Über den Approximationssatz von Weierstraß.

Von

Ch. H. Müntz in München.

Ein sehr bekannter, vielfach bewiesener und untersuchter Satz von Weierstraß[1]) besagt, daß jede stetige Funktion $f(x)$ in einem endlichen reellen Intervalle $a \leqq x \leqq b$ durch die Folge aller natürlichen Potenzen

$$x^0 = 1,\ x,\ x^2,\ \ldots,\ x^N,\ \ldots$$

mit beliebig vorgeschriebener Genauigkeit gleichmäßig approximiert werden kann, indem sich zu jeder positiven Fehlergrenze δ ein $N = N(\delta)$ derart finden läßt, daß

$$|f(x) - \alpha_0 - \alpha_1 x - \alpha_2 x^2 \cdots - \alpha_N x^N| < \delta$$

im ganzen gegebenen Intervalle bleibt, wobei die α_k passende Konstanten bedeuten.

Es ist nun leicht zu sehen, daß zu diesem Zwecke nicht alle jene Potenzen erforderlich sind; und ebenso läßt sich zeigen, daß nicht jedes unendliche Teilsystem derselben zum gleichen Zwecke hinreicht. Man kann daher die Frage nach den Bedingungen aufwerfen, unter

1) Werke, Bd. III, S. 1—37.

Vgl. E. Picard: Traité d'analyse, t. I, S. 275—277 (Paris 1901).

E. Borel: Leçons sur les fonctions de variables réelles, S. 50—66 (Paris 1905).

E. Landau: Über die Approximation einer stetigen Funktion durch ganze rationale Funktionen", Rendiconti d. C. M. di Palermo, XXV (1908), S. 337—345.

H. Lebesgue: „Sur la representation approchée des fonctions", ibd., XXVI (1908), S. 325—328.

C. Carathéodory et al. (eds.), *Mathematische Abhandlungen Hermann Amandus Schwarz*

welchen das System

$$x^{p_0},\ x^{p_1},\ x^{p_2},\ \ldots,\ x^{p_N},\ \ldots;\quad 0 \leqq p_0{}^{1)} < p_1 < p_2 \cdots < p_N < \cdots$$

in diesem Sinne ein vollständiges ist. Der Einfachheit halber möge hier zunächst die Beantwortung dieser Frage für das Intervall $0 \leqq x \leqq 1$ gegeben werden:

Satz. Damit die unendliche Folge der Potenzen $x^{p_0} = 1,\ x^{p_1},\ x^{p_2},\ \ldots,\ x^{p_N},\ \ldots$ mit wachsenden positiven Exponenten imstande sei, **jede** stetige Funktion im Intervalle $0 \ldots 1$ beliebig zu approximieren, ist es notwendig und hinreichend, daß $\sum\limits_1^\infty \frac{1}{p_k}$ divergiere.

Für beliebige Intervalle ist nur zu bemerken, daß jede stetige Funktion als Summe einer geraden und einer ungeraden stetigen Funktion dargestellt werden kann. Sind die positiven Exponenten p_k ganzzahlig, so müssen dann darunter sowohl unendlichviele gerade p'_k als unendlichviele ungerade p''_k vorkommen, und es muß jede der Reihen $\sum \frac{1}{p'_k}$ und $\sum \frac{1}{p''_k}$ für sich divergieren. Das gleiche Kriterium bleibt aber auch bei beliebigen p_k gültig, wenn für positive z allgemein

$$(-z)^{p'_k} = z^{p'_k},\quad (-z)^{p''_k} = -z^{p''_k}$$

festgesetzt, und die Frage nach der Approximation durch die so definierten Funktionen gestellt wird.

Das in Frage stehende allgemeine Problem ist in einer Preisschrift von Herrn S. Bernstein[2]), welche viele andere Probleme der Approximation durch Polynome zu einer vollen Erledigung bringt, insofern unvollständig beantwortet worden, als dort teils nur notwendige, teils nur hinreichende Kriterien für die Folge der p_k angegeben werden; auch die Beweisführung in jener Abhandlung ist von der hier befolgten gänzlich verschieden.

