

Einige Sätze über Reihen von allgemeinen Orthogonalfunktionen.

Von

Hans Rademacher in Hamburg.

Einleitung.

Der Riesz-Fischersche Satz, der besagt, daß es zu jeder Folge von reellen Zahlen $c_1, c_2, c_3, \dots, c_n, \dots$ mit konvergenter Quadratsumme eine höchstens bis auf eine Nullmenge bestimmte Funktion $f(x)$ gibt, deren Fourierkoeffizienten in bezug auf ein vorgegebenes System von Orthogonalfunktionen $\varphi_1(x), \varphi_2(x), \dots, \varphi_n(x), \dots$ die Zahlen c_n sind, hat eine Reihe von Versuchen hervorgebracht, jene Funktion $f(x)$ in rechnerisch einfache Beziehung zu den c_n und den $\varphi_n(x)$ zu setzen. Das allgemeinste Ergebnis in dieser Richtung ist wohl der Satz von Weyl¹⁾, nach dem man aus der Reihe

$$(1) \quad c_1 \varphi_1(x) + c_2 \varphi_2(x) + \dots$$

stets eine fast überall konvergente Folge von Teilsummen auswählen kann, deren Limesfunktion, wie man leicht sieht, die vom Riesz-Fischerschen Satz behaupteten Eigenschaften besitzt. Dieser Weylsche Satz soll im folgenden dahin präzisiert werden, daß zu vorgelegter konvergenter Reihe $\sum_{n=1}^{\infty} c_n^2$ die Indizes der Teilsummen einer konvergenten Folge explizite angegeben werden.

Über die Konvergenz der Reihe $\sum c_n \varphi_n(x)$ selbst hat man bisher nur Aussagen machen können, indem man über die c_n noch mehr als nur die Konvergenz von $\sum c_n^2$ voraussetzte. Jerosch²⁾ und Weyl³⁾ haben hierin den

¹⁾ Über die Konvergenz von Reihen, die nach Orthogonalfunktionen fortschreiten; Math. Ann. 67 (1909), S. 225–245, insbes. S. 243–245.

²⁾ F. Jerosch und H. Weyl, Über die Konvergenz von Reihen, die nach periodischen Funktionen fortschreiten, Math. Ann. 66 (1909), S. 67–80; vgl. auch die unter ¹⁾ zitierte Arbeit von Weyl, S. 229 oben.

³⁾ A. a. O. ¹⁾ S. 226 ff.

Anfang gemacht, jener, indem er unter der Voraussetzung $c_\nu = O\left(\nu^{-\frac{\sqrt{17}-1}{4}-\varepsilon}\right)$, dieser, indem er unter Voraussetzung der Konvergenz von $\sum c_\nu^2 \sqrt{\nu}$ bewies, daß die Reihe (1) fast überall konvergent ist. Der nächste Schritt wurde von Hobson getan⁴⁾, der zeigte, daß zu der Konvergenz bis auf eine Nullmenge (der „fast ausnahmslosen“ Konvergenz, wie wir im folgenden sagen werden) schon die Voraussetzung der Konvergenz von $\sum c_\nu^2 \nu^\alpha$ bei irgendeinem $\alpha > 0$ hinreicht. Endlich hat Plancherel bewiesen⁵⁾, daß für dieses Konvergenzverhalten von (1) schon die Konvergenz von $\sum c_\nu^2 (\log \nu)^3$ hinreicht. Aus einem Satze von Hardy⁶⁾, auf den Hobson⁷⁾ und neuerdings Neder⁸⁾ hingewiesen haben, folgt, daß die Reihe $\sum_{\nu=1}^{\infty} (a_\nu \cos \nu x + b_\nu \sin \nu x)$

schon dann fast überall konvergiert, wenn $\sum_{\nu=1}^{\infty} (a_\nu^2 + b_\nu^2) (\log \nu)^2$ konvergent ist. Die naheliegende Frage, ob sich in dem Plancherelschen Satze diesem Hardyschen Satze entsprechend für allgemeine Orthogonalfunktionen der Exponent 3 des Logarithmus durch 2 ersetzen läßt, wird im folgenden bejahend beantwortet. In diesem Zusammenhange sei endlich noch erwähnt, daß Herr Neder neuerdings durch Beispiele klargestellt hat, daß die Konvergenz von $\sum c_\nu^2 (\log \nu)^\alpha$ für irgendein $\alpha < 0$ gewiß nicht ausreicht, um fast überall die Konvergenz von (1) zu sichern⁹⁾.

Außer der schon erwähnten Ergänzung des Weylschen und der Verschärfung des Plancherelschen Satzes enthält die folgende Untersuchung noch asymptotische Abschätzungen über Summen von Orthogonalfunktionen. Mit Hilfe eines Satzes von I. Schur werden viele Eigenschaften von Orthogonalfunktionen auf eine weitere Klasse von Funktionen übertragen. Außerdem werden die verallgemeinerten „Lebesgueschen Konstanten“ auf ihr infinitäres Verhalten hin untersucht. Ein durchgeführtes Beispiel grenzt

⁴⁾ On the convergence of series of orthogonal functions. Proc. of the London Math. Soc. (2) 12 (1913), S. 297–308.

⁵⁾ Sur la convergence des séries de fonctions orthogonales. Comptes Rendus de l'Acad. des Sciences de Paris 157 (1913), S. 539–542.

⁶⁾ On the summability of Fourier's series, Proc. of the London Math. Soc. (2) 12 (1913), S. 365–372. Der erwähnte Satz lautet genau: „Ist $\frac{a_0}{2} + \sum_{\nu} a_\nu \cos \nu x + b_\nu \sin \nu x$ eine

Fourierreihe, so ist $\sum \frac{a_\nu \cos \nu x + b_\nu \sin \nu x}{\log \nu}$ fast überall konvergent.“ (Theorem 3, S. 370.) Die oben hervorgehobene Konsequenz erhält man in dem Spezialfall, wo die Fourierreihe zu einer samt ihrem Quadrate integrierbaren Funktion gehört.

⁷⁾ Proc. of the London Math. Soc. (2) 14, S. 428–439, insbes. S. 429.

⁸⁾ Math. Ann. 84 (1921) [S. 117–136], S. 131.

⁹⁾ A. a. O. ⁸⁾

dann den Spielraum ab, in dem man die erhaltenen Abschätzungen etwa noch verbessern könnte.

I. Fast überall konvergente Teilfolgen von $\sum c_\nu \varphi_\nu(x)$.

1. Es sei $\varphi_1(x), \varphi_2(x), \dots, \varphi_\nu(x), \dots$ ein System von samt ihrem Quadrate im Lebesgueschen Sinne integrierbaren Funktionen, die im Intervall $(0 \dots 1)$ orthogonal und normiert sind:

$$(2) \quad \int_0^1 \varphi_\mu(x) \varphi_\nu(x) dx = \begin{cases} 1, & \mu = \nu \\ 0, & \mu \neq \nu. \end{cases}$$

Ferner sei $c_1, c_2, \dots, c_\nu, \dots$ eine Folge von reellen Zahlen und $\sum_{\nu=1}^{\infty} c_\nu^2$ sei konvergent. Dann gibt es stets monoton ins Unendliche wachsende Folgen von positiven Zahlen $\lambda(\nu)$, so daß

$$(3) \quad \sum_{\nu=1}^{\infty} c_\nu^2 \lambda(\nu)$$

gleichfalls konvergent ausfällt. Ist $r_n = \sum_{\nu=n}^{\infty} c_\nu^2$, so kann man z. B. für irgendein $0 < \gamma < 1$ setzen

$$\lambda(\nu) = r_\nu^{-\gamma},$$

wofür man die Konvergenz von (3) leicht einsieht. Ferner sei A_ϱ das kleinste ν , für das $\lambda(\nu) \geq \varrho$ ausfällt, $\varrho = 1, 2, 3, \dots$. Dann ist $\lambda(A_\varrho) \geq \varrho$. Unter diesen Voraussetzungen und Bezeichnungen gilt nun der

Satz I: Die Folge von Teilsummen $\sum_{\nu=1}^{A_\varrho} c_\nu \varphi_\nu(x)$ ($\varrho = 1, 2, \dots$) ist fast überall konvergent.

Bemerkenswert ist an diesem Satz, daß er die Auswahl der Teilsummen nur aus der Folge $\{c_\nu\}$ bestimmt, unabhängig von dem zugrunde gelegten Orthogonalsystem.

Wir führen die Bezeichnungen ein

$$s(x; m, n) = \sum_{\nu=m+1}^n c_\nu \varphi_\nu(x), \quad \sigma(m, n) = \sum_{\nu=m+1}^n c_\nu^2, \quad \tau(m, n) = \sum_{\nu=m+1}^n c_\nu^2 \lambda(\nu).$$

Dann ergibt sich unter Benutzung von (2) für $0 < N < n$

$$\begin{aligned} \int_0^1 \sum_{r=N}^{n-1} \{s(x; A_r, A_n)\}^2 dx &= \sum_{r=N}^{n-1} \int_0^1 \{s(x; A_r, A_n)\}^2 dx = \sum_{r=N}^{n-1} \sigma(A_r, A_n) \\ &= \sum_{r=N}^{n-1} \sum_{\varrho=r}^{n-1} \sigma(A_\varrho, A_{\varrho+1}) = \sum_{\varrho=N}^{n-1} (\varrho - N + 1) \sigma(A_\varrho, A_{\varrho+1}) \leq \sum_{\varrho=N}^{n-1} \lambda(A_\varrho) \sigma(A_\varrho, A_{\varrho+1}) \\ &\leq \sum_{\varrho=N}^{n-1} \sum_{\nu=A_\varrho+1}^{A_{\varrho+1}} \lambda(\nu) c_\nu^2 = \sum_{\nu=A_N+1}^{A_n} c_\nu^2 \lambda(\nu) = \tau(A_N, A_n). \end{aligned}$$

