Vorlesungen über Variationsrechnung

Bolza, Oskar

Leipzig [u.a.], 1909

Erstes Kapitel. Die erste Variation ber der einfachsten Klasse von Aufgaben

urn:nbn:at:at-ubi:2-6614

25

Erstes Kapitel.

Die erste Variation bei der einfachsten Klasse von Aufgaben.

§ 1. Vorläufige Orientierung über die wichtigsten Probleme der Variationsrechnung.

Die Variationsrechnung beschäftigt sich mit Aufgaben des *Maximums und Minimums,* die jedoch von wesentlich komplizierterer Natur sind als diejenigen, welche in der aus der Differentialrechnung bekannten Theorie der gewöhnlichen Maxima und Minima behandelt werden. Während es sich nämlich dort darum handelt, diejenigen Werte einer oder mehrerer Variabeln zu bestimmen, für welche eine gegebene Funktion dieser Variabeln den größten oder kleinsten Wert annimmt, hat es die Variationsrechnung mit Aufgaben zu tun, bei denen eine oder mehrere unbekannte Funktionen so zu bestimmen sind, daß ein gegebenes, von der Wahl dieser Funktionen abhängiges *bestimmtes Integral* seinen größten oder kleinsten Wert annimmt.

Dies soll zunächst durch Besprechung einiger typischer Beispiele näher erläutert werden.

Beispiel I. In einer Ebene sind zwei Punkte P_1, P_2 und eine Gerade $\mathfrak{L}$ gegeben. Es wird verlangt, unter allen Kurven, welche in dieser Ebene von P_1 nach P_2 gezogen werden können, diejenige zu bestimmen, welche durch ihre Rotation um die Gerade $\mathfrak{L}$ die Fläche von kleinstem Inhalt erzeugt.

Wir wählen die Gerade $\mathfrak{L}$ zur x-Achse eines rechtwinkligen Koordinatensystems und bezeichnen die Koordinaten der beiden gegebenen Punkte mit x_1, y_1 und x_2, y_2. Ist dann

$$\mathfrak{C}: \qquad y = y(x)$$

irgend eine die beiden Punkte P_1 und P_2 verbindende Kurve, so wird der fragliche Flächeninhalt durch das bestimmte Integral[1])

$$J = 2\pi \int_{x_1}^{x_2} y \sqrt{1 + y'^2}\, dx$$

ausgedrückt, wenn y' die Ableitung der Funktion $y(x)$ bedeutet. Der Wert des Integrals J hängt von der Wahl der Kurve $\mathfrak{C}$, d. h. der Funktion $y(x)$ ab, und unsere Aufgabe lautet also analytisch[2]): Unter allen Funktionen $y(x)$, welche für $x = x_1$ und $x = x_2$ die vorgeschriebenen Werte y_1, bzw. y_2, annehmen, diejenige zu bestimmen, für welche das Integral J seinen kleinsten Wert annimmt.

Dieses Beispiel ist ein Repräsentant der *einfachsten Klasse von Aufgaben* der Variationsrechnung, bei denen es sich darum handelt, ein Integral von der Form

$$J = \int_{x_1}^{x_2} f(x, y, y')\, dx, \qquad y' = \frac{dy}{dx}, \tag{1}$$

durch passende Wahl der unbekannten Funktion y zu einem Extremum[3]) zu machen. Mit dieser einfachsten Klasse von Aufgaben, die zugleich von grundlegender Bedeutung ist, werden wir uns in den drei ersten Kapiteln beschäftigen.

Wir haben der Einfachheit halber bei der analytischen Formulierung von Beispiel I die zu betrachtenden Kurven in der Form: $y = y(x)$ angenommen, unter $y(x)$ eine eindeutige Funktion verstanden, d. h. wir haben vorausgesetzt, daß jede der y-Achse parallele Gerade die betreffende Kurve höchstens in einem Punkte trifft. Dies ist jedoch eine Beschränkung, die durchaus nicht in der Natur der Aufgabe liegt. Wir können uns von derselben befreien, indem wir sämtliche Kurven in Parameterdarstellung ansetzen:

$$x = x(t), \qquad y = y(t).$$

---

[1]) Wir werden in der Folge die positive Quadratwurzel aus einer reellen positiven Größe a stets mit $\sqrt{a}$ bezeichnen, während die Bezeichnung $\sqrt{a}$ andeuten soll, daß das Vorzeichen unbestimmt bleibt.

[2]) Die in diesem Paragraphen gegebene Formulierung der verschiedenen Probleme ist nur als eine vorläufige zu betrachten und bedarf nach verschiedenen Richtungen hin einer schärferen Präzisierung.

[3]) Das Wort „*Extremum*" wird nach dem Vorgang von Du Bois-Reymond (Mathematische Annalen, Bd. 15 (1879), p. 564) in gleicher Weise für „Maximum" und „Minimum" gebraucht in solchen Fällen, wo es nicht nötig ist, zwischen beiden zu unterscheiden.

Unser bestimmtes Integral geht dann über in:

$$J = 2\pi \int_{t_1}^{t_2} y \sqrt{x'^2 + y'^2}\, dt,$$

wo jetzt x', y' die Ableitungen von x und y nach t bedeuten.

Allgemeiner nimmt das Integral (1) bei Parameterdarstellung die Form an:

$$J = \int_{t_1}^{t_2} F(x, y, x', y')\, dt, \tag{2}$$

wobei F in bezug auf x', y' homogen von der ersten Dimension ist. Das Problem, ein Integral von der Form (2) zu einem Extremum zu machen, bildet *die einfachste Klasse von Variationsproblemen in Parameterdarstellung.* Wir werden diese Klasse von Problemen, die für geometrische Anwendungen von größter Wichtigkeit ist, und deren Theorie am eingehendsten ausgebildet worden ist, in Kap. V—IX behandeln.

Einen Typus von wesentlich anderer Art repräsentiert das folgende

Beispiel II[1]*: Unter allen Kurven von vorgeschriebener Länge l, welche zwei gegebene Punkte P_1, P_2 verbinden, diejenige zu bestimmen, welche zusammen mit der Sehne $P_1 P_2$ den größten Flächeninhalt einschließt.*

Wählen wir die Gerade durch die beiden Punkte $P_1(x_1, y_1)$ und $P_2(x_2, y_2)$ zur x-Achse eines rechtwinkligen Koordinatensystems und beschränken uns der Einfachheit halber auf Kurven, die in der Form

$$y = y(x)$$

darstellbar sind, so lautet die Aufgabe in analytischer Formulierung:

Unter allen Funktionen: $y = y(x)$, welche den Anfangsbedingungen

$$y(x_1) = 0, \qquad y(x_2) = 0,$$

und überdies der Bedingung

$$\int_{x_1}^{x_2} \sqrt{1 + y'^2}\, dx = l$$

genügen, diejenige zu finden, welche dem Integral

[1]) Zuerst vorgelegt von Jacob Bernoulli 1697, siehe Stäckel's Übersetzung in Nr. 46 von Ostwald's *Klassiker der exakten Wissenschaften*, p. 19 und Anmerkung 19) auf p. 139; weitere Literatur bei Pascal, *Die Variationsrechnung* (Leipzig, 1899), pp. 127, 128.

$$J = \int_{x_1}^{x_2} y\,dx$$

den größten Wert erteilt.

Das charakteristisch Neue besteht hier also darin, daß den zum Vergleich herangezogenen Kurven außer den Anfangsbedingungen noch die weitere Beschränkung auferlegt wird, daß sie einem gewissen bestimmten Integral einen vorgeschriebenen Wert erteilen sollen.

Unser Beispiel ist also ein spezieller Fall der Aufgabe, ein Integral von der Form

$$J = \int_{x_1}^{x_2} f(x, y, y')\,dx \tag{1}$$

zu einem Extremum zu machen, während gleichzeitig die zulässigen Kurven einem zweiten Integral

$$K = \int_{x_1}^{x_2} g(x, y, y')\,dx \tag{3}$$

einen vorgeschriebenen Wert l erteilen sollen.

Aufgaben dieser Art heißen *„isoperimetrische Probleme"*; wir werden dieselben in Kapitel X betrachten, und zwar, im Hinblick auf die vielen geometrischen Aufgaben dieser Art, in Parameterdarstellung.

Die nächstliegende Verallgemeinerung besteht darin, daß man bestimmte Integrale betrachtet, in denen die Funktion unter dem Integral nicht nur die erste, sondern auch *höhere Ableitungen* der unbekannten Funktion y enthält, also Integrale der Form

$$J = \int_{x_1}^{x_2} f(x, y, y', y'', \ldots, y^{(n)})\,dx\,. \tag{4}$$

Diese Klasse von Aufgaben hat in der Geschichte der Variationsrechnung eine gewisse Rolle gespielt, ist jedoch gegenwärtig, wo die rein formalen Fragen mehr in den Hintergrund getreten sind, von geringerem Interesse, teils weil sie in einer weiter unten zu betrachtenden allgemeineren Klasse als Spezialfall enthalten ist, teils weil kaum irgendwelche interessante geometrische oder mechanische Probleme zu dieser Kategorie gehören.[1])

---

[1]) Wir geben am Ende von Kap. III einige Andeutungen über Aufgaben dieser Art; im übrigen verweisen wir auf Pascal, loc. cit., §§ 2—7, 16, 19—22, und Zermelo, *Untersuchungen zur Variationsrechnung, Dissertation* (Berlin 1894), sowie §§ 1, 4, 6, 12, 13, 14 in Kneser's Artikel über Variationsrechnung in der *Encyklopädie der mathematischen Wissenschaften*, II A 8, und Kneser, *Lehrbuch der Variationsrechnung* (Braunschweig, 1900), 6. Abschnitt.

Weit wichtiger ist eine zweite Verallgemeinerung, bei welcher *mehrere unbekannte Funktionen* unter dem bestimmten Integral vorkommen.

Besonders interessant werden die Aufgaben dieser Art, wenn zwischen den unbekannten Funktionen Relationen vorgeschrieben sind. Hierher gehört das folgende

Beispiel III: Die kürzeste Linie zu bestimmen, welche auf einer in der Form

$$\varphi(x, y, z) = 0$$

gegebenen Fläche zwischen zwei auf der Fläche gegebenen Punkten gezogen werden kann.

Nimmt man die zulässigen Kurven in der Form

$$y = y(x), \qquad z = z(x)$$

darstellbar an, so hat man unter allen der Bedingung

$$\varphi(x, y(x), z(x)) = 0$$

genügenden Kurven, welche durch die beiden gegebenen Punkte gehen, diejenige zu bestimmen, für welche das Integral

$$J = \int_{x_1}^{x_2} \sqrt{1 + y'^2 + z'^2}\, dx$$

den kleinsten Wert annimmt.

Hier haben wir also *zwei unbekannte Funktionen* von x zu bestimmen, welche *durch eine endliche Gleichung verbunden* sind.

Die vorgeschriebenen Relationen zwischen den unbekannten Funktionen können aber auch die Form von Differentialgleichungen haben, wie das folgende, schon von EULER[1]) und LAGRANGE[2]) behandelte Beispiel zeigt.

Beispiel IV: Die Brachistochrone im widerstehenden Medium. Unter allen Raumkurven, welche zwischen zwei gegebenen Punkten P_1 und P_2 gezogen werden können, soll diejenige bestimmt werden, entlang welcher ein schwerer materieller Punkt in der kürzesten Zeit von P_1 nach P_2 gelangt, wenn er den Punkt P_1 mit einer gegebenen Anfangsgeschwindigkeit v_1 verläßt. Die Reibung soll vernachlässigt

---

[1]) Vgl. *Methodus inveniendi lineas curvas maximi minimive proprietate gaudentes* (Lausanne, 1744), pp. 126, 214.

[2]) Vgl. *Œuvres*, Bd. 10, p. 440; vgl. ferner LINDELOEF-MOIGNO, *Calcul des Variations* (Paris, 1861), p. 308, und KNESER, *Lehrbuch*, p. 248.

werden, dagegen soll der Widerstand des Mediums berücksichtigt werden.

Die positive z-Achse eines rechtwinkligen Koordinatensystems werde vertikal nach unten gewählt und die Masse des materiellen Punktes gleich 1 angenommen; v bezeichne die Geschwindigkeit, g die Konstante der Schwere. Der Widerstand sei eine gegebene Funktion der Geschwindigkeit, $R(v)$ sei der absolute Wert derselben. Dann erhält man aus dem Prinzip der lebendigen Kraft[1]):

$$d\frac{v^2}{2} = g\,dz - R(v)\sqrt{dx^2 + dy^2 + dz^2}\,.$$

Nehmen wir daher der Einfachheit halber x als unabhängige Variable, so lautet die Aufgabe in analytischer Fassung: Unter allen Funktionensystemen

$$y = y(x), \qquad z = z(x), \qquad v = v(x),$$

welche den Anfangsbedingungen

$$\begin{aligned} y(x_1) &= y_1, & z(x_1) &= z_1, & v(x_1) &= v_1, \\ y(x_2) &= y_2, & z(x_2) &= z_2, \end{aligned}$$

und der Differentialgleichung

$$vv' = gz' - R(v)\sqrt{1 + y'^2 + z'^2}$$

genügen, dasjenige zu bestimmen, welches das Integral

$$J = \int_{x_1}^{x_2} \frac{\sqrt{1 + y'^2 + z'^2}\,dx}{v}$$

zu einem Minimum macht.

Wir haben hier also *drei unbekannte Funktionen*, welche *durch eine Differentialgleichung verbunden* sind.

In naturgemäßer Verallgemeinerung gelangt man so schließlich zu folgender, gewöhnlich als *Lagrange'sches Problem* bezeichneter Aufgabe: Unter allen Funktionensystemen $y_1, y_2, \ldots, y_n$ einer unabhängigen Variabeln x, welche gewissen Anfangsbedingungen und außerdem einer Anzahl von Nebenbedingungen von der Form

$$\varphi_k(x, y_1, y_2, \ldots, y_n;\ y_1', y_2' \ldots, y_n') = 0 \tag{5}$$
$$k = 1, 2, \ldots, m, \quad (m < n)$$

---

[1]) Vgl. z. B. Appell, *Traité de Mécanique rationelle*, Bd. I, Nr. 220, 251 und Sturm, *Cours de Mécanique* (5e éd.), Bd. I, Nr. 274.

genügen, dasjenige zu bestimmen, für welches das Integral

$$J = \int_{x_1}^{x_2} f(x, y_1, y_2, \ldots, y_n;\, y_1', y_2', \ldots, y_n')\, dx$$

seinen kleinsten oder größten Wert annimmt.

Dabei soll der Fall $m = 0$ mit inbegriffen sein und ebenso der Fall, in welchem einige der Funktionen φ_k, oder alle, die Ableitungen $y_1', y_2', \ldots, y_n'$ nicht enthalten.

Dieses Problem ist von ganz besonderer Wichtigkeit, einmal weil es *das allgemeinste Variationsproblem für einfache bestimmte Integrale* darstellt, insofern alle andern sich auf dieses reduzieren lassen, sodann weil die allgemeinen Variationsprinzipien der Mechanik, wie das Hamilton'sche Prinzip und das Prinzip der kleinsten Aktion auf Probleme dieser Art führen. Diese Klasse von Aufgaben werden wir im XI. und XII. Kapitel behandeln.