Aus dem gegebenen Satz läßt sich sofort der folgende ableiten: Wenn $f(x)$ eine in einem Intervalle $a > 0 \ldots b$ stetige eindeutige Funktion ist, und wenn für eine Folge wachsen-

1) Offenbar muß $p_0 = 0$ sein, sobald $x = 0$ zum betrachteten abgeschlossenen Intervall gehört.

2) Memoires couronnés de l'Académie de Belgique, 1912: „Sur l'ordre de la meilleure approximation des fonctions continues par des polynomes de degré donné". Vgl. S. 78—85.

der positiver p_k mit divergierender Summe $\sum \frac{1}{p_k}$ beständig

$$\int_a^b f(x) \, . \, x^{p_k} . \, dx = 0$$

gilt, so muß $f(x)$ identisch verschwinden.

Auch die Übertragung auf Funktionen mehrerer reellen Variablen ist ohne weiteres möglich.

§ 1. **Ein Determinantensatz.**

Der Wert der Determinante mit dem allgemeinen Gliede

$$a_{ik} = \frac{1}{q_i + r_k}; \qquad (i, k = 1 \ldots n),$$

ist

$$D_n = \frac{\prod\limits^{\beta > \alpha} (q_\beta - q_\alpha)(r_\beta - r_\alpha)}{\prod (q_i + r_k)}; \quad (\alpha, \beta, i, k = 1 \ldots n).$$

Beweis: Wir fassen die Determinante als rationale Funktion aller q_i und r_k auf. Die gegebenen Faktoren werden dann im Zähler notwendig, im Nenner notwendig und hinreichend sein; aber auch der Gesamtgrad ist der richtige: $2\frac{n(n-1)}{2} - n^2 = -n$. Es kann sich daher der Wert der Determinante nur um einen konstanten Zahlenfaktor von dem gegebenen Ausdruck unterscheiden, und dieser Faktor kann daraufhin als 1 nachgewiesen werden: man braucht nur den Schluß von $n-1$ auf n zu vollziehen.

Für $n - 1 = 2$ ist die Richtigkeit des Satzes leicht zu verifizieren. Es sei nun

$$D_{n-1} = \frac{\prod\limits^{\beta > \alpha} (q_\beta - q_\alpha)(r_\beta - r_\alpha)}{\prod (q_i + r_k)}; \quad (\alpha, \beta, i, k = 1 \ldots n-1),$$

als richtig vorausgesetzt; wir wählen dann die absoluten Werte von q_n und r_n sehr groß, doch so, daß der absolute Wert von $q_n + r_n$ dagegen sehr klein ausfällt (z. B. $q_n + r_n = 1$). Das Hauptglied der entwickelten Determinante D_n ist dann $\frac{D_{n-1}}{q_n + r_n}$; ferner ist aber auch

$$D_n = c_n \cdot \frac{\prod\limits^{\beta > \alpha} (q_\beta - q_\alpha)(r_\beta - r_\alpha)}{\prod (q_i + r_k)}; \quad (\alpha, \beta, i, k = 1 \ldots n);$$

und

$$\frac{D_n}{D_{n-1}} = \frac{c_n}{q_n + r_n} \cdot \frac{\prod (q_n - q_\nu)(r_n - r_\nu)}{\prod (q_n + q_\nu)(r_n + r_\nu)}; \quad (\nu = 1 \ldots n-1).$$

Der Vergleich für beliebig wachsende Werte von q_n und r_n liefert also tatsächlich $c_n \doteq 1$.