Also ist der positive Integrand

$$\sum_{r=N}^{n-1} \{s(x; A_r, A_n)\}^2 < \tau(A_N, A_n)^{\frac{2}{3}}$$

in einer Menge $E_{N,n}$, deren Maß¹⁰⁾ größer oder gleich $1 - \tau(A_N, A_n)^{\frac{1}{3}}$. Erst recht ist also

$$m E_{N,n} \geq 1 - \tau(A_N, \infty)^{\frac{1}{3}}$$

und in $E_{N,n}$

$$|s(x; A_r, A_n)| < \tau(A_N, \infty)^{\frac{1}{3}},$$

für jedes r , das $N \leq r \leq n$ genügt. Die Menge $E'_{N,n}$, in der für alle ganzen r und r_1 unter der Bedingung $N \leq r \leq r_1 \leq n$

$$|s(x; A_r, A_{r_1})| < 2\tau(A_N, \infty)^{\frac{1}{3}}$$

stattfindet, enthält $E_{N,n}$ wegen

$$|s(x; A_r, A_{r_1})| \leq |s(x; A_r, A_n)| + |s(x; A_{r_1}, A_n)|.$$

Folglich ist erst recht

$$m E'_{N,n} \geq 1 - \tau(A_N, \infty)^{\frac{1}{3}}.$$

Nun ist

$$(4) \quad E'_{N,N+1} \supset E'_{N,N+2} \supset E'_{N,N+3} \supset \dots,$$

was aus der Definition dieser Mengen unmittelbar folgt. Wir bilden den Durchschnitt

$$D_N = E'_{N,N+1} \cdot E'_{N,N+2} \cdot E'_{N,N+3} \dots$$

Wegen (4) ist dann nach einem Theorem der Maßtheorie

$$(5) \quad m D_N = \lim_{k \rightarrow \infty} m E'_{N+N+k} \geq 1 - \tau(A_N, \infty)^{\frac{1}{3}}.$$

In D_N ist nun

$$(6) \quad |s(x; A_r, A_{r_1})| < 2\tau(A_N, \infty)^{\frac{1}{3}}$$

für alle r und r_1 , die nur $N \leq r \leq r_1$ erfüllen. Aus den D_N bilden wir

$$V_N = D_N + D_{N+1} + D_{N+2} + \dots;$$

wegen (5) ergibt sich

$$(7) \quad m V_N = 1.$$

Endlich sei

$$V = V_1 \cdot V_2 \cdot V_3 \dots;$$

aus

$$V_1 \supset V_2 \supset V_3 \supset \dots$$

¹⁰⁾ Das Maß einer Menge E wird durch $m E$ bezeichnet.

und aus (6) folgt dann auch

$$m V = 1.$$

Jeder Punkt von V gehört aber zu *unendlich vielen* D_N . In jedem Punkt x_0 von V gibt es daher zu beliebigem positiven ε ein N , so daß

$$|s(x_0; A_r, A_{r_1})| < \varepsilon$$

ist, für alle r und r_1 , die $N \leq r \leq r_1$ erfüllen. Wegen (6) braucht man nämlich ein solches N aus den unendlich vielen N , für die x_0 in D_N liegt, nur so groß zu wählen, daß $\tau(A_N, \infty) < \frac{\varepsilon^3}{8}$ ausfällt. Damit ist bewiesen, daß

$$(8) \quad f(x) = \lim_{n \rightarrow \infty} \sum_{\nu=1}^{A_n} c_\nu \varphi_\nu(x)$$

in jedem Punkte von V , also fast überall im Orthogonalitätsintervalle, als endliche Funktion existiert.

Diesen Satz hätte man übrigens auch erhalten können durch Weiterverfolgung der Gedanken, die Weyl a. a. O. für den Beweis seines Auswahl-satzes verwendet.

2. Unter leichter Modifikation einer schon von Weyl ausgeführten Überlegung¹¹⁾ können wir den in der Einleitung erwähnten Zusammenhang unseres Satzes I mit dem Riesz-Fischerschen Satze herstellen, indem wir diesen aus Satz I herleiten.

Zu jedem positiven ε gibt es nämlich offenbar eine Teilmenge A_ε von V und eine Zahl M_ε , so daß in A_ε

$$|f(x)| < M_\varepsilon$$

gilt und zugleich

$$m A_\varepsilon > 1 - \varepsilon$$

ist. Wegen der Konvergenz in (8) gibt es ferner in A_ε eine Teilmenge B_ε , mit

$$m B_\varepsilon > 1 - 2\varepsilon,$$

und einen Index n_ε , so daß für $n \geq n_\varepsilon$

$$\left| \sum_{\nu=1}^{A_n} c_\nu \varphi_\nu(x) \right| < 2 M_\varepsilon$$

bleibt. Aus der letzten Ungleichung kann man schließen, daß

$$\lim_{n \rightarrow \infty} \int_{B_\varepsilon} \left(\sum_{\nu=1}^{A_n} c_\nu \varphi_\nu(x) \right)^2 dx = \int_{B_\varepsilon} f^2(x) dx$$

¹¹⁾ A. a. O. ¹⁾ S. 226–227.

ist. Da nun

$$\int_{B_\varepsilon} \left(\sum_{v=1}^{A_n} c_v \varphi_v(x) \right)^2 dx \leq \int_0^1 \left(\sum_{v=1}^{A_n} c_v \varphi_v(x) \right)^2 dx \leq \sum_{v=1}^{\infty} c_v^2$$

ist, so gilt

$$\int_{B_\varepsilon} f^2(x) dx \leq \sum_{v=1}^{\infty} c_v^2.$$

Setzt man $f(x)$ außerhalb von V im Orthogonalitätsintervalle irgendwie willkürlich fest, so folgt, da die Vereinigungsmenge aller B_ε die Menge V ausfüllen muß,

$$\int_0^1 f^2(x) dx \leq \sum_{v=1}^{\infty} c_v^2.$$

Ebenso wie diese Ungleichung aus (8) folgt, so zieht die Relation

$$f(x) - \sum_{v=1}^{A_m} c_v \varphi_v(x) = \lim_{n \rightarrow \infty} \sum_{v=A_m+1}^{A_n} c_v \varphi_v(x)$$

die Ungleichung

$$\int_0^1 \left\{ f(x) - \sum_{v=1}^{A_m} c_v \varphi_v(x) \right\}^2 dx \leq \sum_{v=A_m+1}^{\infty} c_v^2$$

nach sich. Hieraus schließt man

$$\lim_{m \rightarrow \infty} \int_0^1 \left\{ f(x) - \sum_{v=1}^{A_m} c_v \varphi_v(x) \right\}^2 dx = 0,$$

also auch, zufolge der Schwarzschen Ungleichung, bei festem μ

$$\lim_{m \rightarrow \infty} \int_0^1 \left\{ f(x) - \sum_{v=1}^{A_m} c_v \varphi_v(x) \right\} \varphi_\mu(x) dx = 0.$$

Dies bedeutet aber

$$c_\mu = \int_0^1 f(x) \varphi_\mu(x) dx,$$

worin der Riesz-Fischersche Satz ausgesprochen ist.

II. Fast ausnahmslose Konvergenz von $\sum c_v \varphi_v(x)$.

3. Über die Folge der c_v sei jetzt noch weiter vorausgesetzt, daß $\Sigma (c_v \log v)^2$ konvergent ist. Im folgenden sei der Einfachheit halber unter „log“ der Logarithmus zur Basis 2 verstanden. Unter der neuen

Voraussetzung können wir das $\lambda(\nu)$ des Satzes I gleich $\log \nu$, also $A_0 = 2^0$, setzen und erhalten dann die Konvergenz von

$$(9) \quad \sum_{\nu=1}^{2^n} c_\nu \varphi_\nu(x) \quad n \rightarrow \infty$$

in einer maßgleichen Teilmenge V des Orthogonalitätsintervalls.

Die Bezeichnungen von § 1 mögen beibehalten werden; außerdem setzen wir

$$\omega(m, n) = \sum_{\nu=m+1}^n (c_\nu \log \nu)^2.$$

Es sei nun n eine ganze Zahl $2^r < n < 2^{r+1}$; sie läßt sich auf dyadische Art darstellen durch

$$n = 2^r + \vartheta_1 2^{r-1} + \vartheta_2 2^{r-2} + \dots + \vartheta_{r-1} 2 + \vartheta_r, \quad \vartheta_i = 0 \text{ oder } 1.$$

Hieraus ergibt sich die Zerlegung

$$(10) \quad \begin{aligned} s(x; 2^r, n) &= s(x; 2^r, 2^r + \vartheta_1 2^{r-1}) \\ &+ s(x; 2^r + \vartheta_1 2^{r-1}, 2^r + \vartheta_1 2^{r-1} + \vartheta_2 2^{r-2}) + \dots \\ &+ s(x; 2^r + \vartheta_1 2^{r-1} + \dots + \vartheta_{r-1} 2, 2^r + \vartheta_1 2^{r-1} + \dots + \vartheta_{r-1} 2 + \vartheta_r), \end{aligned}$$

worin diejenigen Klammern, deren letztes ϑ_i gleich 0 ist, selbst gleich 0 zu setzen sind. Die nicht verschwindenden Glieder der rechten Seite sind sämtlich von der Form

$$(11) \quad s(x; \alpha, \alpha') = s(x; 2^r + h \cdot 2^j, 2^r + h \cdot 2^j + 2^{j-1}), \quad \begin{matrix} j=1, 2, \dots, r, \\ h=0, 1, 2, \dots, (2^{r-j}-1), \end{matrix}$$

und aus diesen Ausdrücken lassen sich alle $s(x; 2^r, n)$ für $2^r < n < 2^{r+1}$ zusammenfügen.

Nun folgt aus (10) durch Anwendung der Schwarzschen Ungleichung:

$$\begin{aligned} \{s(x; 2^r, n)\}^2 &\leq \sum_{i=1}^r \vartheta_i \sum_{i=1}^r \{s(x; 2^r + \dots + \vartheta_{i-1} 2^{r-i+1}, 2^r + \dots + \vartheta_{i-1} 2^{r-i+1} + \vartheta_i 2^{r-i})\}^2 \\ &\leq r \sum_{i=1}^r \{s(x; 2^r + \dots + \vartheta_{i-1} 2^{r-i+1}, 2^r + \dots + \vartheta_{i-1} 2^{r-i+1} + \vartheta_i 2^{r-i})\}^2 \end{aligned}$$

und erst recht also

$$(12) \quad \{s(x; 2^r, n)\}^2 \leq r \sum_{\alpha, \alpha'} \{s(x; \alpha, \alpha')\}^2,$$

worin (α, α') alle in (11) angegebenen Zahlenpaare durchläuft. Für $2^r < n < 2^{r+1}$ hängt die rechte Seite der Ungleichung (12) nicht mehr von n ab.