Endlich kann man auch *mehrfache Integrale* in den Kreis der Betrachtung ziehen. Ein klassisches Beispiel dieser Art ist das Problem der Minimalflächen:

Beispiel V: Unter allen Flächen, welche von einer gegebenen geschlossenen Raumkurve $\mathfrak{L}$ begrenzt werden, diejenige zu bestimmen, welche den kleinsten Flächeninhalt besitzt.

Beschränken wir uns der Einfachheit halber auf Flächen, welche, bezogen auf ein rechtwinkliges Koordinatensystem, in der Form

$$z = f(x, y)$$

darstellbar sind, und bezeichnen wir mit $\mathcal{S}$ den Bereich der x, y-Ebene, welcher von der Projektion $\mathfrak{L}'$ der Kurve $\mathfrak{L}$ begrenzt wird, so lautet die Aufgabe in analytischer Formulierung: Unter allen Funktionen von zwei unabhängigen Variabeln

$$z = f(x, y),$$

welche entlang der Kurve $\mathfrak{L}'$ vorgeschriebene, sich stetig aneinanderreihende Werte annehmen, diejenige zu bestimmen, für welche das Doppelintegral

$$J = \iint_{\mathcal{S}} \sqrt{1 + \left(\frac{\partial z}{\partial x}\right)^2 + \left(\frac{\partial z}{\partial y}\right)^2}\, dx\, dy$$

den kleinsten Wert annimmt.

Da beim Übergang zu mehrfachen Integralen die Schwierigkeiten ganz erheblich zunehmen, und da die Theorie hier noch nicht zu einem ähnlichen Abschluß gekommen ist wie bei den einfachen Integralen, so werden wir uns bei der Behandlung dieser Klasse von Aufgaben auf die Entwicklung der einfachsten Sätze beschränken. –

Nach dieser allgemeinen Übersicht wenden wir uns jetzt zur näheren Betrachtung der einfachsten Klasse von Variationsproblemen, wobei es sich zunächst um eine genaue Formulierung der Aufgabe handeln wird.

§ 2. Definitionen und Sätze über „gewöhnliche" Maxima und Minima.

Bei der fundamentalen Rolle, welche die Begriffe des Maximums und Minimums in der Variationsrechnung spielen, empfiehlt es sich, die wichtigsten Definitionen und Sätze über „gewöhnliche" Maxima und Minima vorauszuschicken.

a) **Maximum und Minimum einer linearen Punktmenge:**

Unter einer endlichen Anzahl von Werten einer unabhängigen Variabeln x gibt es stets mindestens ein Maximum, d. h. einen Wert, der größer oder wenigstens nicht kleiner ist als alle übrigen; und ebenso ein Minimum.

Ganz anders verhält es sich dagegen bei Mengen von unendlich vielen Werten (sogenannten unendlichen Wertmengen oder linearen Punktmengen). Hier kann es zunächst vorkommen, daß ein Maximum deshalb nicht existiert, weil es unter den Werten der Menge solche gibt, die jede vorgegebene Zahl übersteigen. Dies ist z. B. der Fall bei der aus der Gesamtheit aller positiven ganzen Zahlen gebildeten Menge. Man sagt in diesem Fall, die Menge sei nach oben nicht beschränkt. Aber auch wenn die Menge nach oben beschränkt ist (borné supérieurement[1]) d. h. wenn sämtliche Werte der Menge unterhalb einer festen Größe („Schranke") L liegen, so braucht es deshalb doch keinen größten Wert der Menge zu geben. So ist z. B. die unendliche Menge

$$0,\ \frac{1}{2},\ \frac{2}{3},\ \frac{3}{4},\ \ldots,\ \frac{\nu-1}{\nu},\ \ldots$$

nach oben beschränkt, da sämtliche Werte der Mengen kleiner als z. B. 2 sind; trotzdem gibt es unter den Werten dieser Menge keinen größten.

---

[1]) Vgl. A I 2; die Abkürzung A bedeutet stets den *Anhang* am Ende des Buches.

Dagegen gilt der Satz[1]), daß jede nach oben begrenzte Wertmenge (lineare Punktmenge) eine *obere Grenze* besitzt, d. h. es gibt eine (und nur eine) Größe G, welche folgende beiden charakteristischen Eigenschaften hat:

1. jeder Wert p der Menge ist $\overline{\gtrless}\, G$;

2. wie klein auch die positive Größe ε gewählt sein mag, so gibt es stets mindestens einen Wert p' der Menge für welchen $p' > G - \varepsilon$.

Diese obere Grenze G gehört nun entweder selbst zur Menge und heißt alsdann das *Maximum der Menge* oder aber sie gehört nicht zur Menge; in diesem Fall sagt man, die Menge besitzt kein Maximum. Die entsprechenden Definitionen und Sätze gelten für die *untere Grenze* und das *Minimum* einer Menge.

Für das obige Beispiel ist die obere Grenze 1, dieselbe gehört selbst nicht zur Menge; dagegen gehört die untere Grenze 0 zur Menge und ist daher zugleich das Minimum.

Der Gleichförmigkeit halber pflegt man von einer nach oben (unten) nicht beschränkten Menge zu sagen ihre obere (untere) Grenze sei $+\infty$ $(-\infty)$.

Bei diesem Sprachgebrauch, welchem wir folgen werden, besitzt also jede lineare Punktmenge eine obere Grenze und eine untere Grenze.

b) **Absolutes Maximum und Minimum einer Funktion in einem Intervall:**[2])

Es sei jetzt

$$y = f(x)$$

eine reelle Funktion einer reellen Variabeln, definiert in einem Intervall[3]) $[a\,b]$. Die dem Intervall $[a\,b]$ von x entsprechenden Werte von y bilden eine lineare Punktmenge, die wir mit Y bezeichnen. Diese Punktmenge hat nach dem vorigen stets eine obere Grenze G und eine untere Grenze K, welche die obere, resp. untere, Grenze der Funktion $f(x)$ in dem Intervall $[a\,b]$ genannt werden.

Sind G und K beide endlich, so heißt die Funktion $f(x)$ *endlich im Intervall* $[ab]$.

---

[1]) Vgl. A I 3.

[2]) Hierzu die *Übungsaufgaben* 1_a, 1_b, 1_c am Ende von Kap. III.

[3]) d. h. für alle x, für welche $a \overline{\overline{<}}\, x \overline{\overline{<}}\, b$, (falls $a < b$), $a \overline{\overline{>}}\, x \overline{\overline{>}}\, b$, (falls $a > b$); im allgemeinen soll mit der Bezeichnung $[a\,b]$ die Annahme $a < b$ verbunden sein. Für eine Funktion, welche nicht für ein stetiges Intervall, sondern für eine beliebige Punktmenge X im Gebiet der Variabeln x definiert ist, (vgl. *Encyclopädie*, II A, p. 11 und Jordan, *Cours d'Analyse* (Paris, 1893), I, Nr. 41), ersetze man in den folgenden Definitionen überall das Intervall $[a\,b]$ durch die Menge X.

Gehört die obere Grenze G zur Menge Y d. h. gibt es wenigstens einen Wert c des Intervalls $[ab]$, so daß: $f(c) = G$, während gleichzeitig: $f(x) \overline{\gtrless} G$ für jedes x in $[ab]$, so heißt[1] G das *absolute Maximum der Funktion $f(x)$ im Intervall $[ab]$*. Gehört dagegen G nicht zu Y, so besitzt die Funktion $f(x)$ im Intervall $[ab]$ kein absolutes Maximum; letzteres ist stets der Fall, wenn $G = +\infty$, da jedes einzelne y als endlich vorausgesetzt wird.

Analoge Definitionen gelten für das absolute Minimum.

Für *stetige* Funktionen gilt der Satz[2]: Ist $f(x)$ stetig im Intervall $[ab]$, so ist sowohl die obere als die untere Grenze endlich und beide werden wirklich erreicht, d. h. $f(x)$ besitzt in $[ab]$ ein absolutes Maximum und ein absolutes Minimum.

Die oben gegebenen Definitionen lassen sich ohne weiteres auf Funktionen mehrerer Variabeln ausdehnen.

c) **Relatives Maximum und Minimum einer Funktion:**

Von dem absoluten Extremum ist das *relative*[3]) zu unterscheiden. Man sagt eine Funktion $f(x)$, welche in einem Intervall $[ab]$ definiert ist, besitzt ein relatives Minimum an einer Stelle $x = x_0$ des Intervalls $[ab]$, wenn eine positive Größe d existiert, derart, daß

$$f(x_0) \overline{\gtrless} f(x)$$

für jedes x des Intervalls $[ab]$, für welches $|x - x_0| < d$; oder, wie wir auch schreiben können,

$$f(x_0 + h) - f(x_0) \overline{\gtrless} 0 \tag{7}$$

für jedes h, für welches $|h| < d$ und $x_0 + h$ in $[ab]$.

Nach Stolz[4]) unterscheidet man dabei das eigentliche von dem uneigentlichen Minimum: Bei dem *eigentlichen Minimum* läßt sich d so klein wählen, daß

$$f(x_0 + h) - f(x_0) > 0 \tag{8}$$

für alle angegebenen Werte von h außer $h = 0$; bei dem uneigentlichen gibt es, wie klein auch d gewählt sein mag, immer noch von Null verschiedene Werte h', für welche

---

[1]) Vgl. *Encyclopädie*, II A, p. 80.

[2]) Vgl. A III 3.

[3]) Ich folge der Terminologie von Voss in *Encyclopädie*, II A, p. 81; vielfach wird „relatives Extremum" für „Extremum mit Nebenbedingungen" gebraucht; vgl. unten, Kap. X.

[4]) *Grundzüge der Differential- und Integralrechnung*, (Leipzig, 1893), Bd. I, p. 199.

$$f(x_0 + h') - f(x_0) = 0, \qquad |h'| < d.$$

Analoge Definitionen gelten für das relative Maximum.

Beispiele von uneigentlichen Extremen:

a) Die Funktion: $y = konst.$ hat in jedem Punkte ein uneigentliches Maximum und zugleich ein uneigentliches Minimum.

b) Die Funktion [1])

$$y = \begin{cases} \left(x \sin \frac{1}{x}\right)^2 & \text{für } x \neq 0 \\ 0 & \text{für } x = 0 \end{cases}$$

hat für $x = 0$ ein uneigentliches relatives Minimum. Denn hier ist

$$f(x) \geqq f(0), \quad \text{für jedes } x;$$

zugleich gibt es in jeder Nähe der Stelle $x = 0$ Werte von x, für welche $f(x) = f(0)$, nämlich die Werte $x = 1/\mu\pi$, wo μ eine ganze Zahl.

Besitzt die Funktion $f(x)$ im Punkt $x = x_0$ eine Ableitung $f'(x_0)$, so gilt der *Satz* [2]):

Für das Eintreten eines relativen Extremums in einem innern Punkt x_0 des Intervalls $[ab]$ ist notwendig, daß

$$f'(x_0) = 0. \tag{9}$$

Denn nach der Definition der Ableitung hat man in diesem Fall:

$$f(x_0 + h) - f(x_0) = h\,[f'(x_0) + (h)], \tag{10}$$

wo (h), wie stets in der Folge, in Weierstraß'scher Bezeichnung eine unendlich kleine Funktion von h bezeichnet, d. h.

$$\underset{h=0}{L}\,(h) = 0. \qquad = \text{Lim}$$

Wäre nun $f'(x_0) \neq 0$, so könnte man eine positive Größe δ angeben, so daß $|(h)| < |f'(x_0)|$, und daher $f'(x_0) + (h)$ von demselben Zeichen wie $f'(x_0)$ wäre für alle $|h| < \delta$. Für solche Werte von h

---

[1]) Vgl. *Encyclopädie*, II A, p. 81.

[2]) Vgl. *Encyclopädie*, II A, p. 82; Jordan, *Cours d'Analyse*, I, Nr. 394; Peano, *Differentialrechnung und Grundzüge der Integralrechnung* (Leipzig, 1899), Nr. 131; unter allgemeineren Voraussetzungen bei Stolz, *Grundzüge*, I, pp. 203, 204, 207. Ist die Funktion $f(x)$ regulär im Punkt x_0, d. h. nach ganzen positiven Potenzen von $x - x_0$ entwickelbar, so kann man die obigen Resultate auch mittels des Satzes beweisen: Für hinreichend kleine Werte von $|h|$ hat die Potenzreihe

$$a_k h^k + a_{k+1} h^{k+1} + \ldots, \quad a_k \neq 0,$$

dasselbe Vorzeichen wie das erste Glied $a_k\, h^k$, vgl. Stolz, *Grundzüge*, I, p. 205; Kneser, *Lehrbuch*, § 7.

hätte daher $f(x_0 + h) - f(x_0)$ dasselbe oder das entgegengesetzte Zeichen wie $f'(x_0)$, jenachdem h positiv oder negativ. Es könnte also weder ein Maximum noch ein Minimum stattfinden.

Derselbe Schluß läßt sich nicht anwenden, wenn x_0 mit einem der Endpunkte des Intervalls $[ab]$ zusammenfällt. Nehmen wir an, daß $a < b$, so kann h für $x_0 = a$ nur positive, für $x_0 = b$ nur negative Werte annehmen. Daher folgt aus (10): Für das Eintreten eines relativen Maximums (Minimums) im unteren Endpunkt a des Intervalls $[ab]$ ist notwendig daß $\overset{+\prime}{f}(a) \leqq 0 \; (\geqq 0)$, im oberen Endpunkt b: $\overset{-\prime}{f}(b) \geqq 0 \; (\leqq 0)$, wobei $\overset{+\prime}{f}$ und $\overset{-\prime}{f}$ die vordere und hintere Derivierte bedeuten.[1])

Wenn in einer gewissen Umgebung eines inneren Punktes x_0 die n ersten Ableitungen von $f(x)$ existieren und stetig sind, so gilt der *Satz*[2]): *Wenn*

$$f'(x_0) = 0, \quad f''(x_0) = 0, \quad \ldots, \quad f^{(n-1)}(x_0) = 0, \quad f^{(n)}(x_0) \neq 0,$$

so besitzt die Funktion $f(x)$ für $x = x_0$ kein Extremum, wenn n ungerade; dagegen besitzt sie ein relatives Extremum, falls n gerade ist, und zwar ein eigentliches Minimum, wenn $f^{(n)}(x_0) > 0$, ein eigentliches Maximum, wenn $f^n(x_0) < 0$.

Denn unter den gemachten Annahmen läßt sich der Taylor'sche Satz anwenden, nach welchem

$$f(x_0 + h) - f(x_0) = \frac{h^n}{n!} f^{(n)}(x_0 + \theta h), \quad 0 < \theta < 1, \tag{11}$$

woraus unter Zuziehung der vorausgesetzten Stetigkeit von $f^{(n)}(x)$ der Satz unmittelbar folgt.