Der gegebene Satz läßt sich sofort für den Fall $a_{ik} = \frac{s_i t_k a^{q_i^* + r_k^*}}{q_i + r_k + c}$ verallgemeinern, worin a und c beliebige Konstanten bedeutet, indem man z. B. $q_i + \frac{c}{2} = q_i'$, $r_k + \frac{c}{2} = r_k'$ setzt. Es ist daraufhin

$$D_n = a^{\Sigma q_i^* + \Sigma r_k^*} \cdot \prod s_i \cdot \prod t_k \cdot \frac{\prod\limits^{\beta > \alpha} (q_\beta - q_\alpha)(r_\beta - r_\alpha)}{\prod (q_i + r_k + c)};$$
$$(\alpha, \beta, i, k = 1 \ldots n)^{1)}.$$

§ 2. Beweis des Hauptsatzes.

Den folgenden Ausführungen möge das im Intervalle $-\pi \cdots + \pi$ vollständige orthogonale, normierte Funktionensystem der Fourierschen Reihen

$$\frac{1}{\sqrt{2\pi}}, \frac{1}{\sqrt{\pi}} \cos x, \cdots, \frac{1}{\sqrt{\pi}} \cos Nx, \ldots$$
$$\frac{1}{\sqrt{\pi}} \sin x, \cdots, \frac{1}{\sqrt{\pi}} \sin Nx, \ldots$$

zugrunde gelegt werden. Wir deuten jede dieser Grundfunktionen als Achse im Raume von unendlichvielen Dimensionen; alle in Betracht kommenden stetigen Funktionen sollen alsdann durch ihre Komponenten inbezug auf diese Achsen, d. h. durch die zugehörigen Fourierschen Konstanten charakterisiert werden.

Es sei zunächst $m > 1$ und allgemein $p_k > 1$ vorausgesetzt $(p_1 < p_2 < \cdots)$; von diesen scheinbaren Einschränkungen werden wir uns später befreien (v. u.).

Dann gelten im betrachteten Intervalle $-\pi \cdots + \pi$ die folgenden Fourierschen Entwicklungen

---

1) Andere Beweise des gleichen Lemmas bei Cauchy: „Exercices d'analyse et de phys. math.", II, S. 151—159; Rosenhain: „Schreiben an Jacobi über hyperelliptische Funktionen", Journal von Crelle, 40 (1850), S. 350—351, und Joachimsthal: „Über den Sturmschen Satz", ibd. 48 (1854), S. 414—415.

$$|x|^m = \varepsilon_m^{(0)} - \varepsilon_m^{(1)} \frac{\cos x}{1} \cdots - \varepsilon_m^{(N)} \frac{\cos Nx}{N} - \cdots$$

$$\cdots\cdots\cdots\cdots\cdots\cdots$$

$$|x|^{p_k} = \varepsilon_{p_k}^{(0)} - \varepsilon_{p_k}^{(1)} \frac{\cos x}{1} \cdots - \varepsilon_{p_k}^{(N)} \frac{\cos Nx}{N} - \cdots$$

$$\cdots\cdots\cdots\cdots\cdots\cdots$$

worin also allgemein

$$-\varepsilon_{p_k}^{(n)} = \frac{n}{\pi}\int_{-\pi}^{+\pi} |x|^{p_k} . \cos nx \, dx; \quad \varepsilon_{p_k}^{(0)} = \frac{1}{2\pi}\int_{-\pi}^{+\pi} |x|^{p_k} dx$$

ist, und alle entsprechenden Formeln auch für m anzusetzen sind. Man kann dafür auch schreiben

$$|x|^{p_k} = \varepsilon_{p_k}^{(1)} \cdot \frac{1-\cos x}{1} + \varepsilon_{p_k}^{(2)} \cdot \frac{1-\cos 2x}{2} \cdots + \varepsilon_{p_k}^{(N)} \cdot \frac{1-\cos Nx}{N} + \cdots.$$