Nun ist

$$\begin{aligned} \int_0^1 \sum_{\alpha, \alpha'} \{s(x; \alpha, \alpha')\}^2 dx &= \int_0^1 \sum_{j=1}^r \sum_{h=0}^{2^{r-j}-1} \{s(x; 2^r + h \cdot 2^j, 2^r + h \cdot 2^j + 2^{j-1})\}^2 dx \\ &= \sum_{j=1}^r \sum_{h=0}^{2^{r-j}-1} \sigma(2^r + h \cdot 2^j, 2^r + h \cdot 2^j + 2^{j-1}) \leq \sum_{j=1}^r \sigma(2^r, 2^{r+1}) = r \cdot \sigma(2^r, 2^{r+1}), \end{aligned}$$

und somit

$$(13) \quad \int_0^1 r \sum_{\alpha, \alpha'} \{s(x; \alpha, \alpha')\}^2 dx \leq r^2 \sigma(2^r, 2^{r+1}) \leq \omega(2^r, 2^{r+1}).$$

Es sei nun $G_r(\delta)$ diejenige Teilmenge des Orthogonalitätsintervalls, in der

$$r \sum_{\alpha, \alpha'} \{s(x; \alpha, \alpha')\}^2 < \delta^2,$$

worin δ eine positive Zahl sei, über die noch verfügt werden soll. Wegen (13) ist

$$(14) \quad m G_r(\delta) \geq 1 - \frac{1}{\delta^2} \omega(2^r, 2^{r+1}),$$

und wegen (12) gilt in $G_r(\delta)$ für alle n zwischen 2^r und 2^{r+1}

$$|s(x; 2^r, n)| < \delta.$$

Wir bilden nun den Durchschnitt

$$H_N(\delta) = G_N(\delta) \cdot G_{N+1}(\delta) \cdot G_{N+2}(\delta) \dots;$$

dann ist infolge von (14)

$$m H_N(\delta) \geq 1 - \frac{1}{\delta^2} \sum_{r=N}^{\infty} \omega(2^r, 2^{r+1}) = 1 - \frac{1}{\delta^2} \omega(2^N, \infty),$$

und für jedes x in $H_N(\delta)$ ist

$$|s(x; 2^r, n)| < \delta,$$

wenn nur $N \leq r$ und $2^r < n < 2^{r+1}$ gewonnen wird. Wählt man nun von vornherein $\delta = \omega(2^N, \infty)^{\frac{1}{2}}$, so führt diese Konstruktion auf eine Menge $H_N(\omega(2^N, \infty)^{\frac{1}{2}})$, die wir kurz mit H_N bezeichnen wollen. Von H_N weiß man

$$(15) \quad m H_N \geq 1 - \omega(2^N, \infty)^{\frac{1}{2}},$$

und in H_N gilt

$$(16) \quad |s(x; 2^r, n)| < \omega(2^N, \infty)^{\frac{1}{2}}.$$

für $N \leq r$, $2^r < n < 2^{r+1}$. Nun sei

$$U_N = H_N + H_{N+1} + H_{N+2} + \dots$$

Wegen (15) ist dann

$$m U_N = 1.$$

Endlich bilden wir

$$U = U_1 \cdot U_2 \cdot U_3 \dots;$$

es ist

$$m U = 1,$$

und jeder Punkt von U kommt in *unendlich vielen* H_N vor.

4. Jetzt betrachten wir den Durchschnitt

$$W = U \cdot V,$$

wo V die in § 1 konstruierte Menge in der ihr am Anfang von § 3 erteilten Spezialisierung ist. Wir haben

$$m W = 1.$$

In W ist nun $\sum_{v=1}^{\infty} c_v \varphi_v(x)$ konvergent. Denn wegen der zu Anfang des § 3 erwähnten Konvergenz von $s(x; 0, 2^n)$ in V gibt es in jedem Punkte x_0 von W zu beliebigem positivem ε ein N_1 , so daß

$$|s(x_0; 2^r, 2^{r_1})| < \frac{\varepsilon}{3}$$

für alle r und r_1 , die $N_1 \leq r \leq r_1$ erfüllen. Der Punkt x_0 von W liegt ferner in unendlich vielen H_N , und unter den hier vorkommenden Indizes N sei N_2 so groß gewählt, daß

$$\omega(2^{N_2}, \infty) < \frac{\varepsilon}{27}.$$

Da x_0 in H_{N_2} angenommen ist, so gilt wegen (16)

$$|s(x; 2^r, n)| < \frac{\varepsilon}{3}$$

für $N_2 \leq r$, $2^r \leq n < 2^{r+1}$. Von den beiden Zahlen N_1 und N_2 sei N_* die größere. Da stets die Zerlegung

$$s(x; m, n) = s(x; 2^r, 2^{r_1}) + s(x; 2^{r_1}, n) - s(x; 2^r, m)$$

mit $2^r \leq m < 2^{r+1}$, $2^{r_1} \leq n < 2^{r_1+1}$ möglich ist, so hat man endlich für alle m und n , die $2^{N_*} \leq m < n$ erfüllen,

$$|s(x_0; m, n)| < \varepsilon,$$

womit wir den folgenden Satz bewiesen haben:

Satz II. Ist $\sum_{v=1}^{\infty} (c_v \log v)^2$ konvergent, so konvergiert $\sum c_v \varphi_v(x)$ fast überall im Orthogonalitätsintervall.

5. Daß die soeben bewiesene Konvergenz nach der Bezeichnung Weyls eine „wesentlich gleichmäßige“ ist, folgt aus einem allgemeinen Satze von Egoroff¹²⁾, wonach eine in einer Menge C konvergente Folge meßbarer Funktionen in gewissen Teilmengen, deren Maß beliebig wenig unter dem von C bleibt, gleichmäßig konvergent ist. Man kann im vorliegenden Falle aber leicht Mengen dieser Art herausheben. Die in § 1 auftretenden Mengen denken wir uns, dem zu Anfang von § 3 Gesagten gemäß, für $\lambda(\nu) = \log \nu$ und $A_\rho = 2^\rho$ gebildet, was unter der Annahme der Konvergenz von $\sum (c_\nu \log \nu)^2$ möglich ist. Nun sei $N_1, N_2, N_3, \dots, N_p, \dots$ eine solche monotone Folge von natürlichen Zahlen, daß

$$(17a) \quad \sum_{p=1}^{\infty} \omega(2^{N_p}, \infty)^{\frac{1}{2}}$$

und also erst recht auch

$$(17b) \quad \sum_{p=1}^{\infty} \tau(2^{N_p}, \infty)^{\frac{1}{2}}$$

konvergent ausfällt, unter $\tau(m, n)$ hierbei $\sum_{\nu=m+1}^n c_\nu^2 \log \nu$ verstanden. Aus den Mengen D_N des § 1 und den Mengen H_N des § 3 bilden wir nun die Durchschnitte

$$K_p = D_{N_p} \cdot H_{N_p} \cdot D_{N_{p+1}} \cdot H_{N_{p+1}} \cdot \dots$$

Da die Mengen D_N und H_N sämtlich im Einheitsintervall liegen, so liegt wegen (5) und (15) und wegen der Konvergenz von (17a) und (17b) das Maß von K_p bei hinreichend hohem p beliebig nahe bei 1. In K_p ist nun die Konvergenz von $\sum c_\nu \varphi_\nu(x)$ gleichmäßig. Denn wegen (6) und (16) ist für $2^{N_p+k} \leq m < n$ stets in K_p

$$\left| \sum_{m+1}^n c_\nu \varphi_\nu(x) \right| < 2 \tau(2^{N_p+k}, \infty)^{\frac{1}{2}} + 2 \omega(2^{N_p+k}, \infty)^{\frac{1}{2}}.$$

III. Abschätzungen endlicher Summen von Orthogonalfunktionen.

6. Ist $\sum c_\nu^2$ konvergent, so ist nach Satz II die Reihe

$$(18) \quad \sum_{\nu=2}^{\infty} \frac{c_\nu}{\log \nu} \varphi_\nu(x)$$

im Intervall $(0 \dots 1)$ fast überall konvergent. Die Folge $\left\{ \frac{1}{\log \nu} \right\}$ ist monoton abnehmend mit dem Limes Null. Daraus schließt man nach

¹²⁾ Sur les suites de fonctions mesurables, Comptes Rendus de l'Acad. des Sciences de Paris 152 (1911), S. 244.

einem Satze von Kronecker¹³⁾ leicht, daß an den Konvergenzstellen von (18)

$$(19) \quad \sum_{\nu=1}^n c_{\nu} \varphi_{\nu}(x) = o(\log n)$$

sein muß, was daher fast überall im Orthogonalitätsintervall gelten muß. Sind die $\varphi_{\nu}(x)$ die trigonometrischen Funktionen und die c_{ν} die gewöhnlichen Fourierschen Konstanten einer integrierbaren Funktion, so findet sich diese Abschätzung schon bei Hardy¹⁴⁾, doch sagt der Hardysche Satz auch hier wieder mehr aus (vgl. Einleitung, Fußnote ⁶⁾), denn die Voraussetzung der Konvergenz von $\sum c_{\nu}^2$ ist äquivalent mit der Annahme, daß die c_{ν} die Fourierkonstanten einer *samt ihrem Quadrat* integrierbaren Funktion seien.

Der eben erwähnte Kroneckersche Satz läßt noch mehrfache Anwendungen folgender Art zu. Es sei

$$c_{\nu} = \nu^{-\frac{1}{2}} (\log \nu)^{-\frac{3}{2}-\varepsilon}, \quad \varepsilon > 0, \quad \nu = 2, 3, \dots;$$

für diese Wahl ist $\sum (c_{\nu} \log \nu)^3$ konvergent. Folglich ist

$$\sum_{\nu=2}^{\infty} \nu^{-\frac{1}{2}} (\log \nu)^{-\frac{3}{2}-\varepsilon} \varphi_{\nu}(x)$$

fast überall im Intervalle $(0 \dots 1)$ konvergent. Die Folge der c_{ν} ist aber monoton abnehmend mit dem Limes Null. An den Konvergenzstellen ist daher notwendig

$$(20) \quad \sum_{\nu=1}^n \varphi_{\nu}(x) = O(n^{\frac{1}{2}} (\log n)^{\frac{3}{2}+\varepsilon}),$$

was somit fast überall im Orthogonalitätsintervalle der $\varphi_{\nu}(x)$ gelten muß.

