Werk

Titel: Mathematische Zeitschrift

Ort: Berlin

Jahr: 1942

Kollektion: Mathematica

Werk Id: PPN266833020_0048

PURL: http://resolver.sub.uni-goettingen.de/purl?PID=PPN266833020_0048|LOG_0048

Die Kettenregel für höhere Ableitungen.

Von

U. T. Bödewadt in Göttingen.

Wenn in einer Funktion die unabhängige Veränderliche ihrerseits doch wiederum von einer anderen Veränderlichen abhängt:

$$y = F(x); \quad x = \varphi(t); \quad y = F(\varphi(t)) = f(t), \tag{1}$$

so ist damit eine unmittelbare oder zusammengesetzte Funktion dieser letzten Veränderlichen gegeben. Ihre Ableitung — hier und im folgenden soll vorausgesetzt sein, daß alle auftretenden Ableitungen vorhanden und stetig sind — setzt sich aus den Ableitungen der beiden einfachen Funktionen nach der Regel $f'(t) = F'(x) \cdot \varphi'(t)$ zusammen, wobei man rechter Hand für x wieder den Wert $x = \varphi(t)$ zu nehmen hat. Verwendet man zur Bezeichnung der Ableitung einer Funktion nach ihrem unmittelbaren Argument das Zeichen D, so lautet die Regel:

$$Df = DF \cdot D\varphi. \tag{2}$$

Die Ausdehnung dieser Regel auf eine mehrfach zusammengesetzte Funktion

$$z_0 \equiv t; \quad z_\nu = F_\nu(z_{\nu-1}) \qquad [\nu = 1, 2, \ldots N-1]; \tag{3}$$
$$y = F_N(z_{N-1}) = \Big(F_N\big(F_{N-1}(\ldots F_2(F_1(t))\ldots)\big)\Big) = f(t),$$

eine sogenannte Funktionskette, erfordert nur mehrfache Anwendung von (2) und gibt die bekannte Kettenregel

$$Df = DF_1 \cdot DF_2 \cdot \ldots \cdot DF_N. \tag{4}$$

Da man von einer richtigen Kette eigentlich nur sprechen kann, wenn sie außer dem Anfangs- und dem Endglied noch wenigstens ein Zwischenglied enthält, so müßte der Name „Kettenregel" eigentlich den Fällen $n \geqq 3$ vorbehalten bleiben. Er wird freilich auch oft für (2) gebraucht, und nicht ohne Recht: Es bedarf kaum einer Überlegung, wie die Regel (2) für den Fall (3) zu verallgemeinern sei, die Antwort (4) bietet sich sozusagen von selbst dar.

Mit den höheren Ableitungen verhält es sich hingegen anders: Hier scheint die Kettenregel bislang noch nicht angegeben worden zu sein. Der Grund dafür liegt vermutlich im Mangel an geeigneten Bezeichnungen, die sich leicht bei mehreren Funktionen gleichzeitig anwenden lassen und offene Zahlenfaktoren in den Formeln vermeiden. Mit der in dieser Mitteilung vorgeschlagenen Bezeichnung durch Operatoren höheren Grades ist nun die Aufstellung der Kettenregel (im oben erläuterten Sinne: für mehrfach zusammengesetzte Funktionen) in übersichtlicher Form möglich. — Um des Zusammenhanges willen möge es gestattet sein, auch Bekanntes zu wiederholen[1]). Dafür sind die Beweise nur in ihren Leitgedanken angegeben worden; es wird an Hand dieser Andeutungen keine besonderen Schwierigkeiten bereiten, das Fehlende zu ergänzen. Solche Kürze ist wohl um so eher erlaubt, als durch die Arbeiten von AMALDI [2]) und MAMBRIANI [3]) alle Wünsche nach ausführlicher und gründlicher Behandlung der höheren Ableitungen einfach zusammengesetzter Funktionen auch von mehreren Veränderlichen als erfüllt gelten können. — Inhaltlich neu sind im folgenden die Formeln: (10) (für $n > 1$) und (11); vielleicht (12) bis (14), (17) bis (20), (21); ferner (26) (für $n > 1$), (34), (38) (für $n > 1$), (39), (41) und (44).

1.

Auch die höheren Ableitungen einer zusammengesetzten Funktion (1) lassen sich durch wiederholte Anwendung der Regel (2) bilden. Dabei zeigt sich, daß die k-te Ableitung der Funktion $f(t)$ für die ersten Werte von k die Gestalt

$$D^k f = \sum_{n=1}^{k} D^n F \cdot D_{kn} \varphi \tag{5}$$

hat. Bildet man hier beiderseits die Ableitung, so sieht man, daß diese Form für alle k gültig bleibt, wenn man für die Ausdrücke $D_{kn}\varphi$ die Rekursionsformel

$$D_{kn}\varphi = D(D_{k-1,n}\varphi) + D\varphi \cdot D_{k-1,n-1}\varphi \quad [k > n > 1] \tag{6}$$

mit den Ausgangswerten

$$D_{k0}\varphi = 0 \;\; [k > 0]; \;\; D_{k1}\varphi = D^k\varphi \;\; [k \geqq 0]; \;\; D_{kk}\varphi = (D\varphi)^k \;\; [k \geqq 0] \tag{7}$$

ansetzt. Ohne das erste Glied auf der rechten Seite gilt die Formel (6) noch für $k = n$, bei Fortlassen des zweiten Gliedes für $n = 1$. — D_{kn} ist also ein Differentialoperator von der Ordnung k (in bezug auf die unabhängige Veränderliche) und vom Grade n (in der abhängigen Veränderlichen); linear ist

---

[1]) Vgl. die am Schlusse aufgeführte Literatur, besonders die zusammenfassenden Darstellungen in [2], [3], [7].

[2]) L [4].

