III. KAPITEL.

Orthogonalreihen in L^2.

Das erste Problem der Theorie der Orthogonalreihen ist die Frage nach der Existenz der Orthogonalsysteme. Die Beispiele des II-ten Kapitels gewähren noch keinen Einblick in die „Freiheitsgrade", über welche man beim Bauen von Orthogonalsystemen verfügt, obwohl ein jedes Beispiel durch bloße Änderung der Reihenfolge oder durch Vorzeichenänderung bereits zu unendlich vielen anderen führt. Wir betonen, daß alle Ausführungen dieses Kapitels den Raum L^2 zum Gegenstand haben, wenn nicht ausdrücklich anders vorausgesetzt wird. Andere Räume kommen im Kapitel VI zur Sprache.

§ 1. Orthogonalisierung.

Es seien $f_1(t), f_2(t), \ldots, f_n(t)$ n in $\langle a, b \rangle$ erklärte Funktionen des Raumes L^2. Es hat bereits im J. 1879 J. P. Gram in einer dänisch geschriebenen Arbeit die Frage nach einer Matrix

$$C = \| c_{ik} \| \qquad (i, k = 1, 2, \ldots, n)$$

aufgeworfen, die vermittels der linearen Substitution

$$(1) \qquad \begin{aligned} \varphi_1(t) &= c_{11} f_1(t) + \ldots + c_{1n} f_n(t) \\ \varphi_2(t) &= c_{21} f_1(t) + \ldots + c_{2n} f_n(t) \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ \varphi_n(t) &= c_{n1} f_1(t) + \ldots + c_{nn} f_n(t) \end{aligned}$$

zu einem normierten Orthogonalsystem $\{\varphi_i(t)\}$ führt.

Wir bemerken hier gleich, daß die lineare Unabhängigkeit der Funktionen $\{f_i\}$ eine notwendige Bedingung für die Existenz von C ist. Wären nämlich die Funktionen f_i linear abhängig, d. h. hätte man in $\langle a, b\rangle$ fast überall $\sum_{i=1}^{n} d_i f_i(t) = 0$ mit nicht sämtlich verschwindenden Konstanten d_i, so hätte man — nach einem bekannten Satze der Theorie der linearen Gleichungen — $\sum_{i=1}^{n} d_i' \varphi_i(t) = 0$ für dieselben t-Werte, d. h. fast überall in $\langle a, b\rangle$, und nicht alle d_i' $(i = 1, 2, \ldots, n)$ wären Null. Dies aber steht im Widerspruch zum folgenden Satz:

[311] *Funktionen eines normierten Orthogonalsystems können keine endliche lineare Abhängigkeit eingehen.*

In der Tat liefert eine Beziehung, wie

$$\sum_{i=1}^{n} d_i' \varphi_i(t) = 0 \qquad \text{f. ü.},$$

nach Multiplikation mit $\varphi_k(t)$ und nach Integration über $\langle a, b\rangle$

$$\sum_{i=1}^{n} d_i' \int_a^b \varphi_i(t)\, \varphi_k(t)\, dt = 0, \qquad \text{d. h. } d_k' = 0 \qquad (k = 1, 2, \ldots, n).$$

Gram hat folgendes Kriterium der linearen Unabhängigkeit angegeben:

[312] *Damit die Funktionen $f_1(t), f_2(t), \ldots, f_n(t)$ — im obigen Sinne — linear unabhängig seien, ist die Bedingung $A_n \neq 0$ notwendig und hinreichend, wobei A_n die Determinante der Matrix $\| a_{ik} \|$ bedeutet, deren Elemente als*

$$(2) \qquad a_{ik} = \int_a^b f_i(t)\, f_k(t)\, dt \qquad (i, k = 1, \ldots, n)$$

definiert sind.

Der Beweis stützt sich auf die augenscheinliche Tatsache, daß A_n die Diskriminante der quadratischen Form

$$(3) \qquad F(x_1, x_2, \ldots, x_n) = \int_a^b [x_1 f_1(t) + x_2 f_2(t) + \ldots + x_n f_n(t)]^2\, dt$$

der n Veränderlichen $x_1, x_2, \ldots, x_n$ ist. Diese Form ist als Integral eines Quadrats stets nichtnegativ und verschwindet für ein anderes x System als $x_1 = x_2 = \ldots = x_n = 0$, wenn dieses andere System die Summe $\sum_{i=1}^{n} x_i f_i(t)$ fast überall in $\langle a, b\rangle$ zu Null macht.

$F(x_1, x_2, \ldots, x_n)$ verschwindet also nichttrivialerweise, wenn die $f_i(t)$ linear abhängig sind. Der Gram'sche Satz wird also bewiesen werden, wenn wir zeigen können, daß $A_n = 0$ für nichttriviales Verschwinden der Form F charakteristisch ist. Nun aber impliziert $A_n = 0$ ein nichttriviales Lösungssystem für die Gleichungen

$$\sum_{k=1}^{n} a_{ik}\, x_k = 0 \qquad (i = 1, 2, \ldots, n).$$

Wenn man die i te Gleichung mit x_i multipliziert und alles addiert, so kommt

$$F(x_1, x_2, \ldots, x_n) = \sum_{i=1}^{n} x_i \sum_{k=1}^{n} a_{ik}\, x_k = 0.$$

Umgekehrt, ist $F(x_1, x_2, \ldots, x_n) = 0$ und $\{x_i\}$ nichttrivial, so ist wegen (3) $\sum_{k=1}^{n} x_k f_k(t) = 0$ f. ü.; wenn man mit $f_i(t)$ multipliziert, integriert und (2) einführt, so wird

$$a_{i1} x_1 + a_{i2} x_2 + \ldots + a_{in} x_n = 0 \qquad (i = 1, 2, \ldots, n),$$

woraus bereits $A_n = 0$ folgt.

Die Determinante A_n ist stets nichtnegativ. Es sei nämlich $A_j = \| a_{ik} \|$ $(i, k = 1, 2, \ldots, j)$, $D_j^{\alpha, \beta}$ die zu $a_{\alpha\beta}$ gehörige Unterdeterminante von A_j. Dann ist

$$B_j = \int_a^b \left[\sum_{\alpha=1}^{j} D_j^{j, \alpha} f_\alpha(t) \right]^2 dt = \sum_{\alpha=1}^{j} D_j^{j, \alpha} \sum_{\beta=1}^{j} D_j^{j, \beta}\, a_{\alpha\beta} = D_j^{j, j} A_j,$$

also $B_j = A_{j-1} . A_j$; es ist $a_{11} \geqslant 0$ und $B_j \geqslant 0$, also $A_2 \geqslant 0$, $A_3 \geqslant 0$ u. s. w. (Null ist A_j nur bei linearer Abhängigkeit; dann ist

$$\sum_{\alpha=1}^{j} D_j^{j, \alpha} f_\alpha(t) = 0 \quad \text{f. ü.}).$$

Erhard Schmidt hat im J. 1905 folgende *Methode zur Berechnung der Koeffizienten* c_{ik} *in* (1) angegeben. Er setzt von vornherein $c_{ik} = 0$ für $k > i$ und verlangt, daß die Ausdrücke [313]

$$(4) \qquad \begin{aligned} \varphi_1(t) &= c_{11} f_1(t), \\ \varphi_2(t) &= c_{21} f_1(t) + c_{22} f_2(t), \\ &\ldots\ldots\ldots\ldots \\ \varphi_n(t) &= c_{n1} f_1(t) + c_{n2} f_2(t) + \ldots + c_{nn} f_n(t) \end{aligned}$$

orthogonal und normiert seien.

Man sieht, daß alle c_{ii} von Null verschieden sein müssen, denn wäre $c_{ii}=0$, so wären $\varphi_1(t), \dots, \varphi_i(t)$ durch $f_1(t), \dots, f_{i-1}(t)$ linear ausgedrückt, also linear abhängig, gegen [311]. $c_{ii} \neq 0$ erlaubt (4) nach f aufzulösen und es kommt sukzessive

$$
\begin{aligned}
f_1(t) &= d_{11}\,\varphi_1(t),\\
f_2(t) &= d_{21}\,\varphi_1(t) + d_{22}\,\varphi_2(t),\\
&\dots\dots\dots\dots\dots\dots\\
f_n(t) &= d_{n1}\,\varphi_1(t) + d_{n2}\,\varphi_2(t) + \dots + d_{nn}\,\varphi_n(t).
\end{aligned} \tag{5}
$$

Führt man in die Formel (4) für $\varphi_i(t)$ $(i>1)$ die Werte (5) für $f_1, f_2, \dots, f_{i-1}$ ein, so erhält man

$$\varphi_i(t) = \lambda_{i1}\,\varphi_1(t) + \lambda_{i2}\,\varphi_2(t) + \dots + \lambda_{i,i-1}\,\varphi_{i-1}(t) + c_{ii} f_i(t),$$

und die Orthogonalität von $\varphi_k(t)$ und $\varphi_i(t)$ $(k<i)$ liefert

$$\lambda_{ik} + c_{ii}\int_a^b \varphi_k(t) f_i(t)\,dt = 0 \qquad (k=1,2,\dots,\ i-1); \tag{6}$$

c_{ii} wird aus der Bedingung $\int_a^b \varphi_i^2(t)\,dt = 1$ bestimmt, nachdem die λ_{ik} den Gleichungen (6) gemäß berechnet wurden.

Wir können aber $\{\varphi\}$ auch direkt berechnen. (5) verlangt, daß $\varphi_i(t)$ zu $f_k(t)$ orthogonal sei, sobald $k<i$. Legen wir also der Funktion $\varphi_i(t)$ die Bedingung der Orthogonalität zu $f_1(t), f_2(t), \dots, f_{i-1}(t)$ auf; dann resultiert bereits aus (4) die Orthogonalität der $\{\varphi\}$ zueinander. Es folgt für $k=1,2,\dots,i-1$

$$0 = \int_a^b f_k(t)\,\varphi_i(t)\,dt = \sum_{j=1}^{i} c_{ij}\int_a^b f_k(t)\,f_j(t)\,dt = \sum_{j=1}^{i} c_{ij}\,a_{kj}.$$

Es soll also $\{c_{ij}\}$ $(j=1,2,\dots,i)$ ein Lösungssystem $\{x_j\}$ des Systems der $i-1$ Gleichungen

$$\sum_{j=1}^{i} a_{kj}\,x_j = 0 \qquad (k=1,2,\dots,\ i-1) \tag{7}$$

mit i Unbekannten sein. Die lineare Unabhängigkeit von $f_1(t)$, $f_2(t), \dots, f_n(t)$, also von $f_1(t), f_2(t), \dots, f_{i-1}(t)$ $(i>1)$, sichert für (7) — nach dem Gram'schen Kriterium [312] — den Rang $i-1$, es ist also

$$x_j = \lambda\mu_j \qquad (j=1,2,\dots,i)$$

mit beliebigem λ, wo μ_j den Faktor von ξ_j in der Entwicklung der Determinante

$$\begin{vmatrix} \xi_1 & \xi_2 & \dots & \xi_i \\ a_{11} & a_{12} & \dots & a_{1i} \\ \dots & \dots & \dots & \dots \\ a_{i-1,1} & a_{i-1,2} & \dots & a_{i-1,i} \end{vmatrix} \tag{8}$$

nach der ersten Zeile bezeichnet. Schreiben wir dasselbe für alle i $(i=1,2,\dots,n)$, so müssen wir obere Indizes i einführen und erhalten

$$
\begin{aligned}
\varphi_1(t) &= \lambda^{(1)}\mu_1^{(1)}(t),\\
\varphi_2(t) &= \lambda^{(2)}\mu_1^{(2)} f_1(t) + \lambda^{(2)}\mu_2^{(2)} f_2(t),\\
&\dots\dots\dots\dots\dots\dots\\
\varphi_n(t) &= \lambda^{(n)}\mu_1^{(n)} f_1(t) + \lambda^{(n)}\mu_2^{(n)} f_2(t) + \dots + \lambda^{(n)}\mu_n^{(n)} f_n(t).
\end{aligned} \tag{9}
$$

Es ist dabei $\mu_i^{(i)} \neq 0$ $(i=1,2,\dots,n)$ und, die Orthogonalität einmal gesichert, können wir durch die Wahl von $\lambda^{(i)}$ $(i=1,\dots,n)$ die Normierung erreichen; es genügt (und es ist notwendig) $\lambda^{(i)}$ aus der Gleichung

$$(\lambda^{(i)})^2 \int_a^b \left(\sum_{k=1}^{i} \mu_k^{(i)} f_k(t)\right)^2 dt = 1$$

zu bestimmen. Das Quadratintegral ist positiv (sonst wären die i ersten Funktionen f abhängig wegen $\mu_i^{(i)} \neq 0$), es läßt sich also $\lambda^{(i)}$ auf zwei Arten berechnen. Es hat also das Schmidt'sche Verfahren genau 2^n Lösungen, die sich durch die Vorzeichenwahl bei den Funktionen $\{\pm\varphi_i\}$ voneinander unterscheiden.

Der Schmidt'sche Prozess hat den wichtigen Vorteil, daß die Formeln (9) ihre Gültigkeit behalten, wenn man zum System $f_1, \dots, f_n$ eine neue Funktion $f_{n+1}(t)$ hinzunimmt. Die Formeln (9) bleiben unverändert und ist $f_{n+1}(t)$ von $f_1, \dots, f_n$ linear unabhängig, so ist der Prozess um einen Schritt weiter zu führen und liefert eine Funktion $\varphi_{n+1}(t)$, die zu den bereits erhaltenen φ orthogonal, selbst aber normiert ist. Daher:

Hat man eine unendliche Folge $\{f_i\}$ von linear unabhängigen [314]
Funktionen in L^2, so läßt sich ein normiertes Orthogonalsystem $\{\varphi_i\}$ bilden, welches mit $\{f_i\}$ durch unendlich viele Formeln (4) *und* (5) *verknüpft ist; die Diagonalkoeffizienten c_{ii}, d_{ii} sind sämtlich von*

Null verschieden und alle c_{ik} *und* d_{ik} *sind eindeutig bestimmt, wenn man verlangt, daß die* c_{ii} *positiv seien.*

[315] *Ist nun* K *irgend ein Raum, so ist die Vollständigkeit des Systems* $\{f_i\}$ *inbezug auf* K *mit der Vollständigkeit von* $\{\varphi_i\}$ *inbezug auf* K *äquivalent.* Denn ist $\|g\|_K = 0$ für ein $g(t) \in K$, sobald alle Integrale $\int_a^b g(t) f_k(t)\, dt$ $(k = 1, 2, \ldots)$ verschwinden, so muß auch $\|g\|_K$ Null sein, wenn $g(t)$ zu allen $\varphi_k(t)$ orthogonal ist, da ja aus

$$\int_a^b g(t)\, \varphi_k(t)\, dt = 0 \qquad (k = 1, 2, \ldots),$$

vermöge (5), $\int_a^b g(t) f_k(t)\, dt = 0$ $(k = 1, 2, \ldots)$ folgt. Der umgekehrte Schluß stützt sich auf (4).