Ferner, ist $a_1, a_2, a_3, \dots, a_{\nu}, \dots$ irgendeine Folge von reellen Zahlen, so ist nach einem Satze von Pringsheim¹⁵⁾ die Reihe

$$\sum_{\nu=1}^{\infty} \frac{a_{\nu}^2}{\left(\sum_{\mu=1}^{\nu} a_{\mu}^2 \right)^{1+2\varepsilon}} \quad (\varepsilon > 0)$$

¹³⁾ Quelques remarques sur la détermination des valeurs moyennes, Comptes Rendus de l'Acad. des Sciences de Paris 103 (1886), S. 980–987. Vgl. auch Rademacher, Über die asymptotische Verteilung gewisser konvergenzerzeugender Faktoren. Math. Zeitschr. 11, S. 276–288.

¹⁴⁾ A. a. O. ⁶⁾ S. 369, Theorem 2.

¹⁵⁾ Über Konvergenz und Divergenz von unendlichen Reihen, Math. Ann. 35 (1890), S. 297–394, insbesondere S. 329.

konvergent. Folglich genügen die Konstanten

$$c_\nu = \frac{a_\nu}{\log \nu \left(\sum_{\mu=1}^{\nu} a_\mu^2 \right)^{\frac{1}{2} + \epsilon}} \quad (\nu \geq 2)$$

der Bedingung des Satzes II, und

$$\sum_{\nu=2}^{\infty} \frac{a_\nu \varphi_\nu(x)}{\log \nu \left(\sum_{\mu=1}^{\nu} a_\mu^2 \right)^{\frac{1}{2} + \epsilon}}$$

ist wieder fast überall im Orthogonalitätsintervall konvergent. Da überdies

$$\frac{1}{\log \nu \left(\sum_{\mu=1}^{\nu} a_\mu^2 \right)^{\frac{1}{2} + \epsilon}}$$

monoton gegen Null strebt, so gilt wegen des Kroneckerschen Satzes bis auf höchstens einer Nullmenge

$$(21) \quad \sum_{\nu=1}^n a_\nu \varphi_\nu(x) = O \left(\log n \left(\sum_{\mu=1}^n a_\mu^2 \right)^{\frac{1}{2} + \epsilon} \right).$$

Diese Formel erreicht für $a_\nu \equiv 1$ die Abschätzung (20) nicht ganz an Schärfe, doch könnte man statt (21) leicht dieselbe Genauigkeit wie in (20) erhalten, wenn man statt des oben angewandten Konvergenzsatzes von Pringsheim gewisse logarithmisch verschärfte Sätze desselben Autors heranzöge¹⁶⁾.

7. Die eben erhaltenen Größenordnungen gelten auch „wesentlich gleichmäßig“, d. h. gleichmäßig in Mengen, deren Maß beliebig wenig unter dem des Orthogonalitätsintervalles liegt. Für (19), d. h. für

$$\lim_{n \rightarrow \infty} \frac{\sum_{\nu=1}^n c_\nu \varphi_\nu(x)}{\log n} = 0$$

folgt dies einfach aus dem schon erwähnten Satze von Egoroff. Weiter möge etwa die wesentliche Gleichmäßigkeit von (20) gezeigt werden. Es sei $A_{k,l}$ die Menge, in der

$$(22) \quad \left| \sum_{\nu=1}^n \varphi_\nu(x) \right| < k \cdot n^{\frac{1}{2}} (\log n)^{\frac{8}{3} + \frac{1}{l}}$$

von einem gewissen $n = n(x, k, l)$ an ist. Es ist

$$A_{k,l} \subset A_{k+1,l},$$

¹⁶⁾ A. a. O. ¹⁵⁾ S. 333.

und wenn

$$A_l = A_{1,l} + A_{2,l} + A_{3,l} + \dots$$

gesetzt wird, so

$$m A_l = 1,$$

da (20) für $\varepsilon = \frac{1}{l}$ fast überall gilt. Bei vorgeschriebenem positiven δ gibt es also zu l ein k_l , so daß

$$m A_{k_l, l} > 1 - \frac{\delta}{2^{l+1}}$$

ist. Zur Abkürzung sei $A_{k_l, l}$ mit B_l bezeichnet. In B_l gilt also (22) für $k = k_l$ von einem gewissen Index $n = n(x, l)$ an. Nun sei $B_{l, m}$ die Teilmenge von B_l , in der (22) von $n = m$ an gilt. Da

$$B_{l1} + B_{l2} + B_{l3} + \dots = B_l$$

und

$$B_{lm} < B_{l, m+1},$$

so gibt es ein $m = m_l$, so daß

$$m B_{l, m_l} > m B_l - \frac{\delta}{2^{l+1}} > 1 - \frac{\delta}{2^l}$$

ist. Sei $C_l = B_{l, m_l}$ gesetzt. Bilden wir endlich den Durchschnitt

$$C = C_1 \cdot C_2 \cdot C_3 \dots,$$

so ist

$$m C > 1 - \delta.$$

Diese Menge C hat dann die Eigenschaft, daß es zu jedem ε ein $k = k(\varepsilon)$ und ein $n = n(k, \varepsilon)$ gibt, so daß

$$\left| \sum_{v=1}^n \varphi_v(x) \right| < k(\varepsilon) \cdot n^{\frac{1}{2}} (\log n)^{\frac{1}{2} + \varepsilon}$$

von $n = n(k, \varepsilon)$ an für alle x , d. h. *gleichmäßig* in C gilt.

IV. Eine allgemeinere Klasse von Funktionensystemen.

8. Herr I. Schur machte mich darauf aufmerksam, daß die bisher erhaltenen Ergebnisse über Konvergenzen und Größenordnungen ja auch in jedem Teilintervall des Orthogonalitätsintervalles fast überall stattfinden und daher auch von einer Funktionsfolge gelten, die in einem gewissen Intervall $(a \dots b)$ definiert ist und deren Funktionen in einem größeren Intervall so ergänzend definiert werden können, daß die ergänzten Funktionen ein normiertes Orthogonalsystem bilden. Die Charakterisierung solcher zu Orthogonalsystemen ergänzbaren Funktionensysteme leistet der folgende

Satz III (von Schur). *Eine Folge von Funktionen $f_1(x), f_2(x), \dots$, die im Intervall $(a \dots b)$, $0 < a < b < 1$, gegeben und dort samt ihrem Quadrate im Lebesgueschen Sinne integrierbar sind, läßt sich dann und*

nur dann zu einem System von in $(0 \dots 1)$ normierten Orthogonalfunktionen $g_1(x), g_2(x), \dots$ ergänzen, derart, daß $f_\nu(x) = g_\nu(x)$ in $(a \dots b)$, wenn die quadratische Form von unendlich vielen Variablen

$$F(z_1, z_2, z_3, \dots) = \int_a^b (z_1 f_1(x) + z_2 f_2(x) + z_3 f_3(x) + \dots)^2 dx$$

im Hilbertschen Sinne beschränkt ist und die Schranke 1 besitzt.

In der Tat, es sei

$$\int_a^b f_\mu(x) f_\nu(x) dx = c_{\mu\nu}.$$

Da nun $f_\mu(x) = g_\mu(x)$ in $(a \dots b)$ und

$$\int_0^1 g_\mu(x) g_\nu(x) dx = \begin{cases} 1 & \text{für } \mu = \nu \\ 0 & \text{für } \mu \neq \nu \end{cases}$$

sein soll, so muß

$$\int_0^a g_\mu(x) g_\nu(x) dx + \int_b^1 g_\mu(x) g_\nu(x) dx = \begin{cases} 1 - c_{\mu\mu} & \text{für } \mu = \nu \\ -c_{\mu\nu} & \text{für } \mu \neq \nu \end{cases}$$

sein. Die Formen

$$F_n(z_1, z_2, \dots, z_n) = \int_a^b (z_1 f_1 + z_2 f_2 + \dots + z_n f_n)^2 dx$$

und

$$G_n(z_1, z_2, \dots, z_n) = \int_0^a + \int_b^1 (z_1 g_1 + z_2 g_2 + \dots + z_n g_n)^2 dx$$

sind nichtnegativ. Ihre Matrizen sind

$$F_n : \begin{pmatrix} c_{11} & c_{12} & \dots & c_{1n} \\ c_{21} & c_{22} & \dots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{n1} & c_{n2} & \dots & c_{nn} \end{pmatrix}, \quad G_n : \begin{pmatrix} 1 - c_{11} & -c_{12} & \dots & -c_{1n} \\ -c_{21} & 1 - c_{22} & \dots & -c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -c_{n1} & -c_{n2} & \dots & 1 - c_{nn} \end{pmatrix}.$$

Also ist

$$F_n(z_1, z_2, \dots, z_n) + G_n(z_1, z_2, \dots, z_n) = z_1^2 + z_2^2 + \dots + z_n^2,$$

und folglich

$$0 \leq F_n(z_1, z_2, \dots, z_n) \leq z_1^2 + z_2^2 + \dots + z_n^2.$$

Da diese Ungleichung für alle n gilt, ist F in der Tat beschränkt und besitzt die Schranke 1.

Ist nun umgekehrt die Form $F = \sum c_{\mu\nu} z_\mu z_\nu$ beschränkt und $\leq \sum z_\mu^2$, so ist auch die Form

$$G = \sum z_\mu^2 - \sum c_{\mu\nu} z_\mu z_\nu$$

nichtnegativ. Dann kann man aber nach einem Satz von I. Schur¹⁷⁾ in den Intervallen $(0 \dots a)$ und $(b \dots 1)$ Funktionen $g_\mu(x)$ bestimmen, so daß

$$\int_0^a g_\mu(x) g_\nu(x) dx = \int_b^1 g_\mu(x) g_\nu(x) dx = \begin{cases} \frac{1 - c_{\mu\mu}}{2} & \text{für } \mu = \nu \\ -\frac{c_{\mu\nu}}{2} & \text{für } \mu \neq \nu \end{cases}$$

ist. Solche Funktionen $g_\mu(x)$ ergänzen aber nach den schon angestellten Überlegungen die Funktionen $f_\mu(x)$ zu einem im Intervall $(0 \dots 1)$ normierten Orthogonalsystem.

