

Series A

I. MATHEMATICA

375

**ZUR THEORIE DER GEWÖHNLICHEN
DIFFERENTIALGLEICHUNGEN**

VON

ALEXANDER DINGHAS

ROLF NEVANLINNA zum 70. Geburtstag

**HELSINKI 1966
SUOMALAINEN TIEDEAKATEMIA**

<https://doi.org/10.5186/aasfm.1966.375>

Am 14. Mai 1965 vorgelegt von P. J. MYRBERG und OLLI LEHTO

KESKUSKIRJAPAINO
HELSINKI 1966

Zur Theorie der gewöhnlichen Differentialgleichungen

1. *Einleitung.* Es bedeute zunächst O_x eine offene (nicht leere) Teilmenge der reellen Geraden

$$\mathbf{R}_x = \{ x \mid -\infty < x < +\infty \}$$

und

$$f = \{ f(x) \mid x \in O_x \}$$

eine reellwertige stetige Abbildung von O_x in \mathbf{R}_x . Man setze für $n = 1, 2, \dots$, $x \in O_x$ und hinreichend kleines $h > 0$

$$(1.1) \quad S_h^n f(x) = \frac{1}{(2h)^n} \sum_{k=0}^n (-1)^k \binom{n}{k} f(x + (n-2k)h).$$

Dann gilt, sofern f an der Stelle x eine n -te Ableitung $f^{(n)}(x)$ hat, die Gleichung ¹⁾

$$(1.2) \quad S^n f(x) = \lim_{h \rightarrow 0} S_h^n f(x) = f^{(n)}(x).$$

Im folgenden wird

$$(1.3) \quad \underline{S}^n f(x) = \lim_{h \rightarrow 0} S_h^n f(x)$$

und

$$(1.4) \quad \bar{S}^n f(x) = \overline{\lim}_{h \rightarrow 0} S_h^n f(x)$$

gesetzt und für $\underline{S}^n f(x) = \bar{S}^n f(x)$ (im Einklang mit (1.2)) $S^n f(x)$ geschrieben.

In der vorliegenden Arbeit wird zunächst die Frage untersucht, inwieweit die Ungleichungen

$$(1.5) \quad \underline{S}^n f(x) \leq g(x, f(x))$$

und

¹⁾ Man vgl. etwa [7], S. 12.

$$(1.6) \quad \bar{S}^n f(x) \geq g(x, f(x)),$$

wobei $g(x, y)$ reellwertig und stetig in $O = O_x \times R_y$ ist, die Existenz von $f^{(n)}(x)$ und das Bestehen der Differentialgleichung

$$(1.7) \quad f^{(n)}(x) = g(x, f(x))$$

zur Folge haben. Diese Frage kann hier nur für $n = 1$ und $n = 2$ behandelt werden. Eigentlich ist der grösste Teil der Arbeit dem Funktional

$$(1.8) \quad D^n(J, f) = \frac{1}{|b-a|^n} \sum_{k=0}^n (-1)^k \binom{n}{k} f \left(\left(1 - \frac{k}{n}\right)a + \frac{k}{n}b \right)$$

und seinen Anwendungen auf die Theorie der gewöhnlichen Differentialgleichungen gewidmet, wobei $J = [a, b]$ ein (beliebiges) Intervall von O_x bedeutet. Diese Intervallfunktion, obwohl für $n > 1$ nicht additiv, gestattet, bei geeigneter Zerlegung von J in $n+1$ kongruente, nicht unbedingt punktfremde Teilintervalle J_0, \dots, J_n , eine Darstellung von der Form

$$(1.9) \quad D^n(J, f) = \sum_{k=0}^n \mu_k(n) D^n(J_k, f)$$

mit $0 < \mu_k(n) < 1$ ($k = 0, 1, \dots, n$), $\mu_0(n) + \dots + \mu_n(n) = 1$ und somit die Bildung von Maximalketten, die für das Goursat-Verfahren charakteristisch sind.

2. Die Fälle $n = 1$ und $n = 2$. Nachfolgende Sätze können einfach bewiesen werden:

Satz 1. *Es sei $g(x, y)$ stetig in $O = O_x \times R_y$ und f stetig in O_x . Gilt dann*

$$(2.1) \quad \bar{S}^1 f(x) \leq g(x, f(x))$$

und

$$(2.2) \quad \underline{S}^1 f(x) \geq g(x, f(x))$$

in jedem Punkt von O_x , so ist f differenzierbar in O_x und

$$(2.3) \quad f'(x) = g(x, f(x)).$$

Aus dem Satz 1 folgt, wenn man dort $g = 0$ voraussetzt, der Satz:

Satz 2. *Ist f in einem offenen Intervall J_0 von R_x stetig und gilt dort*

$$(2.4) \quad \bar{S}^1 f(x) \leq 0$$

und

$$(2.5) \quad \underline{S^1} f(x) \geq 0,$$

so ist f eine Konstante.

Satz 3. Es sei f stetig in O_x , $g(x, y)$ stetig in O und

$$(2.6) \quad \begin{aligned} \underline{S^2} f(x) &\leq g(x, f(x)), \\ \bar{S}^2 f(x) &\geq g(x, f(x)) \end{aligned}$$

in jedem Punkt von O_x . Dann ist $f(x)$ zweimal stetig differenzierbar und genügt der Differentialgleichung

$$(2.7) \quad f''(x) = g(x, f(x)).$$

Nimmt man hier $g = 0$, so erhält man den Satz:

Satz 4. Ist f in einem offenen Intervall J_0 von \mathbf{R}_x stetig und gilt dort

$$(2.8) \quad \underline{S^2} f(x) \leq 0$$

und

$$(2.9) \quad \bar{S}^2 f(x) \geq 0,$$

so ist f eine lineare Funktion.