[316] *Auch ist die Abgeschlossenheit von* $\{f_i\}$ *inbezug auf* K *mit derselben Eigenschaft von* $\{\varphi_i\}$ *äquivalent.* Ist nämlich $g(t) \in K$ und gibt es zu $\varepsilon > 0$ Zahlen $a_1, \ldots, a_n$ mit

$$\text{(10)} \qquad \Big\| g(t) - \sum_{i=1}^{n} a_i f_i(t) \Big\|_K < \varepsilon,$$

so führt die Substitution (5) auf

$$\Big\| g(t) - \sum_{i=1}^{n} b_i \varphi_i(t) \Big\|_K < \varepsilon,$$

d. h. auf die Abgeschlossenheit von $\{\varphi_i\}$. Den umgekehrten Schluß liefert (4).

[317] Ein Beispiel für die Orthogonalisierung. Es sei

$$\langle a, b \rangle = \langle -1, +1 \rangle \quad \text{und} \quad f_i(t) = t^{i-1} \qquad (i = 1, 2, 3, \ldots).$$

Es kommt

$$a_{ik} = \int_{-1}^{+1} t^{i+k-2}\, dt = \begin{cases} \dfrac{2}{i+k-1}, & \text{wenn } i+k \text{ gerade} \\ 0, & \text{wenn } i+k \text{ ungerade.} \end{cases}$$

Es ist $\varphi_1(t) = 1/\sqrt{2}$ und $\mu_1^{(2)} = 0$, $\mu_2^{(2)} = 2$, also $\varphi_2(t) = \lambda^{(2)}.\, 2t$; die Bedingung

$$4\, (\lambda^{(2)})^2 \int_{-1}^{+1} t^2\, dt = 1$$

gibt $\varphi_2(t) = \sqrt{3/2}\; t$. Das in dieser Weise erhaltene ON heißt das *System der Legendre'schen Polynome.*

Allgemeine Orthogonalisierung. Es sei $\varphi_1(t), \ldots, \varphi_n(t)$ ein ON, welches aus $f_1(t), \ldots, f_n(t)$ vermittels des Schmidt'schen Verfahrens gewonnen wurde. Schreiben wir

$$\text{(11)} \qquad \begin{aligned} \psi_1(t) &= e_{11} \varphi_1(t) + e_{12} \varphi_2(t) + \ldots + e_{1n} \varphi_n(t), \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ \psi_n(t) &= e_{n1} \varphi_1(t) + e_{n2} \varphi_2(t) + \ldots + e_{nn} \varphi_n(t), \end{aligned}$$

so kommen wir auf ein System $\{\psi\}$, welches nach Einführung der Werte (9) für $\{\varphi\}$ auch durch $\{f\}$ ausgedrückt erscheint.

Wir werden zeigen, daß wir auf diese Weise die Frage, mit welcher dieser § beginnt, dann und nur dann lösen können, *wenn* [318]
die Matrix $\|e_{ik}\| = E$ *zeilenweise orthogonal und normiert ist; in diesem und nur in diesem Falle wird* $\{\psi\}$ *ein* ON *sein.*

Definition. Eine Matrix $\|e_{ik}\|$ heißt *zeilenweise orthogonal,* wenn $\sum_{k=1}^{n} e_{ik} e_{jk} = 0$ für $i \neq j$; sie heißt *zeilenweise normiert,* wenn $\sum_{k=1}^{n} e_{ik}^2 = 1$ $(i = 1, 2, \ldots, n)$. Analoges gilt für Spalten.

Beweis: Ist $\{\psi\}$ ein ON, so gibt (11) unmittelbar — da $\{\varphi\}$ ein ON ist —

$$\text{(12)} \qquad \sum_{k=1}^{n} e_{ik} e_{jk} = \int_a^b \psi_i(t)\, \psi_j(t)\, dt = \begin{cases} 0 & \text{für } i \neq j \\ 1 & \text{für } i = j, \end{cases}$$

w. z. b. w.

Umgekehrt: ist $\|e_{ik}\|$ zeilenweise orthogonal und normiert, so zeigen die Formeln (12), wenn man die Integrale voranstellt, daß $\{\psi\}$ ein ON ist. Wir behaupten, daß wir durch Verbindung der Substitution (11), wo $\|e_{ik}\|$ eine beliebige zeilenweise orthogonale und normierte Matrix ist, mit den Schmidt'schen Formeln (9) alle Systeme $\{\psi\}$ gewinnen, die das Problem der Orthogonalisierung des Systems $\{f\}$ lösen. Tatsächlich, ist $\{\psi\}$ ein solches System und

$$\psi_i(t) = \sum_{k=1}^{n} h_{ik} f_k(t) \qquad (i = 1, \ldots, n),$$

so kommt wegen (5)

$$\psi_i(t) = \sum_{k=1}^{n} e_{ik} \varphi_k(t) \qquad (i = 1, \ldots, n),$$

wo $\{\varphi\}$ die Schmidt'schen Funktionen sind, und der Schluß (12) zeigt, daß $\|e_{ik}\|$ zeilenweise ON ist.

Die Matrizes $\|e_{ik}\| = E$ sind aus der euklidischen n-dimensionalen Koordinatengeometrie bekannt: ist $x_1, x_2, \dots, x_n$ ein rechtwinkliges Koordinatensystem und $y_1, y_2, \dots, y_n$ ein ebensolches (mit demselben Ursprung), so hat man

$$(13) \qquad y_i = \sum_{k=1}^{n} e_{ik} x_k \qquad (i = 1, \dots, n),$$

und die Koeffizientenmatrix $E = \|e_{ik}\|$ ist zeilenweise normiert-orthogonal. Es ist leicht zu zeigen, daß dann

$$(14) \qquad x_i = \sum_{k=1}^{n} e_{ki} y_k \qquad (i = 1, \dots, n),$$

(was aus der Bedeutung von e_{ik} als Cosinus des Winkels zwischen x_i und y_k ersichtlich ist). Es ist also die zu E inverse Substitution E^{-1} durch bloße Vertauschung der Zeilen mit den Spalten zu gewinnen. Da die Formel (14) dieselbe geometrische Bedeutung hat wie (13), muß E^{-1} als Substitutionsmatrix von (14) ebenfalls zeilenweise orthogonal und normiert sein. Infolgedessen ist E spaltenweise orthogonal und normiert.

Bekanntlich (Kronecker, Hermite, Cayley) lassen sich sämtliche Matrizes E erhalten, indem man für e_{ik} n^2 geeignete rationale Funktionen von $\frac{n(n-1)}{2}$ Parametern einführt und sodann diesen Parametern alle möglichen Werte erteilt.

Der Determinantenwert von E ist ± 1; setzt man nämlich (14) in (13) ein, so erhält man die identische Substitution. Es ist also

$$E \, . \, E^{-1} = L,$$

wo L die Einheitsmatrix bedeutet. Für die Determinantenwerte folgt daraus $|E| . |E^{-1}| = 1$; da E^{-1} durch Vertauschung von Zeilen und Spalten aus E gewonnen wurde, ist $|E| = |E^{-1}|$, also $|E| = \pm 1$, w. z. b. w.

§ 2. **Orthogonalisierung im längeren Intervall.**

Das Problem der Orthogonalisierung eines Systems von linear unabhängigen, in $<a, b>$ erklärten Funktionen $\{f_n\}$ kann man anders als im § 1 auffassen. Man kann nämlich fragen, ob man nicht etwa durch Erweiterung der $\{f_n\}$ auf ein über $<a, b>$

hinausragendes Intervall die Orthogonalität erreichen kann. Dieselbe Frage hat einen Sinn auch in dem Falle, wo $\{f_n\}$ bereits in $<a, b>$ orthogonal ist; die triviale Erweiterung „$f_n(t) = 0$ außerhalb $<a, b>$" hebt die Vollständigkeit auf; eine nichttriviale Erweiterung auf $<a, c>$ mit $a < b < c$ ist stets möglich, und zwar nur so, daß $\{f_n\}$ in $<b, c>$ orthogonal sei, wie aus der Formel

$$0 = \int_a^c f_i(t)\, f_k(t)\, dt = \int_a^b + \int_b^c = \int_b^c f_i(t)\, f_k(t)\, dt$$

unmittelbar folgt; augenscheinlich darf $\{f_n\}$ in $<b, c>$ ein beliebiges O sein.

Ist $\{f_n(t)\}$ irgendein System von in $<a, b>$ erklärten Funktionen aus L^2, so lassen sich die $f_n(t)$ so in der Verlängerung $<b, c>$ von $<a, b>$ definieren, daß das System in $<a, c>$ orthogonal wird. [321]

Wir erinnern an die Matrix (T) in [212] (3). Jede Zeile von (T) wird zur Definition einer Funktion $f_n(t)$ benutzt. Man teile $<b, c>$ in Teilintervalle $\{I_k\}$ vermittels einer Punktfolge, die wachsend gegen c konvergiert; ist ε_k das k-te Element der n ten Zeile von (T), so setze man

$$f_n(t) = 0 \quad \text{für} \quad t \in I_k, \qquad \text{wenn} \quad \varepsilon_k = 0 \qquad (n \geqslant 1),$$

$$f_1(t) = 1 \quad \text{für} \quad t \in I_k, \qquad \text{wenn} \quad \varepsilon_k = 1\,,$$

$$f_n(t) = -\int_a^b f_n(u)\, f_p(u)\, du \cdot \frac{1}{|I_k|} \quad \text{für} \quad t \in I_k \qquad (n > 1),$$

wenn $\varepsilon_k = 1$ und ε_k die p-te Einheit in der Zeile ist, wo $p < n$; ist ε_k eine der weiteren Einheiten und $t \in I_k$, so soll $f_n(t) = 1$ gelten.

Zunächst zeigen wir, daß $\int_a^c f_n^2(t)\, dt$ endlich ist. Es handelt sich bloß um $\int_b^c f_n^2(t)\, dt$. Aber in der n-ten Zeile von (T) gibt es nur endlichviele Einheiten mit $p < n$; das Integral über die entsprechenden I_k ist selbstverständlich endlich; in sonstigen I_k ist $f_n^2 \leqslant 1$, also ist $\int_a^b f_n^2(t)\, dt < \infty$.

B e w e i s der Orthogonalität: Für $n > 1$ und $p < n$ ist

$$\int_a^c f_n(t)\, f_p(t)\, dt = \int_a^b f_n(t)\, f_p(t)\, dt + \int_{I_{k(p)}} f_n(t)\, f_p(t)\, dt + \int_D f_n(t)\, f_p(t)\, dt\,,$$

wo $k(p)$ der Spaltenindex der p-ten Einheit der n-ten Zeile ist, D aber den nach Abzug von $I_{k(p)}$ bleibenden Rest von $<b, c>$ bezeichnet. Die zwei ersten Integrale rechts heben sich auf vermöge der Definition von $f_n(t)$. Da in der n-ten und p-ten Zeile nur in der $k(p)$-ten Spalte gleichzeitig Einser stehen, ist unter $\int\limits_D$ stets der eine oder der andere Faktor Null. Daher: $\int\limits_a^c f_n(t) f_p(t)\, dt = 0$.

Will man das System [321] normieren und dividiert zu diesem Zwecke jedes $f_n(t)$ durch $\sqrt{\int\limits_a^c f_n^2(t)\, dt}$, so ändern sich die $f_n(t)$ auch in $<a, b>$. Es ensteht die Frage nach den Eigenschaften eines Systems $\{f_n(t)\}$ in $<a, b>$, die eine Erweiterung zu einem ON in einem $<a, b>$ umfassenden Intervall ermöglichen. I. Schur beantwortet sie, wie folgt:

[322] *Ist* $\{f_n(t)\}$ *in* $<a, b>$ *erklärt*, $0<a<b<1$, $f_n(t) \in L^2$, *so ist eine notwendige und hinreichende Bedingung für die Existenz eines* ON $\{\varphi_n(t)\}$ *in* $<0, 1>$ *mit* $\varphi_n(t) \equiv f_n(t)$ *in* $<a, b>$, *daß das Maximum von*

$$\int\limits_a^b \left[\sum_{i=1}^{\infty} \xi_i f_i(t)\right]^2 dt$$

unter der Nebenbedingung $\sum\limits_{i=1}^{\infty} \xi_i^2 = 1$ *gleich* 1 *sei.*

Die Erweiterung eines Orthogonalsystems. Es sei $\{\psi_n(t)\}$ ein ON in $<a, b>$; setzt man $\psi_n(t) \equiv 0$ in $<b, c>$ für alle n, so bekommt man ein unvollständiges System ($f(t)=0$ in $<a, b>$, $=1$ in $<b, c>$ ist zu allen ψ in $<a, c>$ orthogonal). Ist nun $\{\chi_n(t)\}$ ein ON in $<b, c>$ und setzt man

$$\varphi_n(t) = \lambda\psi_n(t) \qquad \text{für} \quad t \in <a, b>,$$
$$\varphi_n(t) = \mu\chi_n(t) \qquad \text{für} \quad t \in <b, c>,$$

mit $\lambda^2 + \mu^2 = 1$, so ist $\{\varphi_n(t)\}$ ein ON in $<a, c>$.

Dieses System ist unvollständig. Setzt man nämlich

$$g_m(t) = \mu\psi_m(t) \qquad \text{für} \quad t \in <a, b>,$$
$$g_m(t) = -\lambda\chi_m(t) \qquad \text{für} \quad t \in <b, c>,$$

so kommt

$$\int\limits_a^c g_m^2(t)\, dt = \mu^2 + \lambda^2 = 1,$$

und für $n \neq m$

$$\int\limits_a^c \varphi_n(t)\, g_m(t)\, dt = \int\limits_a^b \lambda\psi_n(t)\, \psi_m(t)\, dt + \int\limits_b^c \mu\chi_n(t)\, \chi_m(t)\, dt = 0,$$

für $n = m$

$$\int\limits_a^c \varphi_n(t)\, g_m(t)\, dt = \int\limits_a^b \lambda\mu\psi_m^2(t)\, dt + \int\limits_b^c (-\lambda\mu)\, \chi_m^2(t)\, dt = 0,$$

also stets $\int\limits_a^c \varphi_n(t)\, g_m(t)\, dt = 0$.

Kostitzin, dem man obige Bemerkung verdankt, wies darauf [323]
hin, daß

$$\varphi_1, g_1, \varphi_2, g_2, \ldots, \varphi_n, g_n, \ldots$$

bereits ein ONV (d. h. ein orthogonales, normiertes und vollständiges System) in $<a, c>$ ergibt, wenn $\{\varphi\}$ und $\{\chi\}$ vollständig ist.