Für die Anwendung des somit bewiesenen Satzes III auf die früheren Ergebnisse braucht offenbar das Intervall $(a \dots b)$ nicht notwendig im Innern von $(0 \dots 1)$ zu liegen und die Schranke der quadratischen Form nicht gerade 1 zu sein. Zusammenfassend erhalten wir:

Satz IV. *Ist im Intervall $(a \dots b)$ eine Folge von samt ihrem Quadrate integrierbaren Funktionen $f_1(x), f_2(x), \dots$ gegeben und ist*

$$F(z_1, z_2, \dots) = \int_a^b (z_1 f_1(x) + z_2 f_2(x) + \dots)^2 dx$$

eine im Hilbertschen Sinne beschränkte quadratische Form von unendlich vielen Variablen, so gelten fast überall in $(a \dots b)$ für die $f_\nu(x)$ die Konvergenzaussagen der Sätze I und II und die Abschätzungen (19), (20) und (21).

V. Lebesguesche Konstanten und Funktionen.

9. Zu einem normierten Orthogonalsystem $\varphi_1(x), \varphi_2(x), \dots, \varphi_\nu(x), \dots$ führt man nach Lebesgue¹⁸⁾ die oberen Grenzen $\varrho_n(x)$ von

$$s_n(x) = \sum_{\nu=1}^n \varphi_\nu(x) \int_0^1 \varphi_\nu(y) f(y) dy$$

ein, wobei als $f(y)$ alle quadratisch integrierbaren Funktionen vom Betrage $|f(y)| \leq 1$ zugelassen sind. Es ist

$$\varrho_n(x) = \text{obere Grenze von } \int_0^1 f(y) \sum_{\nu=1}^n \varphi_\nu(x) \varphi_\nu(y) dy \leq \int_0^1 \left| \sum_{\nu=1}^n \varphi_\nu(x) \varphi_\nu(y) \right| dy,$$

und diese obere Grenze wird auch wirklich erreicht durch die Funktion

$$f(y) = \text{sign} \left(\sum_{\nu=1}^n \varphi_\nu(x) \varphi_\nu(y) \right),$$

¹⁷⁾ Über endliche Gruppen und Hermitesche Formen, Math. Zeitschr. 1 (1918) S. 183–207, Satz VII.

¹⁸⁾ Leçons sur les séries trigonométriques [Paris, 1906], S. 86–87.

in der x als Parameter auftritt. Also ist

$$(23) \quad \varrho_n(x) = \int_0^1 \left| \sum_{v=1}^n \varphi_v(x) \varphi_v(y) \right| dy.$$

Im Falle der trigonometrischen Funktionen erweisen sich die $\varrho_n(x)$ als von x unabhängig, und man spricht daher von den „Lebesgueschen Konstanten“. Allgemein aber hätte man die „Lebesgueschen Funktionen“ $\varrho_n(x)$ zu betrachten. Es sind für spezielle Orthogonalsysteme viele Untersuchungen über das infinitäre Verhalten der $\varrho_n(x)$ für wachsendes n angestellt worden, und zwar hat man, sobald $\varrho_n(x)$ von x abhängig ist, meist die Konstanten

$$\bar{\varrho}_n = \text{obere Grenze } \varrho_n(x) \\ 0 \leq x \leq 1$$

betrachtet. Im folgenden sollen nun mit den Methoden der §§ 1 und 3 die Funktionen $\varrho_n(x)$ selbst abgeschätzt werden, was allerdings nur wieder bis auf eine Nullmenge gelingen wird.

Solche Abschätzungen von $\varrho_n(x)$ sind darum von Wert, weil das Wachstum der $\varrho_n(x)$ aufs engste mit der Darstellungskraft von Reihen, die nach den Orthogonalfunktionen des Systems fortschreiten, zusammenhängt. Wenn $f(x)$ eine Funktion ist, die sich durch lineare Verbindungen der $\varphi_v(x)$ mit konstanten Koeffizienten (z. B. durch trigonometrische Polynome) beliebig gut approximieren läßt, so möge $f(x)$ als zum „Bereich“ des Orthogonalsystems gehörend bezeichnet werden. Dann kann man zeigen: Sind die Lebesgueschen Funktionen eines Orthogonalsystems gleichmäßig beschränkt, so wird jede Funktion seines Bereiches durch die auf Fouriersche Art gebildete Orthogonalfunktionenreihe dargestellt. Wir beweisen gleich noch mehr. Die Funktion $f(x)$ möge sich durch ein lineares Aggregat der ersten n Funktionen des Orthogonalsystems bis auf eine gewisse Genauigkeit approximieren lassen:

$$(24) \quad \left| f(x) - \sum_{v=1}^n a_v \varphi_v(x) \right| < \varepsilon_n.$$

Wir betrachten die Reihe mit den Fourierkoeffizienten

$$c_v = \int_0^1 f(y) \varphi_v(y) dy.$$

Dann ist

$$\begin{aligned} \left| \sum_{v=1}^n c_v \varphi_v(x) - \sum_{v=1}^n a_v \varphi_v(x) \right| &= \left| \int_0^1 \sum_{v=1}^n \varphi_v(x) \varphi_v(y) \left[f(y) - \sum_{v=1}^n a_v \varphi_v(y) \right] dy \right| \\ &\leq \int_0^1 \left| \sum_{v=1}^n \varphi_v(x) \varphi_v(y) \right| \left| f(y) - \sum_{v=1}^n a_v \varphi_v(y) \right| dy < \varepsilon_n \cdot \varrho_n(x), \end{aligned}$$

also zusammen mit (24)

$$\left| f(x) - \sum c_r \varphi_r(x) \right| < \varepsilon_n (\varrho_n(x) + 1).$$

Hieraus ersieht man, daß auch bei wachsendem $\varrho_n(x)$

$$f(x) = \sum_{r=1}^{\infty} c_r \varphi_r(x)$$

sein kann, wenn nur die ε_n derart gegen Null streben, daß $\varepsilon_n \varrho_n(x) \rightarrow 0$. Auf diese Weise läßt sich also die Theorie der Approximation reeller Funktionen durch Orthogonalfunktionen in den Dienst der Theorie der Reihen von Orthogonalfunktionen stellen.

10. Zur Abkürzung werde die Bezeichnung

$$\sum_{r=m+1}^n \varphi_r(x) \varphi_r(y) = K(m, n)$$

eingeführt.

Um zu einer Abschätzung der $\varrho_n(x)$ zu gelangen, gehen wir aus von

$$\begin{aligned} (25) \quad \int_0^1 \left[\sum_{j=1}^{r-1} \sum_{h=0}^{2^{r-j}-1} \int_0^1 K(h \cdot 2^j, h \cdot 2^j + 2^{j-1})^2 dy \right] dx &= \sum_{j=1}^{r-1} \sum_{h=0}^{2^{r-j}-1} 2^{j-1} \\ &= \sum_{j=1}^{r-1} 2^{r-j} 2^{j-1} = (r-1) 2^{r-1}. \end{aligned}$$

In der Doppelsumme unter dem äußeren Integral sind alle möglichen Integrale der Form

$$\int_0^1 K(\alpha, \alpha')^2 dy$$

addiert, wobei (α, α') Zahlenpaare von folgenden Eigenschaften sind:

$$0 \leq \alpha < \alpha' < 2^r, \quad \alpha = h \cdot 2^j, \quad \alpha' = \alpha + 2^{j-1}.$$

Wir schreiben somit kurz für (25)

$$\int_0^1 \left[\sum_{(\alpha, \alpha')} \int_0^1 K(\alpha, \alpha')^2 dy \right] dx = (r-1) 2^{r-1}.$$

Hieraus schließt man nun, daß der positive Integrand

$$(26) \quad \sum_{(\alpha, \alpha')} \int_0^1 K(\alpha, \alpha')^2 dy < (r-1) 2^{r-1} r^{1+2s}$$

in einer Menge S_r sein muß, deren Maß nicht unter $1 - \frac{1}{r^{1+2s}}$ liegt.

Ist nun n eine ganze positive Zahl, $n < 2^r$, so läßt sich n stets als Summe von gewissen der Zweierpotenzen $2^{r-1}, 2^{r-2}, \dots, 2^1, 2^0$ darstellen und somit $K(0, n)$ als Summe von höchstens r der $K(\alpha, \alpha')$. Wir schreiben etwa

$$K(0, n) = \sum_{(\alpha, \alpha')}^* K(\alpha, \alpha'),$$

worin durch $\sum^*$ die Auswahl aus den Paaren (α, α') angedeutet werden möge. Dann ist

$$\begin{aligned} (\varrho_n(x))^2 &= \left(\int_0^1 |K(0, n)| dy \right)^2 \leq \left(\sum_{(\alpha, \alpha')}^* \int_0^1 |K(\alpha, \alpha')| dy \right)^2 \\ &\leq \sum_{(\alpha, \alpha')}^* 1 \sum_{(\alpha, \alpha')}^* \left(\int_0^1 K(\alpha, \alpha') dy \right)^2 \leq r \sum_{(\alpha, \alpha')}^* \int_0^1 dy \int_0^1 K(\alpha, \alpha')^2 dy \\ &= r \sum_{(\alpha, \alpha')}^* \int_0^1 K(\alpha, \alpha')^2 dy \leq r \sum_{(\alpha, \alpha')}^* \int_0^1 K(\alpha, \alpha')^2 dy. \end{aligned}$$