Für die Endpunkte ist der Satz wieder ähnlich wie oben zu modifizieren.

Aus den oben gegebenen Definitionen folgt unmittelbar: Besitzt die Funktion $f(x)$ für $x = x_0$ ein absolutes Maximum (Minimum) in bezug auf ein den Punkt x_0 enthaltendes Intervall $[ab]$, so besitzt sie a fortiori auch ein relatives Maximum (Minimum) für $x = x_0$. Hierdurch reduziert sich die Bestimmung der absoluten Extrema auf die der relativen. Hat man letztere gefunden und weiß man a priori, daß die Funktion $f(x)$ im Intervall $[ab]$ ein absolutes Maximum (Minimum) besitzt, so hat man nur unter den relativen Maximalwerten (Minimalwerten) den

[1]) Vgl. A IV 1.

[2]) Vgl. p. 11, Fußnote [2]).

größten (kleinsten) auszusuchen. Neue Schwierigkeiten treten nur dann auf, wenn die Funktion unendlich viele Maxima und Minima besitzt.[1])

§ 3. Definition des Maximums und Minimums eines bestimmten Integrals.[2])

Ganz analoge Begriffsbildungen treten nun auch bei der Definition des Maximums und Minimums eines bestimmten Integrals auf.

Wir werden uns dabei (sowie stets in der Folge) der folgenden abgekürzten Ausdrucksweise bedienen: Indem wir unter einer *Funktion* stets eine *eindeutige*, *reelle* Funktion einer oder mehrerer *reeller Variabeln* verstehen, sagen wir, eine Funktion $f(x)$ einer Variabeln x, welche in einem Intervall $[ab]$ definiert ist, sei in diesem Intervall *von der Klasse* C', wenn sie in[3]) $[ab]$ stetig ist und eine stetige erste Ableitung $f'(x)$ besitzt; *von der Klasse* C'', wenn außerdem die zweite Ableitung $f''(x)$ existiert und stetig ist in $[ab]$, und so fort, wobei man beachte, daß die Klasse $C^{(n+1)}$ in der Klasse $C^{(n)}$ enthalten ist.

Ebenso soll die *Kurve*

$$y = f(x), \qquad a \overline{\gtrless} x \overline{\gtrless} b,$$

---

[1]) Hierzu *Übungsaufgabe* 1_d, am Ende von Kap. III.

[2]) Bis in das letzte Drittel des vorigen Jahrhunderts hat allgemein große Unklarheit über die Grundlagen der Variationsrechnung geherrscht. Das größte Verdienst um die Klärung der Grundbegriffe und um eine scharfe Formulierung der Aufgaben haben: Du Bois-Reymond, *„Erläuterungen zu den Anfangsgründen der Variationsrechnung“*, Mathematische Annalen, Bd. XV. (1879), p. 283; Scheeffer: *„Über die Bedeutung der Begriffe Maximum und Minimum in der Variationsrechnung“*, ibid. Bd. XXVI (1886), p. 197; und vor allem Weierstrass in seinen an der Berliner Universität gehaltenen *Vorlesungen über Variationsrechnung* (1865—1890). Wertvolle Beiträge nach dieser Richtung haben auch geliefert: Zermelo, *Dissertation*, p. 24; Kneser, *Lehrbuch*, § 17 und Osgood, *„Sufficient conditions in the Calculus of Variations“*, Annals of Mathematics (2), Bd. II (1901), p. 105.

[3]) Da der Begriff der Ableitung, wie er gewöhnlich definiert wird, nur eine Bedeutung hat für *innere* Punkte des Definitionsbereiches einer Funktion, so ist noch eine besondere Festsetzung bezüglich der Endpunkte a und b notwendig. Wir wollen dieselbe dahin formulieren, daß es möglich sein soll, die Definition der Funktion $f(x)$ so über das Intervall $[ab]$ hinaus auszudehnen, daß die erweiterte Funktion für $a' < x < b'$ die genannten Eigenschaften hat, wo $a' < a$, $b' > b$. Dies ist gleichbedeutend mit der Annahme, daß die betreffenden Ableitungen stetig sein sollen im Innern des Intervalls $[ab]$ und sich bestimmten endlichen Grenzen nähern sollen bei Annäherung an die Endpunkte (vgl. A IV 4).

von der Klasse C', resp. C'', ... heißen, wenn $f(x)$ in $[ab]$ von der Klasse C', resp. C'' ... ist.

Endlich sagen wir auch von einer Funktion von m Variabeln: $f(x_1, x_2, \ldots, x_m)$, sie sei in einem Bereich[1]) $\mathfrak{A}$ im Gebiet der Variabeln $x_1, x_2, \ldots, x_m$ von der Klasse $C^{(n)}$, wenn sie selbst samt ihren partiellen Ableitungen bis zur n^{ten} Ordnung inklusive stetig ist in[2]) $\mathfrak{A}$.

a) **Absolutes Extremum eines bestimmten Integrals:**

Es sei jetzt einerseits eine Funktion $f(x, y, y')$ der drei unabhängigen Variabeln x, y, y' gegeben, welche reell und von der Klasse[3]) C''' ist in einem Bereich $\mathfrak{T}$, welcher aus allen Punkten (x, y, y') besteht, für welche (x, y) einem gewissen Bereich $\mathfrak{R}$ der x, y-Ebene angehört, während y' irgend einen endlichen Wert haben kann.

Andererseits sei

$$\mathfrak{C}: \qquad y = y(x), \qquad x_1 \overline{\gtrless} x \overline{\gtrless} x_2,$$

eine ganz im[4]) Bereich $\mathfrak{R}$ gelegene Kurve der Klasse C'.

Alsdann ist die Funktion

$$f(x, y(x), y'(x))$$

---

[1]) Unter einem *„Bereich"* soll stets eine Punktmenge verstanden werden, welche „innere Punkte" enthält, einerlei von welcher Beschaffenheit sie sonst sein mag. Über die Definition eines inneren Punktes vgl. A I 7.

[2]) Dies hat eine unmittelbare Bedeutung wieder nur, wenn der Bereich $\mathfrak{A}$ nur innere Punkte enthält (oder, wie wir sagen, ein *„stetiger Bereich"* ist); enthält er auch *Begrenzungspunkte* (d. h. Punkte, in deren jeder Nähe es Punkte gibt, welche nicht zu $\mathfrak{A}$ gehören), so fügen wir noch die Festsetzung hinzu, daß sich die Definition von $f(x_1, x_2, \ldots, x_m)$ so über den Bereich $\mathfrak{A}$ hinaus auf einen stetigen, den Bereich $\mathfrak{A}$ enthaltenden Bereich $\mathfrak{B}$ fortsetzen lassen soll, daß die erweiterte Funktion die angegebenen Eigenschaften im Bereich $\mathfrak{B}$ besitzt.

[3]) Bei den geometrischen und mechanischen Anwendungen der Variationsrechnung ist die Funktion $f(x, y, y')$ meistens eine *analytische* Funktion einfachster Art. Es wäre daher vollständig genügend, die Untersuchung für analytische Funktionen f durchzuführen, wie es Weierstrass und Kneser getan haben. Dagegen muß man gerade auch vom Standpunkt der Anwendungen bei den meisten Aufgaben auch *nicht-analytische Kurven* zulassen, wenn man die Aufgabe nicht ganz unnatürlich einschränken will. Wenn man es aber doch einmal mit nicht-analytischen Funktionen zu tun hat, so wird die Darstellung einheitlicher, wenn man auch die Funktion f nicht als analytisch voraussetzt, wie dies auch schon Pascal, loc. cit. p. 21 und Osgood, loc. cit. p. 105 getan haben.

[4]) Eine Kurve liegt *„in einem Bereich"* soll stets bedeuten: jeder Punkt der Kurve ist zugleich ein Punkt des Bereiches, nicht notwendig ein innerer Punkt.

nach den Sätzen über zusammengesetzte Funktionen (A IV 9) eine im Intervall $[x_1 x_2]$ stetige Funktion von x, und daher hat das Integral

$$J = \int_{x_1}^{x_2} f(x,\ y(x),\ y'(x))\,dx \tag{12}$$

einen bestimmten endlichen Wert. Wir nennen dieses Integral das *Integral der Funktion $f(x, y, y')$ genommen entlang der Kurve $\mathfrak{C}$* und bezeichnen dasselbe mit

$$J = \int_{\mathfrak{C}} f(x,\ y,\ y')\,dx$$

oder kürzer mit $J_{\mathfrak{C}}$.

Es seien jetzt im Bereich $\mathfrak{R}$ zwei Punkte $P_1(x_1, y_1)$ und $P_2(x_2, y_2)$ gegeben, wobei wir stets $x_1 < x_2$ voraussetzen; wir betrachten die Gesamtheit $\mathfrak{M}$ aller Kurven, welche folgende Bedingungen erfüllen:

1. Sie gehen durch die beiden gegebenen Punkte P_1 und P_2.
2. Sie sind in der Form

$$y = y(x), \qquad x_1 \gtreqless x \gtreqless x_2$$

darstellbar, wo $y(x)$ eine eindeutige Funktion von x bedeutet, d. h. geometrisch, jede Kurve wird von jeder Geraden parallel der y-Achse: $x = c$ in einem und nur einem Punkt geschnitten, wenn $x_1 \gtreqless c \gtreqless x_2$.

3. Sie sind von der Klasse C', d. h. geometrisch, sie sind stetig und besitzen in jedem Punkt eine Tangente, deren Gefälle[1]) sich stetig ändert, und die nie mit der y-Achse parallel ist.
4. Sie liegen ganz im Bereich $\mathfrak{R}$. Wir nennen diese Kurven die „*zulässigen Kurven*“ oder auch die „Vergleichskurven“.

Jede zulässige Kurve $\mathfrak{C}$ liefert einen bestimmten endlichen Wert $J_{\mathfrak{C}}$ für das Integral J. Die Menge dieser Integralwerte $\{J_{\mathfrak{C}}\}$ besitzt eine untere Grenze K und eine obere Grenze G (endlich oder unendlich). Wenn alsdann die untere (obere) Grenze endlich ist und wirklich erreicht wird, d. h. wenn es eine zulässige Kurve $\mathfrak{C}$ gibt, für welche

$$J_{\mathfrak{C}} = K, \qquad (J_{\mathfrak{C}} = G), \tag{13}$$

so sagen wir, die Kurve $\mathfrak{C}$ liefert ein *absolutes Minimum (Maximum) für das Integral J* in bezug auf die Menge $\mathfrak{M}$.

[1]) Unter *Gefälle* (slope) einer Geraden verstehen wir die trigonometrische Tangente des Winkels, welchen die Gerade mit der positiven x-Achse bildet.

Für jede andere zulässige Kurve $\overline{\mathfrak{C}}$ hat man dann

$$J_{\overline{\mathfrak{C}}} \overline{\gtrless} J_{\mathfrak{C}}, \qquad (J_{\overline{\mathfrak{C}}} \overline{\lessgtr} J_{\mathfrak{C}}); \tag{14}$$

und diese Ungleichung kann ebenfalls zur Definition des absoluten Minimums (Maximums) benutzt werden.

Die *Aufgabe der Variationsrechnung* in ihrer einfachsten Form besteht nun darin, diejenige oder diejenigen zulässigen Kurven zu bestimmen, welche in diesem Sinn ein absolutes Minimum oder Maximum für das Integral J liefern.[1])

Die so formulierte Aufgabe läßt sich auf die mannigfachste Weise modifizieren, indem man den zulässigen Kurven andere Bedingungen auferlegt. Wir erwähnen die wichtigsten dieser Modifikationen:

1. Statt die Endpunkte vorzuschreiben, kann man auch nur verlangen, daß sie auf gegebenen Kurven liegen sollen, oder sonst in vorgeschriebener Weise veränderlich sein sollen (vgl. § 7 und Kap. VI).

2. Man kann die Bedingung fallen lassen, daß y sich als eindeutige Funktion von x darstellen lassen soll, indem man die Kurven in Parameterdarstellung annimmt (vgl. Kap. V).

3. Endlich kann man auch die „Klasse" der zulässigen Kurven modifizieren, indem man z. B. Kurven mit „Ecken" zuläßt (Kap. VIII), oder bloß die Existenz der rechtsseitigen Tangente verlangt (vgl. Kneser, *Lehrbuch*, § 17); ja man kann die Aufgabe sogar so erweitern, daß nicht einmal die Existenz einer einseitigen Tangente vorausgesetzt wird (vgl. Kap. IX). Anderseits kann man auch dem Gefälle gewisse Beschränkungen auferlegen, z. B. beständig positiv zu sein (vgl. Kap. VIII), oder dem absoluten Wert nach eine vorgegebene Grenze nicht zu überschreiten (vgl. § 19, c).

Wie man die zulässigen Kurven am besten zu definieren hat, das hängt in jedem einzelnen Fall von der speziellen Natur des vorgelegten Problems ab. Aber wie man auch die Aufgabe formulieren mag, sie muß stets den beiden folgenden Forderungen genügen, wenn sie überhaupt einen bestimmten Sinn haben soll: *1. Die Gesamtheit der als zulässig betrachteten Kurven muß genau definiert werden; 2. für jede zulässige Kurve muß das Integral J, eventuell nach geeigneter Erweiterung seiner Definition, einen bestimmten endlichen Wert haben.*

Was den Bereich $\mathfrak{R}$ betrifft, so ist derselbe für jede einzelne Aufgabe besonders festzulegen; er kann offen oder geschlossen, endlich oder unendlich sein, er kann auch die ganze x, y-Ebene umfassen.

---

[1]) Hilbert hat in seinen Vorlesungen (1904/1905) ein allgemeines Problem des Maximums und Minimums formuliert, welches sowohl die Aufgaben der Variationsrechnung, als die der Theorie der gewöhnlichen Maxima und Minima umfaßt: *Gegeben ist eine unendliche Menge irgend welcher mathematischer Objekte a, b, . . . (Zahlen, Punkte, Kurven, Flächen usw.), und jedem Individuum dieser Menge ist eine reelle Zahl J_a, J_b . . . zugeordnet. Es soll dasjenige Individuum der Menge bestimmt werden, welchem die größte oder kleinste Zahl zugeordnet ist.*

So müssen z. B. bei der Aufgabe, das Integral

$$J = \int_{\mathfrak{C}} \frac{\sqrt{1 + y'^2}\, dx}{\sqrt{y - y_1 + k}}$$

zu einem Minimum zu machen (Brachistochrone, vgl. § 26, b), die zulässigen Kurven notwendig auf den durch die Ungleichung

$$y - y_1 + k > 0$$

definierten Bereich beschränkt werden, weil sonst die Funktion f entweder unstetig oder imaginär werden würde.

Bei der Aufgabe der Rotationsfläche von kleinster Oberfläche, wo

$$J = \int_{\mathfrak{C}} y \sqrt{1 + y'^2}\, dx\,,$$

muß man die zulässigen Kurven auf den Bereich

$$y \geqq 0$$

beschränken, weil nur dann das bestimmte Integral J die gewünschte geometrische Bedeutung besitzt.

