BAYERISCHE AKADEMIE DER WISSENSCHAFTEN
MATHEMATISCH-NATURWISSENSCHAFTLICHE KLASSE

SITZUNGSBERICHTE

JAHRGANG

1955

MÜNCHEN 1956

VERLAG DER BAYERISCHEN AKADEMIE DER WISSENSCHAFTEN

In Kommission bei der C. H. Beck'schen Verlagsbuchhandlung München

Bemerkungen zu Carathéodory's Einführung in Euler's Arbeiten über Variationsrechnung

Von **Heinrich Tietze** in München

Vorgelegt am 14. Januar 1955

1. Es handelt sich um eine kleine Ergänzung, die an Carathéodory's Feststellung anknüpft, daß bei dem Bernoulli-Euler'schen Verfahren, wie es Euler in „Methodus inveniendi lineas curvas maximi minimive proprietate gaudentes" (1744) zur Herleitung seiner Differentialgleichung entwickelt, die Vorstellung einer Zerlegung der Kurve in unendlich viele unendlich kleine Stücke vermieden und von einer Zerlegung in endlich viele Stücke ausgegangen werden kann[1].

Ist (mit $a < b$)

$$\int_a^b f(x, y, y', \ldots, y^{(m)})\, dx \tag{1}$$

das Integral, das für eine geeignet zu bestimmende Kurve $y = y(x)$ nebst $y' = y'(x), \ldots, y^{(m)} = y^{(m)}(x)$ zu einem Extremum

[1] Siehe Leonhardi Euleri Opera Omnia, ed. sub auspiciis Societatis Scientiarum Naturalium Helveticae, series I, opera mathematica, vol. 24, 1952; ebenda als Einleitung zu den der Variationsrechnung gewidmeten Bänden 24, 25: C. Carathéodory, Einführung in Eulers Arbeiten über Variationsrechnung, p. XIII, Nr. 8. – Carathéodory, der sich mit großer Liebe in die Euler'schen Arbeiten vertieft hat und dort, wo er Euler auf irrigen Wegen sieht, ausführliche kritische Erläuterungen gibt, zollt dem Genie Euler's bei der Herleitung der Differentialgleichung der Variationsrechnung hohe Bewunderung (l. c., p. XVI, Nr. 11) und betont, daß wir bei Beurteilung der Methoden früherer Mathematiker hinsichtlich ihrer Strenge nicht die Maßstäbe anlegen dürfen, die sich für uns aus einer Technik ergeben, die wir uns im Laufe der Zeit angeeignet haben. – Die genannte „Einführung" soll auch in C. Carathéodory's Gesammelten mathematischen Schriften (in Band V) zum Wiederabdruck kommen, und es mag die Gelegenheit wahrgenommen werden, Herrn A. Speiser, dem hochverdienten Herausgeber der fraglichen Bände der Opera Euler's, für die schon vor Jahren ausgesprochene Zustimmung zu diesem Wiederabdruck nochmals zu danken.

gemacht werden soll, wobei entsprechend den vorwiegend behandelten Variationsproblemen die Kurvenendpunkte vermöge vorgegebener Werte $y(a)$, $y(b)$ festgehalten (und außerdem etwa noch $y^{(k)}(a)$, $y^{(k)}(b)$ für $k \leqq m-1$ vorgegeben[2]) werden sollen, so werde $b-a$ in eine endliche Anzahl N gleicher Teile $u = \frac{b-a}{N}$ zerlegt. Die zu den eingeschalteten Stellen $x = x_\nu = a + \nu u (\nu = 1, \ldots, N-1)$ gehörigen Ordinatenwerte y_ν werden dann als $N-1$ unabhängige Variable angesehen, die nach der gewöhnlichen Methode der Berechnung der Maxima und Minima einer Funktion von mehreren Variablen so zu bestimmen sind, daß ein den Wert des Integrals (1) approximierender Ausdruck J als Funktion der $y_1, \ldots, y_{N-1}$ ein Extremum wird. Dieser Ausdruck

$$(2) \qquad J = \sum_{\nu=1}^{N} Z_\nu (x_\nu - x_{\nu-1}) = u \sum_{\nu=1}^{N} Z_\nu$$

wird erhalten, indem man die Z_ν vermöge

$$(3) \qquad Z_\nu = f(x_\nu, y_\nu, \overline{y'_\nu}, \ldots, \overline{y^{(m)}_\nu})$$

bestimmt, wo bei dem l. c.[3] auseinandergesetzten Verfahren die $\overline{y^{(k)}_\nu}$ durch