Wir haben gesehen, daß die g zu den φ orthogonal sind; daß sie untereinander orthogonal sind, ist unmittelbar klar. Um die Vollständigkeit zu prüfen, setzen wir

$$\int\limits_a^c h(t)\, \varphi_n(t)\, dt = 0, \qquad \int\limits_a^c h(t)\, g_n(t)\, dt = 0 \qquad (n).$$

Es folgt

$$\lambda \int\limits_a^b h\psi_n\, dt + \mu \int\limits_b^c h\chi_n\, dt = 0, \qquad \mu \int\limits_a^b h\psi_n\, dt - \lambda \int\limits_b^c h\chi_n\, dt = 0$$

und, wegen $\begin{vmatrix} \lambda & \mu \\ \mu & -\lambda \end{vmatrix} = -1$,

$$\int\limits_a^b h(t)\, \psi_n(t)\, dt = 0, \qquad \int\limits_b^c h(t)\, \chi_n(t)\, dt = 0$$

für alle n, also — infolge der Vollständigkeit von $\{\psi_n\}$ in $<a, b>$ und von $\{\chi_n\}$ in $<b, c>$ — $h(t) = 0$ f. ü in $<a, b>$ und $<b, c>$. Dasselbe läßt sich für drei und mehr Intervalle (Mengen), ja sogar für Intervallfolgen durchführen.

Es sei A_p eine Folge von Intervallen (Punktmengen) mit [324]
$A_p \times A_q = 0$ für $p \neq q$. Es sei $\{\varphi_{iq}(t)\}$ $(i = 1, 2, \ldots)$ ein ONV auf A_q. Es sei

$$\| \alpha_{pq} \| \qquad (p, q = 1, 2, \ldots)$$

eine unendliche Matrix, die zeilenweise orthogonal, normiert und vollständig ist; letzteres heißt, daß man zu $\| \alpha_{pq} \|$ keine Zeile hinzuschreiben kann, so daß die Matrix orthogonal und normiert bleibt. Wir bilden ein System $\{f_i^{(p)}\}$:

$$f_i^{(p)}(t)=\begin{cases}\alpha_{p1}\,\varphi_{i1}(t) & \text{für } t\in A_1\\ \dots\dots & \dots\ \dots\dots\\ \alpha_{pq}\,\varphi_{iq}(t) & \text{für } t\in A_q\\ \dots\dots & \dots\ \dots\dots\end{cases}\qquad (i,p,q=1,2,\dots).$$

Das System $\{f_i^{(p)}(t)\}$ $(i,p=1,2,\dots)$ ist ein ONV auf $\Sigma=\sum_{q=1}^{\infty}A_q$.

Orthogonalität: Es ist für $i\neq j$

$$\int_{\Sigma} f_i^{(p)}(t)\,f_j^{(n)}(t)\,dt=\sum_{q=1}^{\infty}\alpha_{pq}\,\alpha_{nq}\int_{A_q}\varphi_{iq}(t)\,\varphi_{jq}(t)\,dt=0,$$

wegen der Orthogonalität von $_i\{\varphi_{iq}(t)\}$ auf A_q. Für $p\neq n$ ist

$$\int_{\Sigma} f_i^{(p)}(t)\,f_i^{(n)}(t)\,dt=\sum_{q=1}^{\infty}\alpha_{pq}\,\alpha_{nq}\int_{A_q}\varphi_{iq}^2(t)\,dt=\sum_{q=1}^{\infty}\alpha_{pq}\,\alpha_{nq}=0,$$

wegen der Normiertheit von $_i\{\varphi_{iq}\}$ und der Orthogonalität von $\|\alpha_{pq}\|$.

Normiertheit: Es ist—wegen der Normiertheit von $_i\{\varphi_{iq}(t)\}$ und $\|\alpha_{pq}\|$ —

$$\int_{\Sigma}(f_i^{(p)}(t))^2\,dt=\sum_{q=1}^{\infty}\alpha_{pq}^2\int_{A_q}\varphi_{iq}^2(t)\,dt=\sum_{q=1}^{\infty}\alpha_{pq}^2=1\,.$$

Vollständigkeit: Es sei $g(t)$ in Σ erklärt, $g\in L^2$ und

$$\int_{\Sigma} g(t)\,f_i^{(p)}(t)\,dt=0$$

für alle i und p. Man setze

$$g_{iq}=\int_{A_q} g(t)\,\varphi_{iq}(t)\,dt\,.$$

Es wird

$$0=\int_{\Sigma} g(t)\,f_i^{(p)}(t)\,dt=\sum_{q=1}^{\infty}\alpha_{pq}\,g_{iq}\qquad (p=1,2,\dots);$$

es folgt $g_{iq}=0$ $(q=1,2,\dots)$, da $\|\alpha_{pq}\|$ vollständig. Dieses gilt für alle i. Es ist also (bei festem q) $g_{iq}=0$ $(i=1,2,\dots)$. Die Vollständigkeit von $_i\{\varphi_{iq}(t)\}$ auf A_q ergibt $g(t)=0$ f. ü. auf A_q, also auch auf Σ.

Obiges erlaubt uns ONV-Systeme für das Intervall $\langle 0,\infty)$ zu bilden. Hat man nämlich ein ONV in $\langle 0,1\rangle$ und nennt man es $_i\{\varphi_{i1}(t)\}$, so kann man $_i\{\varphi_{iq}(t)\}$ in $\langle q-1,q\rangle$ vermittels

$$\varphi_{iq}(t)=\varphi_{i1}(t-q+1)$$

erklären, $\langle q-1,q\rangle$ mit A_q bezeichnen, eine beliebige ONV-Matrix $\|\alpha_{pq}\|$ wählen und die vorhergehende Theorie anwenden.

Wählt man z. B. als $_i\{\varphi_{i1}(t)\}$ das trigonometrische System für $\langle 0,1\rangle$ und als $\|\alpha_{pq}\|$ die Matrix der identischen Substitution ($\alpha_{pp}=1$, $\alpha_{pq}=0$ für $p\neq q$), so kommt (II, § 4)

$$f_i^{(p)}(t)=\begin{cases}\sqrt{2}\sin\pi i t & \text{für } p-1\leqslant t\leqslant p\\ 0 & \text{für andere } t\end{cases}\qquad (p,i=1,2,\dots).$$

Dieses System ist ein ONV in $\langle 0,\infty\rangle$. Läßt man für p die Werte $0,-1,-2,\dots$ zu, so erhält man ein ONV in $(-\infty,+\infty)$, und zwar das einfachste bekannte Beispiel.

In den Anwendungen (harmonische Analyse) braucht man manchmal Orthogonalentwicklungen in unendlichen Intervallen; dabei wird oft die Orthogonalität anders als üblich erklärt, und zwar als

$$\lim_{T=\infty}\frac{1}{2T}\int_{-T}^{+T}\varphi_i(t)\,\varphi_k(t)\,dt=0\qquad (i\neq k,\ T>0).$$

Prüfen wir z. B. das System $\{\cos it\}$ auf diese Orthogonalität hin. Es kommt

$$\left|\frac{1}{2T}\int_{-T}^{+T}\cos it\,.\cos kt\,.\,dt\right|=\left|\frac{1}{2T}\int_{-T}^{+T}\frac{1}{2}(\cos(i+k)t+\cos(i-k)t)\,dt\right|$$
$$=\left|\frac{1}{4T}\left[\frac{\sin(i+k)t}{i+k}+\frac{\sin(i-k)t}{i-k}\right]_{-T}^{T}\right|\leqslant\frac{1}{2T}\left(\frac{1}{|i+k|}+\frac{1}{|i-k|}\right)\to 0.$$

Dabei dürfen i,k beliebige reelle Zahlen sein; es ist also das System $\{\cos it\}$ (i=reell) in diesem speziellen Sinne in $(-\infty,+\infty)$ orthogonal und besteht aus $\mathfrak{c}$ (= Kontinuum) verschiedenen Funktionen.

§ 3. **Über die beste Approximation.**

Ist $\{f_i(t)\}$ $(i=1,2,\dots,n)$ *ein beliebiges System und ist* $\sum_{i=1}^{n}c_if_i(t)$ **[331]**
die beste Annäherung an $g(t)$ *im Sinne der kleinsten Quadrate* (II, § 6), *so muß die Differenz*

$$(15)\qquad g(t) - \sum_{i=1}^{n} c_i f_i(t) = h(t)$$

zu sämtlichen $f_i(t)$ $(i = 1, 2, \dots, n)$ *orthogonal sein.*

Beweis: Es seien die $f_i(t)$ linear unabhängig. Bilden wir nach Schmidt ([313]) ein ON-System $\{\varphi_i(t)\}$ und drücken $\{f_i(t)\}$ durch $\{\varphi_i(t)\}$ aus; (15) geht in

$$(16)\qquad h(t) = g(t) - \sum_{i=1}^{n} d_i \varphi_i(t)$$

über. Da die Norm $\|h(t)\|$ minimal sein soll, muß nach II, § 6

$$d_i = \int_a^b g(t)\, \varphi_i(t)\, dt \qquad (i = 1, 2, \dots, n)$$

sein, woraus wegen (16)

$$\int_a^b h(t)\, \varphi_i(t)\, dt = 0 \qquad (i = 1, 2, \dots, n)$$

unmittelbar folgt, was wieder

$$\int_a^b h(t) f_i(t)\, dt = 0 \qquad (i)$$

zur Folge hat, denn die f_i drücken sich linear durch die φ_i aus. Wären die $f_i(t)$ linear abhängig, so wären darunter k linear unabhängige $\overline{f}_1(t), \overline{f}_2(t) \dots \overline{f}_k(t)$ und die restlichen $(n-k)$ ließen sich linear durch jene ausdrücken; $\sum_{i=1}^{n} c_i f_i(t)$ wäre dann eine lineare Verbindung von $\overline{f}_j(t)$ $(j = 1, 2, \dots, k)$ und man hätte $\int_a^q h(t) \overline{f}_j(t)\, dt = 0$ $(j = 1, 2, \dots, k)$, woraus $\int_a^b h(t) f_i(t)\, dt = 0$ auch für die restlichen f folgt.

[331] bringt die Frage nach dem Minimum der quadratischen Form

$$F(x_1, x_2, \dots, x_n) = \int_a^b \left[g(t) - \sum_{i=1}^{n} x_i f_i(t) \right]^2 dt$$

auf die notwendigen Bedingungen

$$(17)\qquad \int_a^b \left[g(t) - \sum_{i=1}^{n} x_i f_i(t) \right] f_k(t)\, dt = 0 \qquad (k = 1, 2, \dots, n),$$

oder — wenn $\int_a^b f_i(t)\, g(t)\, dt = g_i$ gesetzt wird — auf das Gleichungssystem

$$(18)\qquad \sum_{k=1}^{n} a_{ik}\, x_k = g_i \qquad (i = 1, 2, \dots, n),$$

wo a_{ik} in [312] (2) erklärt wurde. Verschwindet nicht die Gram'sche Determinante A_n, d. h. die Determinante der Matrix $\|a_{ik}\|$, so ist das System (18) eindeutig auflösbar. Wir bemerken aber, daß es auch im Falle $A_n = 0$ auflösbar ist. Denn sind $\lambda, \mu, \nu, \dots, \omega$ nicht alle Null und

$$(19)\qquad \lambda a_{ik} + \mu a_{jk} + \dots + \omega a_{wk} = 0 \qquad (k = 1, 2, \dots, n)$$

eine der (sicher vorhandenen) linearen Relationen zwischen den Formen (18), so ist $L(t) \equiv \lambda f_i(t) + \mu f_j(t) + \dots + \omega f_w(t) \equiv 0$ f. ü., weil ja (19)

$$\int_a^b f_k(t)\, L(t)\, dt = 0 \qquad (k = 1, \dots, n)$$

zur Folge hat, und, wenn man hier $k = i, j, \dots, w$ setzt, mit $\lambda, \mu, \dots, \omega$ multipliziert und addiert, $\int_a^b L^2(t)\, dt = 0$ resultiert. Daraus folgt $\int_a^b L(t)\, g(t)\, dt = 0$, also $\lambda g_i + \mu g_j + \dots + \omega g_w = 0$, d. h. daß dieselbe lineare Beziehung, die laut (19) die i-te, j-te, …, w-te Linearform des Systems (18) verbindet, für die freien Glieder g gilt; dieses ist aber zur Lösbarkeit hinreichend (und notwendig).

Damit ist (18) *als notwendige Bedingung für das Minimum* [332]
von $F(x_1, x_2, \dots, x_n)$ *erkannt. Wir wollen zeigen, daß diese Bedingung auch hinreichend ist.*

Wir unterscheiden zwei Fälle: $A_n \neq 0$, $A_n = 0$.

Im ersten Fall lassen sich die $f_1, f_2, \dots, f_n$ orthogonalisieren; es sei $\{\varphi_i(t)\}$ das so erhaltene ON. Setzt man in dem die Form $F(x_1, x_2, \dots, x_n)$ definierenden Integral für $\{f_i(t)\}$ die Ausdrücke (5) ein, so wird

$$F(x_1, x_2, \dots, x_n) = \int_a^b \left[g(t) - \sum_{i=1}^{n} y_i \varphi_i(t) \right]^2 dt = \Phi(y_1, y_2, \dots, y_n)$$

mit

$$
(20)\qquad
\begin{array}{llll}
y_1 = d_{11}x_1 + d_{21}x_2 + \dots + d_{n1}x_n, & & x_1 = c_{11}y_1 + c_{21}y_2 + \dots + c_{n1}y_n, & \\
y_2 = \phantom{d_{11}x_1 +{}} d_{22}x_2 + \dots + d_{n2}x_n, & & x_2 = \phantom{c_{11}y_1 +{}} c_{22}y_2 + \dots + c_{n2}y_n, & \\
\dots\quad \dots\dots\dots & & \dots\quad \dots\dots\dots & \\
y_n = \phantom{d_{11}x_1 + d_{21}x_2 + \dots +{}} d_{nn}x_n, & & x_n = \phantom{c_{11}y_1 + c_{21}y_2 + \dots +{}} c_{nn}y_n. &
\end{array}
$$

Es genügt also das Minimum der Form $\Phi(y_1, y_2, \dots, y_n)$ zu bestimmen; dieses ist in [261] und [262] schon geschehen: es gibt genau eine Lösung $y_i = \int_a^b g(t)\,\varphi_i(t)\,dt$ $(i = 1, \dots, n)$, von der man mit Hilfe von (20) zu den gesuchten x übergeht. Da es nur eine Lösung gibt, (18) aber eindeutig auflösbar ist, so ist (18) hinreichend.