Neben solchen „natürlichen" Beschränkungen, die sich aus dem Problem mit Notwendigkeit ergeben, kann man aber auch den zulässigen Kurven „künstlich" Beschränkungen auf einen gewissen Bereich auferlegen. So sind z. B. bei der Aufgabe, die kürzeste Linie zwischen zwei Punkten zu finden, wo

$$J = \int_{\mathfrak{C}} \sqrt{1 + y'^2}\, dx\,,$$

die zulässigen Kurven keinerlei derartigen Beschränkung unterworfen, so daß der Bereich $\mathfrak{R}$ die ganze x, y-Ebene umfaßt. Man kann aber auch die Aufgabe dahin modifizieren, unter allen Kurven, welche in einem gewissen gegebenen Bereich gelegen sind, die kürzeste Verbindungslinie zweier gegebener Punkte zu finden (vgl. Kap. VIII); hier ist dann $\mathfrak{R}$ eben dieser vorgegebene Bereich.

b) Relatives Extremum eines bestimmten Integrals:

Ganz analog wie in der Theorie der gewöhnlichen Extrema läßt sich nun auch die im vorhergehenden formulierte Aufgabe des absoluten Extremums eines bestimmten Integrals, (welche das eigentliche Endziel der Variationsrechnung ist), auf diejenige des relativen Extremums zurückführen, bei welchem die gesuchte Kurve nur mit sogenannten „benachbarten" Kurven verglichen wird, und welches folgendermaßen definiert wird: *Eine zulässige Kurve* $\mathfrak{C}$: $y = y(x)$ *liefert ein* *relatives Minimum (Maximum)* *für das Integral* J, *wenn eine positive Größe* ϱ *existiert, derart, daß*

$$J_{\overline{\mathfrak{C}}} \geqq J_{\mathfrak{C}}, \qquad (J_{\overline{\mathfrak{C}}} \leqq J_{\mathfrak{C}}) \tag{14}$$

für jede zulässige Kurve $\overline{\mathfrak{C}}$: $y = \bar{y}(x)$, *für welche*

$$|\,\bar{y}(x) - y(x)\,| < \varrho \quad \textit{für} \quad x_1 \leqq x \leqq x_2. \tag{15}$$

Diese Ungleichung bedeutet geometrisch, daß die Kurve $\overline{\mathfrak{C}}$ (abgesehen von den Endpunkten) im Innern des Streifens liegt, welcher durch die beiden Kurven

$$y = y(x) + \varrho, \qquad y = y(x) - \varrho$$

einerseits und durch die beiden Geraden

$$x = x_1, \qquad x = x_2$$

anderseits begrenzt wird. Diesen Streifen werden wir „*die Nachbarschaft*[1]) (ϱ) *der Kurve* $\mathfrak{C}$“ nennen, wobei von der Begrenzung nur die Punkte P_1 und P_2 als mit zur Nachbarschaft (ϱ) gehörig betrachtet werden sollen.

Wir nennen dann wieder das relative Minimum (Maximum) ein *eigentliches*, wenn ϱ so gewählt werden kann, daß in der Ungleichung (14) das Zeichen $>$ ($<$) für alle von $\mathfrak{C}$ verschiedenen zulässigen Kurven gilt, welche in der Nachbarschaft (ϱ) liegen; dagegen *uneigentlich*, wenn es, wie klein auch ϱ gewählt sein mag, stets mindestens eine von $\mathfrak{C}$ verschiedene zulässige Kurve $\overline{\mathfrak{C}}$ gibt, welche ganz in der Nachbarschaft (ϱ) liegt und für welche: $J_{\overline{\mathfrak{C}}} = J_{\mathfrak{C}}$.

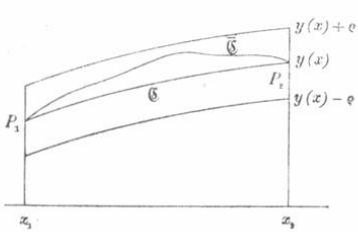

Fig. 1.

Eine Kurve, welche ein absolutes Extremum liefert, liefert a fortiori auch ein relatives, und daher reduziert[2]) sich die ursprüngliche Aufgabe darauf, *alle Kurven zu finden, welche ein relatives Extremum für das Integral J liefern*, und in dieser Form werden wir die Aufgabe in der Folge betrachten.

Wir werden dabei die Worte „Minimum, Maximum“ stets im Sinn von „relatives Minimum, Maximum“ gebrauchen und wir werden

---

[1]) Vgl. Osgood, loc. cit. p. 107. Die Nachbarschaft (ϱ) inklusive ihrer Begrenzung werden wir die *geschlossene Nachbarschaft* [ϱ] von $\mathfrak{C}$ nennen.

[2]) Vgl. die entsprechenden Bemerkungen beim gewöhnlichen Extremum, § 2, c). Für eine *direkte* Behandlung des absoluten Extremums vergleiche man Hilbert's Existenzbeweis (Kap. IX), Darboux, *Théorie des surfaces*, Bd. III, p. 89; und Zermelo, Jahresbericht der Deutschen Mathematiker-Vereinigung, Bd. XI (1902) p. 184.

uns auf den Fall des Minimums beschränken, da jede Kurve, welche ein Minimum für das Integral J liefert, zugleich ein Maximum für das Integral $-J$ liefert, und vice versa.

§ 4. Verschwinden der ersten Variation.

Wir setzen jetzt voraus, wir hätten eine Kurve

$$\mathfrak{C}: \qquad y = y(x), \qquad x_1 \overline{\gtrless} x \overline{\gtrless} x_2,$$

gefunden, welche das Integral

$$J = \int_{\mathfrak{C}} f(x, y, y')\, dx$$

in dem im vorigen Paragraphen erklärten Sinn zu einem Minimum macht. Wir nehmen überdies an, daß die Kurve $\mathfrak{C}$ ganz im Innern[1]) des Bereiches $\mathfrak{R}$ liegt.

Aus der letzten Annahme folgt[2]), daß wir ϱ so klein nehmen können, daß die Nachbarschaft (ϱ) von $\mathfrak{C}$ ebenfalls ganz im Innern von $\mathfrak{R}$ liegt.

Wir *„variieren"* nun die Kurve $\mathfrak{C}$, d. h. wir ersetzen sie durch eine andere zulässige Kurve

$$\overline{\mathfrak{C}}: \qquad y = \overline{y}(x), \qquad x_1 \overline{\gtrless} x \overline{\gtrless} x_2,$$

welche ganz in der Nachbarschaft (ϱ) liegt. Eine solche Kurve pflegt man eine der Kurve $\mathfrak{C}$ *„benachbarte Kurve"* zu nennen.

Das Inkrement

$$\Delta y = \overline{y}(x) - y(x),$$

welches wir mit ω (oder wenn nötig $\omega(x)$) bezeichnen[3]), wird die *„vollständige Variation von y"* genannt. Da die Kurve $\overline{\mathfrak{C}}$ in der Nachbarschaft (ϱ) liegt, so ist

$$|\, \omega(x) \,| < \varrho \text{ in } [x_1 x_2]; \tag{16}$$

überdies ist im Fall fester Endpunkte

$$\omega(x_1) = 0, \qquad \omega(x_2) = 0, \tag{17}$$

da

$$\overline{y}(x_1) = y(x_1) = y_1; \qquad \overline{y}(x_2) = y(x_2) = y_2.$$

---

[1]) Der Fall, wo die Kurve Punkte mit der Begrenzung von $\mathfrak{R}$ gemein hat, wird in Kap. VIII behandelt werden.

[2]) Vgl. § 21, a); die Kurve $\mathfrak{C}$ ist eine begrenzte, abgeschlossene Menge, nach A VII 1.

[3]) Bezeichnung nach LAGRANGE, *Œuvres*, Bd. IX, p. 296.

Das entsprechende Inkrement des Integrals,

$$\Delta J = J_{\overline{\mathfrak{C}}} - J_{\mathfrak{C}},$$

heißt die „*vollständige*[1]) *Variation des Integrals J*“. Da wir die Endpunkte als fest voraussetzen, so sind die Integrationsgrenzen in beiden Integralen dieselben und wir können daher schreiben:

$$\Delta J = \int_{x_1}^{x_2} [f(x,\, y + \omega,\, y' + \omega') - f(x,\, y,\, y')]\, dx.$$

Da wir annehmen, daß die Kurve $\mathfrak{C}$ ein Minimum liefert, so muß nach (14)

$$\Delta J \geqq 0 \tag{18}$$

sein für alle „benachbarten“ Kurven, vorausgesetzt, daß ϱ hinreichend klein gewählt worden ist.

Für die weitere Diskussion dieser Ungleichung betrachten wir mit LAGRANGE[2]) spezielle[3]) Variationen von der Form

$$\omega(x) = \varepsilon \eta(x), \tag{19}$$

wo $\eta(x)$ eine beliebige Funktion der Klasse C' ist, welche in x_1 und x_2 verschwindet, und ε eine Konstante, deren absoluter Wert so klein gewählt ist, daß die Bedingung (16) erfüllt ist.

Alsdann geht das Integral $\overline{J} = J_{\overline{\mathfrak{C}}}$ in eine Funktion von ε über, die wir mit $J(\varepsilon)$ bezeichnen wollen. Insbesondere ist dann $J(0) = J_{\mathfrak{C}}$, so daß wir die Ungleichung (14) schreiben können

$$J(\varepsilon) \geqq J(0)$$

für alle hinreichend kleinen Werte von $|\varepsilon|$. Das heißt aber: Die Funktion $J(\varepsilon)$ muß für $\varepsilon = 0$ ein Minimum besitzen, und daher muß

$$J'(0) = 0, \qquad J''(0) \geqq 0 \tag{20}$$

sein.

---

[1]) Nach WEIERSTRASS, *Vorlesungen*.

[2]) Vgl. LAGRANGE, *Œuvres*, Bd. IX, p. 298; die hier gegebene Methode benutzen LINDELÖF-MOIGNO (loc. cit.), DIENGER, *Grundriß der Variationsrechnung* (Braunschweig, 1867), und OSGOOD (loc. cit.).

[3]) Variationen dieser speziellen Art werden auch schon von EULER erwähnt, *Instit. Calc. Integr.*, Bd. IV, Supplem. XI, § 5. Solange es sich nur um die Herleitung notwendiger Bedingungen handelt, dürfen wir die Variationen nach Belieben spezialisieren; ganz anders verhält es sich bei der Herleitung hinreichender Bedingungen, vgl. § 15.

Aus der Definition der Ableitung folgt, daß

$$\Delta J = J(\varepsilon) - J(0) = \varepsilon J'(0) + \varepsilon(\varepsilon), \tag{21}$$

wenn (ε) wieder eine unendlich kleine Funktion von ε bezeichnet, so daß also $\varepsilon J'(0)$ das Differential der Funktion $J(\varepsilon)$ nach ε bedeutet. Man pflegt dasselbe nach LAGRANGE[1]) mit δJ zu bezeichnen:

$$\delta J = \varepsilon J'(0),$$

und *die erste Variation des Integrals J* zu nennen. Analog definiert man die höheren Variationen

$$\delta^2 J = \varepsilon^2 J''(0), \ldots, \quad \delta^n J = \varepsilon^n J^{(n)}(0).$$

Unter Benutzung dieser Bezeichnungsweise können wir die Bedingungen (20) also auch so aussprechen:

Für ein Minimum des Integrals J ist notwendig, daß die erste Variation verschwindet[2]) *und die zweite Variation nicht negativ ist*[3]), *und zwar für alle zulässigen Funktionen η:*

$$\delta J = 0, \qquad \delta^2 J \gtreqless 0. \tag{22}$$

Wir haben es hier zunächst nur mit der ersten Variation zu tun. Nach der Regel[4]) für die Differentiation eines bestimmten Integrals nach einem Parameter findet man:

$$J'(0) = \int_{x_1}^{x_2} (f_y \eta + f_{y'} \eta')\, dx, \tag{23}$$

wenn wir, wie stets in der Folge, partielle Ableitungen einer Funktion mehrerer Variabeln durch Suffixe bezeichnen:

---

[1]) Das Symbol δ findet sich zum erstenmal in einem Brief von LAGRANGE an EULER vom 12. August 1755, siehe LAGRANGE, *Œuvres,* Bd. XIV, p. 140.

[2]) EULER, *Methodus inveniendi etc.*(1744), Kap. I, § 63.

[3]) Höhere Variationen finden sich zuerst bei LEGENDRE (1786), der jedoch den Faktor $\frac{1}{n!}$ mit zu $\delta^n J$ hinzurechnet.

[4]) Vgl. A V 7. Die Regel ist hier anwendbar; denn aus unseren Annahmen über die Funktionen f, y, η folgt nach A IV 9, daß die partielle Ableitung

$$\frac{\partial}{\partial \varepsilon} f(x, y(x) + \varepsilon \eta(x), y'(x) + \varepsilon \eta'(x))$$

eine stetige Funktion von x und ε ist in dem Bereich

$$x_1 \lesseqgtr x \lesseqgtr x_2, \qquad |\varepsilon| \lesseqgtr \varepsilon_0,$$

wofern die positive Größe ε_0 hinreichend klein gewählt wird.

$$f_y = \frac{\partial f}{\partial y}, \qquad f_{y'} = \frac{\partial f}{\partial y'}, \text{ usw.},$$

und ebenso später:

$$f_{yy} = \frac{\partial^2 f}{\partial y^2}, \qquad f_{yy'} = \frac{\partial^2 f}{\partial y\, \partial y'}, \text{ usw.}$$

Wir erhalten also den Satz:

Für ein Extremum des Integrals J ist es notwendig, daß

$$\int_{x_1}^{x_2} (f_y \eta + f_{y'} \eta')\, dx = 0 \tag{24}$$

für alle Funktionen η der Klasse C', welche in x_1 und x_2 verschwinden.

Dabei sind die Argumente von f_y, $f_{y'}$: x, $y(x)$, $y'(x)$.

Man kann dieses Resultat auch dadurch ableiten, daß man auf die Differenz

$$f(x,\, y + \omega,\, y' + \omega') - f(x,\, y,\, y')$$

den *Taylor'schen Satz mit Restglied*[1]) für $n = 2$ anwendet und dann zwischen den Grenzen x_1 und x_2 integriert. Man erhält so

$$\Delta J = \int_{x_1}^{x_2} (f_y \omega + f_{y'} \omega')\, dx + \tfrac{1}{2} \int_{x_1}^{x_2} (\tilde{f}_{yy} \omega^2 + 2 \tilde{f}_{yy'} \omega \omega' + \tilde{f}_{y'y'} \omega'^2)\, dx,$$

wobei die Argumente von $\tilde{f}_{yy}$, $\tilde{f}_{yy'}$, $\tilde{f}_{y'y'}$ sind: x, $y(x) + \theta \omega(x)$, $y'(x) + \theta \omega'(x)$, unter θ eine Funktion von x verstanden, deren Wert beständig zwischen 0 und 1 liegt.