[3]) L [7]. — Von den Abhandlungen AMALDIs und MAMBRIANIs erhielt Verfasser erst nach Abschluß der vorliegenden Untersuchungen Kenntnis.

nur D_{k1}. Zur unmittelbaren Berechnung der D_{kn} dient die Formel von FAÀ DI BRUNO

$$(8) \qquad D_{kn}\,\varphi(t) = \sum_{(\lambda)} k! \prod_{\nu=1}^{1+k-n} \frac{1}{\lambda_\nu!}\left(\frac{D^\nu\,\varphi(t)}{\nu!}\right)^{\lambda_\nu};$$

die Summe ist zu erstrecken über alle Zusammenstellungen $(\lambda) = (\lambda_1, \ldots \lambda_\nu, \ldots)$ nichtnegativer ganzer Zahlen $\lambda_\nu \geqq 0$, welche die beiden Bedingungen

$$(9) \qquad \Sigma\,\lambda_\nu = n; \quad \Sigma\,\nu\cdot\lambda_\nu = k$$

erfüllen. Der Beweis für diese Formel ergibt sich am schnellsten mit Hilfe der Polynomialformel, indem man in

$$f(t+\tau) = F(\varphi(t+\tau))$$

F und φ (also auch f) als in ihre Taylorreihen entwickelbar annimmt — diese Annahme ist erlaubt, da die algebraische Gestalt der Formel von den besonderen Eigenschaften von F und φ nicht abhängt — und die Koeffizienten von $\frac{\tau^n}{n!}$ vergleicht.

Von den beiden Zeigern k, n des Operators D_{kn} wurde der Ordnungszeiger an die erste Stelle gesetzt, da er in der Formel (5) offensichtlich die Hauptrolle spielt, während der Gradzeiger n zur weiteren Unterscheidung zwischen den zusammen auftretenden Operatoren gleicher Ordnung dient. Das hat allerdings zur Folge, daß man in der Kettenregel (11) die Teilfunktionen der Kette (3) in der angegebenen Weise zu beziffern hat, d. h.: nicht in der Reihenfolge, die sich beim Schreiben einstellt, wenn man $f(t)$ mittels der F_ν anschreibt, sondern in der Reihenfolge, in der man die Teilfunktionen F_ν braucht, wenn man zu gegebenem t den Funktionswert $f(t)$ berechnet.

Bevor man die Kettenregel hinschreiben kann, muß die Formel (5) auf die Operatoren höheren Grades ausgedehnt werden:

$$(10) \qquad D_{kn} f = \sum_{\lambda=n}^{k} D_{k\lambda}\,\varphi\cdot D_{\lambda n} F.$$

Diese Gleichung ist für $n = 1$ mit (5) identisch, für $n = k$ mit der k-ten Potenz von (2); für $k > n > 1$ läßt sie sich leicht mittels (6) durch Induktion nach k beweisen. Die Operatoren D_{kn} beziehen sich (ebenso wie D) stets auf das unmittelbare Argument der dabeistehenden Funktion, in (10) also bei f und φ auf t, bei F auf $x = \varphi(t)$. — Von (10) zur Kettenregel (11) ist jetzt nur noch ein ähnlich kleiner Schritt, wie oben von (2) nach (4):

$$(11) \qquad D_{kn} f = \sum_{(\lambda)} D_{\lambda_0\lambda_1} F_1(z_0)\cdot D_{\lambda_1\lambda_2} F_2(z_1)\cdot\ldots\cdot D_{\lambda_{N-1}\lambda_N} F_N(z_{N-1})$$
$$[k = \lambda_0 \geqq \lambda_1 \geqq \cdots \geqq \lambda_N = n];$$

die Summe ist über alle $\frac{(N-1+k-n)!}{(N-1)!\,(k-n)!}$ Zusammenstellungen $(\lambda) = (\lambda_0, \lambda_1, \ldots, \lambda_N)$ aus nichtnegativen ganzen λ_ν zu erstrecken, welche den angegebenen

Ungleichungen genügen. Formel (11) enthält nach (7) für $n = 1$ auch die Regel für die höheren Ableitungen der durch (3) vorgelegten Funktion $f(t)$.

Die Ausdrücke $D_{kn} f$ lassen sich, statt nach Formel (6), auch als Summen von Produkten zweier Operatoren niedrigeren Grades aufbauen:

$$\binom{n}{\nu} D_{kn} f = \sum_{\varkappa=n-\nu}^{k-\nu} \binom{k}{\varkappa} D_{k-\varkappa,\nu} f \cdot D_{\varkappa,n-\nu} f; \tag{12}$$

$$\binom{n-1}{\nu-1} D_{kn} f = \sum_{\varkappa=n-\nu}^{k-\nu} \binom{k-1}{\varkappa} D_{k-\varkappa,\nu} f \cdot D_{\varkappa,n-\nu} f; \tag{13}$$

$$\binom{n-1}{\nu} D_{kn} f = \sum_{\varkappa=n-\nu}^{k-\nu} \binom{k-1}{\varkappa-1} D_{k-\varkappa,\nu} f \cdot D_{\varkappa,n-\nu} f \tag{14}$$

(gültig für $0 < \nu < n$).

Beweis: (12) bestätigt man durch Einsetzen von (8), indem man die aus der Binomialentwicklung von

$$\prod_{\mu} (1+x)^{r_\mu} = (1+x)^R \quad [\Sigma r_\mu = R]$$

zu erhaltende Gleichung

$$\sum_{(\lambda)} \prod_{\mu} \binom{r_\mu}{\lambda_\mu} = \binom{R}{\Lambda} \quad [0 \leqq \lambda_\mu \leqq r_\mu; \; \Sigma \lambda_\mu = \Lambda; \; \Sigma r_\mu = R]$$

beachtet. Wenn man diese Gleichung mit $\sigma \cdot \lambda_\sigma$ vervielfacht, so gelangt man wegen

$$b \cdot \binom{a}{b} = a \cdot \binom{a-1}{b-1}$$

durch Summation über σ zu der Gleichung

$$\sum_{(\lambda)} (\Sigma \mu \cdot \lambda_\mu) \cdot \prod_{\mu} \binom{r_\mu}{\lambda_\mu} = (\Sigma \mu \cdot r_\mu) \cdot \binom{R-1}{\Lambda-1}$$

$$[0 \leqq \lambda_\mu \leqq r_\mu; \quad \Sigma \lambda_\mu = \Lambda; \quad \Sigma r_\mu = R].$$

Diese dient zusammen mit (8) zur Bestätigung von (13); und (14) ist die Differenz zwischen (12) und (13).

2.