Diese Ungleichung gilt für alle $n < 2^r$. Die letzte Summe der Ungleichung ist nun in (26) schon abgeschätzt worden, so daß man im ganzen erhält: Auf einer Menge S_r , deren Maß nicht kleiner als $1 - \frac{1}{r^{1+2s}}$ ist, gilt

$$(27) \quad \varrho_n(x) < \sqrt{r(r-1)2^{r-1}r^{1+2s}} < \sqrt{r^{3+2s} \cdot 2^{r-1}},$$

sobald $n < 2^r$. Es sei nun

$$T_N = S_N \cdot S_{N+1} \cdot S_{N+2} \cdot \dots,$$

dann ist

$$m T_N \geq 1 - \sum_{r=N}^{\infty} \frac{1}{r^{1+2s}},$$

und in T_N gilt (27) für alle $r \geq N$. Sei nun

$$U = T_1 + T_2 + T_3 + \dots,$$

so ist

$$m U = 1.$$

Jeder Punkt x_0 von U ist in einem T_N mit niedrigstem Index $N = N(x_0)$ enthalten, und für $r \geq N(x_0)$ gilt dann

$$(28) \quad \varrho_n(x_0) < r^{\frac{1}{2}+s} \cdot 2^{\frac{r-1}{2}}$$

bei $n < 2^r$. Zu $n \geq 2^{N(x_0)-1}$ gibt dann stets ein $r \geq N(x_0)$ von der Beschaffenheit, daß $n < 2^r \leq 2n$ ist, so daß also aus (28) folgt:

$$\varrho_n(x_0) < (\log 2n)^{\frac{1}{2}+s} \cdot \sqrt{n}$$

für x_0 in U und $n \geq 2^{N(x_0)-1}$. Dieses Ergebnis sprechen wir aus als

Satz V. *Außer höchstens in einer Nullmenge des Orthogonalitätsintervalles gilt für jedes x die Abschätzung der „Lebesgueschen Funktionen“*

$$(29) \quad \varrho_n(x) = O((\log n)^{\frac{3}{2}+\varepsilon} \cdot n^{\frac{1}{2}}).$$

Ist übrigens irgendwie bekannt, daß die $\varrho_n(x)$ von x unabhängig sind, so kann man leicht eine etwas bessere Abschätzung erhalten:

$$\begin{aligned} \varrho_n^2 &= \left(\int_0^1 |K(0, n)| dy \right)^2 = \left(\int_0^1 \int_0^1 |K(0, n)| dx dy \right)^2 \\ &\leq \int_0^1 \int_0^1 dx dy \int_0^1 \int_0^1 K(0, n)^2 dx dy = n, \\ (30) \quad \varrho_n &= O(\sqrt{n}). \end{aligned}$$

An einem Beispiele werden wir zeigen, daß diese Größenordnung auch wirklich erreicht werden kann.

VI. Ein spezielles Orthogonalsystem.

11. Man bemerkt leicht, daß man in (20) und (29) die Größenordnung auf $O(n^{\frac{1}{2}}(\log n)^{\frac{3}{2}}(\log \log n)^{1+\varepsilon})$ herunterdrücken könnte, wenn man, ohne den Gang der Rechnung zu ändern, die Reihe $\sum \frac{1}{n(\log n)^{1+\varepsilon}}$ als konvergente Reihe von positiven Gliedern anstatt $\sum \frac{1}{n^{1+\varepsilon}}$ benutzte. An einem Beispiel wollen wir nun zeigen, daß in (20) gewiß nicht $O(n^{\frac{1}{2}})$, in (29) aber nichts Besseres als $O(n^{\frac{1}{2}})$ stehen darf.

Es sei $e_\nu(x)$ die ν -te Stelle der dyadischen Bruchentwicklung von x , $0 \leq x \leq 1$. Als Funktion von x aufgefaßt ist $e_\nu(x)$ stückweise konstant in $(0 \dots 1)$, und zwar ist $e_\nu(x)$ in den Stücken $\frac{q}{2^\nu} \leq x < \frac{q+1}{2^\nu}$, $q = 0, 1, \dots, 2^\nu - 1$, abwechselnd 0 und 1. Als System von Orthogonalfunktionen führen wir nun ein

$$(31) \quad \psi_\nu(x) = 2e_\nu(x) - 1 \quad (\nu = 1, 2, 3, \dots).$$

Da $\psi_\nu(x)$ im Intervall $(0 \dots 1)$ abwechselnd -1 und $+1$ ist, so ist $\psi_\nu(x)$ in der Tat normiert

$$\int_0^1 \psi_\nu(x)^2 dx = 1.$$

Ist ferner $\mu < \nu$, so ist

$$\int_0^1 \psi_\mu(x) \psi_\nu(x) dx = 0,$$

denn in jeder der 2^μ gleichlangen Strecken der Konstanz von $\psi_\mu(x)$ gibt $2^{\nu-\mu}$ gleichlange Strecken abwechselnden Vorzeichens von $\psi_\nu(x)$, so daß die Integrale über die Strecken der Konstanz von ψ_μ einzeln verschwinden

und somit auch das von 0 bis 1 erstreckte Integral. (Das System der $\psi_\nu(x)$ ist übrigens *kein vollständiges* Orthogonalsystem, da z. B. $\psi_\lambda(x) \cdot \psi_\mu(x)$ für $\lambda \neq \mu$ zu allen $\psi_\nu(x)$ orthogonal ist. Die Vollständigkeit eines Orthogonalsystems spielt aber in den vorliegenden Untersuchungen überhaupt keine Rolle.)

Nun ist

$$\sum_{\nu=1}^n \psi_\nu(x) = 2 \sum_{\nu=1}^n e_\nu(x) - n$$

gleich der Differenz der Anzahl der Einsen und der Nullen in den ersten n Stellen der dyadischen Bruchentwicklung von x . Nach einem Satz von Hardy und Littlewood¹⁹⁾ kann aber

$$2 \sum_{\nu=1}^n e_\nu(x) - n = O(\sqrt{n})$$

nur in einer Nullmenge der x gelten, womit also gezeigt ist, daß die in einer Menge vom Maße 1 gültige Abschätzung (20) nicht bis auf $O(\sqrt{n})$ herabgedrückt werden kann.

Übrigens kann man aus (20) umgekehrt schließen, daß

$$2 \sum_{\nu=1}^n e_\nu(x) - n = O(n^{\frac{1}{2}} (\log n)^{\frac{3}{2} + \epsilon})$$

außer höchstens in einer Nullmenge der x sein muß, ein Resultat übrigens, das schon auf direktem Wege und schärfer Hardy und Littlewood gefunden haben²⁰⁾ und zu dem ich an anderer Stelle auf einem dritten Wege gelangt bin²¹⁾.

12. Die „Lebesgueschen Funktionen“ $e_n(x)$ des Systems $\{\psi_\nu(x)\}$, definiert durch

$$e_n(x) = \int_0^1 \left| \sum_{\nu=1}^n \psi_\nu(x) \psi_\nu(y) \right| dy$$

werden sich als von x unabhängig herausstellen, so daß man es in diesem Falle mit „Lebesgueschen Konstanten“ zu tun hat.

Da die Funktionen $\psi_\nu(x)$ nur die Werte ± 1 annehmen, so kann $\sum_{\nu=1}^n \psi_\nu(x) \psi_\nu(y)$ nur ganzzahlige Werte annehmen. Und zwar muß stets $\sum_{\nu=1}^n \psi_\nu(x) \psi_\nu(y) \equiv n \pmod{2}$ sein, und man sieht genauer, daß nur die Werte

$$(32) \quad -n, -n+2, -n+4, \dots, +n-2, +n$$

¹⁹⁾ Some Problems of Diophantine Approximation, Acta Mathem. 37 (1914), S. 155–239, insbesondere S. 187, Theorem 1. 47, und S. 189, Theorem 1. 471.

²⁰⁾ A. a. O. Theorem 1. 45, S. 185.

²¹⁾ A. a. O. ¹²⁾.

möglich sind. Es werde nun $\sum^n \psi_\nu(x_0) \psi_\nu(y) = K_n(x_0, y)$ bei festem $x = x_0$ als Funktion von y betrachtet. In gleichlangen Intervallen der Länge $\frac{1}{2^n}$ ist sie konstant und nimmt dort einen der eben gekennzeichneten ganzzahligen Werte an. Wir fragen nun nach der Gesamtlänge der Intervalle eines gewissen Funktionswertes. Wir wollen diese Gesamtlängen einfach hintereinander nach der Größenanordnung der in ihnen angenommenen Funktionswerte aufschreiben, wobei wir diese Funktionswerte als ohnehin bekannt nicht mehr zu erwähnen brauchen. Zu $K_1(x_0, y)$ gehören z. B. die Gesamtlängen $\frac{1}{2}, \frac{1}{2}$.

In jedes der 2^n gleichlangen Teilintervalle, in denen $K_n(x_0, y)$ gewiß konstant ist, fallen gerade zwei gleichlange Teilintervalle von $\psi_{n+1}(x_0) \psi_{n+1}(y)$ von entgegengesetztem Wert. Die Hälfte der Intervalle eines konstanten Wertes k von $K_n(x_0, y)$ wird durch Addition von $\psi_{n+1}(x_0) \psi_{n+1}(y)$ zu den Intervallen des Wertes $k+1$ der Funktion $K_{n+1}(x_0, y)$ geschlagen, die andere Hälfte fällt den Intervallen des Wertes $k-1$ zu. Daher bietet sich folgendes Schema für die Wertverteilung in den $K_n(x_0, y)$ dar:

$$\begin{aligned} K_1: & \quad \frac{1}{2}, \quad \frac{1}{2}, \\ K_2: & \quad \frac{1}{2^2}, \quad 2 \cdot \frac{1}{2^2}, \quad \frac{1}{2^2}, \\ K_3: & \quad \frac{1}{2^3}, \quad 3 \cdot \frac{1}{2^3}, \quad 3 \cdot \frac{1}{2^3}, \quad \frac{1}{2^3} \end{aligned}$$

und allgemein

$$(33) \quad K_n(x_0, y): \quad \binom{n}{0} \frac{1}{2^n}, \quad \binom{n}{1} \frac{1}{2^n}, \quad \binom{n}{2} \frac{1}{2^n}, \quad \dots, \quad \binom{n}{n} \frac{1}{2^n},$$

was man sofort durch vollständige Induktion bestätigt. Diese hiermit charakterisierte Verteilung des Wertevorrats auf die Intervalllängen, oder die *relative Häufigkeit* eines Funktionswertes ist von x_0 völlig unabhängig, der *Verlauf* der Funktionen $K_n(x_0, y)$ ist aber erst durch die Anordnung der Intervalle konstanter Werte gegeben und hängt von x_0 ab.

Wenn man den Wertevorrat (32) ins Auge faßt und seine in (33) angedeutete Verteilung hinzunimmt, so erschließt man ohne weiteres

$$\begin{aligned} (34) \quad q_n(x_0) &= \int_0^1 |K_n(x_0, y)| dy \\ &= \frac{1}{2^{n-1}} \left\{ n \binom{n}{0} + (n-2) \binom{n}{1} + (n-4) \binom{n}{2} + \dots + \left(n - 2 \left\lfloor \frac{n}{2} \right\rfloor \right) \binom{n}{\left\lfloor \frac{n}{2} \right\rfloor} \right\}, \end{aligned}$$

womit die schon behauptete Unabhängigkeit der $\varrho_n(x)$ von x nachgewiesen ist²²⁾.