Wählt man jetzt wieder für ω eine Funktion der speziellen Form (19), so nimmt der Ausdruck für ΔJ die Form (21) an, woraus man, wie in § 2, c), schließt, daß (24) erfüllt sein muß.

Wenn $f(x, y, y')$ eine *analytische* Funktion ist, welche in dem oben mit $\mathfrak{T}$ bezeichneten Bereich regulär ist, so kann man statt der Taylor'schen Formel mit Restglied auch die *Taylor'sche unendliche Reihe*[2]) in Anwendung bringen. Man erhält dann unter der Voraussetzung, daß gliedweise Integration erlaubt ist, für ΔJ eine für hinreichend kleine Werte von $|\omega|$ und $|\omega'|$ konvergente Reihe:

---

[1]) Diese Methode wurde zuerst von Lagrange gebraucht, siehe *Œuvres,* Bd. IX, p. 297. Vgl. auch Du Bois-Reymond, Mathematische Annalen, Bd. XV (1879), p. 292, und Pascal, loc. cit., p. 22.

[2]) Diese Methode benutzen Weierstrass, Kneser (*Lehrbuch,* §§ 2, 8) und Jordan, *Cours d'Analyse,* III, Nr. 350), ohne jedoch einen strengen Detailbeweis zu geben, der hier wesentlich umständlicher ausfallen würde als bei den beiden anderen Methoden.

$$\Delta J = \int_{x_1}^{x_2} (f_y \omega + f_{y'} \omega') dx + \tfrac{1}{2} \int_{x_1}^{x_2} (f_{yy} \omega^2 + 2 f_{yy'} \omega \omega' + f_{y'y'} \omega'^2) dx + \cdots +$$
$$+ \int_{x_1}^{x_2} (\omega, \omega')_n dx + \cdots,$$

wobei in Weierstraß'scher Bezeichnungsweise $(\omega, \omega')_n$ eine homogene Funktion n^{ter} Dimension von ω, ω' bedeutet.

Wählt man dann wieder für ω eine Funktion von der speziellen Form (19), so geht diese Reihe in eine nach Potenzen von ε fortschreitende Reihe über, auf die man dann das auf p. 11, Fußnote [2]) erwähnte Lemma anzuwenden hat.

Von den drei angeführten Methoden liefert unzweifelhaft die erste, den einfachsten strengen Beweis für die beiden Bedingungen (20). Man hat sogar lange geglaubt, daß diese Methode das ganze Problem der Variationsrechnung auf ein Problem der Theorie der gewöhnlichen Maxima und Minima zurückführt. Dem ist aber nicht so; denn wie wir später sehen werden, liefert die Methode nur notwendige Bedingungen, genügt aber nicht einmal für ein sogenanntes schwaches[1]) Extremum zur Herleitung hinreichender Bedingungen, während die auf die Taylor'sche Formel basierte Methode, obgleich weniger elegant, wenigstens für ein schwaches Extremum hinreichende Bedingungen liefert.

§ 5. Die Euler'sche Differentialgleichung.

Wir gehen jetzt dazu über, aus der Bedingung $\delta J = 0$ weitere Folgerungen zu ziehen.

a) Die Lagrange'sche partielle Integration:

Zu diesem Zweck pflegt man nach dem Vorgang von Lagrange[2]) das zweite Glied in dem Ausdruck für δJ durch partielle Integration umzuformen und erhält so:

$$\delta J = \varepsilon \left\{ \Big[\eta f_{y'} \Big]_{x_1}^{x_2} + \int_{x_1}^{x_2} \eta \left(f_y - \frac{d}{dx} f_{y'} \right) dx \right\}, \tag{25}$$

wobei von der Bezeichnung

$$\Big[\Phi(x) \Big]_{x_1}^{x_2} = \Phi(x_2) - \Phi(x_1) \tag{26}$$

Gebrauch gemacht ist.

Da η für $x = x_1$ und $x = x_2$ verschwindet, so ist das vom Integralzeichen freie Glied gleich Null, und die Bedingung $\delta J = 0$ reduziert sich auf

---

[1]) Vgl. § 15, b).

[2]) Zuerst in dem bereits oben erwähnten Brief an Euler vom 12. August 1755. (*Œuvres de Lagrange*, Bd. XIV, p. 141).

$$\int_{x_1}^{x_2} \eta\left(f_y - \frac{d}{dx} f_{y'}\right) dx = 0. \tag{27}$$

Diese Gleichung muß erfüllt sein für alle Funktionen η, welche in x_1 und x_2 verschwinden und in $[x_1 x_2]$ von der Klasse C' sind.

Aus der Willkürlichkeit von η schließt[1]) man dann, daß dies nur möglich ist, wenn der Faktor von η unter dem Integralzeichen für sich verschwindet, und erhält so den

Fundamentalsatz I: Soll die Funktion

$$y = y(x)$$

das Integral

$$J = \int_{x_1}^{x_2} f(x,\, y,\, y')\, dx$$

zu einem Maximum oder Minimum machen, so muß sie der Differentialgleichung genügen:[2])

$$f_y - \frac{d}{dx} f_{y'} = 0. \tag{I}$$

Wir werden diese Differentialgleichung nach ihrem Entdecker[3]) die *Euler'sche Differentialgleichung* nennen.

Man beachte, daß die Argumente der Funktionen f_y, $f_{y'}$ sind: x, $y(x)$, $y'(x)$, und daß die Differentiation totale Differentiation nach x bedeutet, so daß die Differentialgleichung in entwickelter Form lautet:

$$f_y - f_{y'x} - y' f_{y'y} - y'' f_{y'y'} = 0. \tag{28}$$

Die obige Ableitung der Euler'schen Differentialgleichung weist jedoch zwei erhebliche Lücken auf:

Erstens[4]) setzt die partielle Integration zum mindesten die Existenz der Ableitung $\frac{d}{dx} f_{y'}$ voraus, und da wir von der Funktion $y(x)$

---

[1]) Vgl. unter b).

[2]) Wegen der Ausdehnung dieses Satzes auf allgemeinere Variationsprobleme vgl. die *Übungsaufgaben* Nr. 41—47 am Ende von Kap. III, sowie Kap. XI und XII.

[3]) Euler, *Methodus inveniendi etc.* (1744), Kap. II, Art. 21; in Stäckel's Übersetzung in Ostwald's Klass. Nr. 46, p. 54. Die Differentialgleichung ist neuerdings vielfach die „*Lagrange'sche Differentialgleichung*" genannt worden. Lagrange selbst schreibt sie Euler zu, vgl. *Œuvres*, Bd. X, p. 397: „Cette équation est celle qu'Euler a trouvée le premier."

[4]) Dieser Einwand ist zuerst von Du Bois-Reymond in der wichtigen Abhandlung: „*Erläuterungen zu den Anfangsgründen der Variationsrechnung*" Mathematische Annalen, Bd. XV (1879), p. 283 erhoben worden.

nichts weiter als die Existenz und Stetigkeit der **ersten** Ableitung vorausgesetzt haben, so wissen wir von der Funktion $f_{y'}(x, y(x), y'(x))$ nur, daß sie stetig ist, nicht aber, ob sie eine Ableitung besitzt. Um diesem Einwand zu begegnen, machen wir vorläufig die weitere *Annahme*[1]), *daß auch die zweite Ableitung von $y(x)$ existiert und stetig ist* im Intervall $[x_1 x_2]$. Alsdann besitzt die Funktion $f_{y'}(x, y(x), y'(x))$ eine Ableitung[2]), die überdies stetig ist in $[x_1 x_2]$, und es steht nunmehr der Anwendung der partiellen Integration[3]) nichts mehr im Wege.

b) **Das Fundamentallemma der Variationsrechnung:**

Der zweite Einwand bezieht sich auf den Schluß, daß wegen der Willkürlichkeit von η aus der Gleichung (27) die Differentialgleichung (I) folgt; derselbe ist durchaus nicht selbstverständlich,[4]) wie noch bis in die Mitte des 19. Jahrhunderts allgemein angenommen wurde, sondern bedarf eines Beweises. Letzterer beruht auf dem folgenden

Lemma: Ist M eine Funktion von x, welche stetig ist in $[x_1 x_2]$ und ist

$$\int_{x_1}^{x_2} \eta \, M \, dx = 0 \tag{29}$$

für alle Funktionen η, welche in x_1 und x_2 verschwinden und eine stetige Ableitung in $[x_1 x_2]$ besitzen, so ist

$$M \equiv 0 \tag{30}$$

in $[x_1 x_2]$.

Denn[5]) angenommen es sei $M(x') \neq 0$, z. B. > 0, in einem Punkt x' des Intervalles $[x_1 x_2]$; dann können wir

---

[1]) Von dieser Annahme werden wir uns weiter unten, siehe c), befreien.

[2]) Nach A IV 9. [3]) Nach A V 5.

[4]) Der Schluß tritt zuerst bei LAGRANGE in dem oben (p. 21, Fußnote [1])), zitierten Briefe an EULER auf und wird dort als ganz selbstverständlich gegeben.

[5]) Der hier gegebene Beweis rührt von DU BOIS-REYMOND her (Mathematische Annalen, Bd. XV (1879), pp. 297, 300). In derselben Arbeit beweist DU BOIS-REYMOND, daß der Schluß $M \equiv 0$ gültig bleibt, selbst wenn man nur weiß, daß die Gleichung (29) besteht:

1. für alle in x_1 und x_2 verschwindenden Funktionen, welche in $[x_1 x_2]$ von der *Klasse* $C^{(n)}$ sind; man verfahre wie oben, wähle jedoch für das Intervall $[\xi_1 \xi_2]$

$$\eta = (x - \xi_1)^{n+1} (\xi_2 - x)^{n+1};$$

2. für alle in x_1 und x_2 verschwindenden Funktionen, welche in $[x_1 x_2]$ Ableitungen **aller** Ordnungen besitzen.

wegen[1]) der Stetigkeit von M ein den Punkt x' enthaltendes Teilintervall $[\xi_1 \xi_2]$ von $[x_1 x_2]$ angeben, derart daß $M > 0$ im ganzen Intervall $[\xi_1 \xi_2]$. Jetzt wähle man $\eta \equiv 0$ außerhalb $[\xi_1 \xi_2]$ und $\eta = (x - \xi_1)^2 (x - \xi_2)^2$ in $[\xi_1 \xi_2]$; die so definierte Funktion η besitzt eine stetige[2]) Ableitung in $[x_1 x_2]$, und verschwindet in x_1 und x_2. Trotzdem ist für sie

$$\int_{x_1}^{x_2} \eta\, M\, dx > 0,$$

entgegen der Annahme. Es ist also unmöglich, daß $M(x') \neq 0$ und somit ist der Satz bewiesen. Dieses Lemma wird häufig das *Fundamentallemma der Variationsrechnung* genannt.

Die von Du Bois-Reymond gebrauchten speziellen Funktionen η sind alle aus mehreren analytischen Funktionen zusammengesetzt. Es lassen sich aber auch Funktionen konstruieren, welche im ganzen Intervall durch eine einzige reguläre analytische Funktion dargestellt werden und denselben Zweck erfüllen. Eine solche Funktion bildet Zermelo einer Anregung von Weierstrass folgend, in seiner *Dissertation*, p. 35: Der Punkt x' sei ein innerer Punkt; dann setze man

$$\eta = (x - x_1)(x_2 - x)\, e^{-\varrho^2 (x - x')^2}$$

im ganzen Intervall $[x_1 x_2]$, wo ϱ eine hinreichend große Konstante ist.

Eine andere Funktion, welche denselben Zweck erfüllt und die durch ihre geometrische Interpretation interessant ist, hat H. A. Schwarz in seinen Vorlesungen gegeben, vgl. Hancock, *Lectures on the Calculus of Variations* (Cincinnati, 1904), Nr. 78.

Es folgt hieraus, daß es für den Schluß $M \equiv 0$ genügt zu wissen, daß die Gleichung (29) für alle Funktionen erfüllt ist, welche in $[x_1 x_2]$ *regulär* sind und in x_1 und x_2 verschwinden.

Der älteste Beweis des Lemmas rührt von Stegemann her (*Lehrbuch der Variationsrechnung* (1854), § 24).

Er setzt

$$\eta = (x - x_1)(x_2 - x) / M;$$

einfacher jedoch ist es,

$$\eta = (x - x_1)(x_2 - x)\, M$$

zu wählen. Hier müssen jedoch stärkere Annahmen gemacht werden, nämlich entweder, daß die Gleichung (29) für alle *stetigen* in x_1 und x_2 verschwindenden Funktionen η gilt, oder aber, daß *M von der Klasse C'* ist, was für unsere Anwendung bedeuten würde, daß y''' existiert und stetig ist in $[x_1 x_2]$.

Auch der Beweis von Heine (Mathematische Annalen, Bd. II (1870), p. 189) läßt sich nicht ohne weitere beschränkende Annahmen über y auf unsern Fall anwenden.

[1]) Nach A III 2.

[2]) Auch in ξ_1 und ξ_2; denn in jedem dieser Punkte ist sowohl die vordere als auch die hintere Derivierte gleich Null; es existiert also eine Ableitung in ξ_1 und in ξ_2 und ihr Wert in diesen Punkten ist Null, und dies ist zugleich der Grenzwert von η' bei Annäherung an ξ_1 resp. ξ_2. Vgl. A IV 1.

Die Voraussetzungen dieses Lemmas sind für die Gleichung (27) erfüllt, wenn wir annehmen, daß y'' existiert und stetig ist in $[x_1 x_2]$; denn alsdann ist die Funktion

$$M = f_y - \frac{d}{dx} f_{y'}$$

stetig in $[x_1 x_2]$. Somit haben wir in der Tat die Notwendigkeit der Differentialgleichung (I) bewiesen für alle Funktionen y von der Klasse C''.

c) **Du Bois-Reymond's Lemma:**

Die unter a) gegebene Methode der partiellen Integration liefert, wie wir gesehen haben, nur diejenigen Lösungen unserer Aufgabe, welche eine stetige zweite Ableitung besitzen. Es fragt sich nun, ob es außerdem noch andere Lösungen gibt und wie dieselben zu finden sind.

Zur Beantwortung dieser Frage kehren wir zur Gleichung $\delta J = 0$ in der ursprünglichen Form (24) zurück und wenden, nach dem Vorgang von Du Bois-Reymond, die *partielle Integration* nicht auf das zweite, sondern *auf das erste Glied* an. Setzt man in der allgemeinen Formel der partiellen Integration[1]):

$$\int_a^b u \frac{dv}{dx}\, dx = \Big[u v \Big]_a^b - \int_a^b v \frac{du}{dx}\, dx,$$

welche sicher gültig ist, wenn u und v in $[ab]$ von der Klasse C' sind,

$$u = \eta, \qquad v = \int_{x_1}^{x} f_y\, dx,$$

und beachtet, daß η an den beiden Endpunkten verschwindet, so geht die Gleichung (24) über in

$$\int_{x_1}^{x_2} \eta' \left(f_{y'} - \int_{x_1}^{x} f_y\, dx \right) dx = 0. \tag{31}$$

Diese partielle Integration ist erlaubt, selbst wenn y'' nicht existieren sollte, da

$$\eta' = du/dx \qquad \text{und} \qquad f_y = dv/dx$$

stetig sind.