Ersetzt man hier die Ungleichungen (2.8) und (2.9) durch die Gleichung $\underline{S^2} f(x) = 0$ ($x \in J_0$), so erhält man ein klassisches Resultat von H. A. Schwarz²⁾.

Wir beweisen zunächst den Satz 2.

Es sei $J = [a, b]$ ein abgeschlossenes Teilintervall von J_0 . Man setze

$$(2.10) \quad h(x) = f(x) - f(a) + \varepsilon(x-a) + \varepsilon_1 \quad (\varepsilon, \varepsilon_1 > 0).$$

Dann ist wegen (2.4) und (2.5)

$$(2.11) \quad \underline{S^1} h(x) = \varepsilon > 0 \quad (x \in J).$$

Wäre nun $H = \{x \mid x \in J, h(x) < 0\}$ nicht leer, so müsste der Punkt $x_0 = \inf \{x \mid x \in H\}$ in $]a, b[$ liegen. Andererseits müsste mit Rücksicht auf (2.11) die Ungleichung $h(x_0 + \eta) > h(x_0 - \eta)$ für alle hinreichend kleinen $\eta > 0$ gelten. Das widerspricht jedoch der Definition von x_0 . Somit gilt $h \geq 0$ in J . Ähnlich zeigt man, dass die Funktion

$$(2.12) \quad f(x) - f(a) - \varepsilon(x-a) - \varepsilon_1$$

²⁾ Man vgl. [1], S. 81, [9], S. 258 ff., und [10], S. 341 ff.

in J nicht positiv sein kann. Durch Grenzübergang $\varepsilon \rightarrow 0$ und $\varepsilon_1 \rightarrow 0$ erhält man nun $f(x) = f(a)$ in J und durch (monotone) Ausschöpfung von J_0 durch abgeschlossene Intervalle $f(x) = f(a)$ in J_0 .

Zum Beweis des Satzes 4 bilde man mit Schwarz³⁾ die Funktionen

$$h_2(x) = f(x) - Lf(x) - \frac{\varepsilon}{2} (b-x)(x-a),$$

$$h_1(x) = f(x) - Lf(x) + \frac{\varepsilon}{2} (b-x)(x-a),$$

mit festen $a, b \in J_0$ ($a < b$), und

$$Lf(x) = \frac{b-x}{b-a} f(a) + \frac{x-a}{b-a} f(b).$$

Dann wird wegen (2.8) und (2.9)

$$(2.13) \quad \underline{S}^2 h_1(x) = \underline{S}^2 f(x) - \varepsilon \leq -\varepsilon < 0$$

und

$$(2.14) \quad \bar{S}^2 h_2(x) = \bar{S}^2 f(x) + \varepsilon \geq \varepsilon > 0.$$

Es sei jetzt

$$M = \sup \{ h_2(x) \mid x \in [a, b] \}$$

und

$$E_{ab} = \{ x \mid x \in [a, b], h_2(x) = M \}.$$

Dann hat E_{ab} die Eigenschaften:

1. $E_{ab} \neq \emptyset$. (Nach dem Satz von Weierstrass, wonach jede in einem abgeschlossenen Intervall stetige Funktion ein Maximum hat.)

2. $E_{ab} \cap]a, b[= \emptyset$. Gäbe es nämlich einen Punkt x_0 von E_{ab} in $]a, b[$, so müsste dieser Häufungspunkt von links und von rechts von Punkten von E_{ab} sein. Eine leichte Überlegung zeigt dann, dass (2.14) nicht gelten kann.

Somit enthält E_{ab} höchstens die Punkte a, b . Wegen $h_2(a) = h_2(b) = 0$ gilt also $h_2(x) \leq 0$ in $[a, b]$. Ein entsprechender Schluss liefert die Ungleichung $h_1(x) \geq 0$ in $[a, b]$. Lässt man nun $\varepsilon \rightarrow 0$ konvergieren, so erhält man $f(x) \leq Lf(x)$ und $f(x) \geq Lf(x)$, d. h. $f(x) = Lf(x)$. Mithin ist $f(x)$ eine lineare Funktion in jedem Intervall $[a, b]$ von J_0 , d. h. auch in J_0 .

³⁾ [10], S. 341 ff.

3. *Beweis der Sätze 1 und 3. Verallgemeinerungen.* Es sei $x_0 \in O_x$ und J_{x_0} ein offenes Intervall in O_x mit dem Mittelpunkt x_0 , derart dass die Differentialgleichung

$$(3.1) \quad y' = g(x, y)$$

in J_{x_0} eine Lösung $y = y(x)$ hat. Wir bilden die Funktion $q = q(x)$,

$$(3.2) \quad q(x) = \int_{x_0}^x \{g(t, f(t)) - g(t, y(t))\} dt,$$

und beachten, dass

$$(3.3) \quad S^1 q(x) = g(x, f(x)) - g(x, y(x))$$

gilt. Andererseits ist wegen (2.1), (2.2) und

$$(3.4) \quad S^1 y(x) = y'(x) = g(x, y(x)),$$

$$(3.5) \quad \bar{S}^1 (f(x) - y(x)) \leq g(x, f(x)) - g(x, y(x)),$$

$$\underline{S}^1 (f(x) - y(x)) \geq g(x, f(x)) - g(x, y(x)),$$

d. h.

$$(3.6) \quad \bar{S}^1 (f(x) - y(x) - q(x)) \leq 0$$

und

$$(3.7) \quad \underline{S}^1 (f(x) - y(x) - q(x)) \geq 0 \quad (x \in J_{x_0}).$$

Somit ist (nach dem Satz 2) $f - y - q$ eine Konstante, d. h. (wegen $f = y + q + \text{Konstante}$) f differenzierbar. Das beweist den Satz 1.