$$\overline{y'_\nu} = \frac{1}{u}(y_{\nu+1} - y_\nu),$$

$$\overline{y''_\nu} = \frac{1}{u^2}(y_{\nu+2} - 2\, y_{\nu+1} + y_\nu),$$

allgemein

$$(4) \qquad \begin{aligned} \overline{y^{(k)}_\nu} &= \frac{1}{u^k} \sum_{i=0}^{k} (-1)^{k-i} \binom{k}{i} y_{\nu+i} = \\ &= \frac{1}{u^k} \sum_{\lambda=\nu}^{\nu+k} (-1)^{k+\nu-\lambda} \binom{k}{\lambda-\nu} y_\lambda, \qquad (1 \leqq k \leqq m) \end{aligned}$$

---

[2] Bei Euler selbst sind, worauf Carathéodory l. c. Nr. **13**, p. XVII, hinweist, im Falle $m > 1$ die Vorstellungen über Festlegung der Lösungskurve noch nicht so präzisiert, wie man es nunmehr zu tun gewohnt ist.

[3] Vgl. Euler, Methodus, Caput I, Corollarium 3, 4, 5, pag. 24; Carathéodory, Einführung, Nr. 8, p. XIV. Wegen einer gewissen Abweichung in der Bezeichnung vgl. unten, Nr. **2**.

oder, wenn noch $\overline{y_\nu^{(0)}} = y_\nu$ gesetzt wird, rekursorisch durch

$$\overline{y_\nu^{(k+1)}} = \frac{1}{u}\left(\overline{y_{\nu+1}^{(k)}} - \overline{y_\nu^{(k)}}\right) \qquad (0 \leqq k < m)$$

gegeben sind, also – beiläufig gesprochen – als solche Werte der Ableitungen $y', \ldots, y^{(m)}$, wie sie näherungsweise aus Funktionswerten y_λ an Stellen x in der Nähe von x_ν berechnet werden können. Bei dem eben geschilderten Verfahren sind hierfür die Stellen

$$x_\nu, x_{\nu+1}, \ldots, x_{\nu+m} \tag{5}$$

gewählt, sodaß in den – in den y_λ linearen – Ausdrücken $\overline{y_\nu^{(k)}}$ speziell $y_\nu, y_{\nu+1}, \ldots, y_{\nu+m}$ auftreten. Dazu ist nun aber zu sagen, – und das ist die eine Bemerkung, die hier gemacht werden soll, – daß, wenn man, um der Gleichförmigkeit der späteren Formeln willen, an der Wahl dieser Stellen (5) für jedes ν festhält, dann für Werte[4] $\nu > N - m$ nicht nur $y_1, \ldots, y_N$ sondern auch Werte y_ν mit $\nu > N$ (bis zu y_{N+m}) in den Ausdrücken für die $\overline{y_\nu^{(k)}}$ auftreten. Man wird, wenn man J als Funktion von $y_1, \ldots, y_{N-1}$ betrachtet, diese Werte $y_\nu (\nu > N)$ ebenso wie $y_N = y(b)$ als fest gegeben anzusehen haben.

Ist nun $1 \leqq \lambda \leqq N - 1$, so wird

$$\begin{aligned} \frac{\partial J}{\partial y_\lambda} &= u \sum_{\nu=1}^{N} \frac{\partial Z_\nu}{\partial y_\lambda} = u \sum_{\nu=1}^{N} \sum_{k=0}^{m} P_\nu^{(k)} \frac{\partial \overline{y_\nu^{(k)}}}{\partial y_\lambda} = \\ &= u \sum_{k=0}^{m} (-1)^k \frac{1}{u^k} \sum_{\nu=1}^{N} (-1)^{\nu-\lambda} \binom{k}{\lambda - \nu} P_\nu^{(k)} \end{aligned} \tag{6}$$

wobei $P^{(k)}$ (für $0 \leqq k \leqq m$) die partielle Ableitung

$$P^{(k)}(x, y, y', \ldots, y^{(m)}) = \frac{\partial}{\partial y^{(k)}} f(x, y, y', \ldots, y^{(m)})$$

von f nach der $(k+2)$-ten seiner $m+2$ unabhängigen Variablen bedeute, und $P_\nu^{(k)}$ den Wert dieser Ableitung an der Stelle $x_\nu, y_\nu, \overline{y_\nu^{(1)}}, \ldots, \overline{y_\nu^{(m)}}$).