Als Anwendung hierfür seien die Ableitungen der Umkehrfunktion zu einer gegebenen Funktion (mit nicht verschwindender erster Ableitung) berechnet: Es sei

$$y = f(x), \; x = g(y); \quad \text{also } x \equiv g(f(x)); \quad f'(x) \neq 0.$$

Daraus folgen die Gleichungen

$$D_{11} x = 1 = Dg \cdot Df;$$

$$D_{k1} x = 0 = \sum_{\nu=1}^{k} D^\nu g \cdot D_{k\nu} f \quad [k = 2, 3, \ldots, n].$$

Aus der ersten fließt $Dg = (Df)^{-1}$. Berechnet man aus diesen n Gleichungen für die n Unbekannten $Dg, D^2g, \ldots D^ng$ die letzte mittels Determinanten, so erhält man[4])

$$(15)\qquad D^n g = (Df)^{-\binom{n+1}{2}} \cdot (-1)^{n+1} \cdot \mathrm{Det}\,|A_{ik}| \quad [i, k = 1, 2, \ldots, n-1]$$
$$\text{mit } A_{ik} = D_{i+1,k} f \quad [k \leqq i+1]; \quad = 0 \quad [k > i+1].$$

Diese Determinante läßt sich zu einer anderen umformen, in der überall nur eine Ableitung von f steht (nur Operatoren 1. Grades). Man ziehe rechts einen Faktor $\frac{1}{(n-2)!}$ vor die Determinante und vervielfache dafür die zweite, ... k-te, ... $(n-1)$-te Spalte mit $(n-2), \ldots, (n-k), \ldots, 1$. Dann wird also

$$(16)\quad D^n g = \frac{(-1)^{n+1}}{(n-2)!} \cdot (Df)^{-\binom{n+1}{2}} \cdot \mathrm{Det}\,|B^{(n)}_{i,k}| \quad [i, k = 1, 2, \ldots, n-1]$$
$$\text{mit } B^{(n)}_{i,k} = D_{i+1,1} f \; [k = 1]; \; = (n-k) D_{i+1,k} f \; [2 \leqq k \leqq i+1];$$
$$= 0 \; [i+1 < k \leqq n-1].$$

Wenn man in (12) und (13) für $k, n, \varkappa, \nu$ die Werte $i+1, k, \lambda-1, 1$ einsetzt, hierauf (12) mit n vervielfacht und davon (13) abzieht, so kommt man auf die Gleichung

$$(17)\quad (n-k) D_{i+1,k} f = \sum_{\lambda=k}^{i+1} \left\{ n \binom{i}{\lambda-1} - \binom{i+1}{\lambda-1} \right\} \cdot D_{i-\lambda+2,1} f \cdot D_{\lambda-1,k-1} f.$$

Die Klammer rechts verschwindet für $i = n-1$, $\lambda = i+1$; schreibt man daher

$$(18)\qquad C^{(n)}_{i,k} = D^{i+1} f \; [k = 1]; \; = \alpha^{(n)}_{i,k} \cdot D^{i-k+2} f \; [2 \leqq k \leqq i+1];$$
$$= 0 \; [i+1 < k \leqq n-1]$$
$$\text{mit } \alpha^{(n)}_{i,k} = n \cdot \binom{i}{k-1} - \binom{i+1}{k-1},$$

so wird für $k > 1$

$$B^{(n)}_{i,k} = \sum_{\lambda=k}^{n-1} C^{(n)}_{i,\lambda} \cdot D_{\lambda-1,k-1} f.$$

Der zweite Faktor ist für alle Zeilen i derselbe, die Spalten B_{ik} sind also linear aus den Spalten C_{ik} zusammengesetzt, und aus (16) wird mit (18) die Darstellung

$$(19)\quad D^n g = \frac{(-1)^{n+1}}{(n-2)!} \cdot (Df)^{-2n+1} \cdot \mathrm{Det}\,|C^{(n)}_{i,k}| \quad [i, k = 1, \ldots n-1].$$

---

4) L [1], Formel (6). Die dort daraus abgeleitete Formel (7) ist für $n > 3$ *nicht* richtig.

Von den Zahlen $\alpha_{i,k}^{(n)}$ in (18) bilden die in der Schrägreihe $i+1=k$ stehenden und ebenso die in der zweiten Spalte je eine arithmetische Reihe erster Ordnung mit den Differenzen 1 und $n-1$, und die anderen lassen sich bequem rekursiv aus ihnen berechnen:

$$\begin{aligned} &\alpha_{i,2}^{(n)} = i\cdot(n-1)-1; \quad \alpha_{k-1,k}^{(n)} = n-k; \\ &\alpha_{i,k}^{(n)} = \alpha_{i-1,k}^{(n)} + \alpha_{i-1,k-1}^{(n)} \quad [3 \leqq k \leqq i]. \end{aligned} \tag{20}$$

Durch Entwickeln der Determinante (19) entsteht die der Brunoschen Formel (8) entsprechende Darstellung

$$\begin{gathered} D^n g = (Df)^{-2n+1} \sum_{(\lambda)} (2n-2-\lambda_1)!\,(Df)^{\lambda_1} \prod_{\nu=2}^{n} \frac{1}{\lambda_\nu!}\left(-\frac{D^\nu f}{\nu!}\right)^{\lambda_\nu} \\ \left[0 \leqq \lambda_\nu;\ \sum_1^n \lambda_\nu = n-1;\ \sum_1^n \nu\,\lambda_\nu = 2n-2\right]. \end{gathered} \tag{21}$$

Für den Beweis mag es indessen bequemer sein, sie aus dem Potenzreihenansatz für f und g herzuleiten mittels der Bürmann-Lagrangeschen Umkehrformel[5])

$$\frac{d^n x}{d y^n} = \frac{d^{n-1}}{d x^{n-1}}\left(\frac{x}{y}\right)^n.$$

3.