Der Satz 3 lässt sich analog beweisen. Wir setzen hier (unter Beibehaltung der Bedeutung von J_{x_0})

$$p(x) = \int_{x_0}^x (x-t) \{g(t, f(t)) - g(t, y(t))\} dt,$$

wobei y irgendeine (fest gewählte) Lösung der Differentialgleichung

$$(3.8) \quad y'' = g(x, y)$$

in J_{x_0} bedeutet, die durch x_0 geht. Wegen

$$p''(x) = g(x, f(x)) - g(x, y(x))$$

folgt, mit Rücksicht auf (2.6),

$$(3.9) \quad \underline{S}^2 (f(x) - y(x) - p(x)) \leq 0$$

und

$$(3.10) \quad \bar{S}^2(f(x) - y(x) - p(x)) \geq 0.$$

Somit hat $f - y - p$, nach dem Satz 4, die Form $Ax + B$ mit Konstanten A, B in J_{x_0} , und es gilt $f(x) = y(x) + p(x) + Ax + B$ in J_{x_0} . Das beweist die zweimalige Differenzierbarkeit von f und zugleich den Satz 3.

Die Sätze 1–4 können dahin verallgemeinert werden, dass man die Differentialgleichungen (3.1) und (3.8) durch entsprechende Systeme ersetzt. Im folgenden ersetzen wir diese Gleichungen durch

$$(3.11) \quad \eta' = g(x, \eta)$$

und

$$(3.12) \quad \eta'' = g(x, \eta)$$

und fassen η bzw. g als Vektoren des n -dimensionalen euklidischen Raumes E^n auf. Ist $\eta = (y_1, \dots, y_n)$ ein Punkt von E^n , so soll (3.11) bzw. (3.12) das Bestehen der Differentialgleichungssysteme

$$(3.13) \quad y'_k = g_k(x, y_1, \dots, y_n) \quad (k = 1, \dots, n)$$

bzw.

$$(3.14) \quad y''_k = g_k(x, y_1, \dots, y_n)$$

zum Ausdruck bringen.

Satz 5. Die Funktion $g(x, \eta)$ sei in $J \times E^n$, wobei J ein offenes, nicht leeres Intervall von R_x ist, definiert und stetig. Genügt dann die stetige Funktion $\bar{f} = (f_1, \dots, f_n)$ den Bedingungen:

- 1) $\bar{f} \in E^n$,
- 2) $\bar{S}^1 f_k(x) \leq g_k(x, f_1(x), \dots, f_n(x))$

und

- 3) $\underline{S}^1 f_k(x) \geq g_k(x, f_1(x), \dots, f_n(x))$

für alle $k = 1, \dots, n$ und alle $x \in J$, so ist \bar{f} differenzierbar und genügt der Differentialgleichung (3.11).

Satz 6. Gilt (unter denselben Voraussetzungen für \bar{f}, g) für $x \in J$

$$(3.15) \quad \underline{S}^2 f_k(x) \leq g_k(x, f_1(x), \dots, f_n(x))$$

und

$$(3.16) \quad \bar{S}^2 f_k(x) \geq g_k(x, f_1(x), \dots, f_n(x))$$

anstelle von 2) und 3), so ist \mathfrak{f} zweimal stetig differenzierbar in J , und es gilt

$$(3.17) \quad \mathfrak{f}''(x) = g(x, \mathfrak{f}(x))$$

in J .

Der Beweis dieser beiden Sätze läuft wie im Falle $n = 1$ und wird dem Leser überlassen.

Aus den Sätzen 1 und 3 folgen, wenn man dort die Funktion $g(x, y)$ unabhängig von y voraussetzt, die Sätze:

Satz 7. Die reellwertige Funktion f sei in $O \subseteq R_x$ stetig. Existiert dann $S^1 f(x)$ in O und ist dort stetig, so hat $f(x)$ eine Ableitung $f'(x)$ in O , und es gilt $f'(x) = S^1 f(x)$.

Satz 8. Die reellwertige Funktion f sei in $O \subseteq R_x$ stetig. Existiert dann $S^2 f(x)$ in O und ist dort stetig, so hat $f(x)$ eine zweite Ableitung $f''(x)$ gleich $S^2 f(x)$ in jedem Punkt von O .

Es sei in der Tat x_0 ein Punkt von O und J_{x_0} ein offenes Intervall um x_0 in O . Man setze

$$(3.18) \quad f^*(x) = \int_{x_0}^x S^1 f(t) dt \quad (x \in J_{x_0}).$$

Dann ist $f^{*'}(x) = S^1 f(x)$, also auch $S^1 f^*(x) = S^1 f(x)$. Somit ist $f^*(x) - f(x)$ eine Konstante. Das beweist den Satz 7.

Definiert man $f^*(x)$ durch die Gleichung

$$(3.19) \quad f^*(x) = \int_{x_0}^x (x-t) S^2 f(t) dt \quad (x \in J_{x_0}),$$

so erhält man, sofern f ein stetiges $S^2 f(x)$ in J_{x_0} hat, die Gleichung

$$f^{*''}(x) = S^2 f^*(x) = S^2 f(x) \quad (x \in J_{x_0}),$$

d. h.

$$S^2 (f(x) - f^*(x)) = 0 \quad \text{in } J_{x_0}.$$

Somit ist $f(x) - f^*(x)$ linear in J_{x_0} und somit f dort zweimal differenzierbar. Das beweist den Satz 8⁴⁾.