Im zweiten Fall ist der Unterfall $f_i(t) \equiv 0$ f. ü. $(i = 1, \dots, n)$ trivial; alle Systeme $\{x_i\}$ lösen (18) und alle liefern $\int_a^b g^2(t)\,dt$ als Minimalwert. Es seien also $f_1(t), \dots, f_p(t)$ linear unabhängig $(1 \leqslant p < n)$ und die restlichen $n-p$ Funktionen durch jene p ausdrückbar. Wir betrachten

$$\mathring{F}(x_1, x_2, \dots, x_p) = \int_a^b \left[g(t) - \sum_{i=1}^{p} x_i f_i(t)\right]^2 dt$$

und finden, wie oben, ein Minimalsystem $\mathring{x}_1, \mathring{x}_2, \dots, \mathring{x}_p$ für $\mathring{F}$. Da $\sum_{i=1}^{n} x_i' f_i(t)$ auf die Gestalt $\sum_{i=1}^{p} x_i f_i(t)$ (durch Ersetzung von $f_i(t)$ $(i = p+1, \dots, n)$ durch lineare Ausdrücke in $f_1, \dots, f_p$) sich bringen läßt, liefert

$$(21)\qquad \mathring{x}_1, \mathring{x}_2, \dots, \mathring{x}_p, 0, 0, \dots, 0$$

das gesuchte Minimum. Es ist bloß zu prüfen, ob andere Lösungen von (18) denselben Wert für $F(x_1, x_2, \dots, x_n)$ wie (21) liefern. Das ist aber leicht einzusehen; sind $x_1', x_2', \dots, x_n'$, $x_1'', x_2'', \dots, x_n''$, zwei Lösungssysteme von (18), so ist

$$(22)\qquad \sum_{i=1}^{n} x_i' f_i(t) \equiv \sum_{i=1}^{n} x_i'' f_i(t) \quad \text{f. ü.},$$

denn schreibt man $x_i' - x_i'' = \xi_i$, so hat man

$$\sum_{k=1}^{n} a_{ik}\,\xi_k = 0 \quad (i), \qquad \sum_{i=1}^{n} \xi_i \sum_{k=1}^{n} a_{ik}\,\xi_k = 0,$$

also ([312], (3))

$$\int_a^b \left(\sum_{i=1}^{n} \xi_i f_i(t)\right)^2 dt = 0,$$

was (22) zur Folge hat.

Jetzt können wir das Minimum von F unmittelbar (ohne Orthogonalisierung) bestimmen. Das System (18) hat im Falle $A_n \neq 0$ als einziges Lösungssystem

$$(23)\qquad x_k = \frac{\underset{k}{A_n}}{A_n} \qquad (k = 1, 2, \dots, n),$$

wobei die Determinante $\underset{k}{A_n}$ aus A_n gewonnen wird, indem man die k-te Spalte in A_n durch $g_1, g_2, \dots, g_n$ ersetzt. Nun gilt aber in (23) nach (17)

$$F(x_1, x_2, \dots, x_n) = \int_a^b \left[g(t) - \sum_{i=1}^{n} x_i f_i(t)\right]\left[g(t) - \sum_{i=1}^{n} x_i f_i(t)\right] dt$$

$$= \int_a^b g(t)\left[g(t) - \sum_{i=1}^{n} x_i f_i(t)\right] dt = \int_a^b g^2(t)\,dt - \sum_{k=1}^{n} g_k x_k.$$

Schreibt man hier $\int_a^b g^2(t)\,dt = g_0$, so kommt

$$\operatorname{Min} F(x_1, x_2, \dots, x_n) = g_0 - \frac{1}{A_n}\sum_{k=1}^{n} g_k \underset{k}{A_n} = \frac{1}{A_n}\begin{vmatrix} g_0 & g_1 & \dots & g_n \\ g_1 & a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots & \dots \\ g_n & a_{n1} & \dots & a_{nn} \end{vmatrix}.$$

Diese Determinante ist nichts anderes, als die Gram'sche Determinante $G(f_0, f_1, \dots, f_n)$ des Funktionensystems $f_0(t), f_1(t), \dots, f_n(t)$, wo $f_0(t)$ für $g(t)$ geschrieben wurde; dieses führt zur Formel

$$\operatorname{Min} F(x_1, x_2, \dots, x_n) = \frac{G(g, f_1, f_2, \dots, f_n)}{G(f_1, f_2, \dots, f_n)},$$

oder

$$\operatorname*{Min}_{\{x_i\}} \left\| g(t) - \sum_{i=1}^{n} x_i f_i(t) \right\| = \sqrt{\frac{G(g, f_1, f_2, \dots, f_n)}{G(f_1, f_2, \dots, f_n)}}. \qquad [333]$$

Damit haben wir einen Ausdruck für die beste Annäherung an eine gegebene Funktion in L^2 durch lineare Verbindung von n gegebenen, linear unabhängigen Funktionen.

Ist $f_{n+1}(t)$ durch $f_1(t), \dots, f_n(t)$ linear ausdrückbar, so läßt sich die Approximation durch Heranziehung von $f_{n+1}(t)$ nicht verbessern; würde nämlich das Minimum kleiner werden, so könnte man $f_{n+1}(t)$ durch den linearen Ausdruck ersetzen und das Minimum [333]
verkleinern.

[334] *Eine notwendige und hinreichende Bedingung dafür, daß die Hinzunahme von $f_{n+1}(t)$ die Annäherung nicht beeinfluße, ist die Orthogonalität von $f_{n+1}(t)$ zu der Differenz*

$$r(t) = g(t) - \sum_{k=1}^{n} \overset{n}{x}_k f_k(t),$$

(wo $\{\overset{n}{x}_k\}$ das Lösungssystem der Approximationsaufgabe für n Funktionen $f_k(t)$ bedeutet).

Ist nämlich $\int_a^b r(t) f_{n+1}(t)\, dt = 0$, so ist

$$\int_a^b [r(t) - \lambda f_{n+1}(t)]^2\, dt = \int_a^b r^2(t)\, dt + \lambda^2 \int_a^b f_{n+1}^2(t)\, dt \geqslant \int_a^b r^2(t)\, dt, \quad \text{d. h.}$$

$$\| g(t) - \sum_{k=1}^{n} \overset{n}{x}_k f_k(t) - \lambda f_{n+1}(t) \| \geqslant \| g(t) - \sum_{k=1}^{n} \overset{n}{x}_k f_k(t) \|.$$

Ist aber $\int_a^b r(t) f_{n+1}(t)\, dt \neq 0$, so ist $\int_a^b [r(t) - \lambda f_{n+1}(t)]^2\, dt < \int_a^b r^2(t)\, dt$, wenn λ das Vorzeichen von $\int_a^b r(t) f_{n+1}(t)\, dt$ hat und

$$2|\lambda| \cdot \left| \int_a^b r(t) f_{n+1}(t)\, dt \right| > \lambda^2 \int_a^b f_{n+1}^2(t)\, dt, \quad |\lambda| < \frac{2 \left| \int_a^b r(t) f_{n+1}(t)\, dt \right|}{\int_a^b f_{n+1}^2(t)\, dt} \text{ ist.}$$

§ 4. Abzählbarkeit.

Eine normierte Orthogonalmenge in L^2 kann abzählbar-unendlich sein, wie wir an den Beispielen des II-ten Kapitels gesehen haben, es kann aber keine größere Mächtigkeit besitzen.
[341] M. a. W.: *Ein* ON *in* L^2 *besteht aus höchstens abzählbar-unendlich-vielen Funktionen.*

Beweis: Am Schluß von I, § 2 gaben wir ein abzählbares, in L^2 überalldichtes Funktionensystem an; bezeichnen wir es mit

$\{w_i(t)\}$. Ferner sei $\{\varphi(t)\}$ ein ON; sind $\varphi_\alpha(t)$, $\varphi_\beta(t)$ zwei verschiedene Elemente von $\{\varphi\}$, so ist

$$(24) \qquad \|\varphi_\alpha - \varphi_\beta\| = \sqrt{\int_a^b (\varphi_\alpha(t) - \varphi_\beta(t))^2\, dt} = \sqrt{2}.$$

Es sei $i(\alpha)$ die kleinste natürliche Zahl mit

$$\|\varphi_\alpha - w_{i(\alpha)}\| < \frac{\sqrt{2}}{2}.$$

Für $\alpha \neq \beta$ (α, β sind Zeichen, keine Zahlen) ist $i(\alpha) \neq i(\beta)$, denn sonst wäre $\|\varphi_\alpha - w_{i(\alpha)}\| < \frac{\sqrt{2}}{2}$, $\|\varphi_\beta - w_{i(\alpha)}\| < \frac{\sqrt{2}}{2}$, $\|\varphi_\alpha - \varphi_\beta\| < \sqrt{2}$ gegen (24). Die Funktion $i(\alpha)$ bildet also die Menge $\{\alpha\}$ eineindeutig auf einen Teil der natürlichen Zahlenreihe ab, woraus die Endlichkeit oder Abzählbarkeit der Menge $\{\alpha\}$ aller Indizes resultiert.

Es ist klar, daß jedes endliche O erweitert werden kann; zu den Funktionen $\varphi_1(t), \varphi_2(t), \dots, \varphi_n(t)$ braucht man nur ein unabhängiges $f(t)$ hinzuzufügen und aus den $n+1$ Funktionen durch Orthogonalisierung $\varphi_{n+1}(t)$ zu bilden.

Die Eigenschaft der Vollständigkeit oder der Abgeschlossenheit ist von Einfluß auf die Anzahl:

Ein vollständiges System ist stets unendlich. [342]

Besteht das System aus $f_1(t), f_2(t), \dots, f_n(t)$, so kann man ein Polynom n-ten Grades $P(t) \equiv 0$

$$P(t) = \alpha_0 + \alpha_1 t + \dots + \alpha_n t^n$$

so bestimmen, daß

$$\int_a^b P(t) f_i(t)\, dt = 0 \qquad (i = 1, 2, \dots, n)$$

sei; dieses führt auf n homogene Gleichungen mit $n+1$ Unbekannten $\alpha_0, \alpha_1, \dots, \alpha_n$, die man bekanntlich nichttrivial lösen kann; $P(t)$ ist dann nicht Null, und die Unvollständigkeit von $\{f\}$ erwiesen. [341] und [342] liefern das Resultat:

Ein ONV *ist stets abzählbar-unendlich.* [343]

Ein abgeschlossenes System (kurz: „A“) ist stets unendlich. [344]

Ist nämlich $\int_a^b \left[f(t) - \sum_{i=1}^{n} c_i f_i(t) \right]^2 dt < \varepsilon$ für gewisse $\{c_i\}$, so liefert das Häufungspunktverfahren ein $\{d_i\}$ mit

$$\int_a^b \left[f(t) - \sum_{i=1}^n d_i f_i(t)\right]^2 dt = 0,$$

also $f(t) = \sum_{i=1}^n d_i f_i(t)$ f. ü. Wählt man für $f(t)$ nacheinander $1, t, \ldots, t^n$, so resultiert eine lineare Abhängigkeit der Potenzen $\{t^k\}$ $(k=0,1,\ldots,n)$ untereinander, gegen einen elementaren Satz der Algebra.

[345] *Ein* ONA *ist stets abzählbar-unendlich.*

Dieser Satz folgt aus [341] und [344].

Jetzt bringen wir Sätze, die vollständige oder abgeschlossene Systeme auf abzählbare reduzieren:

[346] *Jedes abgeschlossene System enthält einen abzählbaren abgeschlossenen Teil.*

Beweis: Sei $\{w_i(t)\}$ das in [341] so bezeichnete und $\{f(t)\}$ das gegebene, inbezug auf L^2 abgeschlossene System; wir bezeichnen mit $h(t)$ eine lineare Form $\sum_{i=1}^n c_i f_i(t)$, wo die f_i beliebig aus $\{f\}$ zu wählen und die c_i beliebige Konstanten sind. Wir bilden das Schema

$$\begin{array}{l} h_{11} \\ h_{21},\ \ h_{22} \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ h_{p1},\ \ h_{p2},\ \ \ldots,\ \ h_{pp}, \end{array} \tag{25}$$

wo h_{pq} $(p, q = 1, 2, \ldots)$ ein $h(t)$ ist, welches die Ungleichung

$$\| w_q - h_{pq} \| < 1/p$$

erfüllt. Zu jedem Zahlenpaar p, q $(p \geqslant q)$ gibt es ein h_{pq}, denn $\{f_i\}$ ist abgeschlossen. Ist n natürlich und $\varepsilon > 0$, so ist für $p > n$, $p > 1/\varepsilon$, $q = n$

$$\| w_n - h_{pq} \| < \varepsilon.$$

Betrachtet man jetzt bloß jene $f_i(t)$, die zum Bau der h_{pq} des Schemas (25) gebraucht wurden, so bilden sie einen höchstens abzählbaren Teil von $\{f\}$, denn $\{h_{pq}\}$ ist abzählbar und jedes h_{pq} setzt sich aus endlichvielen f_i zusammen. Lineare Verbindungen der Elemente dieses Teiles $\{f_i\}$ nähern jedes $w_i(t)$ beliebig genau an. Da $\{w_i\}$ überalldicht ist, sind es jene linearen Verbindungen auch. $\{f_i\}$ ist also abgeschlossen; es ist höchstens abzählbar, also — laut [344] — abzählbar.

Jedes vollständige System enthält einen abzählbaren vollständigen Teil. [347]

Beweis: Es sei $\{f\}$ inbezug auf L^2 vollständig. $\{f\}$ ist ein Teil von L^2. L^2 ist aber separabel (vgl. [139]); es geht darum, zu zeigen, daß $\{f\}$ separabel ist; dies ist der Hauptpunkt des Beweises. Sei $\{g_i\}$ ein abzählbarer, überalldichter Teil von L^2. Es sei δ_i die Distanz $(g_i, \{f\})$, d. h. die untere Grenze der Distanzen (g_i, f) für alle $f \in \{f\}$. Es gibt also ein f_i, so daß

$$\| g_i - f_i \| < \delta_i + 1/i.$$

Wir behaupten, daß die höchstens abzählbare Menge $\{f_i\}$ in $\{f\}$ überalldicht ist. Es sei nämlich f ein Element von $\{f\}$ und $\varepsilon > 0$. Es gibt dann ein g_i mit $i > 3/\varepsilon$ und

$$\| f - g_i \| < \varepsilon/3;$$

infolgedessen ist $\delta_i < \varepsilon/3$; es folgt

$$\| g_i - f_i \| < \delta_i + 1/i < 2\varepsilon/3, \qquad \text{also} \qquad \| f - f_i \| < \varepsilon.$$

Jetzt zeigen wir, daß $\{f_i\}$ vollständig ist. Ist nämlich für ein $h(t) \in L^2$

$$\int_a^b h(t) f_i(t)\, dt = 0 \qquad (i) \tag{26}$$

und $f(t)$ irgend eine Funktion aus $\{f\}$, so enthält, wie eben bewiesen, $\{f_i\}$ eine gegen $f(t)$ stark konvergente Folge und (26) liefert

$$\int_a^b h(t) f(t)\, dt = 0 \qquad (f)$$

für alle $f \in \{f\}$; es ist also $h(t) \equiv 0$ f. ü. Daher ist $\{f_i\}$ vollständig; da es höchstens abzählbar ist, so ist es nach [342] abzählbar und damit [347] bewiesen.