---

[1]) Vgl. A V 5.

Die weiteren Folgerungen aus der Gleichung (31) stützen sich auf das folgende von Du Bois-Reymond[1]) herrührende

Lemma: Ist N eine Funktion von x, welche stetig ist in $[x_1 x_2]$ und ist

$$\int_{x_1}^{x_2} \eta' N \, dx = 0 \tag{32}$$

für alle Funktionen η, welche in x_1 und x_2 verschwinden und eine stetige Ableitung in $[x_1 x_2]$ besitzen, so ist

$$N = konst. \tag{33}$$

in $[x_1 x_2]$.

Den folgenden einfachen Beweis hat Hilbert[2]) in seinen Vorlesungen (Sommer 1899) gegeben:

Man wähle willkürlich vier den Ungleichungen

$$x_1 < \alpha < \beta < \alpha' < \beta' < x_2$$

genügende Größen α, β, α', β' und konstruiere eine Funktion der Klasse C', welche folgende Bedingungen erfüllt:

$\eta \equiv 0$ in $[x_1 \alpha]$;

η wächst beständig von 0 bis zu einem positiven Wert k, während x von α bis β wächst;

η bleibt konstant $= k$ in $[\beta \alpha']$;

η nimmt beständig ab von k bis 0, während x von α' bis β' wächst;

$\eta \equiv 0$ in $[\beta' x_2]$.

Die Existenz einer solchen Funktion — und das ist alles, was zum Beweis erforderlich ist — ist a priori klar[3]), (vgl. Fig. 2).

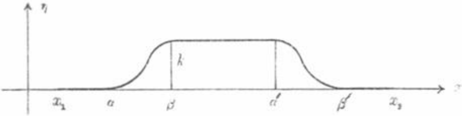

Fig. 2.

Setzt man eine solche Funktion η in (32) ein, so erhält man

$$\int_{\alpha}^{\beta} \eta' N \, dx + \int_{\alpha'}^{\beta'} \eta' N \, dx = 0 .$$

---

[1]) Mathematische Annalen, Bd. XV (1879), p. 313. Du Bois-Reymond's Beweis findet sich bei Bolza, *Lectures on the Calculus of Variations* (Chicago 1904) § 6, reproduziert. Eine interessante Verallgemeinerung dieses Satzes ist kürzlich von Zermelo gegeben worden, Mathematische Annalen, Bd. 58 (1904), p. 558. Vgl. *Übungsaufgabe* Nr. 47 am Ende von Kap. III.

[2]) Siehe Whittemore, Annals of Mathematics (2), Bd. II (1901), p. 132.

[3]) Hilbert gibt ein einfaches Beispiel einer solchen Funktion, siehe Whittemore's Darstellung. Er bildet zunächst eine den Anforderungen genügende

Da $\eta' \gtreqless 0$ in $[\alpha\beta]$ und $\eta' \lesseqgtr 0$ in $[\alpha'\beta']$, so können wir den ersten Mittelwertsatz[1]) anwenden und erhalten, da

$$\eta(\beta) - \eta(\alpha) = k, \qquad \eta(\beta') - \eta(\alpha') = -k:$$

$$k[N(\alpha + \theta(\beta - \alpha)) - N(\alpha' + \theta'(\beta' - \alpha'))] = 0,$$

wo
$$0 < \theta < 1, \qquad 0 < \theta' < 1.$$

Lassen wir jetzt β und β' sich resp. den Grenzen α und α' nähern, so folgt, da $N(x)$ stetig ist, daß

$$N(\alpha) = N(\alpha'),$$

und da α und α' irgend zwei Werte zwischen x_1 und x_2 waren, so bedeutet dies, daß $N(x)$ zunächst im Innern von $[x_1 x_2]$ konstant ist, was sich dann wegen der Stetigkeit von $N(x)$ sofort auf die Endpunkte x_1 und x_2 ausdehnt.[2])

Funktion η'; die sie darstellende Kurve fällt von x_1 bis α, von β bis α' und von β' bis x_2 mit der x-Achse zusammen; zwischen α und β liegt sie oberhalb, zwischen α' und β' unterhalb der x-Achse, und man hat nun nur dafür zu sorgen, daß die Flächen $\alpha\gamma\beta\alpha$ und $\alpha'\gamma'\beta'\alpha'$ dem absoluten Wert nach gleich sind; denn dann ist

$$\eta(\beta) - \eta(\alpha) = -[\eta(\beta') - \eta(\alpha')],$$

und die Funktion

$$\eta = \int_{x_1}^{x} \eta'\, dx$$

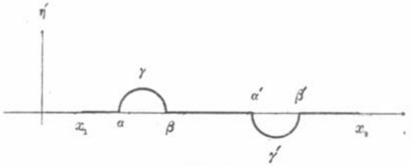

Fig. 3.

hat die verlangten Eigenschaften. Am einfachsten ist es mit HILBERT: $\beta' - \alpha' = \beta - \alpha$ zu nehmen; man kann dann z. B. für die η'-Kurve zwischen α und β einen Halbkreis über dem Segment $\alpha\beta$ und zwischen α' und β' einen solchen unter dem Segment $\alpha'\beta'$ wählen.

[1]) Vgl. A V 6.

[2]) HILBERT's Beweis läßt sich leicht auf den Fall ausdehnen, wo die Funktion N in $[x_1 x_2]$ endlich, aber nur „im allgemeinen stetig" ist, d. h. eine endliche Anzahl von Unstetigkeiten besitzt. Auch in diesem Fall ist N integrabel in $[x_1 x_2]$ (A V 2). Sind nun α und α' Stetigkeitspunkte von N, so können wir stets β und β' so nahe an α, beziehungsweise α' wählen, daß N in $[\alpha\beta]$ und $[\alpha'\beta']$ stetig ist.

Es folgt dann wie oben, daß $N(\alpha) = N(\alpha')$, d. h. *auch unter den gegenwärtigen Voraussetzungen hat N den nämlichen konstanten Wert in allen Stetigkeitspunkten.* Hieraus folgt dann noch weiter, daß in einem Unstetigkeitspunkt c die Grenzwerte $N(c-0)$ und $N(c+0)$ existieren und gleich sind, nämlich gleich dem konstanten Wert in den Stetigkeitspunkten, vgl. WHITTEMORE, loc. cit.

d) **Anwendung des Du Bois-Reymond'schen Lemmas:**

Wir wenden jetzt das Du Bois-Reymond'sche Lemma auf die Gleichung (31) an, was gestattet ist, da die Funktion

$$N \equiv f_{y'} - \int_{x_1}^{x} f_y \, dx$$

nach unsern Annahmen über die Kurve $\mathfrak{C}$ und die Funktion f (siehe § 3, a) und § 4, Anfang) stetig[1]) ist in $[x_1 x_2]$.

Wir schließen daher, daß

$$f_{y'} - \int_{x_1}^{x} f_y \, dx = \lambda$$

sein muß, wo λ eine Konstante bedeutet; hieraus folgt

$$f_{y'} = \lambda + \int_{x_1}^{x} f_y \, dx. \tag{34}$$

Die rechte Seite dieser Gleichung ist differentiierbar[2]) und ihre Ableitung ist f_y; also muß auch die linke Seite, d. h. die Funktion

$$f_{y'}(x, y(x), y'(x)) \equiv f_{y'}[x]$$

differentiierbar sein, und überdies ist

$$\frac{d}{dx} f_{y'} = f_y. \tag{35}$$

Hiermit erst ist der erste Fundamentalsatz vollständig bewiesen, nämlich, daß *jede Funktion y der Klasse C', welche das Integral J zu einem Extremum macht*, und welche durch eine ganz im Innern von $\mathfrak{R}$ gelegene Kurve dargestellt wird, der *Euler'schen Differentialgleichung genügen muß* — einerlei ob sie eine zweite Ableitung besitzt oder nicht.[3]) —

Hilbert[4]) hat hieran die wichtige Bemerkung geknüpft, daß aus der Differentiierbarkeit von $f_{y'}$ *die Existenz der zweiten Ableitung y''*

---

[1]) Nach A III 4 und A V 4. [2]) Nach A V 4.

[3]) Hahn ist neuerdings in dieser Richtung noch einen Schritt weiter gegangen und hat bewiesen, daß jede rektifizierbare Kurve, welche in jedem Punkt eine bestimmte Tangente besitzt, der Euler'schen Differentialgleichung genügen muß, wenn sie ein Minimum für das Integral J liefert (Mathematische Annalen, Bd. 63 (1906), p. 254). Dabei muß allerdings zunächst die Definition des bestimmten Integrals erweitert werden, vgl. ~~Kap. IX.~~ Über sogenannte diskontinuierliche Lösungen, siehe Kap. VIII.

[4]) Vgl. Whittemore, loc. cit.

folgt für alle Werte von x, für welche

$$f_{y'y'}(x, y(x), y'(x)) \neq 0. \tag{36}$$

Denn setzen wir

$$y(x+h) - y(x) = k, \qquad y'(x+h) - y'(x) = l,$$

so ist in der obigen Bezeichnung

$$\frac{f_{y'}[x+h] - f_{y'}[x]}{h} = \frac{f_{y'}(x+h, y+k, y'+l) - f_{y'}(x, y, y')}{h}.$$

Da $f_{y'}$ als Funktion von x, y, y' von der Klasse C' ist in der Umgebung der Stelle x, $y(x)$, $y'(x)$, so kann man auf den Zähler den Satz vom vollständigen Differential[1]) anwenden, und erhält, da überdies y und y' stetig sind, also k und l mit h unendlich klein werden:

$$\frac{f_{y'}[x+h] - f_{y'}[x]}{h} = (f_{y'x} + \alpha) + \frac{k}{h}(f_{y'y} + \beta) + \frac{l}{h}(f_{y'y'} + \gamma),$$

wo α, β, γ mit h unendlich klein werden. Löst man jetzt nach $\frac{l}{h}$ auf und geht zur Grenze $h \doteq 0$ über, so folgt aus der Differentiierbarkeit von $f_{y'}[x]$ und der Gleichung (35), daß

$$\underset{h=0}{L} \frac{l}{h}, \quad \text{d. h. } y''$$

existiert, wenn die Bedingung (36) erfüllt ist, und daß alsdann

$$y'' = \frac{f_y - f_{y'x} - y' f_{y'y}}{f_{y'y'}}. \tag{37}$$

Hieraus folgt weiter, daß y'' stetig ist in allen Punkten von $[x_1 x_2]$, in welchen (36) erfüllt ist, und hieraus endlich, daß auch y''' in denselben Punkten existiert und stetig ist, wie man aus der Betrachtung der rechten Seite von (37) unmittelbar ersieht.

§ 6. Bemerkungen zur Integration der Euler'schen Differentialgleichung.

Wir stellen in diesem Paragraphen verschiedene für die Folge wichtige Bemerkungen über die Integration der Euler'schen Differentialgleichung zusammen.

a) **Die Extremalen:**

Die Euler'sche Differentialgleichung (I) ist im allgemeinen von der zweiten Ordnung wie die entwickelte Form (28) zeigt. Daher

---

[1]) Vgl. A IV 6.

enthält ihr allgemeines Integral zwei willkürliche Konstanten α, β; wir bezeichnen dasselbe mit

$$y = g(x, \alpha, \beta). \tag{38}$$

Die beiden Konstanten α, β sind im Fall fester Endpunkte aus der Bedingung zu bestimmen[1]), daß die gesuchte Kurve durch die beiden gegebenen Punkte $P_1(x_1, y_1)$ und $P_2(x_2, y_2)$ hindurchgehen soll. Es muß also sein

$$y_1 = g(x_1, \alpha, \beta), \qquad y_2 = g(x_2, \alpha, \beta). \tag{39}$$

Jede Lösung der Euler'schen Differentialgleichung (Kurve sowohl als Funktion) wird nach Kneser eine *Extremale* genannt; es gibt also eine doppelt unendliche Schar von Extremalen in der Ebene.

In dem speziellen Fall, wo die Funktion f die Variable y nicht enthält, erhält man sofort ein erstes Integral der Differentialgleichung (I), nämlich

$$f_{y'} = \text{konst.} \tag{40}$$

Aber auch, *wenn f die Variable x nicht explizite enthält, läßt sich ein erstes Integral angeben*[2]). Denn dann ist wegen $f_x \equiv 0$:

$$\frac{d}{dx}(f - y' f_{y'}) = y'\left(f_y - \frac{d}{dx} f_{y'}\right),$$

und daher genügt jede Lösung von (I) auch der Gleichung

$$f - y' f_{y'} = \text{konst.}; \tag{41}$$

und umgekehrt genügt jede Lösung von (41) — mit Ausnahme von: $y = \text{konst.}$ — auch der Differentialgleichung (I).

Beispiel VI: $f = G(y')$, eine Funktion von y' allein.

Hier erhält man nach (40)

$$G'(y') = \text{konst.},$$

daraus

$$y' = \alpha$$

$$y = \alpha x + \beta. \tag{42}$$

---

[1]) Es kann vorkommen, daß diese Bestimmung unmöglich ist — man beachte, daß α und β reell sein müssen —; in diesem Fall existiert keine Lösung der Aufgabe, welche von der Klasse C' ist und im Innern von $\mathfrak{R}$ liegt, vgl. z. B. die Rotationsfläche kleinsten Flächeninhalts (§ 13, c).

[2]) Schon von Euler bemerkt (*Methodus inveniendi etc.*, Kap. II, § 30). Die Existenz des ersten Integrals hängt damit zusammen, daß das Integral J hier bei der kontinuierlichen Gruppe: $\xi = x + a$ invariant bleibt, vgl. Guldberg, *Über Maxima und Minima der Integrale, die eine kontinuierliche Gruppe gestatten*, Videnskabsselskabets Skrifter 1902, Christiania.

Die Extremalen sind also die Geraden der Ebene. Die Bestimmung der Konstanten ist eindeutig.

Ein spezieller Fall hiervon ist die Aufgabe der kürzesten Kurve zwischen zwei gegebenen Punkten, wo

$$f=\sqrt{1+y'^2}\,.$$

Beispiel VII[1]: $f=\dfrac{\sqrt{1+y'^2}}{y}$.

Hier muß y auf den Bereich

$$\mathfrak{R}:\qquad y>0\qquad\qquad (\text{oder } y<0)$$

beschränkt werden.

Nach (41) erhält man ein erstes Integral

$$\frac{1}{y\sqrt{1+y'^2}}=\frac{1}{\beta}$$

und daraus das allgemeine Integral

$$y=\sqrt{\beta^2-(x-\alpha)^2}\,. \tag{43}$$

Die Extremalen sind also Halbkreise, die ihre Mittelpunkte auf der x-Achse haben. Die Konstantenbestimmung ist eindeutig, wie geometrisch ersichtlich.