Die Erweiterung der Regel (5) auf den Fall mehrerer Zwischenveränderlichen

$$\begin{aligned} &y = F(x_1, \ldots, x_\nu, \ldots); \quad x_\nu = \varphi_\nu(t); \\ &y = F\left(\varphi_1(t), \ldots, \varphi_\nu(t), \ldots\right) = f(t) \end{aligned} \tag{22}$$

läßt sich ebenfalls mit Hilfe der Operatoren D_{kn} anschreiben. Um Zahlenfaktoren in den Formeln zu vermeiden, ist es indessen zweckmäßig, neue Bezeichnungen zu wählen:

$$D^{kn} f = \frac{n!}{k!} D_{kn} f. \tag{23}$$

Statt (6) und der zweiten Gleichung von (7) hat man dann

$$D^{kn} f = \frac{1}{k}\cdot D(D^{k-1,n} f) + \frac{n}{k}\cdot Df \cdot D^{k-1,n-1} f, \tag{24}$$

$$D^{k1} f = \frac{1}{k!} D^k f; \tag{25}$$

in (8) tritt $n!$ an Stelle des $k!$ Die erste und dritte Gleichung von (7) und die Formeln (10), (11) bleiben sogar unverändert gültig, wenn die Operatoren-

---

5) L [5]; L [6], S. 138.

zeiger sämtlich nach oben gerückt werden. Insofern hätten zwar von vornherein die D^{kn} eingeführt werden können; bei Funktionen nur je einer Veränderlichen bieten indessen die D_{kn} den Vorzug, daß in den Formeln (6), (7), (8) keine Brüche auftreten. Bei mehreren Veränderlichen sind diese Brüche jedoch noch angenehmer als die Polynomialfaktoren, die sich sonst (insbesondere für $n = 1$) in (26) und den späteren Formeln einstellen würden. Ferner ist durch die Hochstellung der Zeiger jetzt unten Platz, um bei partiellen Ableitungen die Veränderliche anzudeuten, wie es beim Operator D üblich ist: es möge also $D_{(\nu)}$ die partielle Ableitung nach der ν-ten Veränderlichen bezeichnen, dementsprechend $D^{k,1}_{(\nu)} F = (1:k!)\, D^k_{(\nu)} F$; ebenso $D^{0,1}_{(\nu)} F = F$.

Eine der Gleichung (10) entsprechende Regel[6]) hat nun die Form

$$D^{kn} f = \sum_{\Lambda = n}^{k} \sum_{(\lambda)} D^{k(\lambda)} \varphi \cdot D^{(\lambda)n} F \quad [0 \leqq \lambda_\nu;\ \Sigma \lambda_\nu = \Lambda]. \tag{26}$$

Hierin bedeutet der erste Faktor folgendes:

$$D^{k,(\lambda)} \varphi = D^{k,(\lambda_1,\ldots,\lambda_\nu,\ldots)} \varphi = \sum_{(\alpha)} \prod_{\nu} D^{\alpha_\nu,\lambda_\nu} \varphi_\nu(t) \tag{27}$$
$$[0 \leqq \lambda_\nu \leqq \alpha_\nu;\ \Sigma \alpha_\nu = k].$$

Er läßt sich auch rekursiv berechnen. Schreiben wir die Zeigerzusammenstellung (λ), in welcher alle Zeiger λ_ν verschwinden mit Ausnahme desjenigen, welcher zur Funktion $\varphi_\tau(t)$ gehört und den Wert 1 haben soll, kurz als (δ_τ); $(\lambda - \delta_\tau)$ soll somit dieselbe Zusammenstellung bezeichnen wie (λ), nur daß der zur Funktion $\varphi_\tau(t)$ gehörige Zeiger um 1 verkleinert worden ist. Dann geht der Ausdruck (27) aus den Werten

$$D^{k,(\delta_\tau)} \varphi = D^{k1} \varphi_\tau; \quad D^{\Lambda,(\lambda)} \varphi = \prod_\nu (D\,\varphi_\nu)^{\lambda_\nu} \quad [\Sigma \lambda_\nu = \Lambda] \tag{28}$$

nach der Formel

$$D^{k,(\lambda)} \varphi = \frac{1}{k} D(D^{k-1,(\lambda)} \varphi) + \sum_\tau \frac{\lambda_\tau}{k} D\,\varphi_\tau \cdot D^{k-1,(\lambda-\delta_\tau)} \varphi \tag{29}$$

hervor, sobald $k > \Lambda > 1$ ist; für $k = \Sigma \lambda$ hat man nur das erste Glied wegzulassen, um die Formel noch anwenden zu können. Die Summation in (29) braucht nur auf diejenigen τ ausgedehnt zu werden, für die $\lambda_\tau > 0$ ist.

Der Operator $D^{(\lambda),n}$ des zweiten Faktors in (26) ist von der Ordnung λ_ν in x_ν und vom Grade n in F; der Grad ist höchstens gleich der Gesamtordnung Λ. In den Randfällen $n = 0$, $= 1$ und $= \Lambda$ ist sein Sinn:

$$\begin{aligned} &D^{(\lambda)0} F = 1\ [\Lambda = 0],\ = 0\ [\Lambda > 0];\\ &D^{(\lambda)1} F = \Big(\prod_\nu D^{\lambda_\nu 1}_{(\nu)}\Big) F;\\ &D^{(\lambda)\Lambda} F = \Lambda!\, \prod_\nu \frac{(D_{(\nu)} F)^{\lambda_\nu}}{\lambda_\nu!}\,; \end{aligned} \tag{30}$$

6) Der Fall $n = 1$ findet sich bei TEIXEIRA [1].