4. *Heranziehung des Funktional* $D^n(J, f)$. Zwischen den Funktionalen $S_h^n f(x)$ und $D^n(J, f)$ besteht ausser der Tatsache, dass beim ersteren Funktional der Mittelpunkt x bei Grenzübergängen fest

⁴⁾ Man vgl. in diesem Zusammenhang auch die Arbeiten von R. D. James, etwa [5]. Einen Einblick in diese Arbeit konnte ich allerdings nur über das Buch von R. L. Jeffery [6] erhalten.

bleibt, strukturell kein Unterschied. Vielmehr liegen die Vorteile des zweiten Funktional in der n -Additivität (1.9) desselben, die, wie bereits erwähnt, die Anwendbarkeit des Goursatschen Verfahrens ermöglicht.

Es bezeichne X eine nicht leere, konvexe Teilmenge eines (reellen oder komplexen) Banach-Raumes \mathbf{B}_0 , [3]. Es seien ferner $a_0, b_0, x \in \mathbf{B}_0$, $a_0 \neq b_0$ und

$$(4.1) \quad I_0 = \{ x \mid x = (1-\lambda) a_0 + \lambda b_0, 0 \leq \lambda \leq 1 \}$$

ein festes (Ausgangs-) Intervall von X . Wir verstehen unter $J = [a, b]$ ein (beliebiges) Teilintervall von I_0 mit dem Anfangspunkt $a = (1-\lambda') a_0 + \lambda' b_0$ und dem Endpunkt $b = (1-\lambda'') a_0 + \lambda'' b_0$ ($0 \leq \lambda' < \lambda'' \leq 1$) und definieren das in 1 eingeführte Funktional $D^n(J, f)$ folgendermassen:

Definition 1. *Es sei $f: I_0 \rightarrow \mathbf{B}$ eine stetige Abbildung von I_0 in einen Banach-Raum \mathbf{B} . Dann soll der Ausdruck*

$$(4.2) \quad D^n(J, f) = \frac{1}{\|J\|^n} \sum_{k=0}^n (-1)^k \binom{n}{k} f\left(\left(1 - \frac{k}{n}\right)a + \frac{k}{n}b\right)$$

die n -te Differenz von f auf J heissen. Dabei ist $\|J\| = \|b-a\|$ die Norm von $b-a$.

Ist $\|a_0 - b_0\| = 1$ und $a = (1-\lambda_a) a_0 + \lambda_a b_0$, $b = (1-\lambda_b) a_0 + \lambda_b b_0$ ($0 \leq \lambda_a < \lambda_b \leq 1$), so ist $\|J\| = \lambda_b - \lambda_a$. Für ein Intervall J_0 von \mathbf{R} und somit reellen a_0, b_0 wird $D^n(J, f)$ durch die Gleichung

$$(4.3) \quad D^n(J, f) = \frac{1}{(b-a)^n} \sum_{k=0}^n (-1)^k \binom{n}{k} f\left(\left(1 - \frac{k}{n}\right)a + \frac{k}{n}b\right)$$

gegeben.

Definition 2. *Es sei N eine (abzählbare) Basis des numerischen Intervalls $\Lambda = [0, 1]$. Dann soll $\mathfrak{R}(N)$ die Basis von I_0 bedeuten, deren Punkte x durch $(1-\lambda) a_0 + \lambda b_0$ ($\lambda \in N$) gegeben werden.*

Definition 3. *Unter $\{J_x\}$ wird im folgenden, bei gegebenem x in I_0 , die Gesamtheit aller (geordneten) Teilintervalle J_x mit Endpunkten in I_0 verstanden, die den Punkt x enthalten.*

Definition 4. *Gibt es ein Element $D^n f(x)$ von \mathbf{B} derart, dass*

$$(4.4) \quad \lim_{\|J_x\| \rightarrow 0} \|D^n f(x) - (-n)^n D^n(J_x, f)\| = 0$$

gilt, so soll $D^n f(x)$ die n -te Intervall-Derivierte auf I_0 von f in x heissen.

Ist $\mathbf{B}_0 = \mathbf{B} = \mathbf{R}$, also f eine reellwertige, stetige Abbildung von I_0 in \mathbf{R} , so kann man neben $D^n f(x)$ eine obere und eine untere n -te Intervall-Derivierte definieren.

Definition 5. Ist f eine stetige Abbildung von $I_0 \subseteq \mathbb{R}$ in \mathbb{R} und $x \in I_0$, so sollen die Grössen

$$(4.5) \quad \underline{D}^n f(x) = \lim_{\|J_x\| \rightarrow 0} (-n)^n D^n(J_x, f)$$

und

$$(4.6) \quad \bar{D}^n f(x) = \overline{\lim}_{\|J_x\| \rightarrow 0} (-n)^n D^n(J_x, f)$$

entsprechend die obere bzw. die untere n -te Intervall-Derivierte von f in x heissen⁵⁾.

5. Eine notwendige und hinreichende Bedingung für Polynom-Abbildungen. Wir betrachten zunächst das Intervall $J_0 = [0, 1]$ und nehmen als Basis die Menge N_0 aller dyadischen Brüche in J_0 , d. h. alle rationalen Zahlen von der Form

$$(5.1) \quad r = \frac{a_1}{2} + \frac{a_2}{2^2} + \dots + \frac{a_k}{2^k}$$

mit $a_r = 0$ oder $= 1$. Im folgenden werden zur Abschätzung der Grössen $\underline{D}^n f(x)$, $\bar{D}^n f(x)$ und $D^n f(x)$ lediglich Intervalle J_x verwendet werden, deren Endpunkte a, b dyadische Brüche sind.

Hilfssatz 1. Es bedeute f eine stetige Abbildung von $J = [0, 1]$ in \mathbb{R} mit der Eigenschaft $D^n f(x) = 0$ ($x \in J$). Dann ist f ein Polynom höchstens vom Grade $n-1$.