Setzt man jetzt

$$\begin{aligned} \psi_{p_k}(x) &= p_k . |x|^{p_k - 1} \quad \text{für } x \geqq 0, \\ &= -p_k . |x|^{p_k - 1} \quad \text{für } x \leqq 0, \end{aligned}$$

so ist durch teilweise Integration

$$\varepsilon_{p_k}^{(n)} = \frac{1}{\pi}\int_{-\pi}^{+\pi} \psi_{p_k}(x) . \sin nx \, . \, dx.$$

Man hat also, da alle ψ_{p_k} stetige ungerade Funktionen ihres Argumentes bedeuten, identisch[1])

$$\begin{aligned} \sum_{n=1}^{\infty} \varepsilon_{p_k}^{(n)2} &= \frac{1}{\pi}\int_{-\pi}^{+\pi} \psi_{p_k}^2(x) . dx = \frac{2}{\pi}\int_0^{\pi} p_k^2 . x^{2p_k - 2} . dx \\ &= 2p_k^2 \cdot \frac{\pi^{2p_k - 2}}{2p_k - 1}; \end{aligned}$$

$$\sum_{n=1}^{\infty} \varepsilon_{p_i}^{(n)} \varepsilon_{p_k}^{(n)} = \frac{1}{\pi}\int_{-\pi}^{+\pi} \psi_{p_i}(x) . \psi_{p_k}(x) . dx = 2p_i p_k \cdot \frac{\pi^{p_i + p_k - 2}}{p_i + p_k - 1}.$$

Wir stellen uns nunmehr das Problem, unter allen Funktionen der ∞^N-Schar

$$R_N(x) = |x|^m - \alpha_1 |x|^{p_1} - \alpha_2 |x|^{p_2} \cdots - \alpha_N |x|^{p_N}$$

bei gegebenem N diejenige Funktion aufzusuchen, für welche das „Perpendikelquadrat"

1) Vgl. A. Hurwitz: „Über die Fourierschen Konstanten integrierbarer Funktionen", Math. Ann., Bd. 57 (1903), S. 425—446.

$$l_N^2 = \int_{-\pi}^{+\pi} R_N^2(x)\,.\,dx$$

ein Minimum sein wird. Dieses Minimum läßt sich nach bekannten Resultaten[1]) explizit angegeben, was auch direkt aus der Theorie der quadratischen Formen gefolgert werden kann. Es ist gleich dem Quotienten zweier symmetrischen Determinanten $\frac{Q_{N+1}}{S_N}$, worin das allgemeine Glied a_{ik} von S_N gegeben ist durch das Integral

$$a_{ik} = \frac{1}{\pi}\int_{-\pi}^{+\pi} |x|^{p_i}\,.\,|x|^{p_k}\,.\,dx = 2\cdot\frac{\pi^{p_i+p_k}}{p_i+p_k+1},$$

während Q_{N+1} aus S_N entsteht durch Ränderung mit

$$a_{0k} = a_{k0} = \frac{1}{\pi}\int_{-\pi}^{+\pi} |x|^{m}\,.\,|x|^{p_k}\,.\,dx = 2\cdot\frac{\pi^{m+p_k}}{m+p_k+1};$$

$$a_{00} = \frac{1}{\pi}\int_{-\pi}^{+\pi} |x|^{2m}\,dx = 2\cdot\frac{\pi^{2m}}{2m+1}.$$

Nach dem gegebenen Lemma (§ 1) lassen sich aber die beiden Determinanten explizit angeben. Es ist

$$S_N = 2^N\,.\,\pi^{2\Sigma p_\alpha}\cdot\frac{\prod\limits^{\beta>\alpha}(p_\beta-p_\alpha)^2}{\prod(p_i+p_k+1)}; \qquad (\alpha,\beta,i,k = 1\ldots N);$$

$$Q_{N+1} = 2^{N+1}\,.\,\frac{\pi^{2m+2\Sigma p_\alpha}}{2m+1}\cdot\frac{\prod\limits^{\beta>\alpha}(p_\beta-p_\alpha)^2\cdot\prod(p_\nu-m)^2}{\prod(p_i+p_k+1)\cdot\prod(p_\nu+m+1)^2}; \quad (\nu = 1\ldots N);$$

$$l_N^2 = \frac{Q_{N+1}}{S_N} = 2\,\frac{\pi^{2m}}{2m+1}\cdot\prod\frac{\left(1-\frac{m}{p_\nu}\right)^2}{\left(1+\frac{m+1}{p_\nu}\right)^2}.$$