für die Zwischenfälle ist er in folgender Weise zu bilden:

$$\lambda_\tau \cdot D^{(\lambda),\, n} F = D_{(\tau)} (D^{(\lambda - \delta_\tau),\, n} F) + n \cdot D_{(\tau)} F \cdot D^{(\lambda - \delta_\tau),\, n-1} F \tag{31}$$

$$[1 < n < \Lambda = \Sigma \lambda_\nu;\ \lambda_\tau > 0].$$

Die Rekursion kann nach jeder Veränderlichen x_τ vorgenommen werden, für die $\lambda_\tau > 0$ ist. Im Fall $n = \Lambda$ ist rechts das erste Glied fortzulassen, damit die Formel gültig bleibe, für $n = 1$ das zweite Glied. Zur unmittelbaren Berechnung aus Operatoren ersten Grades ist die nachstehende Formel tauglich, die (8) und (30) als Sonderfälle einschließt:

$$D^{(k_\sigma),\, n} F = \sum_{(\lambda_\tau)} \sum_{(\nu_{\sigma\tau})} n! \prod_\tau \frac{1}{\lambda_\tau !} (D^{(\nu_{\sigma\tau}),\, 1} F)^{\lambda_\tau}; \tag{32}$$

die Summe ist zu erstrecken über alle möglichen Zusammenstellungen von Zahlengruppen (λ_τ), $(\nu_{1\tau}, \ldots \nu_{\sigma\tau}, \ldots)$ — der Zeiger τ hat hier nur den Sinn, die verschiedenen Gruppen in einer solchen Zusammenstellung zu unterscheiden —, welche die Bedingungen

$$0 \leqq \lambda_\tau;\ 0 \leqq \nu_{\sigma\tau};\ 1 \leqq \sum_\sigma \nu_{\sigma\tau};\ \sum_\tau \lambda_\tau = n;\ \sum_\tau \nu_{\sigma\tau} \cdot \lambda_\tau = k_\sigma \tag{33}$$

befriedigen. Ein Beispiel möge die Anwendung von (31) und (32) zeigen: Nach (31) ist

$$D^{(41)\, 3} F(x, y) = \tfrac{1}{4} D_x (D^{(31)\, 3} F) + \tfrac{3}{4} F_x \cdot D^{(31)\, 2} F$$

$$= \tfrac{1}{4} D_x (3 F_x^2 F_{xy} + 3 F_x F_{xx} F_y) + \tfrac{3}{4} F_x (F_x F_{xxy} + F_{xx} F_{xy} + \tfrac{1}{3} F_{xxx} F_y)$$

oder ebensogut

$$D^{(41)\, 3} F(x, y) = D_y (D^{(40)\, 3} F) + 3 F_y \cdot D^{(40)\, 2} F$$

$$= D_y (\tfrac{3}{2} F_x^2 F_{xx}) + 3 F_y (\tfrac{1}{3} F_x F_{xxx} + \tfrac{1}{4} F_{xx}^2),$$

nach (32) ist

$$D^{(41)\, 3} F(x, y) = 3!\ \{D^{(10)\, 1} F \cdot D^{(20)\, 1} F \cdot D^{(11)\, 1} F + D^{(10)\, 1} F \cdot D^{(30)\, 1} F \cdot D^{(01)\, 1} F$$

$$+ \tfrac{1}{2} (D^{(10)\, 1} F)^2 \cdot D^{(21)\, 1} F + \tfrac{1}{2} (D^{(20)\, 1} F)^2 \cdot D^{(01)\, 1} F\}$$

$$= 6 \{F_x \cdot \tfrac{1}{2} F_{xx} \cdot F_{xy} + F_x \cdot \tfrac{1}{6} F_{xxx} \cdot F_y + \tfrac{1}{2} F_x^2 \cdot \tfrac{1}{2} F_{xxy} + \tfrac{1}{2} (\tfrac{1}{2} F_{xx})^2 \cdot F_y\}$$

$$= 3 F_x F_{xx} F_{xy} + F_x F_{xxx} F_y + \tfrac{3}{2} F_x^2 F_{xxy} + \tfrac{3}{4} F_{xx}^2 F_y.$$

Der Beweis für (26) kann auf folgendem Wege geführt werden: Für $k = n = 1$ ist (26) die bekannte Regel für die totale Ableitung

$$\frac{df}{dt} = \sum_\nu D \varphi_\nu \cdot D_{(\nu)} F.$$

Setzen wir voraus, (26) sei für alle n richtig mit $k - 1$ an Stelle von k; zugleich soll (31) für alle $\Lambda < k$ in Ordnung befunden sein. Durch (27) ist dann $D^{k\,(\lambda)} \varphi$ erklärt; (29) folgt aus (27) vermöge (24). Für $\Lambda = k$ werden die $D^{(\lambda)\, n} F$ jetzt durch (31) erklärt, und zwar unabhängig von der Wahl der zur Rekursion benutzten Veränderlichen: dies erkennt man, indem man $D^{(\lambda)\,n}F$ einmal zuerst nach λ_1 reduziert, die übriggebliebenen Ausdrücke $D^{(\lambda-\delta_1)\,n}\,F$ und $D^{(\lambda-\delta_1),\,n-1}\,F$ darauf nach λ_2; ein andermal zuerst nach λ_2, darauf nach λ_1. Wenn man nun in (24) für $D^{k-1,\,n}f$ und $D^{k-1,\,n-1}f$ die Darstellung (26) einsetzt, so ergibt sich wegen (29) und (31) ihre Gültigkeit auch für $D^{k\,n}f$. — (32) beweist man wie (8) durch Potenzreihenansatz und Polynomialformel.

Die Regel (26) ist nun aber nicht die einzige, welche bei mehreren Veränderlichen der Gleichung (10) entspricht. Hat man es z. B. mit mehreren Funktionen F_ν zu tun, welche von den gleichen Veränderlichen x_μ abhängen, die wiederum als Funktionen φ_μ einer Unveränderlichen t gegeben sind, dann entsteht ein System mittelbarer Funktionen $f_\nu(t) = F_\nu(\varphi_\mu(t))$. Hier lassen sich für die linke Seite die Ausdrücke (27) bilden; und wenn man die einfachen Fälle (28) zuerst nachprüft und dann auf diese mittels (29) zurückgeht, so kommt man zu der Regel

$$(34)\quad D^{k\,(n)}f = \sum_{\Lambda=N}^{k}\sum_{(\lambda)} D^{k\,(\lambda)}\varphi\cdot D^{(\lambda)\,(n)}F \quad [0\leqq\lambda_\mu;\ \Sigma\lambda_\mu=\Lambda;\ \Sigma n_\nu = N].$$

Hier wurden entsprechend (27) gewisse Glieder vorweg vereinigt:

$$(35)\qquad D^{(\lambda)\,(n)}F = \sum_{(\lambda_{\mu\nu})}\prod_\nu D^{(\lambda_{1\nu},\,\ldots\,\lambda_{\mu\nu},\,\ldots)\,n_\nu}F_\nu$$

$$[0\leqq\lambda_{\mu\nu};\ \sum_\nu\lambda_{\mu\nu}=\lambda_\mu;\ \sum_\mu\lambda_{\mu\nu}\geqq n_\nu].$$