Die Gleichungen

$$\frac{\partial J}{\partial y_\lambda} = 0 \qquad (1 \leqq \lambda \leqq N - 1)$$

---

[4] N werde natürlich von vorneherein $> m$ angenommen.

dienen dann zur Bestimmung der ein Extremum liefernden Werte von $y_1, \ldots, y_{N-1}$. Diese Gleichungen haben den l. c.[5] hervorgehobenen gleichmäßigen Bau, zumindest wenn man sich auf Werte $\lambda \geqq m + 1$ beschränkt (was, wenn N gegenüber m groß gewählt ist, die überwiegende Mehrheit unter allen $N - 1$ Werten λ ist); in dem Ausdruck (6) für $\frac{\partial J}{\partial y_\lambda}$ kommen dann nämlich in der Summe über ν alle Werte ν in Betracht, für welche $0 \leqq \lambda - \nu \leqq k$ oder $\lambda - k \leqq \nu \leqq \lambda$ ist; und diese Werte ν gehören dem Summationsbereich $1 \leqq \nu \leqq N$ nur dann alle an, wenn $1 \leqq \lambda - k$ oder $\lambda \geqq k + 1$ ist, was für alle k mit $0 \leqq k \leqq m$ nur für $\lambda \geqq m + 1$ zutrifft. Man erhält so

$$\frac{\partial J}{\partial y_\lambda} = u \sum_{k=0}^{m} (-1)^k K_\lambda^{(k)}, \tag{7}$$

wobei

$$K_\lambda^{(k)} = \frac{1}{u^k} \sum_{\nu=\lambda-k}^{\lambda} (-1)^{\nu-\lambda} \binom{k}{\lambda-\nu} P_\nu^{(k)} = \frac{1}{u^k} \sum_{i=0}^{k} (-1)^i \binom{k}{i} P_{\lambda-i}^{(k)} \tag{8}$$

bei entsprechenden Differenzierbarkeits-Annahmen mit abnehmendem u gegen den Wert von $\frac{d^k}{dx^k} P^{(k)}$ an der Stelle x_λ konvergiert. Damit ist, wie Carathéodory l. c. hervorhebt, mittels einer Zerlegung in endlich viele Stücke und nachträglichen Grenzübergangs der Weg zur „Euler'schen" Differentialgleichung

$$\sum_{k=0}^{m} (-1)^k \frac{d^k}{dx^k} \frac{\partial f}{\partial y^{(k)}} = 0 \tag{9}$$

gegeben.

2. Für unsere Rechnungen schien es erwünscht, die Ordnung m der höchsten in (1) unter dem Integral auftretenden Ableitung nicht zu spezialisieren. Dafür war eine stärkere Belastung der Schreibweise mit Indizes in Kauf zu nehmen, als l. c., wo Carathéodory die Dinge am Fall $m = 3$ erläutert und dabei, im Anschluß an Euler, für die x_ν bzw. y_ν die Bezeichnungen

$$\ldots, x_{,,}, x_{,}, x, x', x'', \ldots$$

bzw.

$$\ldots, y_{,,}, y_{,}, y, y', y'', \ldots$$

---

[5] Vgl. Carathéodory, Einführung, Nr. **10**, p. XVI.

gebraucht, ferner

$$p = \frac{y'-y}{u}, \quad q = \frac{p'-p}{u}, \quad r = \frac{q'-q}{u},$$

$$p' = \frac{y''-y'}{u}, \quad q' = \frac{p''-p'}{u}, \quad r' = \frac{q''-q'}{u}, \quad \text{usw.}$$

und wo anstelle unserer $P^{(0)}$, $P^{(1)}$, $P^{(2)}$, $P^{(3)}$ die Bezeichnungen $N = \frac{\partial f}{\partial y}$, $P = \frac{\partial f}{\partial p}$, $Q = \frac{\partial f}{\partial q}$, $R = \frac{\partial f}{\partial r}$ auftreten.