Wenn also (bei festem m und $m \neq p_k$) $\sum\frac{1}{p_k}$ konvergiert, so nähert sich l_N^2 einer von 0 verschiedenen endlichen positiven Größe, die Approximation von x^m durch $x^{p_1}, x^{p_2}, \ldots x^{p_k}, \ldots$ ist nicht einmal im

1) Vgl. P. Gram: „Über die Entwicklung reeller Funktionen in Reihen mittels der Methode der kleinsten Quadrate", Journal f. d. reine u. angew. Mathematik, Bd. 94 (1883), S. 41—73.

E. Schmidt: „Entwicklung willkürlicher Funktionen nach Systemen vorgeschriebener", Math. Ann., Bd. 63 (1906), S. 433—476.

E. Schmidt: Über die Auflösung linearer Gleichungen mit unendlich vielen Unbekannten", Rendiconti d. C. M. di Palermo, Bd. 25 (1908), S. 53—77.

Gramschen Sinne des unbegrenzten Abnehmens des Integrals der Fehlerquadrate möglich, a fortiori also auch nicht im gewöhnlichen Sinne. Divergiert aber $\sum \frac{1}{p_k}$, so nähert sich l_N^2 beliebig der Null[1]), und es bleibt dann zu zeigen, daß sich auch für die Approximation im gewöhnlichen Sinne, d. h. für die absolute Größe des Fehlers an einer beliebigen Stelle des Intervalls, das gleiche erreichen läßt. Zu diesem Zwecke betrachten wir in der gleichen Weise wie vorhin das Minimum des Integrals

$$l_N'^2 = \frac{2}{\pi}\int_0^\pi R_N'^2(x)\,.\,dx = \frac{2}{\pi}\int_0^\pi \left\{mx^{m-1} - \alpha_1 p_1 x^{p_1-1} \cdots - \alpha_N p_N x^{p_N-1}\right\}^2 .\,dx.$$

Sowohl mit Hilfe der Komponenten $\varepsilon_{p_k}^{(n)}$ (v. o.) als auch direkt läßt sich dieses Minimum wiederum berechnen als ein Quotient zweier symmetrischen Determinanten $\frac{Q'_{n+1}}{S'_n}$. Das allgemeine Glied a'_{ik} von S'_n ist dabei

$$a'_{ik} = \sum_{n=1}^{\infty} \varepsilon_{p_i}^{(n)} \varepsilon_{p_k}^{(n)} = \frac{2}{\pi}\int_0^\pi p_i x^{p_i-1}\,.\,p_k x^{p_k-1}\,.\,dx = 2p_i p_k \cdot \frac{\pi^{p_i+p_k-2}}{p_i+p_k-1},$$

und Q'_{n+1} wird aus S'_n erhalten durch Ränderung mit

$$a'_{0k} = a'_{k0} = \sum_{n=1}^{\infty} \varepsilon_m^{(n)} \varepsilon_{p_k}^{(n)} = \frac{2}{\pi}\int_0^\pi mx^{m-1}\,.\,p_k x^{p_k-1}\,.\,dx = 2\,mp_k \cdot \frac{\pi^{m+p_k-2}}{m+p_k-1},$$

$$a'_{00} = \sum_{n=1}^{\infty} \varepsilon_m^{(n)2} = \frac{2}{\pi}\int_0^\pi m^2 x^{m-2}\,dx = 2m^2 \cdot \frac{\pi^{2m-2}}{2m-1}.$$