Für jedes verschwindende n_ν sind hierbei nach (30) alle zugehörigen $\lambda_{\mu\nu}=0$ zu nehmen, solch ein Faktor ist also fortzulassen. Auch diese Größen (35) lassen sich stufenweise bilden nach der Gleichung (36), die aus (31) und (35) folgt[7]):

$$(36)\quad k_\mu\cdot D^{(k)\,(n)}F = D_{(\mu)}\left(D^{(k-\delta_\mu)\,(n)}F\right)+\sum_\nu n_\nu\cdot D_{(\mu)}F\cdot D^{(k-\delta_\mu)\,(n-\delta_\nu)}F$$

$$[1<N<K;\ \Sigma n = N;\ \Sigma k = K;\ k_\mu>0],$$

wo wie in (31) die Reduktion nach jeder Veränderlichen x_μ erfolgen kann, deren Ordnungszeiger k_μ nicht verschwindet. Ohne das erste Glied gilt die Formel (36), von der (24), (29) und (31) Sonderfälle sind, auch noch für $N=K$. Für $N=1$, d. h. $(n)=(\delta_\nu)$ benutze man

$$(37)\qquad D^{(k)\,(n)}F = D^{(k)\,m}F \quad [n_\nu = m, \text{ alle übrigen } n=0].$$

Setzt man nun (34) in (26) ein, so ergibt sich (denn da die auftretenden Funktionen nur stetige Ableitungen zu haben brauchen und im übrigen ganz willkürlich sind, so ist ein gliedweiser Vergleich zulässig) eine Regel für den Fall, daß eine Funktion F von mehreren Zwischenveränderlichen x_ν vorliegt,

[7]) In anderer Bezeichnung von AMALDI [4] angegeben. Dort und bei MAMBRIANI [7] sind auch weitere Untersuchungen über den kombinatorischen Aufbau der Ausdrücke (35) zu finden.

welche Funktionen φ_ν von mehreren Urveränderlichen t_ϱ sind: also $f(t_\varrho) = F(\varphi_\nu^*(t_\varrho))$. Man kann sie auch mittels (30) und (31) aus (26) ableiten, worauf Einsetzen von (38) in (26) eben (34) liefern würde. Die Regel[8]) lautet:

$$(38)\quad D^{(k)\,n} f = \sum_{\Lambda=n}^{K} \sum_{(\lambda)} D^{(k)\,(\lambda)}_{\varrho} \varphi \cdot D^{(\lambda)\,n} F \quad [0 \leqq \lambda_\nu;\ \Sigma \lambda_\nu = \Lambda;\ \Sigma k_\varrho = K].$$

Setzt man schließlich sowohl (34) wie (38) in (26) ein, so findet man die Regel für den Fall, daß ein System von Funktionen F_ν vorliegt, welche von einem System Zwischenveränderlicher x_μ abhängen, die ihrerseits als Funktionen φ_μ mehrerer Urveränderlichen t_ϱ gegeben sind, so daß ein System mittelbarer Funktionen von mehreren Veränderlichen $f_\nu(t_\varrho) = F_\nu(\varphi_\mu(t_\varrho))$ entsteht:

$$(39)\quad D^{(k)\,(n)} f = \sum_{\Lambda=N}^{K} \sum_{(\lambda)} D^{(k)\,(\lambda)} \varphi \cdot D^{(\lambda)\,(n)} F$$
$$[\Sigma k = K;\quad \Sigma n = N;\quad 0 \leqq \lambda_u;\quad \Sigma \lambda = \Lambda].$$

Erst diese Formel ist die vollständige Verallgemeinerung zu (10); sie umfaßt (34) und (38), aber auch (26) und (10) als Sonderfälle, und sie läßt sich ebenso leicht wie (10) zur Kettenregel verallgemeinern.

Die Kettenregel für mehrere Gruppen von Zwischenveränderlichen in voller Allgemeinheit hinzuschreiben, würde wegen zu vieler Zeiger an den Hauptbuchstaben doch wohl ein unklares Bild geben. Es seien darum nur drei Zwischengruppen angenommen; das Aussehen der Regel bei mehr Zwischengruppen ist aus diesem Beispiel dann mühelos zu entnehmen:

$$(40)\quad \begin{aligned} x_\lambda &= \varphi_\lambda(t); \quad y_\mu = \psi_u(\ldots x_\lambda \ldots); \\ z_\nu &= \chi_\nu(\ldots y_\mu \ldots); \quad F(\ldots z_\nu \ldots) = f(t). \end{aligned}$$

An die Stelle der Zeiger λ_ν von Formel (11) tritt hier je ein ganzer Satz solcher Zeiger: $\alpha_\lambda, \beta_\mu, \gamma_\nu$, die die Ordnung (wenn sie an dem Operator in der vorderen Gruppe stehen) oder den Grad (wenn sie in der zweiten Zeigergruppe stehen) des betreffenden Faktors in bezug auf die zugehörige Zwischenveränderliche x_λ, y_μ, z_ν anzeigen. Nur zur Erhöhung der Übersichtlichkeit werden die Summanden, wie schon in (26), (34), (38) und (39), noch nach gleichen Summenwerten (Λ) zusammengefaßt. Die Kettenregel sieht dann so aus:

$$(41)\quad D^{kn} f = \sum_{(\Lambda)} \sum_{(\alpha\beta\gamma)} D^{k\,(\alpha)} \varphi \cdot D^{(\alpha)\,(\beta)} \psi \cdot D^{(\beta)\,(\gamma)} \chi \cdot D^{(\gamma)\,n} F.$$

Die Summe ist zu erstrecken über alle Zusammenstellungen

$$(\Lambda) = (\Lambda_1, \Lambda_2, \Lambda_3) \text{ und } (\alpha, \beta, \gamma) = (\ldots \alpha_\lambda \ldots, \ldots \beta_\mu \ldots, \ldots \gamma_\nu \ldots),$$

die folgende Bedingungen erfüllen:

$$(42)\quad \begin{aligned} k = \Lambda_0 \geqq \Lambda_1 \geqq \Lambda_2 \geqq \Lambda_3 \geqq \Lambda_4 = n;\ 0 \leqq \alpha_\lambda,\ 0 \leqq \beta_\mu,\ 0 \leqq \gamma_\nu; \\ \Sigma \alpha_\lambda = \Lambda_1,\ \Sigma \beta_\mu = \Lambda_2,\ \Sigma \gamma_\nu = \Lambda_3. \end{aligned}$$

---

8) Der Fall $n = 1$ ist von Amaldi [4] und Mambriani [7] gegeben worden.