13. Was nun das Wachstum der ϱ_n betrifft, so zeigt sich zunächst, daß $\varrho_{2m} = \varrho_{2m-1}$ ist, denn

$$\begin{aligned}\varrho_{2m} &= \frac{1}{2^{2m-1}} \left\{ 2m \binom{2m}{0} + (2m-2) \binom{2m}{1} + (2m-4) \binom{2m}{2} + \dots + 2 \binom{2m}{m-1} \right\} \\ &= \frac{1}{2^{2m-2}} \left\{ m \binom{2m-1}{0} + (m-1) \left(\binom{2m-1}{0} + \binom{2m-1}{1} \right) \right. \\ &\quad \left. + (m-2) \left(\binom{2m-1}{1} + \binom{2m-1}{2} \right) + \dots + \left(\binom{2m-1}{m-2} + \binom{2m-1}{m-1} \right) \right\} \\ &= \frac{1}{2^{2m-2}} \left\{ (2m-1) \binom{2m-1}{0} + (2m-3) \binom{2m-1}{1} + \dots + 3 \binom{2m-1}{m-2} \right. \\ &\quad \left. + 1 \binom{2m-1}{m-1} \right\} = \varrho_{2m-1}.\end{aligned}$$

Dagegen ist

$$\begin{aligned}\varrho_{2m+1} &= \frac{1}{2^{2m}} \left\{ (2m+1) \binom{2m+1}{0} + (2m-1) \binom{2m+1}{1} + (2m-3) \binom{2m+1}{2} \right. \\ &\quad \left. + \dots + 1 \binom{2m+1}{m} \right\} \\ &= \frac{1}{2^{2m}} \left\{ (2m+1) \binom{2m}{0} + (2m-1) \left(\binom{2m}{0} + \binom{2m}{1} \right) \right. \\ &\quad \left. + (2m-3) \left(\binom{2m}{1} + \binom{2m}{2} \right) + \dots + 1 \left(\binom{2m}{m-1} + \binom{2m}{m} \right) \right\} \\ &= \frac{1}{2^{2m}} \left\{ 4m \binom{2m}{0} + (4m-4) \binom{2m}{1} + \dots + 4 \binom{2m}{m-1} \right\} + \frac{1}{2^{2m}} \binom{2m}{m} \\ &= \varrho_{2m} + \frac{1}{2^{2m}} \binom{2m}{m}.\end{aligned}$$

Da außerdem $\varrho_1 = 1$, so ergibt diese Rechnung die neue Darstellung

$$\varrho_{2m-1} = \varrho_{2m} = 1 + \frac{1}{2^2} \binom{2}{1} + \frac{1}{2^4} \binom{4}{2} + \dots + \frac{1}{2^{2m-2}} \binom{2m-2}{m-1},$$

aus der unmittelbar hervorgeht, daß die ϱ_n nie abnehmen bei wachsendem n .

²²⁾ Diese Eigenschaft der Lebesgueschen Funktionen scheint unter den bisher daraufhin untersuchten Systemen nur dem System der trigonometrischen Funktionen [vgl. Lebesgue a. a. O.¹⁷⁾ und neuerdings vor allem G. Szegő, Über die Lebesgueschen Konstanten bei den Fourierschen Reihen, Math. Zeitschr. 9, 161–166] und dem Haarschen Orthogonalsystem zuzukommen [Haar, Zur Theorie der orthogonalen Funktionensysteme, Diss. Göttingen 1909, auch Math. Ann. 69 (1910), S. 331–371]. Es ist dort nicht besonders hervorgehoben, ergibt sich aber leicht aus den Betrachtungen Kap. III, § 2, daß die Lebesgueschen Funktionen $\varrho_n(x)$ des Haarschen Systems von x und n unabhängig und sämtlich gleich 1 sind.

Für den Zweck einer Abschätzung werde jedoch (34) noch auf eine andere Form gebracht. Es ist

$$\varrho_{2m+1} = \frac{1}{2^{2m}} \left\{ (2m+1) \binom{2m+1}{0} + (2m-1) \binom{2m+1}{1} + \dots + 1 \binom{2m+1}{m} \right\},$$

woraus durch Addition von

$$2m+1 = \frac{2m+1}{2^{2m}} \cdot \frac{2^{2m+1}}{2}$$

$$= \frac{1}{2^{2m}} \left\{ (2m+1) \binom{2m+1}{0} + (2m+1) \binom{2m+1}{1} + \dots + (2m+1) \binom{2m+1}{m} \right\}$$

folgt

$$\begin{aligned} \varrho_{2m+1} + 2m+1 &= \frac{1}{2^{2m}} \left\{ (4m+2) \binom{2m+1}{0} + 4m \binom{2m+1}{1} + \dots + (2m+2) \binom{2m+2}{m} \right\} \\ &= \frac{1}{2^{2m-1}} \left\{ (2m+1) \binom{2m+1}{0} + 2m \binom{2m+1}{1} + \dots + (m+1) \binom{2m+1}{m} \right\} \\ &= \frac{2m+1}{2^{2m-1}} \left\{ \binom{2m}{0} + \binom{2m}{1} + \binom{2m}{2} + \dots + \binom{2m}{m} \right\} \\ &= \frac{2m+1}{2^{2m}} \left\{ \binom{2m}{0} + \binom{2m}{1} + \dots + \binom{2m}{m} + \binom{2m}{m+1} + \dots + \binom{2m}{2m} \right\} \\ &\quad + \frac{2m+1}{2^{2m}} \binom{2m}{m} \\ &= 2m+1 + \frac{2m+1}{2^{2m}} \binom{2m}{m}, \end{aligned}$$

also

$$\varrho_{2m+1} = \frac{2m+1}{2^{2m}} \binom{2m}{m}.$$

Dieser Ausdruck für ϱ_n eignet sich gut für eine asymptotische Approximation. Mit Hilfe der Stirlingschen Formel erhält man

$$\varrho_{2m+1} = \frac{2m+1}{2^{2m}} \cdot \frac{(2m)!}{m!m!} = \frac{2m+1}{2^{2m}} \cdot \frac{1}{\sqrt{2\pi}} \cdot \frac{(2m)^{2m+\frac{1}{2}}}{m^{2m+1}} \cdot e^{\frac{\vartheta}{24m} - \frac{\vartheta'}{6m}} \quad (0 \leq \vartheta, \vartheta' \leq 1),$$

$$\varrho_{2m+1} = \frac{1}{\sqrt{\pi}} \cdot \frac{2m+1}{\sqrt{m}} \cdot e^{\frac{\vartheta}{24m} - \frac{\vartheta'}{6m}} \sim \frac{2}{\sqrt{\pi}} \sqrt{m},$$

und endlich, da $\varrho_{2m+1} = \varrho_{2m+2}$,

$$\varrho_n \sim \sqrt{\frac{2n}{\pi}}.$$

Hiermit ist in der Tat gezeigt, daß die Abschätzung (30) nicht verbessert werden kann.

14. Das betrachtete Orthogonalsystem besitzt noch eine weitere bemerkenswerte Eigenschaft. Es gilt nämlich in diesem Falle, hinausgehend über unsern allgemeinen Satz II:

Die Reihe $\sum_{v=1}^{\infty} c_v \psi_v(x)$ konvergiert fast überall in $(0 \dots 1)$, sobald nur $\sum c_v^2$ konvergent ist²³⁾.

Der Beweis für diese Tatsache findet sich eigentlich schon in meiner in Anmerkung ¹³⁾ zitierten Arbeit (S. 283 f.), er möge hier jedoch, auf den gegenwärtigen Zweck zugeschnitten, in neuer Fassung dargestellt werden.

Ohne Einschränkung der Allgemeinheit kann man $|c_v| < 1$ für alle v annehmen. Nun sei

$$g_v(x) = \int_0^x (1 + c_v \psi_v(t)) dt \quad (v = 1, 2, 3, \dots).$$

Da der Integrand stets positiv ist, so sind die $g_v(x)$ monoton wachsende Funktionen, und zwar ist

$$g_v(0) = 0, \quad g_v(1) = 1.$$

Sind $0 = x_0^{(v)}, x_1^{(v)}, x_2^{(v)}, \dots, x_{2^v}^{(v)} = 1$ äquidistante Teilpunkte (die Endpunkte mitgezählt) des Intervalls $(0 \dots 1)$ von der Anzahl $2^v + 1$, so folgt aus der Definition der $\psi_v(x)$ unmittelbar

$$\int_0^{x_i^{(v)}} \psi_{v+1}(t) dt = 0 \quad (i = 0, 1, 2, \dots, 2^v),$$

was

$$(35) \quad g_{v+1}(x_i^{(v)}) = \int_0^{x_i^{(v)}} (1 + c_{v+1} \psi_{v+1}(t)) dt = x_i^{(v)}$$

zur Folge hat. Zwischen aufeinanderfolgenden Teilpunkten $x_i^{(v)}$ und $x_{i+1}^{(v)}$ ist ferner $g_v(x)$ linear, für $g_v(x)$ sind also die Teilpunkte $x_i^{(v)}$ Knickpunkte der Funktion. Innerhalb eines der 2^v gleichlangen Teilintervalle gilt

$$g'_v(x) = 1 + c_v \psi_v(x).$$

Es soll nun mittels der $g_v(x)$ eine Folge von gleichfalls monotonen und stetigen Funktionen $f_v(x)$ bestimmt werden. Es sei $f_1(x) = g_1(x)$. Ferner soll $f_{v+1}(x)$ mit $f_v(x)$ an den $2^v + 1$ Teilpunkten $x_i^{(v)}$ übereinstimmen, deren innere übrigens Knickpunkte von $f_v(x)$ sein werden.

²³⁾ Es dürfte nicht genügend bekannt sein, daß auch dem Haarschen Orthogonalsystem diese Eigenschaft zukommt. Denn ist $\sum c_v^2$ konvergent, so sind nach dem Riesz-Fischerschen Satze die c_v die Fourierkoeffizienten einer integrierbaren Funktion $f(x)$ in bezug auf das Haarsche System. Die Fourierentwicklung von $f(x)$ nach den Haarschen Funktionen konvergiert aber überall da, wo $f(x)$ gleich der Ableitung seines unbestimmten Integrals ist, also, nach Lebesgue, fast überall (vgl. Haar, a. a. O.²²⁾, Kap. III, § 2 am Schluß).