3. Eine zweite Bemerkung betrifft die Wahl der Stellen (5). Statt, wie in Nr. **1**, Näherungswerte der Ableitungen $y^{(k)}$ an der Stelle x_ν mit Hilfe derjenigen Werte zu berechnen, die die Funktion y an den Stellen (5) annimmt, könnte man auch andere Stellen in der Nähe von x_ν, etwa

$$(10) \qquad x_{\nu-m},\ x_{\nu-m+1},\ \ldots,\ x_{\nu-1},\ x_\nu$$

wählen. Die Formel (4) für die $\overline{y_\nu^{(k)}}$ wäre dann zu ersetzen durch

$$(11) \quad \overline{y_\nu^{(k)}} = \frac{1}{u^k} \sum_{i=0}^{k} (-1)^i \binom{k}{i} y_{\nu-i} = \frac{1}{u^k} \sum_{\lambda=\nu-k}^{\nu} (-1)^{\nu-\lambda} \binom{k}{\nu-\lambda} y_\lambda,$$

und anstelle von (6) erhielte man

$$(12) \quad \frac{\partial J}{\partial y_\lambda} = u \sum_{\nu=1}^{N} \sum_{k=0}^{m} P_\nu^{(k)} \frac{\partial \overline{y_\nu^{(k)}}}{\partial y_\lambda} = u \sum_{k=0}^{m} \frac{1}{u^k} \sum_{\nu=1}^{N} (-1)^{\nu-\lambda} \binom{k}{\nu-\lambda} P_\nu^{(k)},$$

wobei wieder

$$P_\nu^{(k)} = P^{(k)}(x_\nu, y_\nu, \overline{y_\nu^{(1)}}, \ldots, \overline{y_\nu^{(m)}})$$

ist, aber mit der nunmehrigen Bedeutung der $\overline{y_\nu^{(k)}}$; beschränkt man sich jetzt auf Werte $\lambda \leqq N - m$ (was bei genügend groß gewähltem N wieder die überwiegende Mehrheit der Werte λ ist), so ist die zuletzt angeschriebene Summe über alle ν aus $\lambda \leqq \nu \leqq \lambda + k$ zu erstrecken und es gilt wieder die Formel (7), wenn jetzt für $K_\lambda^{(k)}$ genommen wird:

$$(13) \quad \begin{aligned} K_\lambda^{(k)} &= \frac{1}{u^k} \sum_{\nu=\lambda}^{\lambda+k} (-1)^{k+\nu-\lambda} \binom{k}{\nu-\lambda} P_\nu^{(k)} = \\ &= \frac{1}{u^k} \sum_{i=0}^{k} (-1)^{k-i} \binom{k}{i} P_{\lambda+i}^{(k)}, \end{aligned}$$

also wieder ein Ausdruck, den man für abnehmendes u als Approximationswert für $\frac{d^k}{dx^k} P^{(k)}$ an der Stelle $x = x_\lambda$ ansehen kann.

Es treten also gegenüber dem in Nr. **1** geschilderten Verfahren keine wesentlichen Änderungen ein, nur daß für unsere zusätzlichen Bemerkungen über Teilungspunkte x_ν in der Nähe der Intervall-Endpunkte a und b, – einmal, was ein Hinausgreifen über dieses Intervall und die Festlegung zugehöriger y-Werte, das anderemal, was Abweichungen im gleichmäßigen Bau der Gleichungen $\frac{\partial J}{\partial y_\lambda} = 0$ in der Nähe eines Intervall-Endpunktes betrifft, – die Rolle des linken und rechten Intervallendes sich vertauschen.

Noch in einer anderen Hinsicht kann man beim Vergleich des Ansatzes in Nr. **3** mit dem in Nr. **1** von einer Vertauschung der Rollen sprechen: Wenn man in Nr. **1** – unter entsprechenden Annahmen über Differenzierbarkeit – die Ausdrücke (4) als Näherungswerte der Ableitungen $y_\nu^{(k)}$ auffaßt, so kann man dabei an die unten unter (14) angegebene Grenzwertformel denken, und analog an Formel (15), wenn man den Ausdruck (8) als Näherungswert für $\frac{d^k}{dx^k} P^{(k)}(x, y(x), \ldots, y^{(m)}(x))$ an der Stelle x_λ ansieht. Umgekehrt entspricht in Nr. **3** der Ausdruck (11) – als Näherungswert von $y_\nu^{(k)}$ aufgefaßt – der Formel (15) und der Ausdruck (13), angesehen als Näherungswert für $\frac{d^k}{dx^k} P^{(k)}(x, y(x), \ldots, y^{(m)}(x))$ an der Stelle x_λ der Formel (14).