Allgemeiner gilt der Satz:

Hilfssatz 2. Es bedeute I_0 das Intervall (4.1) der konvexen Teilmenge X des Banach-Raumes \mathbf{B}_0 und $f: I_0 \rightarrow \mathbf{B}$ eine stetige Abbildung von I_0 in einen Banach-Raum \mathbf{B} . Gilt dann $D^n f(x) = 0$ in jedem Punkt $x = (1-\lambda)a_0 + \lambda b_0$ ($0 \leq \lambda \leq 1$) von I_0 , so hat f die Form

$$(5.2) \quad 1 f_0 + \lambda f_1 + \dots + \lambda^{n-1} f_{n-1}$$

mit f_0, f_1, \dots, f_{n-1} aus \mathbf{B} .

Da für $f = f(x)$, $x = x(\lambda) = (1-\lambda)a_0 + \lambda b_0$ und $f(x(\lambda)) = 1 f_0 + \lambda f_1 + \dots + \lambda^{n-1} f_{n-1}$ (f_0, f_1, \dots, f_{n-1} fest in \mathbf{B}) stets $D^n f(x) = 0$ gilt, ist die Bedingung $D^n f(x) = 0$ notwendig und hinreichend dafür,

⁵⁾ Der Faktor $(-n)^n$ fällt fort, wenn man die Summe rechts in (4.3) durch $(a-b)^n/n^n$ statt durch $(b-a)^n$ dividiert. Diese Definition, d. h. die Heranziehung des Funktionals $\Delta^n(J, f) = (-n)^n D^n(J, f)$, entspricht mehr der Definition von $S_h^n f(x)$. In den nachfolgenden Entwicklungen wird das Funktional $D^n(J, f)$ lediglich aus technischen Gründen vorgezogen.

dass f auf I_0 ein Polynom in λ mit Koeffizienten aus \mathbf{B} ist. Entsprechendes gilt für den numerischen Fall des Hilfssatzes 1.

Zum Beweis des Hilfssatzes 1 bilden wir neben der Grösse

$$F(J) = \sum_{k=0}^n (-1)^k \binom{n}{k} f\left(\frac{k}{n}\right)$$

die Grössen

$$(5.3) \quad F(J_k) = \sum_{l=0}^n (-1)^l \binom{n}{l} f\left(\frac{k+l}{2n}\right)$$

mit

$$J_k = \left[\frac{k}{2n}, \frac{k}{2n} + \frac{1}{2} \right] \quad (k = 0, 1, \dots, n)$$

und beachten, dass dann

$$D^n(J_k, f) = \frac{F(J_k)}{|J_k|^n} = 2^n \frac{F(J_k)}{|J|^n}.$$

Nun ist, wenn $q = k+l$ gesetzt wird,

$$\sum_{k=0}^n \binom{n}{k} F(J_k) = \sum_{q=0}^{2n} \left\{ \sum_{l=0}^n (-1)^l \binom{n}{q-l} \binom{n}{l} f\left(\frac{q}{2n}\right) \right\}$$

und somit mit Rücksicht auf die bekannte Identität

$$\sum_{l=0}^n (-1)^l \binom{n}{q-l} \binom{n}{l} = \begin{cases} 0 & (q \text{ ungerade}) \\ (-1)^{\frac{q}{2}} \binom{n}{\frac{q}{2}} & (q \text{ gerade}) \end{cases}$$

der Kombinatorik ⁶⁾

$$(5.4) \quad \sum_{k=0}^n \binom{n}{k} F(J_k) = \sum_{k=0}^n (-1)^k \binom{n}{k} f\left(\frac{k}{n}\right) = F(J).$$

Es ist also

$$(5.5) \quad D^n(J, f) = \sum_{k=0}^n \mu_k(n) D^n(J_k, f)$$

mit

$$(5.6) \quad \mu_k(n) = \frac{1}{2^n} \binom{n}{k} \quad (k = 0, 1, \dots, n).$$

⁶⁾ Man vgl. [2] sowie auch [8], S. 256.

Hat man anstelle von $J = [0, 1]$ das Intervall I_0 , so ist mit $I = [a, b] \subseteq I_0$ und

$$D^n(I, f) = \frac{1}{\|I\|^n} \sum_{k=0}^n (-1)^k \binom{n}{k} f\left(\left(1 - \frac{k}{n}\right)a + \frac{k}{n}b\right)$$

sowie

$$D^n(I_k, f) = \frac{1}{\|I_k\|^n} \sum_{l=0}^n (-1)^l \binom{n}{l} f\left(a + \frac{k+l}{2n}(b-a)\right),$$

$$(5.7) \quad D^n(I, f) = \sum_{k=0}^n \mu_k(n) D^n(I_k, f).$$

Aus (5.5) und (5.7) erhält man nun wegen $\mu_0(n) + \dots + \mu_n(n) = 1$ die Ungleichungen

$$(5.8) \quad |D^n(J, f)| \leq \max_{0 \leq k \leq n} |D^n(J_k, f)|$$

und

$$(5.9) \quad \|D^n(I, f)\| \leq \max \|D^n(I_k, f)\|.$$

Liefert f eine Abbildung von $J = [0, 1]$ in \mathbf{R} , so findet man entsprechend

$$(5.10) \quad \min_{0 \leq k \leq n} D^n(J_k, f) \leq D^n(J, f) \leq \max_{0 \leq k \leq n} D^n(J_k, f).$$

Mit Hilfe von (5.5) kann nun gezeigt werden, dass $F(J)$ gleich Null sein muss, falls $D^n f(x)$ (unter Zugrundelegung von N_0) in jedem Punkt x von $J = [0, 1]$ verschwindet.