Dagegen braucht ein System von kontinuierlich-unendlichvielen linear unabhängigen Funktionen nicht vollständig zu sein. Es sei S die Menge aller Funktionen aus L^2 im Intervall $\langle 0, 1 \rangle$; definiert man jede Funktion in $(1, 2\rangle$ als Null, so hat man ein Kontinuum von Funktionen in $\langle 0, 2\rangle$ erklärt. Dieses Kontinuum enthält einen Teil T von gleicher Mächtigkeit mit lauter linear-unabhängigen Elementen. Setzt man $h(t)=0$ in $<0,1>$ und $=1$ in $(1,2>$, so ist $h(t)$ zu allen Funktionen des Kontinuums T orthogonal und nicht Null f. ü.

§ 5. Vollständigkeit und Abgeschlossenheit.

In diesem § sollen einige Eigenschaften der vollständigen und der abgeschlossenen Systeme angegeben werden.

[351] *Ist ein System in $<a,b>$ vollständig und ist $a<b'<b$, so ist es auch in $<a,b'>$ vollständig.*

Diese Eigenschaft ist sofort ersichtlich, denn ist

$$\int_a^{b'} f(t)\,g(t)\,dt=0 \qquad (f)$$

für alle $f(t)$ des Systems und $g(t)$ nicht f. ü. Null in $<a,b'>$, so kann man $g(t)$ in $(b',b>$ als Null erklären und man hat $\int_a^b f(t)\,g(t)\,dt=0$ für dieses g und alle f des Systems. Die Unvollständigkeit in $<a,b'>$ hat also die Unvollständigkeit in $<a,b>$ zur Folge, w. z. b. w.

Aus dem Beweise sieht man auch, daß der umgekehrte Satz nicht gilt (so ist z. B. $\{\sin ix\}$ in $<0,\pi>$ aber nicht in $<-\pi,\pi>$ vollständig). Analog zu [351] ist der Satz:

[352] *Ein in $<a,b>$ abgeschlossenes System ist auch in $<a,b'>$ abgeschlossen* ($a<b'<b$). Es ist nämlich

$$\int_a^{b'}\left[f(t)-\sum_{i=1}^n c_i f_i(t)\right]^2 dt \leqslant \int_a^b \left[f(t)-\sum_{i=1}^n c_i f_i(t)\right]^2 dt.$$

Vervollständigung von Systemen. Es ist trivial, daß jedes unvollständige System sich vervollständigen läßt; es genügt nämlich ein Beispiel eines vollständigen Systems zu haben und dieses hinzuzufügen. Nichttrivial ist aber:

[353] *Jedes* ON *läßt sich zu einem* ONV *erweitern durch Hinzunahme von geeigneten Funktionen.*

Beweis: $\{\varphi_i\}$ sei ON, aber nicht ONV. Die Menge $\{f\}$ aller normierten, zu sämtlichen φ_i orthogonalen Funktionen ist nicht leer. $\{\varphi_i\}+\{f\}$ ist bereits vollständig, wie unmittelbar ersichtlich, und enthält nach [347] einen vollständigen abzählbaren Teil T; T enthält einen höchstens abzählbaren Teil $\{f_i\}$ von $\{f\}$ und ist daher in $\{\varphi_i\}+\{f_i\}$ enthalten; $\{\varphi_i\}+\{f_i\}$ ist also vollständig und abzählbar. Schreibt man $\{f_i\}$ als (endliche oder unendliche) Folge: $f_1, f_2, \ldots$ und streicht sukzessive jedes f, welches von den vorhergehenden linear abhängig ist, so erhält man ein System $\{f_i^*\}$ von linear unabhängigen, normierten und zu allen φ_i orthogonalen Funktionen. $\{\varphi_i\}+\{f_i^*\}$ ist noch immer vollständig. Orthogonalisiert man $\{f_i^*\}$ und ist $\{\psi_i\}$ das Ergebnis, so bildet $\{\varphi_i\}+\{\psi_i\}$ ein durch Erweiterung von $\{\varphi_i\}$ gewonnenes ONV.

Eine wichtige Eigenschaft ist die Identität der Begriffe „vollständig" und „abgeschlossen" inbezug auf L^2. Sie beruht auf zwei Sätzen, deren erster weniger tief liegt als der zweite.

Jedes abgeschlossene System ist vollständig. [354]

Beweis: Ein abgeschlossenes System $\{f\}$ enthält nach [346] einen abgeschlossenen abzählbaren Teil $\{f_i\}$. Ist $\{\varphi_i\}$ das aus $\{f_i\}$ durch Orthogonalisierung gewonnene ON, so ist ([316]) $\{\varphi_i\}$ abgeschlossen. Es sei nun $\int_a^b g(t) f_i(t)\,dt=0$ (i); es folgt $\int_a^b g(t)\,\varphi_i(t)\,dt=0$ (i). Ist $s_n(t)$ die n-te Partialsumme der Orthogonalentwicklung von $g(t)$ nach $\{\varphi_i\}$ ([253]), so ist — wegen [267] —

$$\lim_{n\to\infty}\int_a^b [g(t)-s_n(t)]^2\,dt=0.$$

Nun sind hier aber die Koeffizienten der Entwicklung Null, es verschwindet also $s_n(t)$ und es folgt $g(t)=0$ f. ü.

Jedes vollständige System ist abgeschlossen. [355]

Beweis: Laut [347] enthält ein vollständiges System $\{f\}$ einen vollständigen abzählbaren Teil $\{f_i\}$. Daraus gewinnt man durch Orthogonalisierung ein ON $\{\varphi_i\}$, welches nach [315] ebenfalls vollständig ist. Wir werden zeigen, daß $\{\varphi_i\}$ abgeschlossen ist. Denn sonst gäbe es ein $g(t)\in L^2$ und ein $\alpha>0$ mit

$$\int_a^b [g(t)-s_n(t)]^2\,dt>\alpha \qquad (n) \tag{27}$$

(wo $s_n(t)$ wie in [354] erklärt ist); schreibt man $c_i=\int_a^b g(t)\,\varphi_i(t)\,dt$, so ist nach [264]

$$\sum_{i=1}^{\infty} c_i^2<\infty,$$

also $\lim\limits_{m\geqslant n\to\infty} \| s_m - s_n \| = 0$, d. h. $\{s_n(t)\}$ ist ([161]) stark konvergent; L^2 ist — wie wir in [131] gesehen haben — ein kompletter Raum, es gibt also ein $f(t) = \operatorname{Lim}\limits_{n\to\infty} s_n(t)$ in L^2; man hat $\lim\limits_{n\to\infty} \| f - s_n \| = 0$ und die Schwarz'sche Ungleichung gibt

$$\left| \int_a^b f(t)\,\varphi_i(t)\,dt - c_i \right| = \left| \int_a^b [f(t) - s_n(t)]\,\varphi_i(t)\,dt \right| \leqslant \|f - s_n\| \,.\, \|\varphi_i\| = \|f - s_n\| \to 0, \qquad (n > i)$$

d. h.

$$\int_a^b f(t)\,\varphi_i(t)\,dt = c_i \qquad (i). \tag{28}$$

Es folgt $\int_a^b [f(t) - g(t)]\,\varphi_i(t)\,dt = 0$ (i) und die Vollständigkeit von $\{\varphi_i\}$ liefert $f(t) = g(t)$ f. ü. Man darf also in (27) f für g schreiben und erhält $\|f - s_n\| > \sqrt{\alpha}$ für alle n, gegen das bewiesene $\lim\limits_{n\to\infty} \|f - s_n\| = 0$. Das System $\{\varphi_i\}$ ist also abgeschlossen; so ist es auch $\{f_i\}$ ([316]) und somit $\{f\}$, w. z. b. w.

[355] ist tiefer als [354], denn 1° ist das Lemma [347] versteckter als [346], 2° ist die Eigenschaft [131], eine der wichtigsten in der Theorie des Raumes L^2, die selbst einen nichttrivialen Beweis erfordert hat, für den Beweis von [355] wesentlich.

Der Beweis für [355] enthält die Begründung eines berühmten, ungefähr gleichzeitig von Friedrich Riesz und Ernst Fischer (1907) entdeckten Satzes:

[356] *Ist $\{\varphi_i(t)\}$ orthogonal und normiert und $\{c_i\}$ eine Zahlenfolge von konvergenter Quadratsumme, so gibt es ein $f(t) \in L^2$, dessen Orthogonalentwicklung nach $\{\varphi_i(t)\}$ die Reihe*

$$\sum_{i=1}^{\infty} c_i\,\varphi_i(t)$$

liefert.

Denn [356] behauptet nichts anderes, als daß man unter der Voraussetzung $\sum\limits_{i=1}^{\infty} c_i^2 < \infty$ ein $f(t) \in L^2$ mit (28) finden kann, was wir im Beweise von [355] lesen können, wenn wir von der Entstehungsweise der c_i absehen, $s_n(t)$ als $\sum\limits_{i=1}^{n} c_i\,\varphi_i(t)$ verstehen und mit (28) den Gedankengang abschließen.

Anwendungen. Das System $\{\sin it\}$ ist in $\langle 0, \pi \rangle$ nach [234] vollständig inbezug auf L^2, es ist also nach [355] abgeschlossen inbezug auf L^2. Daraus folgt aber die Abgeschlossenheit inbezug auf L. Denn für ein $f(t) \in L$ gibt es — nach [129] — ein $g(t) \in C$ mit

$$\int_a^b |f(t) - g(t)|\,dt < \frac{\varepsilon}{2}.$$

Ist $w_n(t)$ ein Sinuspolynom n-ter Ordnung mit

$$\int_a^b [g(t) - w_n(t)]^2\,dt < \frac{\varepsilon^2}{4(b-a)},$$

so ist $\int_a^b |g(t) - w_n(t)|\,dt < \frac{\varepsilon}{2}$, also $\int_a^b |f(t) - w_n(t)|\,dt < \varepsilon$, w. z. b. w.

Ebenso erhält man die Abgeschlossenheit von $\{\cos it\}$ in $\langle 0, \pi \rangle$ inbezug auf L^2 und L.

Der Weierstraß'sche Satz [122] zeigt die Abgeschlossenheit der Polynome inbezug auf C; obiger Schluß überträgt diese Eigenschaft auf L und L^2. Daraus folgt nach [354] die Vollständigkeit der Polynome inbezug auf L^2. Wir wollen nun, indem wir einen klassischen Lerch'schen Satz verschärfen, die Vollständigkeit inbezug auf L zeigen. Der Lerch'sche Satz lautet:

Ist $f(t)$ in $\langle 0, 1 \rangle$ stetig und [357]

$$\int_0^1 f(t)\,t^n\,dt = 0 \qquad (n = 0, 1, 2, \ldots),$$

so ist $f(t) \equiv 0$.

In dieser Form haben wir eigentlich den Satz bereits bewiesen, denn ist W die Menge aller Polynome $w(t)$, so impliziert die Lerch'sche Voraussetzung, daß

$$\int_a^b f(t)\,w(t)\,dt = 0 \qquad (w \in W)$$

und — da W inbezug auf L^2, also umsomehr inbezug auf C vollständig — daß $f(t) = 0$ f. ü., d. h. daß $f(t) \equiv 0$ ist.

Es sei jetzt $f(t) \in L$ und $\int_0^1 f(t)\,.\,t^n\,dt = 0$ (n). Setzen wir [358]

$$F(t) = \int_t^1 f(u)\,du;$$

es kommt

$$\int_0^1 f(t)\,t^n\,dt = \Big[-F(t)\,t^n\Big]_0^1 + \int_0^1 nF(t)\,t^{n-1}\,dt = n\int_0^1 F(t)\,t^{n-1}\,dt \qquad (n \geqslant 1),$$

d. h. $\int_0^1 F(t)\,t^n\,dt = 0$ $(n = 0, 1, 2, \ldots)$, und der Lerch'sche Satz liefert $F(t) \equiv 0$, woraus $f(t) = \frac{d}{dt}F(t) = 0$ f. ü. folgt.

Der Satz [358] ist für ein beliebiges Intervall $<a, b>$ gültig. Ist nämlich $\int_a^b f(t)\,t^n\,dt = 0$ und setzt man $\frac{t-a}{b-a} = \tau$, $f(t) = \varphi(\tau)$, so ist $\varphi(\tau) \in L$, wenn $f(t) \in L$, und wegen $\tau^n = \left(\frac{t-a}{b-a}\right)^n$ ergibt die Voraussetzung $\int_a^b f(t)\,\tau^n\,dt = 0$, d. h. $(b-a)\int_0^1 \varphi(\tau)\,\tau^n\,d\tau = 0$, woraus $\varphi(\tau) = 0$ f. ü. nach [358] und $f(t) = 0$ f. ü. resultiert. A fortiori ist W in $<a, b>$ vollständig inbezug auf L.

Die Orthogonalisierung der Potenzen $\{t^n\}$ in $<-1, +1>$ führt
[359] auf Legendre'sche Polynome [317]. Auf Grund des eben Bewiesenen und auf Grund des allgemeinen Satzes [315] bilden diese Polynome ein ONV; [355] zeigt, daß sie auch ein ONA bilden (was sich übrigens auch leicht aus [122] folgern läßt).

§ 6. **Der Müntz'sche Satz.**

Dieser Satz ist eine Vervollkommnung des Weierstraß'schen und des Lerch'schen Satzes. Allgemein gesprochen: in der Folge $\{t^n\}$ sind nicht alle Glieder nötig, um ein abgeschlossenes System zu erhalten. Genau formuliert lautet das Problem: Wie muß die Exponentenfolge $\{p_i\}$ beschaffen sein, damit das Funktionensystem $\{t^{p_i}\}$ in $<0, 1>$ inbezug auf L^2 abgeschlossen sei?