Es ist daher

$$l_N'^2 = \frac{Q'_{n+1}}{S'_n} = \frac{2m^2 \pi^{2m-2}}{2m-1} \prod \frac{\left(1-\frac{m}{p_k}\right)^2}{\left(1+\frac{m-1}{p_k}\right)^2}.$$

Wegen der vorausgesetzten Divergenz von $\sum \frac{1}{p_k}$ läßt sich für jedes positive δ^2 ein hinreichend hohes $N = N(\delta)$ so finden, daß für die zugehörigen Konstanten α_k

$$l_N'^2 = \frac{2}{\pi}\int_0^\pi \left\{m x^{m-1} - \alpha_1 p_1 x^{p_1-1} \cdots - \alpha_N p_N x^{p_N-1}\right\}^2 dx < \delta^2$$

1) Diese und die folgenden Ausführungen behalten ihre Gültigkeit auch dann, wenn m für $k > k_0$ größer als alle p_k vorausgesetzt wird, in welchem Falle natürlich eine endliche obere Grenze für die Exponenten existiert.

wird. Zerlegt man also

$$R_N(x) = |x|^m - \alpha_1 |x|^{p_1} \cdots - \alpha_N |x|^{p_N} = \delta_0 - \sum_1^\infty \delta_n \frac{\cos nx}{n} = \sum_1^\infty \delta_n \frac{1-\cos nx}{n},$$

so ist daraufhin

$$\sum_1^\infty \delta_n^2 = l_N'^2 < \delta^2,$$

also

$$R_N(x)^2 \leqq \sum_1^\infty \delta_n^2 \sum_1^\infty \left(\frac{1-\cos nx}{n}\right)^2 < \delta^2 \cdot \frac{4\pi^2}{6}; \quad |R_N(x)| < \frac{2\pi}{\sqrt{6}}\,\delta.$$

Eine noch bessere und direkte Abschätzung erhält man durch die bekannte Ungleichheitsbeziehung von Herrn H. A. Schwarz [1]), aus der als Spezialfall

$$[R_N(x)]^2 = \left\{\int_0^x R_N'(z)\,dz\right\}^2 \leqq \int_0^x dz \,.\int_0^x R_N'^2(z)\,.\,dz$$

gefunden werden kann. Es ist demnach im Intervalle $0 \ldots \pi$ beständig

$$|R_N(x)| < \frac{\pi\delta}{\sqrt{2}}.$$

Für wachsende positive p_k — und auch für $\lim p_k \neq 0$ überhaupt — ist auf diese Weise die Approximation von x^m durch die Folge der Potenzen x^{p_k} im Intervalle $0 \ldots \pi$ gewährleistet, sobald $\sum \frac{1}{p_k}$ divergiert. Durch die Substitution $x = \pi\zeta$ erreicht man das Gleiche für das Intervall $0 \ldots 1$, und durch die weitere Substitution $\zeta = \xi^n$ lassen sich auch die Einschränkungen $m > 1$, $p_k > 1$ aufheben: denn die Konvergenz bezw. Divergenz der unendlichen Reihe der reziproken Exponenten wird durch die fragliche Substitution natürlich nicht geändert.

Die Möglichkeit der Approximation einer beliebigen stetigen Funktion ist aber gesichert, sobald diese Möglichkeit für jede ganze positive Potenz x^n feststeht. Der Hauptsatz dieser Untersuchung ist damit vollständig bewiesen.

Es möge noch bemerkt werden, daß auch die Konstante 1 im Intervalle $0 \ldots \pi$ im Gramschen Sinne beliebige Approximation durch die x^{p_k} zuläßt, sobald $\sum \frac{1}{p_k}$ divergiert; aber eine Entwicklung von

1) Festschrift zu Ehren von K. Weierstraß, Mathem. Abhdl., Bd. I, S. 251.

der Form

$$1 = \varepsilon_0^{(1)} \frac{1-\cos x}{1} + \cdots + \varepsilon_0^{(N)} \frac{1-\cos n x}{n} + \cdots$$

existiert hier **nicht**, und natürlich ist auch die gewöhnliche Approximation im Punkte 0 hier unmöglich.