Es bezeichne in der Tat J'_1 ein Intervall aus der Menge $\{J_0, J_1, \dots, J_n\}$, für das

$$(5.11) \quad D^n(J'_1, f) = \max_{0 \leq k \leq n} D^n(J_k, f)$$

ist. Wir nehmen dieses Intervall zum Ausgangsintervall und bilden die Kette J'_1, J'_2, \dots (kurz: (J'_σ)) durch folgendes Verfahren:

Jedes J'_σ ($\sigma \geq 2$) gehört der Unterteilungsmenge von $J'_{\sigma-1}$ in kongruente Teilintervalle nach dem Vorgang $J \rightarrow \{J_0, \dots, J_n\}$ an und genügt der entsprechenden Bedingung (5.11). Danach gilt stets $J'_{\sigma+1} \subseteq J'_\sigma$ ($\sigma = 1, 2, \dots$), $2|J'_{\sigma+1}| = |J'_\sigma|$ und

$$(5.12) \quad D^n(J'_\sigma, f) \leq D^n(J'_{\sigma+1}, f) \quad (\sigma = 1, 2, \dots).$$

Man setze jetzt $x'_0 = \bigcap_{\sigma=1}^{\infty} J'_\sigma$. Dann ist $x'_0 \in J$, und es gilt

$$(5.13) \quad D^n(J, f) \leq \lim_{\sigma \rightarrow +\infty} D^n(J'_\sigma, f) = 0.$$

Ersetzt man in (5.11) max durch min, so erhält man eine Folge (J''_σ) von ineinandergeschachtelten Intervallen mit

$$(5.14) \quad D^n(J''_\sigma, f) \geq D^n(J''_{\sigma+1}, f)$$

anstelle von (5.12). Wird dann noch $x''_0 = \bigcap_{\sigma=1}^{\infty} J''_\sigma$ gesetzt, so wird

$$(5.15) \quad D^n(J, f) \geq \lim_{\sigma \rightarrow +\infty} D^n(J''_\sigma, f) = 0$$

und somit auch $D^n(J, f) = 0$.

Durch ein entsprechendes Verfahren, das zu den Ungleichungen

$$\|D^n(I, f)\| \leq \lim_{\sigma \rightarrow \infty} \|D^n(I'_\sigma, f)\| = n^{-n} \|D^n f(x'_0)\| = 0$$

(bei entsprechender Definition und Bedeutung der Intervalle I'_σ) führt, erhält man folgendes Teilergebnis in Zusammenhang mit dem Beweis des Hilfssatzes 1:

Gilt $D^n f(x) = 0$ in jedem Punkt von I , so gilt $F(I) = 0$. Dabei ist I ein beliebiges Teilintervall von I_0 .

6. *Ende des Beweises der Hilfssätze 1 und 2.* Der Beweis des Hilfssatzes 1 lässt sich nun folgendermassen zu Ende führen: Man gehe von den Gleichungen

$$F(J_k) = \sum_{l=0}^n (-1)^l \binom{n}{l} f\left(\frac{k+l}{2n}\right) = 0 \quad (k = 0, 1, 2, \dots, n)$$

aus und bilde die Summe

$$(6.1) \quad \sum_{k=0}^n (-1)^k \alpha_k F(J_k) = 0$$

mit zunächst willkürlichen reellen α_k . Daraus folgt, wenn man

$$\gamma_s = \sum_{k=0}^n \binom{n}{s-k} \alpha_k$$

setzt,

$$(6.2) \quad \sum_{s=0}^{2n} (-1)^s f\left(\frac{s}{2n}\right) \gamma_s = 0.$$

Man betrachte jetzt die Punkte

$$x_k = \frac{k}{n} \quad (k = 0, 1, \dots, n-1)$$

von J , setze

$$(6.3) \quad \psi_k(x) = \prod_{v \neq k} (x - x_v) \quad (k = 0, 1, \dots, n-1)$$

und bilde das Polynom

$$(6.4) \quad P_{n-1}(x) = \sum_{k=0}^{n-1} \frac{\psi_k(x)}{\psi_k(x_k)} f\left(\frac{k}{n}\right).$$

Ist dann

$$\bar{f}(x) = f(x) - P_{n-1}(x) \quad (x \in J),$$

so gilt offenbar, wegen $P_{n-1}^{(n)}(x) \equiv 0$ in J ,

$$\bar{F}(J) = \bar{F}(J_1) = \dots = \bar{F}(J_n) = 0,$$

sofern $\bar{F}(J), \bar{F}(J_1), \dots, \bar{F}(J_n)$ die mit Hilfe von \bar{f} gebildeten Grössen $F(J), F(J_1), \dots, F(J_n)$ bedeuten.

Nun folgt aus $\bar{F}(J) = 0$ und

$$(6.5) \quad \begin{aligned} \bar{f}\left(\frac{k}{n}\right) &= 0 & (k = 0, 1, 2, \dots, n-1), \\ \bar{f}\left(\frac{s}{2n}\right) &= 0 & (s = 0, 2, 4, \dots, 2n), \end{aligned}$$

und somit gilt

$$(6.6) \quad \sum_{s=1}^n \bar{f}\left(\frac{2s-1}{2n}\right) \gamma_{2s-1} = 0.$$

Es soll nun gezeigt werden, dass man durch geeignete Wahl der α_k sämtliche Werte $\bar{f}((2s-1)/(2n))$ zum Verschwinden bringen kann.

Hilfssatz 3. Die Determinante des Gleichungssystems

$$(6.7) \quad \sum_{k=0}^{n-1} \binom{n}{2s-k-1} \alpha_k = 0 \quad (s = 1, 2, \dots, n)$$

ist von Null verschieden.