Hängen die x_λ und damit die ganze mittelbare Funktion f statt von einer Urveränderlichen t von mehreren $(\ldots t_\varrho \ldots)$ ab:

$$\begin{aligned}(43)\qquad & x_\lambda = \varphi_\lambda(\ldots t_\varrho \ldots); \quad y_\mu = \psi_\mu(\ldots x_\lambda \ldots); \\ & z_\nu = \chi_\nu(\ldots y_\mu \ldots); \quad F(\ldots z_\nu \ldots) = f(\ldots t_\varrho \ldots),\end{aligned}$$

so braucht man an Stelle von (41) die Formel

$$(44)\qquad D^{(k)\,n} f = \sum_{(A)} \sum_{(\alpha\beta\gamma)} D^{(k)\,(\alpha)} \varphi \cdot D^{(\alpha)\,(\beta)} \psi \cdot D^{(\beta)\,(\gamma)} \chi \cdot D^{(\gamma)\,n} F$$

mit denselben Bedingungen (42); nur ist jetzt $\Sigma k_\varrho = K = A_0$ einzusetzen. — Für den Fall, daß statt einer Funktion $f(\ldots t_\varrho \ldots) = F(\ldots z_\nu \ldots)$ deren mehrere vorhanden sind und für dieses System einer der in (26) oder (31) eingeführten Ausdrücke $D^{k\,(n)} f$ oder $D^{(k)\,(n)} f$ — je nach der Anzahl der Urveränderlichen kommt der eine oder der andere in Betracht — mittels der Zwischenveränderlichen darzustellen ist, läßt sich die erforderliche Änderung der Formel (41) oder (44) leicht vornehmen.

Die Anwendung der Kettenregel (41) soll zum Schluß an dem einfachsten nicht trivialen Beispiel gezeigt werden:

$$x_\mu = \varphi_\mu(t)\ [\mu = 1, 2]; \quad y_\nu = \psi_\nu(x_1, x_2)\ [\nu = 1, 2]; \quad F(y_1, y_2) = f(t).$$

Verlangt sei, $f''(t)$ durch die Ableitungen von F, ψ_1, ψ_2, φ_1, φ_2 auszudrücken. Es ist mithin $k = 2$, $n = 1$ zu nehmen, da $D^{21} f = \frac{1}{2} f''$. Den Veränderlichen $x_1, x_2; y_1, y_2$ sollen die Zeiger $\alpha_1, \alpha_2; \beta_1, \beta_2$ zugeordnet sein. Dann ist nach (41)

$$D^{21} f = \sum_{\mathsf{A},\,\mathsf{B}} \sum_{(\alpha,\,\beta)} D^{2\,(\alpha)} \varphi \cdot D^{(\alpha)\,(\beta)} \psi \cdot D^{(\beta)\,1} F,$$

wobei die nichtnegativen ganzen Zahlen α_1, α_2, β_1, β_2, A, B auf jede mögliche Weise einmal so einzusetzen sind, daß dabei

$$2 \geqq \mathsf{A} \geqq \mathsf{B} \geqq 1; \ \alpha_1 + \alpha_2 = \mathsf{A}, \ \beta_1 + \beta_2 = \mathsf{B}.$$

Für die Zeigersummen A, B gibt es hiernach drei Möglichkeiten: $\mathsf{A} = \mathsf{B} = 1$; $\mathsf{A} = 2$, $\mathsf{B} = 1$; $\mathsf{A} = \mathsf{B} = 2$. Zu $\mathsf{A} = 1$ gibt es zwei Möglichkeiten für die Wahl der (α): entweder ist $(\alpha_1, \alpha_2) = (1, 0)$ oder $= (0, 1)$; zu $\mathsf{A} = 2$ gibt es drei: $(\alpha_1, \alpha_2) = (2, 0)$ oder $(1, 1)$ oder $(0, 2)$. Mit den entsprechenden Möglichkeiten für (β) erhält man im ganzen $2 \cdot 2 + 3 \cdot 2 + 3 \cdot 3 = 19$ verschiedene Zusammenstellungen (α, β). Es wird also

$$\begin{aligned}
D^{21} f = {} & D^{2\,(10)}\varphi \cdot D^{(10)\,(10)}\psi \cdot D^{(10)\,1}F + D^{2\,(01)}\varphi \cdot D^{(01)\,(10)}\psi \cdot D^{(10)\,1}F \\
& + D^{2\,(10)}\varphi \cdot D^{(10)\,(01)}\psi \cdot D^{(01)\,1}F + D^{2\,(01)}\varphi \cdot D^{(01)\,(01)}\psi \cdot D^{(01)\,1}F \\
& + D^{2\,(20)}\varphi \cdot D^{(20)\,(10)}\psi \cdot D^{(10)\,1}F + D^{2\,(11)}\varphi \cdot D^{(11)\,(10)}\psi \cdot D^{(10)\,1}F + D^{2\,(02)}\varphi \cdot D^{(02)\,(10)}\psi \cdot D^{(10)\,1}F \\
& + D^{2\,(20)}\varphi \cdot D^{(20)\,(01)}\psi \cdot D^{(01)\,1}F + D^{2\,(11)}\varphi \cdot D^{(11)\,(01)}\psi \cdot D^{(01)\,1}F + D^{2\,(02)}\varphi \cdot D^{(02)\,(01)}\psi \cdot D^{(01)\,1}F \\
& + D^{2\,(20)}\varphi \cdot D^{(20)\,(20)}\psi \cdot D^{(20)\,1}F + D^{2\,(11)}\varphi \cdot D^{(11)\,(20)}\psi \cdot D^{(20)\,1}F + D^{2\,(02)}\varphi \cdot D^{(02)\,(20)}\psi \cdot D^{(20)\,1}F \\
& + D^{2\,(20)}\varphi \cdot D^{(20)\,(11)}\psi \cdot D^{(11)\,1}F + D^{2\,(11)}\varphi \cdot D^{(11)\,(11)}\psi \cdot D^{(11)\,1}F + D^{2\,(02)}\varphi \cdot D^{(02)\,(11)}\psi \cdot D^{(11)\,1}F \\
& + D^{2\,(20)}\varphi \cdot D^{(20)\,(02)}\psi \cdot D^{(02)\,1}F + D^{2\,(11)}\varphi \cdot D^{(11)\,(02)}\psi \cdot D^{(02)\,1}F + D^{2\,(02)}\varphi \cdot D^{(02)\,(02)}\psi \cdot D^{(02)\,1}F.
\end{aligned}$$