§ 3. Explizite Darstellung.

Man kann die Frage nach der besten Approximation von x^m durch eine vorgeschriebene Anzahl N von Potenzen x^{p_k} der gegebenen Folge aufwerfen. Für die gewöhnliche Approximation sind hier sehr komplizierte algebraische Gleichungen aufzulösen, sogar schon für $N = 1$. Dagegen läßt sich die gleiche Frage für die Gramsche Approximation vollständig beantworten, womit auch zugleich für den gewöhnlichen Fall brauchbare Formeln geliefert werden. Der Einfachheit halber möge hier die Lösung des Problems für das Intervall $0 \ldots 1$ gegeben werden.

Ist jetzt für die beliebige positive Potenz x^m bei vorgeschriebenem N

$$x^m \cong \alpha^{(N)} x^{p_1} + \alpha_2^{(N)} x^{p_2} \cdots + \alpha_N^{(N)} x^{p_N}$$

die verlangte beste Approximation, so muß bekanntlich[1]) die Restfunktion

$$R_N(x) = x^m - \alpha_1^{(N)} x^{p_1} - \alpha_2^{(N)} x^{p_2} \cdots - \alpha_N^{(N)} x^{p_N}$$

zu allen x^{p_k} im betrachteten Intervalle orthogonal sein.

Es ist daher identisch

$$\int_0^1 R_N(x) \,.\, x^{p_k} \,.\, dx = 0, \qquad (k = 1 \ldots N),$$

oder auch

$$\alpha_1^{(N)} \cdot \frac{1}{p_1 + p_k + 1} + \alpha_2^{(N)} \cdot \frac{1}{p_2 + p_k + 1} \cdots + \alpha_N^{(N)} \cdot \frac{1}{p_N + p_k + 1} = \frac{1}{m + p_k + 1},$$

folglich allgemein

$$(-1)^{i-1} \,.\, \alpha_i^{(N)} = \frac{\prod\limits^{\beta > \alpha} (p_\beta - p_\alpha)^2 \cdot \prod \frac{(m - p_\nu)}{(p_i - p_\nu)}}{\prod (p_\lambda + p_\mu + 1)^2 \cdot \prod \frac{(m + p_\nu + 1)}{(p_i + p_\nu + 1)}} : \frac{\prod\limits^{\beta > \alpha} (p_\beta - p_\alpha)^2}{\prod (p_\lambda + p_\mu + 1)^2}$$

$$= \prod \frac{(p_\nu - m)(p_i + p_\nu + 1)}{(p_\nu - p_i)(m + p_\nu + 1)}; \qquad (\alpha, \beta, \lambda, \mu, \nu = 1 \ldots N).$$

1) Vgl. die wiederholt zitierten Arbeiten von Gram und Schmidt.

Die so gewonnenen Formeln gelten auch für $p = 0$, ferner für alle positiven Potenzenfolgen x^{p_k} überhaupt. Für divergierende $\sum \frac{1}{p_k}$ wird hier ($\lim p_k \neq 0$) das Integral der Fehlerquadrate beliebig klein, es wachsen aber die Koeffizienten $\alpha_{p_i}^{(N)}$ für $p_i > m$ mit N über alle Grenzen, während sie für $p_i < m$ gegen Null konvergieren. Bei der Konvergenz von $\sum \frac{1}{p_k}$ verhalten sich dagegen die Approximationen umgekehrt: d. h. es läßt sich ein endlicher Fehler ε nicht im gesamten Intervall gleichmäßig unterbieten, während die Koeffizienten der Darstellung stark gegen bestimmte Werte konvergieren. Die Restfunktion $R_N(x)$ konvergiert dann gleichmäßig gegen eine bestimmte stetige Funktion $R(x)$, welche die beste Gramsche Approximation von x^m durch alle x^{p_k} darstellt.