Der Beweis dieses Hilfssatzes stützt sich weitgehend auf folgendes berühmte Kriterium von Hurwitz ⁷⁾ über Polynome, deren Nullstellen einen negativen Realteil besitzen:

Es sei

$$(6.8) \quad P(x) = a_0 x^n + a_1 x^{n-1} + \dots + a_n \quad (a_0 = 1, a_n \neq 0)$$

ein Polynom n -ten Grades mit reellen Koeffizienten. Man setze $\Delta_1 = a_1$ und definiere für $\lambda = 2, 3, \dots, n$ die Determinanten Δ_λ durch die Gleichung

⁷⁾ Man vgl. [4] und [11], S. 335 ff.

$$(6.9) \quad \Delta_\lambda = \begin{vmatrix} a_1 & a_0 & 0 & 0 & 0 & \dots & 0 \\ a_3 & a_2 & a_1 & a_0 & 0 & \dots & 0 \\ a_5 & a_4 & a_3 & a_2 & a_1 & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{2\lambda-1} & a_{2\lambda-2} & a_{2\lambda-3} & a_{2\lambda-4} & a_{2\lambda-5} & \dots & a_\lambda \end{vmatrix}.$$

Dann haben sämtliche Wurzeln von (6.8) dann und nur dann negativen Realteil, wenn alle Determinanten $\Delta_1, \Delta_2, \dots, \Delta_n$ positiv sind.

Ausführlich geschrieben lauten die Determinanten $\Delta_1, \Delta_2, \Delta_3, \dots$

$$a_1, \quad \begin{vmatrix} a_1 & a_0 \\ a_3 & a_2 \end{vmatrix}, \quad \begin{vmatrix} a_1 & a_0 & 0 \\ a_3 & a_2 & a_1 \\ a_5 & a_4 & a_3 \end{vmatrix}, \quad \begin{vmatrix} a_1 & a_0 & 0 & 0 \\ a_3 & a_2 & a_1 & a_0 \\ a_5 & a_4 & a_3 & a_2 \\ a_7 & a_6 & a_5 & a_4 \end{vmatrix}, \quad \dots$$

Mit Hilfe dieses Kriteriums, für dessen Beweis wir auf die Hurwitzsche Abhandlung verweisen, kann nun der Hilfssatz 3 ohne Schwierigkeit bewiesen werden:

Man betrachte in der Tat das Polynom

$$(6.10) \quad (x+1)^n = x^n + \binom{n}{1} x^{n-1} + \binom{n}{2} x^{n-2} + \dots + \binom{n}{n}$$

und beachte, dass dieses die (negative) Wurzel -1 als n -fache Wurzel hat. Daraus folgt, dass die mit Hilfe der Koeffizienten $\binom{n}{0}, \binom{n}{1}, \dots, \binom{n}{n}$ gebildeten Determinanten Δ_λ ($\lambda = 1, 2, \dots, n$) positiv sind. Da jetzt

$$(6.11) \quad \Delta_n = \left| \binom{n}{2s-k-1} \right|$$

ist, so ist $\Delta_n > 0$ und somit $\neq 0$.

Ist nun die durch (6.9) definierte Grösse Δ_n ungleich Null, so müssen alle $\bar{f}((2s-1)/(2n))$ ($s = 1, 2, \dots, n$) verschwinden. Denn für jedes (feste) $\mu = 1, 2, \dots, n$ ist das System

$$(6.12) \quad \sum_{k=0}^n \binom{n}{2s-k-1} \alpha_k = \delta_{s\mu} \quad (s = 1, 2, \dots, n)$$

(wobei $\delta_{s\mu}$ das bekannte Kronecker-Symbol bedeutet) lösbar, und man kann durch die Wahl der α_k ($\alpha_k = \alpha_k(\mu)$) erreichen, dass jedes $\bar{f}((2s-1)/(2n))$ ($s = 1, 2, \dots, n$) verschwindet. Somit ist

$$(6.13) \quad \bar{f}\left(\frac{k}{2n}\right) = 0 \quad (k = 0, 1, \dots, 2n).$$

Wiederholt man dieses Verfahren q -mal, so erhält man die Gleichungen

$$(6.14) \quad \bar{f}\left(\frac{k}{2^q n}\right) = 0 \quad (k = 0, 1, 2, \dots, 2^q n).$$

Danach verschwindet $\bar{f}(x)$ auf N_0 und mithin (da f noch stetig ist) auch auf $J = [0, 1]$.

Der Beweis des Hilfssatzes 2 verläuft analog und wird hier nicht explizit durchgeführt.

7. *Anwendung auf die Theorie der gewöhnlichen Differentialgleichungen.*
Die Anwendung der Hilfssätze 1 und 2 auf die Differentialgleichungen

$$(7.1) \quad D^n f(x) = g(x, f(x)),$$

bei geeignetem Wertevorrat von f und g , bringt den Nachweis, dass $D^n f(x)$, sofern $f(x)$ eine n -te Ableitung $f^{(n)}(x)$ hat, gleich $f^{(n)}(x)$ ist. Dabei wird angenommen, dass x in einem Intervall $J_0 = [a_0, b_0]$ ($a_0 < b_0$) von \mathbf{R} variiert, und dass f in einem Banach-Raum \mathbf{B} liegt. Grundlegend für nachfolgende Entwicklungen ist der Satz:

Hilfssatz 4. *Hat die stetige Abbildung $f: J_0 \rightarrow \mathbf{R}$ eine n -te Ableitung $f^{(n)}(x_0)$ in einem Punkt $x_0 \in]a_0, b_0[$, so existiert $D^n f(x_0)$, und es gilt,*

$$(7.2) \quad D^n f(x_0) = f^{(n)}(x_0).$$

Beweis. Für $m \geq 0$, ganz, ist

$$A_m = \sum_{k=0}^n (-1)^k \binom{n}{k} \frac{k^m}{n^m} = \frac{1}{n^m} \frac{d^m}{d\alpha^m} (1 - e^\alpha)^n \Big|_{\alpha=0},$$

und somit

$$A_m = \begin{cases} 0 & (0 \leq m < n) \\ (-1)^n \frac{n!}{n^n} & (m = n) \end{cases}.$$

Daraus folgt für $J_{x_0} = [a, b] \subset]a_0, b_0[$ ($a < b$, $x_0 \in J_{x_0}$), wegen

$$(b-a)^n D^n(J_{x_0}, x^m) = \sum_{k=0}^n (-1)^k \binom{n}{k} \left[a + (b-a) \frac{k}{n} \right]^m$$

also

$$(7.3) \quad \begin{aligned} D^n(J_{x_0}, x^m) &= \sum_{\lambda=0}^m \binom{m}{\lambda} A_\lambda a^{m-\lambda} (b-a)^{\lambda-n}, \\ D^n x^m &= (x^m)^{(n)}. \end{aligned}$$

Hat nun f in $x_0 \in]a_0, b_0[$ eine n -te Ableitung $f^{(n)}(x_0)$, so gilt nach der Taylor-Formel für alle J_{x_0} mit hinreichend kleiner Länge $|J_{x_0}|$

$$(7.4) \quad f(x) = P_{n-1}(x) + (x-x_0)^n \left\{ \frac{f^{(n)}(x_0)}{n!} + o(1) \right\},$$

wobei $P_{n-1}(x)$ ein Polynom vom Grade $\leq n-1$ ist und $|o(1)|$ mit $|J_{x_0}| \rightarrow 0$ gegen Null strebt. Das liefert in Verbindung mit (7.3) die Gleichung (7.2).

Der Hilfssatz 4 gilt, in ähnlicher Form, auch dann, wenn f eine stetige Abbildung von $I_0 \subseteq B_0$ in einen Banach-Raum B liefert und $f = f((1-\lambda)a_0 + \lambda b_0)$ als Funktion von λ in einem Punkt λ_0 von $(0, 1)$ eine n -te Derivierte in gewöhnlichem Sinne besitzt.

Die Übertragung der Hauptergebnisse dieser Arbeit⁸⁾ auf partielle Differentialgleichungen von der Form

$$\frac{\partial^{m+n} V}{\partial x^m \partial y^n} = f(x, y; V)$$

(und entsprechend für mehrere Veränderliche) geschieht durch Heranziehung der Entwicklungskoeffizienten von $(1-x)^m (1-y)^n$ und Anwendung eines entsprechend verallgemeinerten Goursatschen Zerstückelungsverfahrens. Die ausführliche Darstellung der gewonnenen Zusammenhänge soll später gegeben werden.

Freie Universität Berlin
Deutschland

⁸⁾ Auf eine Reihe von Ergebnissen »in der Norm«, deren Existenz ich durch Herrn Professor P. L. BUTZER (Aachen) während einer Tagung über Funktionalanalysis im August 1965 in Oberwolfach erfahren habe, will ich hier um so weniger eingehen, als dadurch Hauptziel der vorliegenden Arbeit, die zunächst ohne jeglichen Zusammenhang mit vorhandener Literatur geschrieben wurde, nicht allzu sehr berührt wird.

Literatur

- [1] CANTOR, G.: Beweis, dass eine für jeden reellen Wert von x durch eine trigonometrische Reihe gegebene Funktion $f(x)$ sich nur auf eine einzige Weise in dieser Form darstellen lässt. - Gesammelte Abhandlungen, Verlag von Julius Springer, Berlin, 1932, S. 80–83.
- [2] HAGEN, J. G.: Synopsis der höheren Mathematik. I. - Dames Verlag, Berlin, 1891.
- [3] HILLE, E., und R. S. PHILLIPS: Functional analysis and semi-groups. - Colloquium Publications 31, American Mathematical Society, Providence (R. I.), 1957. [Zweite revidierte Auflage des gleichnamigen Buches von E. HILLE.]
- [4] HURWITZ, A.: Ueber die Bedingungen, unter welchen eine Gleichung nur Wurzeln mit negativen reellen Theilen besitzt. - Math. Ann. 46, 1895, S. 273–284.
- [5] JAMES, R. D.: A generalized integral. II. - Canad. J. Math. 2, 1950, S. 297–306.
- [6] JEFFERY, R. L.: Trigonometric series. - Canadian Mathematical Congress Lecture Series 2, University of Toronto Press, Toronto, 1956.
- [7] LORENTZ, G. G.: Bernstein polynomials. - Mathematical Expositions 8, University of Toronto Press, Toronto, 1953.
- [8] NETTO, E.: Lehrbuch der Combinatorik. - Druck und Verlag von B. G. Teubner, Leipzig/Berlin, 1927.
- [9] RIEMANN, B.: Ueber die Darstellbarkeit einer Funktion durch eine trigonometrische Reihe. - Gesammelte mathematische Werke und wissenschaftlicher Nachlass, Zweite Auflage, Druck und Verlag von B. G. Teubner, Leipzig, 1892, S. 227–265.
- [10] SCHWARTZ, H. A.: Beweis eines für die Theorie der trigonometrischen Reihen in Betracht kommenden Hilfssatzes. - Gesammelte mathematische Abhandlungen. II, Verlag von Julius Springer, Berlin, 1890, S. 341–343.
- [11] WEBER, H.: Lehrbuch der Algebra. I. - Zweite Auflage, Druck und Verlag von Friedrich Vieweg und Sohn, Braunschweig, 1898.