Zunächst darf man alle $p_i > -1/2$ voraussetzen; sonst wären bei der Annäherung an irgend eine Funktion in L^2 die (etwa vorhandenen) t^{p_i} $(p_i \leqslant -1/2)$ stets mit Nullen als Koeffizienten zu versehen, damit die Summe $\sum c_i t^{p_i}$ zu L^2 gehöre; sie spielen also keine Rolle bei der Annäherung. Wir wenden die Formel [333] an, um die beste Annäherung an $g(t) = t^q$ $(q > -1/2)$ vermittels einer aus den n ersten t^{p_i} gebildeten Linearform zu berechnen. In dortigen Bezeichnungen ist jetzt

$$a_{ik} = \int_0^1 t^{p_i}\,t^{p_k}\,dt = \frac{1}{p_i + p_k + 1},$$

$$g_0 = \int_0^1 (t^q)^2\,dt = \frac{1}{q+q+1}, \quad g_i = \int_0^1 t^q \,.\, t^{p_i}\,.\,dt = \frac{1}{q+p_i+1} \quad (i=1,2,\ldots,n).$$

Ist also $w_n(t)$ die gesuchte Linearform, so wird

$$\| t^q - w_n(t) \| = \sqrt{\frac{G(t^q, t^{p_1}, t^{p_2}, \ldots, t^{p_n})}{G(t^{p_1}, t^{p_2}, \ldots, t^{p_n})}}$$

mit

$$G(t^{p_1}, t^{p_2}, \ldots, t^{p_n}) = \begin{vmatrix} \frac{1}{p_1+p_1+1} & \frac{1}{p_1+p_2+1} & \cdots & \frac{1}{p_1+p_n+1} \\ \frac{1}{p_2+p_1+1} & \frac{1}{p_2+p_2+1} & \cdots & \frac{1}{p_2+p_n+1} \\ \cdots\cdots & \cdots\cdots & & \cdots\cdots \\ \frac{1}{p_n+p_1+1} & \frac{1}{p_n+p_2+1} & \cdots & \frac{1}{p_n+p_n+1} \end{vmatrix}$$

$$G(t^q, t^{p_1}, \ldots, t^{p_n}) = \begin{vmatrix} \frac{1}{q+q+1} & \frac{1}{q+p_1+1} & \cdots & \frac{1}{q+p_n+1} \\ \frac{1}{p_1+q+1} & \frac{1}{p_1+p_1+1} & \cdots & \frac{1}{p_1+p_n+1} \\ \cdots\cdots & \cdots\cdots & & \cdots\cdots \\ \frac{1}{p_n+q+1} & \frac{1}{p_n+p_1+1} & \cdots & \frac{1}{p_n+p_n+1} \end{vmatrix}.$$

Zur Auswertung dieser Determinanten brauchen wir einen Cauchy'schen Satz über Determinanten der Form

$$D = \left| \frac{1}{p_i + q_k} \right| \qquad (i, k = 1, 2, \ldots, n),$$

nach welchem der Wert von D gleich

$$\frac{\prod\limits_{i>k} (p_i - p_k)(q_i - q_k)}{\prod\limits_{i,k=1,1}^{n,n} (p_i + q_k)} \tag{29}$$

ist. Hier ist das Produkt im Nenner über alle n^2 Paare (i, k) von

(1, 1) bis (n, n) erstreckt, während im Zähler der Vorbehalt $i > k$ die Anzahl der Faktoren vermindert.

Beweis des Cauchy'schen Satzes: Entwickelt man D und summiert alle Glieder, so sieht man, daß D eine rationale Funktion $(-n)$-ten Grades der p_i, q_k ist, denn jedes $\frac{1}{p_i+q_k}$ hat den Grad -1. Bringt man also D auf den (gemeinsamen) Nenner $\prod_{i,k=1,1}^{n,n}(p_i+q_k)$ vom Grade n^2, so erhält der Zähler höchstens den Grad n^2-n. Nun muß aber der Zähler für $p_i=p_k$ $(i\neq k)$ und $q_i=q_k$ $(i\neq k)$ verschwinden, denn dann sind zwei Zeilen (bzw. Spalten) von D miteinander identisch. Es muß also der Zähler alle Faktoren p_i-p_k, q_i-q_k $(i>k)$ enthalten, deren es n^2-n gibt. Diese Anzahl stimmt mit dem Grade des Zählers überein, es kann also bloß noch ein konstanter Faktor das Produkt $\prod_{i>k}(p_i-p_k)(q_i-q_k)$ multiplizieren. Man sieht aber, daß bei der Reduktion auf den Nenner $\prod_{1,1}^{n,n}(p_i+q_k)$ das Zählerpolynom lauter Glieder mit ± 1 als Koeffizienten erhält. Es handelt sich also bloß um das Vorzeichen; nun ist aber D für $p_i=m^{i-1}$, $q_i=1-m^{i-1}$ (i) für große positive m-Werte nahe an 1 und (29) ebenfalls; demnach ist das Vorzeichen $+$ und D mit dem Bruch (29) als identisch erwiesen.

Wir können den Cauchy'schen Satz auf $G(t^{p_1}, t^{p_2}, \ldots, t^{p_n})$ anwenden, indem wir $p_i+1/2$ für p_i und $p_k+1/2$ für q_k einsetzen. Es kommt

$$G(t^{p_1}, t^{p_2}, \ldots, t^{p_n}) = \frac{\prod\limits_{i>k}(p_i-p_k)^2}{\prod\limits_{1,1}^{n,n}(p_i+p_k+1)}$$

und ebenso

$$G(t^q, t^{p_1}, \ldots, t^{p_n}) = \frac{\prod\limits_{i>k}(p_i-p_k)^2 . \prod\limits_{1}^{n}(p_i-q)^2}{\prod\limits_{1,1}^{n,n}(p_i+p_k+1) . \prod\limits_{1}^{n}(p_i+q+1)^2 . (2q+1)},$$

also

$$\| t^q - w_n(t) \|^2 = \frac{1}{2q+1}\prod_{i=1}^{n}\left(\frac{p_i-q}{p_i+q+1}\right)^2. \tag{30}$$

Damit das System $\{t^p\}$ inbezug auf L^2 abgeschlossen sei, muß — bei jedem $q\geqq 0$ — der Ausdruck $\| t^q - w_n(t) \|$ beliebig klein gemacht werden können durch entsprechende Wahl von n und $t^{p_1}, t^{p_2}, \ldots, t^{p_n}$ aus $\{t^p\}$. Andererseits ist dieses hinreichend, denn die t^q $(q\geqq 0)$ bilden gewiß ein abgeschlossenes System und man kann wie beim Beweise von [245] schließen. Nach (30) verlangt diese notwendige und hinreichende Bedingung, daß das Exponentensystem $\{p\}$ zu jedem q eine Folge $\{p_i\}$ mit

$$\prod_{i=1}^{\infty}\left(\frac{p_i-q}{p_i+q+1}\right)^2 = 0 \tag{31}$$

enthalte (dabei sind Produkte mit Nullfaktoren als konvergent anzusehen).

1° Hat $\{p\}$ einen Häufungspunkt p_∞ im Endlichen und $p_\infty \neq -1/2$, so gibt es eine Folge $\{p_i\}$ mit $\lim\limits_{i\to\infty} p_i = p_\infty$; dann ist

$$\lim_{i\to\infty}\frac{p_i-q}{p_i+q+1} = \frac{p_\infty-q}{p_\infty+q+1}$$

und, wegen $q\geqq 0$, $p_\infty > -1/2$,

$$\left|\frac{p_\infty-q}{p_\infty+q+1}\right| < 1,$$

woraus (31) folgt.

2° Ist $-1/2$ Häufungspunkt von $\{p\}$, so schreibe man

$$\prod_{1}^{n}\left(\frac{p_i-q}{p_i+q+1}\right)^2 = \frac{\prod\limits_{1}^{n}\left(1-\frac{p_i+1/2}{q+1/2}\right)^2}{\prod\limits_{1}^{n}\left(1+\frac{p_i+1/2}{q+1/2}\right)^2} \qquad (\lim_{i\to\infty} p_i = -1/2) \tag{32}$$

und unterscheide zwei Fälle:

$\alpha)$ $\sum\limits_{i=1}^{\infty}|p_i+1/2| < \infty$, $\qquad\beta)$ $\sum\limits_{i=1}^{\infty}|p_i+1/2| = \infty$.

Im Falle $\alpha)$ konvergiert das Zählerprodukt (32) gegen Null dann und nur dann, wenn ein Faktor verschwindet, d. h. wenn $q=p_i$ für ein i ist. Ist aber q von allen p_i verschieden, so ist der Grenzwert des Zählerproduktes positiv; da

$$\prod_{1}^{\infty}\left(1+\frac{p_i+1/2}{q+1/2}\right)^2 < \infty,$$

so ist in diesem Falle für gewisse q und die betrachtete Folge $\{p_i\}$ das Kriterium (31) nicht erfüllt.

Im Falle β) ist, wegen $\lim\limits_{i\to\infty} p_i = -1/2$, $p_i > -1/2$, von einem i an

$$0 < \frac{p_i + 1/2}{q + 1/2} < 1,$$

also das Zählerprodukt (wegen β)) gegen Null konvergent; das Nennerprodukt ist divergent gegen ∞, also das Kriterium (31) erfüllt.

3° Ist $\lim p_i = +\infty$ und $q \neq p_i$ (i), so schreibt man

$$\prod_1^n \left(\frac{p_i - q}{p_i + q + 1}\right)^2 = \left(\frac{\prod\limits_1^n \left(1 - \dfrac{q}{p_i}\right)}{\prod\limits_1^n \left(1 + \dfrac{q+1}{p_i}\right)}\right)^2 \tag{33}$$

und unterscheidet die Fälle

α) $\sum\limits_{\substack{i=1\\ p_i \neq 0}}^{\infty} \dfrac{1}{p_i} < +\infty$, β) $\sum\limits_{\substack{i=1\\ p_i \neq 0}}^{\infty} \dfrac{1}{p_i} = +\infty$.

Im Falle α) konvergiert das Zählerprodukt rechts in (33) gegen eine endliche von Null verschiedene Zahl und dasselbe gilt vom Nenner. Im Falle β) konvergiert der Zähler gegen Null, der Nenner divergiert gegen ∞. Es ist also das Kriterium (31) im Falle α) nicht erfüllt, wohl aber im Falle β).

Das Gesagte ist im folgenden Satz enthalten:

[361] *Damit das System $\{t^p\}$ $(0 \leqslant t \leqslant 1)$ inbezug auf L^2 abgeschlossen sei, ist es notwendig und hinreichend, daß die Exponentenmenge $\{p\}$ eine Folge $\{p_i\}$ mit einem endlichen Grenzwert $p_\infty > -1/2$ enthalte, oder eine Folge $\{p_i\}$ mit*

$$\lim_{i\to\infty} p_i = -1/2, \qquad \sum_{i=1}^{\infty} |p_i + 1/2| = \infty,$$

oder eine Folge $\{p_i\}$ mit

$$p_i \neq 0, \qquad \lim_{i\to\infty} p_i = +\infty, \qquad \sum_{i=1}^{\infty} \frac{1}{p_i} = +\infty.$$

Um die Notwendigkeit einzusehen, beachte man, daß im Falle, wo $\{p\}$ keine Folge $\{p_i\}$ einer der drei im Satze genannten Arten enthält, $\{p\}$ nur endlich sein, oder aus einer Folge $\{p_i\}$ mit

$\lim\limits_{i\to\infty} p_i = -1/2$ und $\sum\limits_{i=1} |p_i + 1/2| < \infty$, oder aus einer Folge $\{p_i'\}$ mit $\lim\limits_{i\to\infty} p_i' = +\infty$ und $\sum\limits_{i=1}^{\infty} 1/p_i' < +\infty$ $(p_i' \neq 0)$, oder — schließlich — aus zwei solchen elementfremden Folgen $\{p_i\}$ $\{p_i'\}$ bestehen kann. Ist dann q von allen p_i, p_i' verschieden und ist

$$\prod_1^{\infty} \left(\frac{p_i - q}{p_i + q + 1}\right)^2 = a, \qquad \prod_1^{\infty} \left(\frac{p_i' - q}{p_i' + q + 1}\right)^2 = b,$$

so ist stets $\|t^q - w_n(t)\| \geqslant \sqrt{a}/\sqrt{2q+1}$ bzw. $\geqslant \sqrt{b}/\sqrt{2q+1}$ bzw. $\geqslant \sqrt{ab}/\sqrt{2q+1}$, wenn $\{p\}$ unendlich ist; ist $\{p\} = (p_1, p_2, \ldots, p_n)$, so ist

$$\|t^q - w_n(t)\| \geqslant \frac{1}{\sqrt{2q+1}} \prod_1^n \left|\frac{p_i - q}{p_i + q + 1}\right|,$$

also hat in allen Fällen die Annäherung eine untere positive, effektiv bestimmte Grenze.

Beispiele. Ist $\{p_i\}$ die Folge aller (positiven) Primzahlen, so ist $\lim\limits_{i\to\infty} p_i = +\infty$, $\sum\limits_{i=1}^{\infty} 1/p_i = +\infty$; das System $\{t^{p_i}\}$ ist also abgeschlossen. Das System $\{t^{\frac{n}{n+1}}\}$ $(n = 1, 2, \ldots)$ ist abgeschlossen, denn $\lim\limits_{n\to\infty} \dfrac{n}{n+1} = 1$. Ebenso ist $\{\sqrt[n]{t}\}$ ein abgeschlossenes System. Diese beiden Systeme sind dicht-abgeschlossen im Sinne von [247], wie man sofort einsieht. Dagegen ist $\{t^{-1/2+1/n}\}$ abgeschlossen, aber nicht dicht-abgeschlossen, denn $\{t^{-1/2+1/n^2}\}$ ist überhaupt nicht mehr abgeschlossen.

[354] mit [361] geben ein Kriterium für die Vollständigkeit von $\{t^p\}$ inbezug auf L^2. Es kann also der Lerch'sche Satz in einer anderen Richtung als [358] verschärft werden:

Ist $f(t) \in L^2$ und [362]

$$\int_0^1 f(t) \,.\, t^p \,.\, dt = 0 \qquad (p),$$

wo die Exponentenmenge $\{p\}$ dem Kriterium [361] *genügt, so ist $f(t) \equiv 0$* f. ü. (z. B. kann p alle Primzahlen durchlaufen).

Auch dieser Satz kann teilweise auf L ausgedehnt werden:

Ist $f(t) \in L$ *und*

$$\int_0^1 f(t)\, t^{p_i}\, dt = 0 \qquad (i),$$

$$(34) \qquad p_i > 0, \qquad \lim_{i\to\infty} p_i = +\infty, \qquad \sum_{i=1}^{\infty} \frac{1}{p_i} = +\infty,$$

so ist $f(t) \equiv 0$ *f. ü.*

Man setze $F(t) = \int_t^1 f(u)\, du$; es folgt

$$\int_0^1 f(t)\, t^{p_i}\, dt = p_i \int_0^1 F(t)\, t^{p_i - 1}\, dt = 0 \qquad (i).$$

Nun ist aber $F(t) \in L^2$, $\lim_{i\to\infty}(p_i - 1) = +\infty$, $\sum_{i=1}^{\infty}{}' \frac{1}{p_i - 1} = +\infty$ (wo $\sum'$ bedeutet, daß $p_i = 1$ gestrichen wurde), also ist nach [362] $F(t) \equiv 0$ und $f(t) \equiv 0$ f. ü.