An erster Stelle treten in diesen 19 Produkten 5 verschiedene Faktoren auf:

$$D^{2\,(10)}\,\varphi = \frac{1}{2}\cdot\frac{d^2\,\varphi_1}{dt^2};\qquad D^{2\,(01)}\,\varphi = \frac{1}{2}\,\frac{d^2\,\varphi_2}{dt^2};$$

$$D^{2\,(20)}\,\varphi = \left(\frac{d\,\varphi_1}{dt}\right)^2;\qquad D^{2\,(11)}\,\varphi = \frac{d\,\varphi_1}{dt}\cdot\frac{d\,\varphi_2}{dt};\qquad D^{2\,(02)}\,\varphi = \left(\frac{d\,\varphi_2}{dt}\right)^2.$$

Der mittlere Faktor hat in jedem Gliede einen anderen Wert:

$$D^{(10)\,(10)}\,\psi = \frac{\partial\,\psi_1}{\partial\,x_1};\qquad D^{(01)\,(10)}\,\psi = \frac{\partial\,\psi_1}{\partial\,x_2};\qquad D^{(10)\,(01)}\,\psi = \frac{\partial\,\psi_2}{\partial\,x_1};\qquad D^{(01)\,(01)}\,\psi = \frac{\partial\,\psi_2}{\partial\,x_2}$$

$$D^{(20)\,(10)}\,\psi = \frac{1}{2}\,\frac{\partial^2\,\psi_1}{\partial\,x_1^2};\qquad D^{(11)\,(10)}\,\psi = \frac{\partial^2\,\psi_1}{\partial\,x_1\,\partial\,x_2};\qquad D^{(02)\,(10)}\,\psi = \frac{1}{2}\,\frac{\partial^2\,\psi_1}{\partial\,x_2^2};$$

$$D^{(20)\,(01)}\,\psi = \frac{1}{2}\,\frac{\partial^2\,\psi_2}{\partial\,x_1^2};\qquad D^{(11)\,(01)}\,\psi = \frac{\partial^2\,\psi_2}{\partial\,x_1\,\partial\,x_2};\qquad D^{(02)\,(01)}\,\psi = \frac{1}{2}\,\frac{\partial^2\,\psi_2}{\partial\,x_2^2};$$

$$D^{(20)\,(20)}\,\psi = \left(\frac{\partial\,\psi_1}{\partial\,x_1}\right)^2;\qquad D^{(11)\,(20)}\,\psi = 2\,\frac{\partial\,\psi_1}{\partial\,x_1}\cdot\frac{\partial\,\psi_1}{\partial\,x_2};\qquad D^{(02)\,(20)}\,\psi = \left(\frac{\partial\,\psi_1}{\partial\,x_2}\right)^2;$$

$$D^{(20)\,(11)}\,\psi = \frac{\partial\,\psi_1}{\partial\,x_1}\cdot\frac{\partial\,\psi_2}{\partial\,x_1};\qquad D^{(11)\,(11)}\,\psi = \frac{\partial\,\psi_1}{\partial\,x_1}\,\frac{\partial\,\psi_2}{\partial\,x_2} + \frac{\partial\,\psi_1}{\partial\,x_2}\,\frac{\partial\,\psi_2}{\partial\,x_1};\qquad D^{(02)\,(11)}\,\psi = \frac{\partial\,\psi_1}{\partial\,x_2}\cdot\frac{\partial\,\psi_2}{\partial\,x_2}$$

$$D^{(20)\,(02)}\,\psi = \left(\frac{\partial\,\psi_2}{\partial\,x_1}\right)^2;\qquad D^{(11)\,(02)}\,\psi = 2\,\frac{\partial\,\psi_2}{\partial\,x_1}\cdot\frac{\partial\,\psi_2}{\partial\,x_2};\qquad D^{(02)\,(02)}\,\psi = \left(\frac{\partial\,\psi_2}{\partial\,x_2}\right)^2.$$

An dritter Stelle steht eine dieser fünf Größen:

$$D^{(10)\,1}\,F = \frac{\partial\,F}{\partial\,y_1};\qquad D^{(01)\,1}\,F = \frac{\partial\,F}{\partial\,y_2};$$

$$D^{(20)\,1}\,F = \frac{1}{2}\,\frac{\partial^2 F}{\partial\,y_1^2};\qquad D^{(11)\,1}\,F = \frac{\partial^2 F}{\partial\,y_1\,\partial\,y_2};\qquad D^{(02)\,1}\,F = \frac{1}{2}\,\frac{\partial^2 F}{\partial\,y_2^2}.$$

Literaturverzeichnis.

[1] F. G. Teixeira, Giorn. mat. Battaglini **18** (1880), S. 301–309.
[2] A. Voss, Enz. Math. Wiss. II, 1a, S. 87–88.
[3] J. Molk, Enc. Sci. Math. II, 1a, S. 332–336.
[4] U. Amaldi, Rend. Palermo **42** (1917), S. 94–115.
[5] Hass, Mitteilungen Math. Ges. Hamburg **6** (1926), S. 163–189.
[6] Hurwitz-Courant, Funktionentheorie, 3. Aufl. Berlin 1929.
[7] A. Mambriani, Boll. Un. Mat. Ital. **13** (1934), S. 217–222, 284–288; **14** (1935), S. 10–16.

(Eingegangen am 20. September 1942.)