenschaft C 117
- Eigenwert eines Operators 274, 335, 377f., 483
- — —; Rang 274, 335
- Einbettungssätze 292ff., 304
- Einsoperator 332
- Elementarbereich 89
- Elemente, orthogonale 256, 319
- ; Projektion in einen Unterraum 324
- eines Raumes 46
- Erweiterung eines Funktionalen 252
- eines Operators 452, 510ff.
- — — nach FRIEDRICHS 507
- — — bezüglich der Stetigkeit 250
- eines abgeschlossenen symmetrischen Operators 495ff.
- symmetrischer halbbeschränkter Operatoren 504ff.
- , minimale abgeschlossene, eines Operators 452

- Faltungssatz 76
- Familie der B -Funktionen 120
- der Lebesgue-meßbaren Mengen 102
- fast überall 115
- Folge, Grenzwert (Limes) einer 183, 224
- , in sich konvergente 226
- von Mengen 84
- , reguläre, von Zerlegungen 56
- Fortsetzung einer Zerlegung 22, 121
- Fourierkoeffizienten, Fourierreihe 146, 320
- Fouriertransformationen 440ff.
- FRIEDRICHS, K. 291, 506
- , Erweiterung eines Operators nach 507
- Friedrichssche Ungleichung 306
- FUBINI, G. 168
- Fundamentalfolge 226
- Funktion, „abgeschnittene“ 128, 485
- , verallgemeinerte Ableitung einer 277, 281
- , absolut stetige, in zwei Veränderlichen 201

- Funktion, äquivalente 110
 —, charakteristische, einer Menge 117
 —; kanonische Darstellung 38
 —, integrierbare 56
 —, —, bezüglich einer Mengenfunktion 121
 —, —, von beschränkter Variation 39
 —, auf einer abgeschlossenen Menge stetige 117
 —, meßbare 109, 112f.
 —, —, bezüglich einer anderen Funktion 112
 —, im Mittel gleichgradig stetige 239
 —; Mittelwert 73
 —, (monoton) wachsende 20
 — eines selbstadjungierten Operators 395ff.
 —, schwach stetige, bezüglich eines Parameters 53
 —, positiv definite 72
 —, in einem Punkt stetige 117
 —, stetige 153
 —, —, eines selbstadjungierten Operators 480
 — von beschränkter Schwankung (Variation) 33
 —; Sprung 29
 —, stückweise konstante 61, 111, 117
 —, summierbare 128, 132
 —, — komplexe 135
 —; Teil 131
 — von beschränkter Variation in der Ebene 68
 —, verallgemeinerte 308
 — *siehe auch* Intervallfunktion, Mengenfunktion, Punktfunktion
 Funktional 47, 231
 —, beschränktes 47
 —, bilineares 329
 —, distributives 47
 —; Erweiterung 252
 —, halbstetiges 243
 —, lineares 47
 —, —, erster Art (zweiter Art) 325
 —, —, im Raum H 324f.
 —, —, im Raum vom Typ B 251ff.
 —, —, in C , L_p und l_p 262ff.
 — auf einer in sich kompakten Menge 243
 —; Norm 48
 —, quadratisches 330
 —, stetiges 47, 242
 — vom Typ einer Funktion 308
 Funktionenfolge, Grenzwert einer 113
 Funktionensystem, abgeschlossenes 148
 —, orthonormiertes 146
 —, vollständiges 148
 Gebiet, sternförmiges 283
 GELFAND, I. M. 312
 GLIWENKO, W. I. 53
 Graph eines Operators 455
 Grenze, obere (untere), einer Menge 21
 Grenzelement einer Menge 225
 Grenzspektrum 388
 Grenzwert, schwacher 259
 Häufungspunkt einer Menge 85
 Heine-Borel-Lebesguescher Überdeckungsatz *siehe* Satz von BOREL
 HELLINGER, E. 436
 Hellingersches Integral 215, 220ff.
 Hermitesche Funktionen 443
 — Polynome 525
 HILBERT, D. 166, 291, 317, 428, 436
 Hilbertscher Raum 144ff., 317
 — —; Axiome 245f., 317, 319
 Höldersche Ungleichung 156
 HÖRMANDER, L. 515
 Hülle, abgeschlossene 86, 225
 —, lineare 247
 —, — abgeschlossene 379
 Integral, Cauchy-Stieltjessches 78
 —, Fourier-Stieltjessches 71
 —, Hellingersches 215, 220ff.
 —, Lebesguesches 122
 —, —; Volladditivität 124
 —, Lebesgue-Stieltjessches 121
 — über eine Menge unendlichen Maßes 160
 —, Stieltjessches 17, 475ff.
 Integraldarstellungen, Sobolewsche 300ff.
 Integraloperatoren 492ff.
 — mit polaren Kernen 295
 — im Raum L_2 431
 Intervall; α -Dilatation 90
 —; α -Kontraktion 90
 Intervallfunktion 53
 —, additive 54
 — in der Ebene 62ff.
 —, nicht negative 54
 —, normale 54
 —, stetige 63
 —, Unstetigkeitslinie einer 63
 —, Unstetigkeitspunkt einer 63
 Isometrie von Räumen 226
 Jacobische Matrix 423, 522ff., 525
 Kern eines Spektrums 513
 Klasse L_2 139
 — $W_p^{(l)}(D)$ 286ff.
 Kompaktheit 236