[363] *Es hat sich also das System* $\{t^{p_i}\}$ *bei den Voraussetzungen* (34) *als vollständig inbezug auf* L *herausgestellt*. Dieses System ist inbezug auf L abgeschlossen, da es inbezug auf L^2 diese Eigenschaft besitzt (vgl. [356] Anw.). Nun aber hat (34) zur Folge: $p_i - 1 > 0$ (von einem i_1 an), $\lim_{i\to\infty}(p_i - 1) = +\infty$, $\sum_{i=i_1}^{\infty} 1/p_i = \infty$, woraus für $\{t^{p_i - 1}\}$ die Abgeschlossenheit inbezug auf L resultiert.

[364] Zieht man jetzt [245] heran, so folgt: *Unter den Voraussetzungen*

$$\lim_{i\to\infty} p_i = +\infty, \qquad \sum_{i=1}{}' \frac{1}{p_i} = +\infty \qquad (p_i \neq 0)$$

ist $\{t^{p_i}\}$ *zusammen mit der Einheit ein inbezug auf* C *abgeschlossenes System*. (Ist die Bedingung $\sum_{i=1}^{\infty}{}' 1/p_i = +\infty$ nicht erfüllt, so ist das System nicht einmal inbezug auf L^2, also sicher nicht inbezug auf C abgeschlossen. Die Einheit ist auch notwendig; sonst könnte man keine in $t = 0$ von Null verschiedene Funktion gleichmäßig annähern, da alle t^{p_i} $(p_i > 0)$ für $t = 0$ verschwinden).

[364] ist die Müntz'sche Verschärfung des Weierstraß'schen Satzes [122]: Für Folgen $\{p_i\}$ mit $\lim_{i\to\infty} p_i = +\infty$ erweist sich die Divergenz von $\sum' 1/p_i$ zusammen mit $0 \in \{p_i\}$ als notwendig und hinreichend, um im Weierstraß'schen Satz mit den t^{p_i}-Potenzen allein auszukommen.

§ 7. Die Parseval'sche Gleichung.

Wir haben bereits im II-ten Kapitel die sogenannte Bessel'sche Ungleichung [264] abgeleitet:

$$\sum_{i=1}^{\infty} c_i^2 \leq \int_a^b f^2(t)\, dt.$$

Hierin sind c_i die Koeffizienten der Entwicklung von $f(t)$ nach einem ON $\{\varphi_i\}$. Im Falle eines ONA wird daraus die P a r s e v a l'- [371]
s c h e G l e i c h u n g

$$(35) \qquad \sum_{i=1}^{\infty} c_i^2 = \int_a^b f^2(t)\, dt.$$

B e w e i s: Ist $s_n(t)$ die Partialsumme $\sum_{i=1}^{n} c_i \varphi_i(t)$, so ist (II, 6, (30))

$$(36) \qquad \int_a^b [f(t) - s_n(t)]^2\, dt = \int_a^b f^2(t)\, dt - \sum_{i=1}^{n} c_i^2.$$

Nach [267] konvergiert für $n \to \infty$ das linkerhand in (36) stehende Integral gegen Null, wenn $\{\varphi_i\}$ ein ONA ist; daraus folgt aber (35).

Für das trigonometrische System wurde (35) allgemein (d. h. für $f \in L^2$) von P. Fatou bewiesen, bevor noch der Riesz-Fischer'sche Satz bekannt war.

Die Gleichung (35) führt zu

$$(37) \qquad \int_a^b f(t)\, g(t)\, dt = \sum_{n=1}^{\infty} c_n d_n \qquad [372]$$

(mit absoluter Konvergenz der rechtsstehenden Reihe), wenn $\{\varphi_i\}$ ein ONA ist und (II, 5, (25))

$$f(t) \sim \sum_{n=1}^{\infty} c_n \varphi_n(t), \qquad g(t) \sim \sum_{n=1}^{\infty} d_n \varphi_n(t).$$

Denn es gilt $f(t) \pm g(t) \sim \sum_{n=1}^{\infty} (c_n \pm d_n)\, \varphi_n(t)$, also nach (35)

$$\int_a^b (f(t) \pm g(t))^2\, dt = \sum_{n=1}^{\infty} (c_n \pm d_n)^2,$$

woraus durch Substraktion die Formel (37) folgt.

Geometrisch gesprochen gibt (35) die Länge des Vektors $f(t)$, ausgedrückt durch seine Komponenten inbezug auf das rechtwinklige Koordinatensystem $\{\varphi_i\}$; die Vektoren φ_i sind hier die Einheitsvektoren, die das Koordinatensystem definieren. (37) ist die Formel für das Skalarprodukt zweier Vektoren f, g.

Setzt man in (37) $f(t)$ für $g(t)$, so folgt (35) aus (37). Andererseits zeigt (36), daß aus (35)

$$\lim_{n\to\infty} \int_a^b [f(t) - s_n(t)]^2\, dt = 0$$

folgt. Wenn wir noch [354] und [355] heranziehen, so folgt unmittelbar:

[373] *Damit ein* ON $\{\varphi_i(t)\}$ *inbezug auf* L^2 *vollständig sei, ist eine jede der drei Bedingungen* $1^0\ 2^0\ 3^0$ *notwendig und hinreichend:*

1^0 $\{\varphi_i(t)\}$ *ist abgeschlossen inbezug auf* L^2;

2^0 $\sum_{i=1}^{\infty} c_i^2 = \int_a^b f^2(t)\, dt$ *für alle* f *in* L^2;

3^0 $\sum_{i=1}^{\infty} c_i d_i = \int_a^b f(t)\, g(t)\, dt$ *für alle* f, g *in* L^2.

Die Bedingung 2^0 kann abgeschwächt werden: es genügt die Parseval'sche Gleichung für alle f eines vollständigen Systems $\{f\}$ vorauszusetzen. Die Notwendigkeit ist evident. Setzen wir voraus, 2^0 sei in der neuen Fassung richtig; $\{f\}$ ist vollständig, also abgeschlossen; ist $g \in L^2$, so gibt es zu $\varepsilon > 0$ eine Linearform $\sum_{i=1}^{n} \gamma_i f_i(t)$ mit

$$\left\| g(t) - \sum_1^n \gamma_i f_i(t) \right\| < \frac{\varepsilon}{2}.$$

Nun aber folgt aus 2^0 $\lim_{n\to\infty} \|f - s_n(t)\| = 0$ für alle f, also auch

$$\|\gamma_i f_i(t) - \gamma_i s_{n_i}^{(i)}\| < \frac{\varepsilon}{2n} \quad \text{für } n_i > N_i \quad (i = 1, 2, \ldots, n),$$

woraus

$$\left\| g(t) - \sum_{i=1}^{n} \gamma_i s_{n_i}^{(i)}(t) \right\| < \varepsilon \tag{38}$$

resultiert. Dabei sind die $s_m^{(i)}(t)$ Partialsummen der Entwicklungen

von $f_i(t)$ $(i = 1, 2, \ldots, n)$ nach $\{\varphi_k(t)\}$ und $\sum_{i=1}^{n} \gamma_i s_{n_i}^{(i)}(t)$ eine Linearform in endlichvielen $\varphi_k(t)$; (38) zeigt, daß $\{\varphi_k(t)\}$ abgeschlossen, also auch vollständig ist, w. z. b. w.

Eine Folgerung aus dem Vorhergehenden sind die Sätze:

Eine (notwendige und) hinreichende Bedingung, damit ein ON $\{\varphi_k(t)\}$ vollständig sei, ist die Beziehung

$$\int_a^b t^{2n}\, dt = \sum_{k=1}^{\infty} \left(\int_a^b t^n \varphi_k(t)\, dt \right)^2 \tag{n}$$

für alle $n (= 1, 2, \ldots)$. Eine ebensolche Bedingung ist die Beziehung

$$\beta - \alpha = \sum_{k=1}^{\infty} \left(\int_\alpha^\beta \varphi_k(t)\, dt \right)^2 \tag{α, β}$$

für alle α, β mit $a \leqq \alpha \leqq \beta \leqq b$. Hier ist als das System $\{f\}$ im ersten Falle $\{t^n\}$, im zweiten die Gesamtheit der Funktionen $f(t)$, die $= 1$ in $\langle \alpha, \beta \rangle$ und $= 0$ sonst in $\langle a, b \rangle$, gewählt worden.

§ 8. Der Riesz-Fischer'sche Satz.

Wir haben diesen Satz bereits in [356] ausgesprochen und bewiesen. Hier wollen wir einen anderen Beweis bringen. [381]

Es sei $\{\varphi_i(t)\}$ ein ON, $\sum_{i=1}^{\infty} c_i^2 < \infty$. Wir schreiben

$$\psi_i(t) = \int_a^t \varphi_i(u)\, du, \qquad F(t) = \sum_{i=1}^{\infty} c_i \psi_i(t)$$

und bemerken, daß die letztgeschriebene Reihe gleichmäßig konvergiert, wie aus dem Beweise von [269] sofort ersichtlich. Es ist also $F(t)$ vorhanden und stetig. Wir werden zeigen, daß $F(t)$ total-stetig ist. Schreibt man nämlich

$$S_n(t) = \sum_{i=1}^{n} c_i \psi_i(t), \qquad s_n(t) = \sum_{i=1}^{n} c_i \varphi_i(t),$$

so ist

$$S_n'(t) = s_n(t) \text{ f. ü.}, \qquad \lim_{n\to\infty} S_n(t) = F(t) \text{ gleichmäßig.} \tag{39}$$

Ist nun E eine t-Menge, $\gamma(t)$ ihre charakteristische Funktion, so ist

$$\int_E |S'_n(t)|\,dt = \int_E |s_n(t)|\,dt = \int_a^b \gamma(t)\,|s_n(t)|\,dt$$

$$\leqslant \sqrt{\int_a^b \gamma^2(t)\,dt}\,.\sqrt{\int_a^b s_n^2(t)\,dt} \leqslant \sqrt{|E|}\,.\sqrt{\sum_{i=1}^{\infty} c_i^2}\,.$$

Es ist also die Schwankung von $S_n(t)$ über einer Menge E zugleich mit $|E|$ beliebig klein, wie auch n gewählt werden mag. Damit ist aber die Totalstetigkeit von $F(t)$ bewiesen. Es hat also $F(t)$ fast überall eine Ableitung $f(t)$ und es ist $(S_n(a)=0,\ F(a)=0)$

(40) $$F(t) = \int_a^t f(u)\,du.$$

Jetzt zeigen wir, daß $\{s_n(t)\}$ gegen $f(t)$ schwach konvergiert. Dazu ist, nach [164] mit $p=2$, notwendig und hinreichend:

$$\int_a^b |s_k(t)|^2\,dt \leqslant \lambda \qquad (\lambda \text{ frei von } k),$$

$$\lim_{k\to\infty} \int_a^t s_k(u)\,du = \int_a^t f(u)\,du \qquad (a \leqslant t \leqslant b).$$

Die erste Bedingung ist mit $\lambda = \sum_{i=1}^{\infty} c_i^2$ erfüllt, die zweite mit (39) und (40) gleichbedeutend. Die Eigenschaft (23) des § 6 des I-ten Kapitels, die einer schwachkonvergenten Folge zukommt $(p=p'=2)$, gibt

(41) $$c_i = \int_a^b s_n(t)\,\varphi_i(t)\,dt \qquad (n > i),$$

$$\lim_{n\to\infty} \int_a^b s_n(t)\,\varphi_i(t)\,dt = \int_a^b f(t)\,\varphi_i(t)\,dt.$$

Da die schwache Grenze einer Folge in L^2 auch zu L^2 gehört, zeigt (41), daß wir eine Funktion $f(t)$ in L^2 gefunden haben, deren Entwicklung die vorgeschriebenen Koeffizienten $\{c_i\}$ aufweist. Das ist aber der Inhalt des Riesz-Fischer'schen Satzes.

Dieser Beweis erlaubt die Funktion $f(t)$ auf dem Umweg über die gleichmäßig konvergente Reihe für $F(t)$ wirklich zu bestimmen. Aber auch der erste Beweis entbehrt nicht dieses Vorzugs. Er stützt sich auf die starke Konvergenz der Folge $\{s_n(t)\}$

und auf das Cauchy'sche Postulat, d. h. auf die Eigenschaft von L^2, die in [131] ausgesprochen und bewiesen wurde. Wegen

$$\| s_m - s_n \| = \left(\sum_{i=n+1}^{m} c_i^2\right)^{\frac{1}{2}} \qquad (m > n)$$

sind die zur Konstruktion einer im gewöhnlichen Sinne (f. ü.) konvergenten Teilfolge $\{s_{n_k}\}$ brauchbaren Indizes $\{n_k\}$ (die wir in [131] als vorhanden erwiesen) jetzt aus der Reihe $\sum_{i=1}^{\infty} c_i^2$ zu finden:

Ist nämlich $\sum_{i=1}^{\infty} \varepsilon_i$ eine konvergente Reihe mit positiven Gliedern und [382]

$$r_n = \sum_{n+1}^{\infty} c_i^2,$$

so bestimme man $n_1, n_2, \dots$ aus $r_{n_1} < \varepsilon_1^2,\ \ r_{n_2} < \varepsilon_2^2,\ \ \dots,\ \ r_{n_k} < \varepsilon_k^2,\ \ \dots$; es kommt

$$\| s_{n_2} - s_{n_1} \| < \varepsilon_1, \qquad \| s_{n_3} - s_{n_2} \| < \varepsilon_2, \qquad \dots, \qquad \| s_{n_{k+1}} - s_{n_k} \| < \varepsilon_k, \qquad \dots$$

Nach [131] (wo $p=2$ und s statt x gesetzt werde) wird dann $\lim_{k\to\infty} s_{n_k}(t) = f(t)$ f. ü. und die gewöhnliche Grenze der Teilfolge ist mit der starken Grenze der ganzen Folge identisch.