Beispiel I: (Siehe p. 1).

$$f=y\sqrt{1+y'^2},\qquad y\geqq 0\,.$$

Nach (41) erhält man ein erstes Integral

$$\frac{y}{\sqrt{1+y'^2}}=\alpha\,.$$

Für $\alpha>0$, erhält man das allgemeine Integral[2]

$$y=\alpha\,\mathrm{Ch}\,\frac{x-\beta}{\alpha}\,. \tag{44}$$

Die Extremalen sind also Kettenlinien mit der x-Achse als Direktrix. Wegen der Bestimmung der Konstanten verweisen wir auf § 13, c).

Für $\alpha=0$ erhält man: $y=0$; dies ist zwar eine Lösung von (41), aber nicht von (I)[3].

---

[1] Vgl. Osgood, loc. cit. p. 109. Das Beispiel erscheint zuerst bei L'Hospital (Acta Eruditorum, 1697, p. 217) als Brachistochrone für ein Fallgesetz, bei welchem die Geschwindigkeit der Höhe proportional ist.

[2] Ich bediene mich für die hyperbolischen Funktionen der bequemen Bezeichnungsweise von Laisant, *Essai sur les fonctions hyperboliques.*

[3] Hierzu die *Übungsaufgaben* Nr. 2—12, Nr. 18, 19, und Nr. 35—40 am Ende von Kap. III.

Beispiel VIII: $f = \sqrt{y}\,\sqrt{1-y'^2}$.

Hier ist zu beachten, daß unsere allgemeinen Voraussetzungen über die Funktion $f(x, y, y')$ nicht erfüllt sind, da neben der Gebietsbeschränkung

$$y > 0$$

auch eine *Gefällbeschränkung* vorliegt, nämlich

$$-1 < y' < 1 .$$

Trotzdem bleiben unsere Schlüsse von §§ 4 und 5 gültig. Denn erfüllt die Kurve $\mathfrak{C}$ die beiden angegebenen Bedingungen, so können wir $|\varepsilon|$ so klein wählen, daß auch die Kurve $\overline{\mathfrak{C}}$ dieselben beiden Bedingungen erfüllt; es wird also durch die Gefällbeschränkung keine weitere Beschränkung der Funktion η eingeführt[1]).

Aus der Ungleichung für y' folgt durch Integration nach x von x_1 bis zu einem beliebigen größeren Wert von x:

$$-1 < \frac{y - y_1}{x - x_1} < 1 .$$

Ziehen wir also vom Punkt P_1 zwei Halbstrahlen vom Gefälle $+1$ und -1, so müssen alle zulässigen Kurven in dem Winkelraum zwischen diesen beiden Halbstrahlen und zugleich in der oberen Halbebene enthalten sein.

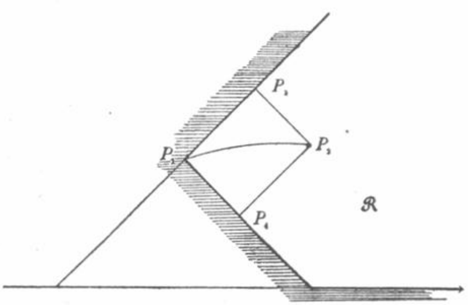

Fig. 4.

Für das allgemeine Integral der Euler'schen Differentialgleichung erhält man leicht

$$y = \beta - \frac{(x-\alpha)^2}{4\beta}, \quad (45)$$

wobei stets $\beta > 0$. *Die Extremalen sind also Parabeln, deren Brennpunkte auf der x-Achse liegen und deren Achsen der negativen y-Achse parallel sind.*

Durch irgend zwei Punkte P_1, P_2 der oberen Halbebene, welche die Bedingung

---

[1]) Ganz anders verhält es sich, wenn wir auch die Werte ± 1 für y' zulassen. Denn wenn entlang einem Segment der Kurve $\mathfrak{C}: y' = 1$ ist, so können wir dieses Segment *überhaupt nicht variieren,* ohne die Ungleichung für y' zu verletzen. Es bleibt also die Möglichkeit, daß es außer den Extremalen noch Lösungen gibt, welche geradlinige Segmente vom Gefälle ± 1 enthalten. In der Tat liefert eine gebrochene Linie, welche aus zwei solchen Segmenten besteht ($P_1 P_3 P_2$ oder $P_1 P_4 P_2$ in Fig. 4) für das Integral J den Wert Null, und daher sicher ein absolutes Minimum, wofern wir Kurven mit Ecken zulassen.

$$-1 < \frac{y_2 - y_1}{x_2 - x_1} < 1$$

erfüllen, geht eine und nur eine dieser Parabeln[1]).

b) Ausartungen[2]) der Euler'schen Differentialgleichung:

Wir betrachten weiter den speziellen Fall, wo die Ordnung der Euler'schen Differentialgleichung sich erniedrigt. Dies tritt dann und nur dann ein, wenn

$$f_{y'y'}(x, y, y') \equiv 0 \tag{46}$$

für alle x, y, y'. In diesem Fall reduziert sich die Euler'sche Differentialgleichung entweder auf eine endliche Gleichung, oder auf die Identität $0 = 0$, aber niemals auf eine Differentialgleichung erster Ordnung[3]).

Denn wenn $f_{y'y'}$ identisch verschwindet, so folgt durch zweimalige Integration der Gleichung (46) nach y', daß f selbst eine ganze lineare Funktion von y' sein muß, also von der Form

$$f = M(x, y) + N(x, y)\, y'.$$

Setzt man dies aber in (I) ein, so kommt

$$M_y - N_x = 0. \tag{47}$$

Hier sind nun zwei Fälle zu unterscheiden:

Entweder die Gleichung (47) ist keine Identität; dann stellt sie eine *endliche Gleichung* zwischen x und y dar. Wir erhalten also nur eine einzige Kurve als mögliche Lösung, und es wird dann im allgemeinen nicht möglich sein, die weitere Bedingung zu erfüllen, daß die Kurve durch die beiden gegebenen Punkte P_1 und P_2 geht.

Beispiel:

$$f = x^2 + y^2 + yy'.$$

Die Differentialgleichung (I) reduziert sich hier auf die endliche Gleichung: $y = 0$. Die Aufgabe ist also nur lösbar, wenn die beiden gegebenen Punkte auf der x-Achse liegen.

---

[1]) Dies folgt unmittelbar aus einem allgemeinen Satz von E. H. Moore (siehe § 26, b)), läßt sich aber auch leicht direkt beweisen.

[2]) Schon von Euler behandelt, *Methodus inveniendi etc.*, Kap. II, § 32.

[3]) Auch in dem allgemeineren Fall, wenn f höhere Ableitungen von y enthält, kann die Euler'sche Differentialgleichung sich nie auf eine Differentialgleichung *ungerader* Ordnung reduzieren, vgl. Frobenius, Journal für Mathematik, Bd. LXXXV (1878), p. 206, und Hirsch, Mathematische Annalen, Bd. XLIX (1897), p. 49.

Oder die Gleichung (47) ist eine *Identität*, gültig für alle Werte von x und y:

$$M_y \equiv N_x. \tag{48}$$

Dies ist aber die bekannte Integrabilitätsbedingung [1]) für den Differentialausdruck

$$M dx + N dy,$$

d. h. sind die Funktionen M und N nebst den partiellen Ableitungen M_y und N_x stetig und genügen der Identität (48) in einem einfach zusammenhängenden [2]) Bereich $\mathcal{S}$ der x, y-Ebene, so existiert eine Funktion $V(x, y)$, welche eindeutig und von der Klasse C' ist in $\mathcal{S}$, und für welche

$$V_x = M, \qquad V_y = N; \tag{49}$$

daher ist

$$f(x, y, y') = V_x + V_y y' = \frac{d}{dx} V(x, y). \tag{50}$$

Ist daher $\mathfrak{C}: y = y(x)$ irgend eine Kurve der Klasse C', welche die beiden Punkte $P_1(x_1, y_1)$ und $P_2(x_2, y_2)$ verbindet, und welche ganz in $\mathcal{S}$ liegt, so hat das Integral $J_{\mathfrak{C}}$ den Wert

$$J_{\mathfrak{C}} = \int_{x_1}^{x_2} \frac{d}{dx} V(x, y)\, dx = V(x_2, y_2) - V(x_1, y_1); \tag{51}$$

sein Wert ist also *unabhängig vom Integrationsweg* und hängt nur von der Lage der Endpunkte ab [3]).

Es ist klar, daß in diesem Falle ein „eigentliches“ Extremum des Integrals nicht stattfinden kann.

---

[1]) Vgl. *Encyclopädie*, II A, p. 112—114; Picard, *Traité d'Analyse*, (2^me éd.) Bd. I, p. 93; Goursat, *Cours d'Analyse*, Bd. I, p. 358. Der Beweis beruht auf der Betrachtung des Integrals

$$\int_{(x_1, y_1)}^{(x, y)} M dx + N dy.$$

[2]) Darunter soll das Innere einer stetigen geschlossenen Kurve ohne vielfache Punkte, zusammen mit dieser Kurve selbst, verstanden werden (einer sogenannten „Jordan'schen Kurve“); vgl. A VI 2.

[3]) Wegen der Stetigkeit der Funktion $V(x, y)$ bleibt der Satz auch richtig für stetige Kurven, welche sich aus einer endlichen Anzahl von Bogen der Klasse C' zusammensetzen (Kurven der Klasse D' in der Terminologie von § 10, c), wovon man sich leicht durch Zerlegung des Integrals $J_{\mathfrak{C}}$ überzeugt.

Beispiel:

$$f = 3x^2y^2 + 2x^3yy'.$$

Hier ist:

$$f_y = 6x^2y + 2x^3y', \qquad f_{y'} = 2x^3y,$$

also

$$\frac{d}{dx}f_{y'} = 6x^2y + 2x^3y' \equiv f_y.$$

Umgekehrt: Wenn das Integral $J_{\mathfrak{C}}$ denselben Wert hat für alle zulässigen Kurven $\mathfrak{C}$, welche durch P_1 und P_2 gehen, und welche im Innern eines einfach zusammenhängenden Bereiches $\mathcal{S}$ liegen, (der in dem in § 3, a) eingeführten Bereich $\mathfrak{R}$ enthalten ist), *dann muß die Euler'sche Differentialgleichung identisch erfüllt sein*

$$f_y - \frac{d}{dx}f_{y'} \equiv 0.$$

Denn sei (x_0, y_0) irgend ein innerer Punkt von $\mathcal{S}$, dessen Abszisse x_0 zwischen x_1 und x_2 liegt und seien y_0' und y_0'' zwei willkürlich vorgeschriebene endliche Werte. Dann können wir stets im Innern von $\mathcal{S}$ eine zulässige Kurve $\mathfrak{C} : y = y(x)$ von der Klasse C'' konstruieren, welche durch die Punkte (x_1, y_1), (x_2, y_2) und (x_0, y_0) hindurchgeht und für welche $y'(x_0) = y_0'$, $y''(x_0) = y_0''$ ist. Variieren wir diese Kurve $\mathfrak{C}$, so muß nach unserer Annahme $\Delta J = 0$ sein für jede zulässige Variation, insbesondere also für Variationen der speziellen Form (19), also unter Benutzung der Bezeichnung von § 4, c)

$$J(\varepsilon) \equiv J(0)$$

für alle hinreichend kleinen $|\varepsilon|$.

Also muß sein: $J'(\varepsilon) \equiv 0$, insbesondere $J'(0) = 0$, woraus nach §§ 4 und 5 folgt, daß $y(x)$ der Euler'schen Differentialgleichung genügen muß. Die linke Seite derselben muß also für das willkürliche Wertsystem $x = x_0$, $y = y_0$, $y' = y_0'$, $y'' = y_0''$, also identisch verschwinden[1]).

c) Das inverse Problem der Variationsrechnung:[2])

Wir betrachten schließlich noch kurz das folgende inverse Problem: *Es sei gegeben eine doppelt unendliche Schar von Kurven (Funktionen)*

$$y = g(x, \alpha, \beta).$$

[1]) Wegen der Verallgemeinerung des Satzes vgl. *Übungsaufgabe* Nr. 46 am Ende von Kap. III.

[2]) Hierzu die *Übungsaufgaben* Nr. 15—16 am Ende von Kap. III.

Es soll eine Funktion $f(x, y, y')$ bestimmt werden, derart, daß das gegebene System von Kurven das System der Extremalen für das Integral

$$J = \int_{x_1}^{x_2} f(x, y, y')\, dx$$

bildet.

Darboux[1]) hat gezeigt, daß dies Problem stets *unendlich viele Lösungen* besitzt, welche durch *Quadraturen* erhalten werden können.

Denn wenn

$$y'' = G(x, y, y') \tag{52}$$

die durch Elimination[2]) von α, β zwischen den drei Gleichungen

$$y = g(x, \alpha, \beta), \qquad y' = g_x(x, \alpha, \beta), \qquad y'' = g_{xx}(x, \alpha, \beta)$$

zu erhaltende Differentialgleichung zweiter Ordnung ist, deren allgemeine Lösung die gegebene Funktion: $y = g(x, \alpha, \beta)$ ist (mit α, β als Integrationskonstanten), so handelt es sich darum, die Funktion $f(x, y, y')$ so zu bestimmen, daß (52) mit der aus f abgeleiteten Euler'schen Differentialgleichung identisch wird, also nach (37) so, daß

$$f_y - f_{y'x} - y' f_{y'y} = G f_{y'y'} \tag{53}$$

für alle x, y, y'.

Differentiiert man jetzt (53) nach y', so erhält man für die Funktion $M = f_{y'y'}$ eine lineare partielle Differentialgleichung erster Ordnung, nämlich

$$\frac{\partial M}{\partial x} + y' \frac{\partial M}{\partial y} + G \frac{\partial M}{\partial y'} + G_{y'} M = 0\,. \tag{54}$$

Bezeichnen

$$\alpha = \varphi(x, y, y'), \qquad \beta = \psi(x, y, y')$$

die Lösung der beiden Gleichungen

$$y = g(x, \alpha, \beta), \qquad y' = g_x(x, \alpha, \beta)$$

nach α und β, und setzt man ferner

$$\theta(x, \alpha, \beta) = e^{\int G_{y'}(x,\, g(x, \alpha, \beta),\, g_x(x, \alpha, \beta))\, dx},$$

so findet man nach der allgemeinen Theorie[3]) der linearen partiellen Differentialgleichungen erster Ordnung für das allgemeine Integral von (54) den Ausdruck

---

[1]) *Théorie des surfaces*, Bd. III, Nr. 604, 605.

[2]) Vgl. z. B. Jordan, *Cours d'Analyse*, I, Nr. 166.

[3]) Vgl. z. B. Jordan, *Cours d'Analyse*, III, Nr. 242.

$$M = \frac{\Phi(\varphi(x, y, y'), \psi(x, y, y'))}{\theta(x, \varphi(x, y, y'), \psi(x, y, y'))},$$

wo Φ eine willkürliche Funktion von φ und ψ bedeutet.

Nachdem M gefunden ist, erhält man f durch zwei sukzessive Quadraturen aus der Gleichung

$$\frac{\partial^2 f}{\partial y'^2} = M(x, y, y'),$$

wobei zwei, noch von x und y abhängige Integrationskonstanten λ, μ eingeführt werden. Schließlich müssen die letzteren noch so bestimmt werden, daß f der ursprünglichen partiellen Differentialgleichung (53) genügt, aus welcher (54) durch Differentiation abgeleitet war.