Kompaktheit im Raum C 238

— — — l_p 241f.

— — — L_p 238ff.

—, schwache 258, 260

Komplementärmenge 84

Kontraktion eines Intervalls 90

Konvergenz fast überall 115

—, gleichmäßige 270

— in sich 145, 152

— dem Maß nach 115

— im Mittel 141

— — — in sich 142

— der Norm nach 255, 269, 341

— von Operatorfolgen 269ff.

— im Raum C 47

— einer unendlichen Reihe 319

—, schwache, von Elementen 259, 342

—, —, von Funktionalen 257ff.

—, —, von Operatoren 270

—, —, in C , L_p und l_p 268ff.

Konvergenz, schwache, in l_2 430

—, —, in L_2 450

—, starke 259, 270, 341

Körper B 107

LADYSHENSKAJA, O. A. 312, 512

Lagrangesche Formel 20

Lebesguesche Summe 122

— Zerlegung 122

—s Integral 122, 124

Lebesgue-Stieltjessches Integral 121

Limes inferior (superior) 113

Ljapunoffsche Fläche 235

LUSIN, N. N. 117

Mächtigkeit des Kontinuums 17

— von Mengen 15

Mannigfaltigkeit, lineare 247

—, —, der finiten stetigen Funktionen 165

—, —, in L_2 152

Maß einer Menge 94

—, äußeres, einer Menge 92

—; Unabhängigkeit von der Wahl des Koordinatensystems 106

Matrix, transformierte 528

Matrizen, Jacobische 423ff., 522ff.

— und Operatoren 526ff.

—, selbstadjungierte 417ff.

—, unendliche 520f.

—, unitäre 416f.

C -Matrizen 528

Menge 15

—, abgeschlossene 85, 225

—, abzählbare 15

—, Borel-meßbare 107

—; Dimension 247

—, diskrete 86

—; charakteristische Funktion 117

—; obere (untere) Grenze 21

—; Grenzelement 225

—; Häufungspunkt 85

—; abgeschlossene Hülle 86, 225

—; lineare Hülle 247

—; — abgeschlossene Hülle 379

—, in sich kompakte 236

—, kompakte 236

—; Komplement 84

—, leere 83, 85

—; Maß 94

—; äußeres Maß 92

— vom Maß Null 94

—, meßbare 94

—; endliches ε -Netz 236

—, offene 85

—; Rand 85

—, schwach kompakte 258

G_δ -Menge 108

Mengen; Abstand 86

—; Differenz 83

— gleicher Mächtigkeit 15

—; Produkt 82

—; Summe 15, 82

—; Vereinigung 15, 82

Mengenfamilie 101

Mengenfolge 84

Mengenfunktion; Ableitung 189

—, absolut stetige 189, 194ff., 209

—, additive 185ff.

—; Hauptsatz 188, 211ff.

—, stückweise konstante 205

—; absolut stetiger Teil 189

—; singulärer Teil 189

—; Unstetigkeitsstelle 188

—, volladditive 105

—, wesentlich beschränkte 205

Mengenkörper 104

Meßbarkeitskriterien 102ff.