Die Funktion $f(t)$, deren Existenz der Riesz-Fischer'sche Satz verbürgt und die in [381] und [382] eindeutig und übereinstimmend berechnet wurde, ist selbstverständlich die einzige Lösung des Problems, aus der Entwicklung die Funktion zu bestimmen, wenn das System $\{\varphi_i(t)\}$ vollständig ist. Ist es unvollständig, so gibt es nichtverschwindende Funktionen $g(t)$, deren Entwicklung nach $\{\varphi_i(t)\}$ lauter Nullen zu Koeffizienten hat. Dann hat aber $f(t)+g(t)$ dieselbe Entwicklung wie $f(t)$ und die Summe $f(t)+g(t)$ — wo $g(t)$ alle oben bezeichneten Funktionen durchläuft — liefert auch alle Funktionen mit der gegebenen Entwicklung. Unter ihnen zeichnet sich das in [381] und [382] gefundene $f(t)$ durch die kleinste Norm aus. Ist nämlich $g(t)$ orthogonal zu allen $\varphi_i(t)$, aber nicht f. ü. Null, so ist

$$\int_a^b s_n(t)\,g(t)\,dt = 0 \qquad (n)$$

und, da $f(t)$ als die schwache Grenze von $s_n(t)$ erkannt wurde,

$$\int_a^b f(t)\, g(t)\, dt = 0,$$

also

$$\int_a^b [f(t)+g(t)]^2\, dt = \int_a^b f^2(t)\, dt + \int_a^b g^2(t)\, dt > \int_a^b f^2(t)\, dt.$$

Wir haben in der Unizität des Riesz-Fischer'schen Problems eine Eigenschaft, die mit der Vollständigkeit, also auch mit den drei Bedingungen in [373] einzeln äquivalent ist.

Ein Beispiel. In

$$(42) \qquad \sum_{i=1}^{\infty} \frac{\varphi_i(t)}{i}$$

ist $c_i = 1/i$. Setzt man $\varepsilon_i = 1/i^2$, so hat man $n_1, n_2, \ldots$, aus $r_{n_1} < 1/1^4$, $r_{n_2} < 1/2^4$, ... zu bestimmen, wo $r_{n_k} = \sum\limits_{n_k+1}^{\infty} 1/i^2 < 1/n_k$. Es genügt also $1/n_k \leqslant 1/k^4$, d. h. $n_k \geqslant k^4$ zu verlangen. Es sind also z. B. die Partialsummen $s_{k^4}(t)$ einer Reihe wie (42) stets f. ü. konvergent gegen ein $f(t)$, dessen Entwicklung sie ist.

[383] *Der Riesz-Fischer'sche Satz schafft eine eineindeutige Beziehung zwischen allen Funktionen aus L^2 und allen Zahlenfolgen von konvergenter Quadratsumme durch Vermittlung eines* ONV.

Es seien $\{\varphi_i\}$, $\{\psi_i\}$ zwei ONV; man setze

$$e_{ik} = \int_a^b \varphi_i(t)\, \psi_k(t)\, dt;$$

die Parseval'schen Formeln [371], [372] liefern

$$(43) \qquad \sum_{k=1}^{\infty} e_{ik}^2 = \int_a^b \varphi_i^2(t)\, dt = 1, \qquad \sum_{\substack{k=1\\ i\neq j}}^{\infty} e_{ik}\, e_{jk} = \int_a^b \varphi_i(t)\, \varphi_j(t)\, dt = 0,$$

und die Matrix

$$E = \| e_{ik} \|$$

erweist sich als zeilenweise normiert-orthogonal; sie ist aber auch zeilenweise vollständig: wäre sie es nicht (Df. [324]), und wäre $\sum\limits_{k=1}^{\infty} e_{0k}^2 = 1$, $\sum\limits_{k=1}^{\infty} e_{0k}\, e_{ik} = 0$ $(i=1, 2, \ldots)$, so würde nach dem Riesz-

Fischer'schen Satze eine Funktion $\varphi_0(t)$ vorhanden sein, deren Entwicklung nach dem ONV $\{\psi_k(t)\}$ $\{e_{0k}\}$ zu Koeffizienten hätte; die Parseval'schen Formeln gäben dann

$$(44) \qquad \begin{aligned} &\int_a^b \varphi_0^2(t)\, dt = \sum_{k=1}^{\infty} e_{0k}^2 = 1, \\ &\int_a^b \varphi_0(t)\, \varphi_i(t)\, dt = \sum_{k=1}^{\infty} e_{0k}\, e_{ik} = 0 \qquad (i=1, 2, \ldots), \end{aligned}$$

gegen die Voraussetzung der Vollständigkeit von $\{\varphi_i\}$. Dieselbe Überlegung zeigt, daß E auch spaltenweise ein ONV ist.

Wir können aber die Matrix E als gegeben annehmen; ist sie zeilenweise normiert-orthogonal und ist $\{\psi_k(t)\}$ ein ONV, so kann jede Zeile als Koeffizientenfolge einer (eindeutig bestimmten) Funktion $\varphi_i(t)$ nach $\{\psi_k(t)\}$ betrachtet werden; die Formeln (43) (nach Umstellung der zwei ersten Glieder) beweisen die Orthogonalität und Normiertheit von $\{\varphi_i(t)\}$; ebenso besagen die Formeln (44), daß $\{\varphi_i(t)\}$ im Falle der Zeilenvollständigkeit von E vollständig ist. Daraus folgt:

Damit die Matrix $\| e_{ik} \|$ *als Koeffizientenschema eines* ON $\{\varphi_i\}$ *nach einem* ONV $\{\psi_k\}$ *betrachtet werden könne, ist die Orthogonalität und Normiertheit ihrer Zeilen notwendig und hinreichend. Damit dann noch* $\{\varphi_i(t)\}$ *vollständig sei, ist die Zeilenvollständigkeit der Matrix notwendig und hinreichend.* [384]

In Verbindung mit dem Riesz-Fischer'schen Satz steht die Frage, ob die Reihenfolge in einem ONV wesentlich ist. M. a. W.: Wenn $\{\varphi_n(t)\}$ ein ONV ist, $\{\varphi_{n_k}(t)\}$ eine Umordnung von $\{\varphi_n(t)\}$, wenn $\sum\limits_{n=1}^{\infty} c_n^2 < \infty$ derselben Umordnung unterworfen wird, ob dann die beiden Reihen [385]

$$\sum_{n=1}^{\infty} c_n\, \varphi_n(t), \qquad \sum_{k=1}^{\infty} c_{n_k}\, \varphi_{n_k}(t)$$

Entwicklungen derselben Funktion sind?

Die bejahende Antwort ist leicht aus den Ausführungen [356] abzulesen. Es ist nämlich bloß

$$\operatorname*{Lim}_{m\to\infty} \sum_{i=1}^{m} c_i\, \varphi_i(t) = \operatorname*{Lim}_{m\to\infty} \sum_{i=1}^{m} c_{n_i}\, \varphi_{n_i}(t)$$

zu beweisen, denn die beiden starken Limites geben die gesuchten Funktionen. Nun ist aber — für genügend große m —

$$\left\| \sum_{i=1}^{m} c_i \varphi_i(t) - \sum_{i=1}^{m} c_{n_i} \varphi_{n_i}(t) \right\| \leq \sqrt{\sum_{k=i(m)}^{\infty} c_k^2} < \varepsilon \qquad (\varepsilon > 0),$$

wenn $i(m)$ den kleinsten Index bedeutet, für welchen $c_i \varphi_i$ in der Differenz nicht aufgehoben wird, denn $i(m)$ wächst mit m unbegrenzt.

[386] Wir haben auch ein Mittel gewonnen, nichttriviale lineare Operationen zu definieren; ist $\{\lambda_i\}$ irgend eine beschränkte Zahlenfolge mit beschränkten Reziproken, sind $\{\varphi_i\}$, $\{\psi_i\}$ zwei ONV und setzt man

$$x(\tau) \sim \sum_{i=1}^{\infty} c_i \varphi_i(\tau), \qquad y(\tau) \sim \sum_{i=1}^{\infty} \lambda_i c_i \psi_i(\tau),$$

so hat man eine Operation $y = U(x)$ definiert, indem man x nach $\{\varphi\}$ entwickelt, formell die ψ-Reihe bildet, und kraft des Riesz-Fischer'schen Satzes y bestimmt. Die Linearität ist klar und die Parseval'sche Gleichung gibt für $y_1 = U(x_1)$, $y_2 = U(x_2)$, $|\lambda_i| < \lambda$, $\|y_1 - y_2\| \leqslant \lambda \|x_1 - x_2\|$, d. h. die Stetigkeit. Die Voraussetzung $|\lambda_i| < \lambda$ wurde hier benutzt, um aus $\sum_{i=1}^{\infty} c_i^2 < \infty$ auf $\sum_{i=1}^{\infty} \lambda_i^2 c_i^2 < \infty$ zu schließen; die Voraussetzung $|1/\lambda_i| < \mu$ genügt zur Umkehrung dieses Schlusses, so daß man eine eindeutige Operation $x = U_{-1}(y)$ erhält mit $\|x_1 - x_2\| \leqslant \mu \|y_1 - y_2\|$, also eine lineare Umkehrung von $y = U(x)$ in ganz L^2.

Das wichtigste an dem Riesz-Fischer'schen Satze ist die eineindeutige und stetige, lineare Zuordnung von L^2 und l^2. Diese Zuordnung erlaubt Probleme, die man in der Sprache der Funktionen formuliert hat, in solche der Analysis der unendlichvielen Variablen zu übersetzen. Hat man ein solches Problem einmal in der neuen Gestalt gelöst, so gestattet wieder der Riesz-Fischer'sche Satz die Antwort in der Originalsprache auszudrücken.

§ 9. Unendliche Intervalle.

Mit Ausnahme des Kostitzin'schen Satzes bezog sich alles bisher Gesagte auf endliche Intervalle. Unendliche Intervalle $(a, +\infty)$, $(-\infty, b)$ oder $(-\infty, +\infty)$ können ebenfalls zur Konstruktion von ON-Systemen dienen; der Kürze halber betrachten wir bloß $(-\infty, +\infty)$.

Es soll nun $f(t) \in L^p$ $(p \geqslant 1)$ soviel als $\lim\limits_{n\to\infty} \int_{-n}^{+n} |f(t)|^p \, dt < \infty$ bedeuten. Infolgedessen ist es nicht mehr wahr, daß L^p in L^q für $p \geqslant q$ enthalten sein muß. Ist z. B. $f(t) = 1/(n+1)$ in $\langle n, n+1)$ $(n = 0, 1, \ldots)$, $f(-t) = f(t)$, so ist $f(t) \in L^2$, aber nicht $f(t) \in L^1$. Es gibt auch Funktionen, die zu L^2 aber zu keinem L^p mit $p \neq 2$ gehören. Ist nämlich

$$a_k > 0, \qquad b_k > 0 \qquad (k),$$

$$\sum_{k=1}^{\infty} a_k^2 < \infty, \qquad \sum_{k=1}^{\infty} a_k^{2-\varepsilon} = \infty,$$

$$\sum_{k=1}^{\infty} \frac{b_k^2}{k^2} < \infty, \qquad \sum_{k=1}^{\infty} \frac{b_k^{2+\varepsilon}}{k^2} = \infty$$

für alle $\varepsilon > 0$ und setzt man

$$f(t) = \begin{cases} a_n & \text{in } \langle n, n+1-1/n^2) \\ b_n & \text{in } \langle n+1-1/n^2, n+1) \end{cases} \qquad (n = 1, 2, \ldots),$$

sonst aber $f(t) = 0$, so gehört $f(t)$ zu L^2 und zu keinem L^p mit $p \neq 2$.

Es sei $f(t) \in L^p$ und [391]

$$f_n(t) = \begin{cases} f(t) & \text{für } |t| \leqslant n \\ 0 & \text{für } |t| > n; \end{cases}$$

man bestätigt mühelos, daß

$$\lim_{n\to\infty} \int_{-\infty}^{+\infty} |f(t) - f_n(t)|^p \, dt = 0,$$

wie im endlichen Intervall. Die Hölder'sche und die Minkowski'sche Ungleichung gilt auch im unendlichen Intervall und die Definitionen der Orthogonalität, der Vollständigkeit und Abgeschlossenheit lassen sich infolgedessen ohne weiteres übertragen; dasselbe bezieht sich auf die Schmidt'sche Orthogonalisierung. Die Sätze des II-ten und III-ten Kapitels bleiben demnach richtig, insbesondere die Bessel'sche Ungleichung, die Parseval'sche Gleichung, der Riesz-Fischer'sche Satz und die Äquivalenz der L^2-Vollständigkeit und L^2-Abgeschlossenheit.

Anders verhält es sich mit den Sätzen von Weierstraß, Lerch und Müntz ([122], [357], [358], [361], [362]); wir werden zeigen, daß

[392] sie nicht mehr richtig bleiben. Es genügt dazu eine Funktion $f(t)$ anzugeben, die zu allen $L^p(p>0)$ gehört und die Eigenschaft

$$\int_{-\infty}^{+\infty} f(t)\, t^n\, dt = 0 \qquad (n = 0, 1, 2, \dots)$$

aufweist ohne f. ü. Null zu sein.

$$f(t) = \begin{cases} t^{-\log t} \sin(2\pi \log t) & \text{für } t > 0 \\ 0 & \text{für } t < 0 \end{cases}$$

ist eine solche Funktion. Es ist nämlich

$$\int_{-\infty}^{+\infty} |f(t)|^p\, dt < \int_0^{\infty} t^{-p\log t}\, dt < \infty,$$

da der Exponent $-p\log t$ für $t \to +0$ gegen $+\infty$, für $t \to +\infty$ gegen $-\infty$ strebt. Berechnen wir jetzt

$$I_n = \int_0^{\infty} t^{-\log t} \sin(2\pi \log t)\,.\, t^n\,.\, dt;$$

die Substitution $u = -\frac{n+1}{2} + \log t$ liefert:

$$I_n = \int_{-\infty}^{+\infty} e^{-\left(u+\frac{n+1}{2}\right)^2}.\sin 2\pi\left(u+\frac{n+1}{2}\right).e^{\left(u+\frac{n+1}{2}\right)n}.\,e^{u+\frac{n+1}{2}}.\,du$$

$$= \pm e^{\left(\frac{n+1}{2}\right)^2} \int_{-\infty}^{+\infty} e^{-u^2} \sin 2\pi u\, du = 0,$$

denn $\sin 2\pi u$ ist ungerade.

[393] *Nichtdestoweniger gibt es abzählbare Systeme in* $(-\infty, +\infty)$, *die inbezug auf* $L^p(p \geqslant 1)$ *vollständig sind.* Ist $\{I_n\}$ die Folge aller Intervalle mit rationalen Endpunkten und $f_n(t)$ die charakteristische Funktion von I_n, so liefert die Voraussetzung $g(t) \in L^p(p \geqslant 1)$ in $(-\infty, +\infty)$: $g(t) \in L^p$ in I_n. Ist dann für alle $n\ (= 1, 2, \dots)$

$$\int_{-\infty}^{+\infty} g(t)\, f_n(t)\, dt = 0,$$

so folgt $\int_{I_n} g(t)\, dt = 0$ und $\int_I g(t)\, dt = 0$, wo I ein beliebiges endliches Intervall bedeutet. Infolgedessen ist

$$G(t) = \int_a^t g(u)\, du = 0, \qquad g(t) = 0 \ \text{ f. ü.}$$

Es ist also $\{f_n(t)\}$ vollständig in $(-\infty, +\infty)$ inbezug auf L^p für alle $p \geqslant 1$.