Zwischen je zwei aufeinanderfolgende der so festgelegten Punkte von $f_{\nu+1}(x)$ aber soll das zwischen denselben Abszissen liegende Stück von $g_{\nu+1}(x)$ eingespannt werden, indem es in seiner Ordinatenrichtung passend affin verzerrt und verschoben wird. Genauer ausgedrückt: Sind $x_i^{(\nu)}$ und $x_{i+1}^{(\nu)}$ zwei aufeinanderfolgende Teilpunkte aus der Teilung des Intervalls $(0 \dots 1)$ in 2^ν gleiche Teile, und ist $x_i^{(\nu)} \leq x \leq x_{i+1}^{(\nu)}$, so soll sein

$$f_{\nu+1}(x) = A g_{\nu+1}(x) + B,$$

worin die Konstanten A und B so bestimmt werden, daß der Anschluß an die schon festgelegten Punkte von $f_{\nu+1}(x)$ also an die Punkte $(x_i^{(\nu)}, f_\nu(x_i^{(\nu)}))$ und $(x_{i+1}^{(\nu)}, f_\nu(x_{i+1}^{(\nu)}))$ stetig erfolgt, was

$$f_{\nu+1}(x) = g_{\nu+1}(x) \frac{f_\nu(x_{i+1}^{(\nu)}) - f_\nu(x_i^{(\nu)})}{g_{\nu+1}(x_{i+1}^{(\nu)}) - g_{\nu+1}(x_i^{(\nu)})} + \frac{f_\nu(x_i^{(\nu)}) g_{\nu+1}(x_{i+1}^{(\nu)}) - f_\nu(x_{i+1}^{(\nu)}) g_{\nu+1}(x_i^{(\nu)})}{g_{\nu+1}(x_{i+1}^{(\nu)}) - g_{\nu+1}(x_i^{(\nu)})}$$

oder wegen (35)

$$(36) \quad f_{\nu+1}(x) = g_{\nu+1}(x) \frac{f_\nu(x_{i+1}^{(\nu)}) - f_\nu(x_i^{(\nu)})}{x_{i+1}^{(\nu)} - x_i^{(\nu)}} + \frac{f_\nu(x_i^{(\nu)}) x_{i+1}^{(\nu)} - f_\nu(x_{i+1}^{(\nu)}) x_i^{(\nu)}}{x_{i+1}^{(\nu)} - x_i^{(\nu)}}$$

ergibt. Diese für $x_i^{(\nu)} \leq x \leq x_{i+1}^{(\nu)}$ definierten Funktionsstücke schließen sich zu einer im ganzen Intervall $(0 \dots 1)$ stetigen Funktion $f_{\nu+1}(x)$ zusammen. Zwischen den $2^{\nu+1} + 1$ Knick- und Endpunkten von $g_{\nu+1}(x)$ ist $f_{\nu+1}(x)$ differenzierbar, und zwar ist dort

$$f'_{\nu+1}(x) = g'_{\nu+1}(x) \cdot \frac{f_\nu(x_{i+1}^{(\nu)}) - f_\nu(x_i^{(\nu)})}{x_{i+1}^{(\nu)} - x_i^{(\nu)}} = g'_{\nu+1}(x) \cdot f'_\nu(x).$$

Da noch $f_1(x) = g_1(x)$ festgesetzt war, so ist

$$(37) \quad f'_\nu(x) = \prod_{\mu=1}^{\nu} g'_\mu(x) = \prod_{\mu=1}^{\nu} (1 + c_\mu \psi_\mu(x)),$$

woraus sich auch ergibt, daß alle $f_\nu(x)$ monoton wachsende Funktionen sind. Die $f_\nu(x)$ konvergieren gleichmäßig gegen einen Limes. Zwischen $x_i^{(\nu)}$ und $x_{i+1}^{(\nu)}$ ist nämlich $f_\nu(x)$ linear, da $g_1(x), g_2(x), \dots, g_\nu(x)$ in diesem Intervall linear sind; daher läßt sich schreiben für $x_i^{(\nu)} \leq x \leq x_{i+1}^{(\nu)}$

$$f_\nu(x) = x \cdot \frac{f_\nu(x_{i+1}^{(\nu)}) - f_\nu(x_i^{(\nu)})}{x_{i+1}^{(\nu)} - x_i^{(\nu)}} + \frac{f_\nu(x_i^{(\nu)}) x_{i+1}^{(\nu)} - f_\nu(x_{i+1}^{(\nu)}) x_i^{(\nu)}}{x_{i+1}^{(\nu)} - x_i^{(\nu)}}.$$

Also ist wegen (36)

$$f_{\nu+1}(x) - f_\nu(x) = (g_{\nu+1}(x) - x) \frac{f_\nu(x_{i+1}^{(\nu)}) - f_\nu(x_i^{(\nu)})}{x_{i+1}^{(\nu)} - x_i^{(\nu)}}$$

$$(38) \quad f_{\nu+1}(x) - f_\nu(x) = (g_{\nu+1}(x) - x) \cdot f'_\nu(x).$$

Einerseits ist nun

$$g_{\nu+1}(x) - x = c_{\nu+1} \int_0^x \psi_{\nu+1}(t) dt,$$

und somit, wegen der Definition der $\varphi_\nu^{(x)}$:

$$|g_{\nu+1}(x) - x| \leq |c_\nu| \frac{1}{2^{\nu+1}} < \frac{1}{2^{\nu+1}}.$$

Andererseits gibt es, da $|c_\nu| < 1$ und $c_\nu \rightarrow 0$, eine Zahl k , $0 < k < 1$, so daß $|c_\nu| < k$ für alle ν . Aus (37) folgt daher, daß $0 < f'_\nu(x) < (1+k)^\nu$ ist, so daß (38) ergibt

$$|f_{\nu+1}(x) - f_\nu(x)| < \frac{(1+k)^\nu}{2^{\nu+1}},$$

woraus man durch Vergleich mit der geometrischen Reihe die gleichmäßige Konvergenz der $f_\nu(x)$ sofort abliest. Es sei

$$\lim_{\nu \rightarrow \infty} f_\nu(x) = f(x);$$

$f(x)$ ist eine monotone und stetige Funktion. Für diese Funktion ist es wesentlich, daß sie mit $f_\nu(x)$ an den Teilpunkten $x_i^{(\nu)}$ übereinstimmt, da die Funktionswerte an diesen Teilpunkten für alle auf $f_\nu(x)$ folgenden Funktionen festbleiben.

Nun gilt der Lebesguesche Satz, daß jede monotone Funktion außer höchstens in einer Nullmenge eine *endliche* Ableitung besitzt. Schließt man nun einen Punkt x in eine Intervallschachtelung ein, die nur ihn als gemeinsamen Punkt besitzt, so kann in x nur dann Differenzierbarkeit herrschen, wenn die Folge der zu jedem Intervall der Schachtelung gebildeten Differenzenquotienten konvergent ist. Jeden Punkt x nun, der nicht gerade der Nullmenge der endlichen dyadischen Brüche angehört, kann man in eine Intervallschachtelung einschließen, deren ν -tes Intervall von zwei Teilpunkten $x_i^{(\nu)}$ und $x_{i+1}^{(\nu)}$ begrenzt wird. Der hierfür gebildete Differenzenquotient von $f(x)$ stimmt aber mit dem von $f_\nu(x)$ überein, der seinerseits gleich $f'_\nu(x)$ ist. Also muß, außer höchstens in einer Nullmenge der x , der Limes von $f'_\nu(x)$ bei $\nu \rightarrow \infty$ existieren. Nach (37) heißt das aber, daß

$$(39) \quad \lim_{n \rightarrow \infty} \prod_{\nu=1}^n (1 + c_\nu \psi_\nu(x)) = \prod_{\nu=1}^{\infty} (1 + c_\nu \psi_\nu(x))$$

außer höchstens in einer Nullmenge existiert. Dabei wäre dieses unendliche Produkt, entgegen einer gebräuchlichen Terminologie, auch dann konvergent zu nennen, wenn sein Limes Null ist.

Nun war aber $\sum c_v^2$ oder, was dasselbe ist, $\sum_{v=1}^{\infty} c_v^2 \psi_v^2(x)$ als konvergent vorausgesetzt. Überall, wo (39) konvergent ist, muß daher

$$(40) \quad \sum_{v=1}^{\infty} c_v \psi_v(x)$$

entweder konvergieren oder bestimmt gegen $-\infty$ divergieren. Letzteres kann aber auch nur auf einer Nullmenge geschehen. Denn $\sum_{v=1}^{\infty} (-c_v) \psi_v(x)$ ist eine Reihe von demselben Typ wie (40) und kann daher höchstens auf einer Nullmenge gegen $+\infty$ divergieren. *Somit ist (40) fast überall im Intervall $(0 \dots 1)$ konvergent.*

Das hier betrachtete Orthogonalsystem teilt mit dem Haarschen die Eigenschaft, aus unstetigen, stückweise konstanten Funktionen zu bestehen. Im übrigen zeigen aber die beiden Systeme ein durchaus verschiedenes Verhalten: Das Haarsche System ist vollständig, seine Funktionen verschwinden auf beliebig langen Teilstrecken des Intervalls und besitzen sämtlich nur einen Vorzeichenwechsel, die Funktionen sind bei hohem Index beliebig großer Werte fähig und die Lebesgueschen Konstanten des Systems sind sämtlich gleich 1; unser System dagegen ist unvollständig, seine Funktionen nehmen den Wert Null nicht an, sind beliebig vieler Zeichenwechsel fähig, bleiben gleichmäßig beschränkt, und die Lebesgueschen Konstanten wachsen in der größtmöglichen Ordnung. Die Tatsache, daß zwei so verschiedene Systeme schon bei Konvergenz von $\sum c_v^2$ fast überall konvergente Reihen liefern, legt aufs neue die Frage nahe, ob nicht unser Satz II einer weiteren Verschärfung fähig sei, und insbesondere auch die Frage nach dem maximalen Maß der Divergenzmenge einer Fourierreihe.

(Eingegangen am 8. 10. 1921.)