4. Es war eine äquidistante Zerlegung, die wir, im Anschluß an Euler-Carathéodory, in Nr. **1** (und ebenso in Nr. **3**) betrachtet haben. Man könnte an die Möglichkeit denken, die Ueberlegungen auch auf nicht-äquidistante Zerlegungen auszudehnen, etwa indem man n positive Zahlen $r_1, \ldots, r_n$ mit der Summe 1 wählt, dann das Intervall $[a, b]$ zunächst in N Teilintervalle $\frac{b-a}{N} = u$ zerlegt und nunmehr jedes dieser Teilintervalle in n Teile unterteilt, die der Reihe nach die Längen $r_1 u, r_2 u, \ldots, r_n u$ haben, sodaß Nn die Gesamtanzahl aller Teile ist. Wieder mögen mit $x_\nu (\nu = 1, \ldots, Nn - 1)$ die zwischen a und b eingeschalteten Teilungspunkte, mit y_ν die Werte einer Funktion $y(x)$ an diesen

Stellen x_ν, und mit $J = \sum Z_\nu(x_\nu - x_{\nu-1})$ ein Näherungswert des Integrals (1) bezeichnet werden, wobei $Z_\nu = f(x_\nu, y_\nu, \overline{y'_\nu}, \ldots, \overline{y_\nu^{(m)}})$ ist und für die $\overline{y'_\nu}, \ldots, \overline{y_\nu^{(m)}}$ Näherungswerte für die Werte von $y', \ldots, y^{(m)}$ an der Stelle x_ν einzusetzen sind, die sich aus Funktionswerten $y(x)$ an Stellen x in der Nähe von x_ν berechnen lassen. Für $k+1$ aufeinander folgende Stellen $x_h, x_{h+1}, \ldots, x_{h+k}$ kann man ja dasjenige Lagrange'sche Interpolationspolynom von einem Grad $\leqq k$ bilden, das an diesen Stellen der Reihe nach die Werte $y_h, y_{h+1}, \ldots, y_{h+k}$ annimmt; und der mit $k!$ multiplizierte Koeffizient von x^k in diesem Polynom kann dann als Näherungswert für $y^{(k)}(x)$ an einer Stelle x_ν angesehen werden, in deren Nähe $x_h, x_{h+1}, \ldots, x_{h+k}$ gewählt wurden[6]. Dabei spielt, was die Auffassung dieser Koeffizienten als Näherungswerte der $y^{(k)}(x_\nu)$ betrifft, ein Grenzwertsatz herein, auf den an anderer Stelle hingewiesen werde[7]. Zur Bestimmung eines Maximums oder Minimums kann man dann wieder die Gleichungen $\frac{\partial J}{\partial y_\lambda} = 0$ $(\lambda = 1, \ldots, Nn - 1)$ aufstellen. Während sich aber im äquidistanten Fall in diesen Gleichungen Ausdrücke einstellen – wir haben sie oben mit $K_\lambda^{(k)}$ bezeichnet – die als Näherungswerte für $\frac{d^k}{dx^k} \frac{\partial f}{\partial y^k}$ angesprochen werden können, – und zwar nicht nur bei der durch (5) gegebenen Wahl der Stellen $x_h, x_{h+1}, \ldots, x_{h+k}$, sondern ebenso bei der Wahl gemäß (10), – scheint es mir zweifelhaft, ob im allgemeinen Fall verschieden großer $x_{\nu+1} - x_\nu$ ein analoges Verfahren funktioniert.

---

[6] Die Formeln (4) fallen im äquidistanten Fall $(n = 1)$ unter dieses Verfahren.

[7] „Zwei Sonderfälle eines Grenzwertsatzes", Mathematische Nachrichten, Band 13. Speziell im äquidistanten Fall (vgl. l. c. Formeln (12a), (12b)) handelt es sich um die – unter entsprechenden Voraussetzungen über $\varphi(x)$ geltende Aussage

$$(14) \qquad \lim_{u \to 0} \frac{1}{u^k} \sum_{i=0}^{k} (-1)^{k-i} \binom{k}{i} \varphi(\xi + iu) = \varphi^{(k)}(\xi)$$

oder die ihr gleichwertige, durch Vertauschung von u mit $-u$ daraus erhältliche

$$(15) \qquad \lim_{u \to 0} \frac{1}{u^k} \sum_{i=0}^{k} (-1)^{i} \binom{k}{i} \varphi(\xi - iu) = \varphi^{(k)}(\xi).$$

ZOBODAT - www.zobodat.at

Zoologisch-Botanische Datenbank/Zoological-Botanical Database

Digitale Literatur/Digital Literature

Zeitschrift/Journal: Sitzungsberichte der mathematisch-physikalischen Klasse der Bayerischen Akademie der Wissenschaften München

Jahr/Year: 1956

Band/Volume: 1955

Autor(en)/Author(s): Tietze Heinrich

Artikel/Article: Bemerkungen zu Carathéodory's Einführung in Euler's Arbeiten über Variationsrechnung 1-7