MICHLIN, S. G. 506

Minkowskische Ungleichung 159

Mittelfunktion 179

Mittelungsradius 180

Mittelwert einer Funktion 73

ε -Netz, endliches, einer Menge 236

NEUMANN, J. v. 400, 530

Norm eines Elements 144, 150

— einer Funktion 47

— eines Funktionalen 48, 324

—, absolute, eines Operators 359

Normen, äquivalente 285

Nullelement 46, 144, 245

Nullfunktional 324

Operator 230

- , abgeschlossener 452
- ; Abschließung 452
- , adjungierter 271, 327, 453ff., 501f.
- , –, im Raum L_2 432ff.
- , beschränkter 249, 412ff.
- , – inverser 332
- ; Definitionsbereich 230
- mit Differenzkern 494
- , distributiver 248
- ; Eigenelemente 335
- ; Eigenwerte 274, 335, 483
- , elementarer symmetrischer 503
- , endlichdimensionaler 345
- ; Erweiterung 452, 495ff., 504ff.
- ; – bezüglich der Stetigkeit 250
- ; –, minimale abgeschlossene 452
- ; Graph 455
- , halbbeschränkter 457, 508ff.
- , hypermaximaler 500
- , identischer 332
- , inverser 250, 333
- , irreduzibler 385
- , isometrischer 358, 496
- , linearer 52, 250
- , –, im Raum H 326ff.
- , –, im Raum l_2 411ff.
- , linksinverser 333
- , maximaler 500, 503ff.
- der Multiplikation mit der unabhängigen Veränderlichen 392ff., 444ff., 489
- ; absolute Norm 359
- , normal lösbarer abgeschlossener 454
- , normaler 402
- , positiv definit 457
- ; Potenzreihe 407
- ; Punkt regulären Typs 512
- ; regulärer Punkt 336, 338, 512
- im Raum C 52
- in einem normierten Raum 248
- , rechtsinverser 333
- , reduzierter 384, 473
- ; Resolvente 336, 341, 375, 481ff.
- ; untere Schranke des Spektrums 480
- , selbstadjungierter 328, 457
- , –, Eigenelemente 377f.
- , –, Eigenwerte 377f.
- , –, Grenzen 330
- , –, positiver 331
- , –, Punktspektrum 471
- , –, reines Punktspektrum 378, 471
- , –, Spektralschar 371ff.

Operator, selbstadjungierter; Spektrum 468ff.

- , –, einfaches kontinuierliches Spektrum 384
- , –, rein kontinuierliches Spektrum 379
- , –, Störung des Spektrums 400ff.
- , –, vollstetiger 352ff.
- ; Spektralschar 408ff., 480
- ; Spektrum 336
- ; gemischtes Spektrum 483
- , stetiger 52, 248
- ; kleine Störungen des Spektrums 487ff.
- , symmetrischer 457
- , –, Spektrum 512ff.
- , unitärer 356
- , –, Spektralzerlegung 395f.
- , vollstetiger 271f., 344, 348
- , –, im Raum C 274
- , –, im Raum l_2 430f.
- , –, im Raum L_2 434ff.
- , –, im Raum l_p 274
- , –, im Raum L_p 274
- ; Wertebereich 230
- Operatoren, gleiche 452
- und Matrizen 526ff.
- , Produkt von 327
- , unitär äquivalente 358, 394
- über Unterräumen 361
- , vertauschbare 327, 399, 472, 486
- Operatorfolgen 341f.
- Orthonormalsystem 146, 320
- , abgeschlossenes 148, 321
- , vollständiges 148, 155, 321

Parsevalsche Gleichung 147, 321

PETROWSKI, I. G. 312

Poincarésche Ungleichung 306

Potenzreihe eines Operators 407

Produkt, inneres 144, 150, 256, 317

– von Mengen 82

– von Operatoren 327

–, skalares 144, 151, 256, 317

– von Zerlegungen 22, 147

Prinzip der kontrahierenden Abbildungen

Projektion 324 [231]

Projektionsmatrizen 416f.

Projektionsoperator (Projektor) 326, 362ff.

Punkt, regulärer 336, 338, 512

– regulären Typs 512

Punktfunktion 65ff.

–, absolut stetige 191

Punktmenge 85

Punktspektrum 388

– eines selbstadjungierten Operators 471

Punktspektrum, reines 378
 —, —, eines selbstadjungierten Operators 471

RADON, J. 71

Rand einer Menge 85

Rang eines Eigenwertes 274, 335

Raum, bikonjugierter 255

—, dichter 228

— der stetigen Funktionen 46

—, gleichmäßig konvexer 261

—, konjugierter 255

—, linearer*normierter 245

—, metrischer 224

— der linearen Operatoren 269ff.

—, reduzierender 473

—, regulärer 256

—, separabler metrischer 243

— vom Typ *B* 246

—, vollständiger metrischer 225

— der reellen rationalen Zahlen 229

Raum *B* 246

— *C* 46, 199, 229

— $\dot{C}^{(l)}(D)$ 308

— *F* 229

— *H* 317, 346

— $\tilde{H}$ 455, 459

— *K* 315

— l_2 150, 317, 346, 411ff.

— l_p 232

— L_2 ; Separabilität 150

— $L_{2,m}$ 450

— $L_2^{(x)}$ 392

— L_p 156

— $L_p(\xi)$ 233

— *m* 232, 244

— *M* 232

— R_n 231, 248

— *s* 232, 244

— *S* 233f.

— *V* [*a*, *b*] 233

— $W_p^{(l)}$ 284

— $\tilde{W}_p^{(l)}$ 284

— $W_2^{(l)}(D)$ 451

— $W_2^{(l)}(D)$ 451

Räume isometrische 226

—, —, vom Typ *B* 247

Resolvente 336, 341, 375, 481f.

Riemann-Stieltjessche Summe 18

RIESZ, F. 71, 254, 348, 398, 438

Satz von BOREL 88

— von COURANT 354

— von FUBINI 168

Satz von FUBINI für charakteristische Funktionen 171

— — — für Integrale beliebiger Vielfachheit 175

— — — im Fall des Lebesgue-Stieltjesschen Integrals 176

— — — im Fall nichtbeschränkter Mengen 174

— von HELLY 43ff.

— von F. RIESZ 49

— von RIESZ-FISCHER 147

SCHILOW, G. E. 312

SCHWARTZ, L. 312

Schwarzsche Ungleichung 318

Separabilität eines Raumes 150

— eines metrischen Raumes 243

SOBOLEW, W. I. 261, 304, 313

Sobolewsche Integraldarstellungen 300ff.

Spektralschar 371ff., 408ff., 435ff., 480, 530ff.

Spektralzerlegung unitärer Operatoren 395f.

Spektrum 336, 378, 468ff., 512ff.

—, einfaches kontinuierliches 384

— einer selbstadjungierten Erweiterung
 —, gemischtes 388, 483 [518f.]

—; Kern 513

—, kontinuierliches 419ff.

—; untere Schranke 480

—; Störung 400ff.

—; kleine Störungen 487ff.

Sphäre 225

Sprung einer Funktion 29

Sprungfunktion 29, 188

Stetigkeit im Mittel 177f.

— — —, gleichgradige 239

STIELTJES, T. 18, 81

Stieltjessche Formel 78, 81

—s Integral 18, 475ff.

—s —, absolut konvergentes 27

—s —, allgemeines 55ff., 60

—s — in der Ebene 67ff.

—s —; Existenzsatz 24

—s — einer stetigen Funktion 24ff.

—s —; physikalische Interpretation 197

—s —, uneigentliches 27

STONE, M. H. 198, 436

SUCHOMLINOW, G. A. 254

Summe, Darbouxsche 21

—, Lebesguesche 122

— von Mengen 15, 82

—, orthogonale, von Unterräumen 361

—, Riemann-Stieltjessche 18

Symmetrieaxiom 224

Sz.-NAGY, B. 71, 254, 438

- Teil, absolut stetiger, einer Mengenfunktion 189
 —, singulärer, einer Mengenfunktion 188
 — eines Unterraumes 362
 Teilmenge 82
 Totalvariation 33, 35, 187
 Transformation, unitäre, im Raum L_2 438ff.
 TYCHONOFF, A. N. 312
- Umgebung eines Punktes 85
 Umkehrformel für das Fouriersche Integral 74
 Unabhängigkeit, lineare 246
 Untermenge 82
 Unterraum 247
 — der zu einem Eigenwert gehörigen Eigenelemente 335
 —, invarianter 384, 473
 —, reduzierender 384
 Unterräume, komplementäre 324
 —, orthogonale 322
 —; orthogonale Summe 361
- Variation 33
 Vektor eines Raumes 46
 Verdichtungsunkt 389, 400
 Vervollständigung eines metrischen Raumes 226ff.
 Vielfachheit eines Eigenwerts 274, 335
 Vollständigkeit eines Raumes 145
 Vollständigkeitsrelation 147, 320
 —, verallgemeinerte 149, 321
- WATSON, G. N. 440
 WILENKIN, N. J. 312
 WINTNER, A. 530
 WISCHIK, M. I. 515
- Zerlegung 22
 — der Einheit 366, 475
 —, Fortsetzung einer 22, 121
 —, Lebesguesche 122
 —en, reguläre Folge von 56
 —en, Produkt von 22, 121