Wir können das erhaltene Resultat auch dahin aussprechen, *daß jede Differentialgleichung zweiter Ordnung (auf unendlich viele Arten) als Euler'sche Differentialgleichung eines Problems der Variationsrechnung von dem einfachsten hier betrachteten Typus aufgefaßt werden kann.*[1])

Beispiel[2]): Alle Funktionen f zu bestimmen, für welche *die Extremalen gerade Linien* sind:

$$y = \alpha x + \beta.$$

Die Differentialgleichung (52) wird in diesem Fall

$$y'' = 0.$$

Wir erhalten daher

$$G \equiv 0, \qquad \varphi = y', \qquad \psi = y - xy'$$

und daraus

$$M = \Phi(y', y - xy')$$

und weiterhin

$$f = \int_0^{y'} (y' - t)\, \Phi(t, y - xt)\, dt + y' \lambda(x, y) + \mu(x, y).$$

Die Bedingung für λ und μ wird in diesem Fall

$$\frac{\partial \lambda}{\partial x} = \frac{\partial \mu}{\partial y};$$

der allgemeinste Ausdruck für λ und μ ist daher

$$\lambda = \frac{\partial \nu}{\partial y}, \qquad \mu = \frac{\partial \nu}{\partial x},$$

wo ν eine willkürliche Funktion von x und y ist.[3])

---

[1]) Dies findet nicht mehr statt bei der entsprechenden Aufgabe für den allgemeineren Typus, wo f höhere Ableitungen enthält; vgl. darüber Hirsch, Mathematische Annalen, Bd. XLIX (1897), p. 49 und Kasner, Bulletin of the Am. Math. Soc., Bd. XIII (1907), p. 289.

[2]) Vgl. Darboux, loc. cit. Nr. 606.

[3]) Die analoge Aufgabe für den Fall, wo die Extremalen Kreise sind, deren Mittelpunkte auf der x-Achse liegen, hat Stromquist gelöst (Transactions of the American Mathematical Society, Bd. 7 (1906), p. 175).

§ 7. Der Fall beweglicher Endpunkte.

Wir haben bei unseren bisherigen Entwicklungen stets die beiden Endpunkte der zulässigen Kurven als fest angenommen. Wir wollen in diesem Paragraphen nun auch den Fall betrachten, in welchem einer der beiden Endpunkte auf einer gegebenen Kurve beweglich[1]) ist.

a) Der Endpunkt P_1 ist auf der Geraden $x = x_1$ beweglich:

Wir betrachten zunächst den speziellen Fall, wo der Punkt P_1 auf der gegebenen Geraden

$$x = x_1 \tag{55}$$

beweglich ist, während der Punkt P_2 fest ist. Die Gesamtheit der zulässigen Kurven besteht jetzt also aus allen Kurven, welche von der gegebenen Geraden (55) nach dem gegebenen Punkt P_2 gezogen werden können, und welche im übrigen den in § 3, a) unter 2) bis 4) aufgezählten Bedingungen genügen.

Wir nehmen wieder an, wir hätten eine Kurve $\mathfrak{C}$ gefunden, welche in bezug auf diese Gesamtheit von zulässigen Kurven ein Minimum für das Integral J liefert; ihre Gleichung sei:

$$\mathfrak{C}: \qquad y = y(x), \qquad x_1 \overline{\gtrless} x \overline{\gtrless} x_2 .$$

Dann muß

$$\Delta J \geqq 0$$

sein für alle zulässigen Kurven, welche in einer gewissen Nachbarschaft der Kurve $\mathfrak{C}$ liegen, also insbesondere auch für alle diejenigen darunter, welche mit der Kurve $\mathfrak{C}$ den Anfangspunkt P_1 gemeinsam haben. Das heißt aber: die Kurve $\mathfrak{C}$ muß auch noch ein Minimum liefern, wenn der Endpunkt P_1 als fest betrachtet wird, und daraus folgt nach der früheren Theorie, daß die Funktion $y(x)$ der Euler'schen Differentialgleichung

$$f_y - \frac{d}{dx} f_{y'} = 0$$

genügen muß, daß somit auch im Falle variabler Endpunkte die Kurve $\mathfrak{C}$ eine *Extremale* sein muß. Wir setzen in der weiteren Diskussion voraus, daß diese Bedingung erfüllt ist.

---

[1]) Das älteste Beispiel dieser Art rührt von Jacob Bernoulli her (1697); es ist die Aufgabe der Brachistochrone, wenn der zweite Endpunkt auf einer vertikalen Geraden beweglich ist. Allgemein sind Aufgaben mit variabeln Endpunkten zuerst von Lagrange behandelt worden (1760), vgl. Œuvres, Bd. I, p. 338, 345.

Um nun weitere Bedingungen zu erhalten, betrachten wir jetzt in zweiter Linie eine Variation der Form

$$\overline{\mathfrak{C}}: \qquad y = y(x) + \varepsilon\eta(x) \equiv \overline{y}(x),$$

bei welcher der Endpunkt P_1 variiert wird. Die Funktion η ist dabei eine willkürliche Funktion der Klasse C', für welche

$$\eta(x_1) \neq 0, \qquad \eta(x_2) = 0.$$

Für diese Variation gelten die Schlüsse von § 4, wonach

$$\delta J \equiv \varepsilon \int_{x_1}^{x_2} (f_y \eta + f_{y'} \eta') dx = 0$$

sein muß. Wendet man jetzt die partielle Integration[1]) von § 5, a) an, so verschwindet in der Gleichung (25) das Integral, weil die Kurve $\mathfrak{C}$ eine Extremale ist, und es bleibt nur

$$\delta J = \varepsilon f_{y'}(x_1, y_1, y_1') \eta(x_1) = 0;$$

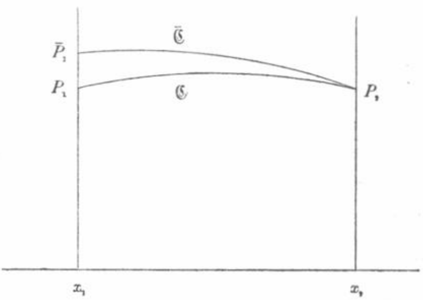

Fig. 5.

daraus folgt aber, da $\eta(x_1) \neq 0$:

Im Punkt P_1 muß die Bedingung

$$f_{y'}(x_1, y_1, y_1') = 0 \tag{56}$$

erfüllt sein.

Diese Bedingung, zusammen mit der Bedingung, daß der Punkt P_1 auf der Geraden $x = x_1$ liegen soll, und daß die Kurve $\mathfrak{C}$ durch den Punkt P_2 gehen soll, bestimmt im allgemeinen die beiden Integrationskonstanten in dem allgemeinen Integral der Euler'schen Differentialgleichung und die unbekannte Ordinate y_1 des Punktes P_1.

Beispiel I (siehe pp. 1, 33):

$$f = y\sqrt{1 + y'^2}.$$

Hier ist: $f_{y'} = \dfrac{yy'}{\sqrt{1+y'^2}}$. Da stets $y > 0$, so folgt

$$y_1' = 0,$$

d. h. die Kettenlinie muß im Punkt P_1 senkrecht auf der Geraden $x = x_1$ stehen.

---

[1]) Dieselbe ist statthaft, da wir annahmen, daß $\mathfrak{C}$ eine Extremale und zwar von der Klasse C' ist. Darin ist enthalten, daß $\dfrac{d}{dx} f_{y'}$ existiert und stetig ist.

b) Der Endpunkt P_1 ist auf einer beliebigen Kurve beweglich:

Wir wenden uns jetzt zu dem allgemeinen Fall, wo der Punkt P_1 auf einer beliebigen im Bereich $\mathfrak{R}$ gelegenen Kurve:

$$\tilde{\mathfrak{C}}: \qquad y = \tilde{y}(x)$$

der Klasse C' beweglich ist, während der Punkt P_2 fest ist.

Dazu müssen wir unsere frühere, in § 3, b) gegebene Definition des (relativen) Minimums etwas verallgemeinern. In den bisher behandelten Fällen lagen nämlich alle zulässigen Kurven in dem Streifen der Ebene zwischen den beiden festen Geraden $x = x_1$, $x = x_2$. Dies ist jetzt, wo die untere Grenze unseres Integrals nicht gegeben ist, nicht mehr der Fall. Wir werden daher jetzt sagen, eine Kurve $\mathfrak{C}$ liefere ein (relatives) Minimum für das Integral J, wenn

$$\Delta J \geqq 0$$

für jede zulässige Kurve $\bar{\mathfrak{C}}$, welche *in einer gewissen Umgebung* $\mathfrak{A}$ *der Kurve* $\mathfrak{C}$ gelegen ist. Dabei soll unter einer „Umgebung $\mathfrak{A}$ der Kurve $\mathfrak{C}$" jeder Bereich verstanden werden, welcher die Kurve $\mathfrak{C}$ in seinem Innern enthält, so daß also jeder Punkt von $\mathfrak{C}$ ein „innerer Punkt" von $\mathfrak{A}$ ist.¹)

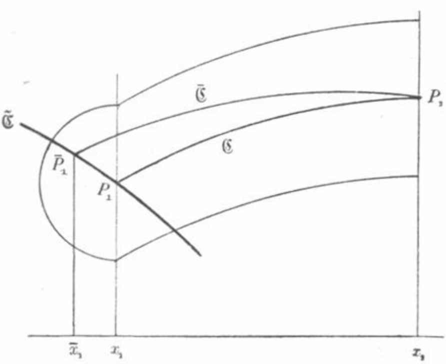

Fig. 6.

Dann schließen wir zunächst wieder, ganz wie unter a), daß die gesuchte Kurve

$$\mathfrak{C}: \quad y = y(x), \ x_1 \leqq x \leqq x_2,$$

eine *Extremale* sein muß.

Alsdann konstruieren wir folgendermaßen eine zulässige Variation, welche den Endpunkt P_1 variiert. Es sei $\bar{P}_1$ derjenige Punkt der Kurve $\tilde{\mathfrak{C}}$, dessen Abszisse

¹) Vgl. A I 7. Mit Hilfe des in § 21, a) bewiesenen Lemmas läßt sich leicht zeigen, daß für die Probleme mit festen Grenzen x_1, x_2, die jetzige Definition mit der früheren äquivalent ist, sowie daß wir bei der vorliegenden Aufgabe ohne Einschränkung der Allgemeinheit für $\mathfrak{A}$, z. B. den Bereich wählen können, der dadurch entsteht, daß man der Nachbarschaft (ϱ) der Kurve $\mathfrak{C}$ noch einen Halbkreis mit dem Mittelpunkt P_1 und dem Radius ϱ hinzufügt (siehe Fig. 6).

$$\bar{x}_1 = x_1 + \varepsilon$$

ist. Dann wählen wir willkürlich eine Funktion $\eta(x)$ von der Klasse C', welche den Bedingungen

$$\eta(x_1) \neq 0, \qquad \eta(x_2) = 0$$

genügt, und betrachten die Schar von Kurven

$$y = y(x) + k\eta(x),$$

welche sämtlich durch den Punkt P_2 gehen. Eine Kurve dieser Schar wird dann durch den Punkt $\overline{P}_1$ gehen, nämlich diejenige, für welche der Parameter k aus der Gleichung

$$y(x_1 + \varepsilon) + k\eta(x_1 + \varepsilon) = \tilde{y}(x_1 + \varepsilon)$$

bestimmt wird, woraus sich

$$k = \frac{\tilde{y}(x_1 + \varepsilon) - y(x_1 + \varepsilon)}{\eta(x_1 + \varepsilon)} \equiv k(\varepsilon)$$

ergibt. Dann stellt die Kurve

$$\overline{\mathfrak{C}}: \qquad y = y(x) + k(\varepsilon)\,\eta(x), \qquad \bar{x}_1 \gtreqqless x \gtreqqless x_2$$

für jeden hinreichend kleinen Wert von $|\varepsilon|$ eine zulässige Variation von $\mathfrak{C}$ dar. Da

$$\tilde{y}(x_1) = y(x_1),$$

so ist $k(0) = 0$; die Kurve $\overline{\mathfrak{C}}$ reduziert sich also für $\varepsilon = 0$ auf die Kurve $\mathfrak{C}$. Ferner merken wir noch an, daß

$$\left(\frac{dk(\varepsilon)}{d\varepsilon}\right)_{\varepsilon=0} = \frac{\tilde{y}_1' - y_1'}{\eta_1}, \tag{57}$$

wenn wir zur Abkürzung schreiben

$$y_1' = y'(x_1), \qquad \tilde{y}_1' = \tilde{y}'(x_1), \qquad \eta_1 = \eta(x_1).$$

Wir bilden jetzt das Integral $J_{\overline{\mathfrak{C}}}$. Dasselbe ist eine eindeutige Funktion von ε, die wir mit $J(\varepsilon)$ bezeichnen, so daß

$$J(\varepsilon) = \int_{\bar{x}_1}^{x_2} f(x,\, y + k(\varepsilon)\eta,\, y' + k(\varepsilon)\eta')\,dx,$$

wobei wir besonders hervorheben, daß jetzt die untere Grenze des Integrals von ε abhängt.

Die Funktion $J(\varepsilon)$, die sich für $\varepsilon = 0$ auf $J_{\mathfrak{C}}$ reduziert, muß nun wieder für $\varepsilon = 0$ ein Minimum besitzen; also muß

$$J'(0) = 0$$

sein. Nach den Regeln für die Differentiation eines bestimmten Integrals nach einem Parameter erhält man aber[1]):

$$J'(0) = -f(x_1, y_1, y_1') + \left(\frac{dk}{d\varepsilon}\right)_0 \int_{x_1}^{x_2} (f_y \eta + f_{y'} \eta')\, dx.$$

Wendet man nun auf das zweite Glied unter dem Integral die Lagrange'sche partielle Integration an, und beachtet, daß die Kurve $\mathfrak{C}$ eine Extremale ist, so erhält man unter Benutzung von (57)

$$J'(0) = -\left[f(x_1, y_1, y_1') + (\tilde{y}_1' - y_1') f_{y'}(x_1, y_1, y_1')\right], \tag{58}$$

und somit das Resultat:

Im Punkt P_1 muß zwischen den Gefällen der beiden Kurven $\mathfrak{C}$ und $\tilde{\mathfrak{C}}$ die Relation stattfinden

$$f(x_1, y_1, y_1') + (\tilde{y}_1' - y_1') f_{y'}(x_1, y_1, y_1') = 0. \tag{59}$$

Wenn diese Bedingung erfüllt ist, so sagt man, *die Kurve $\tilde{\mathfrak{C}}$ schneide die Extremale $\mathfrak{C}$ im Punkte P_1 transversal.*[2])

Die Gleichung (59), zusammen mit den beiden Gleichungen:

$$g(x_2, \alpha, \beta) = y_2, \qquad g(x_1, \alpha, \beta) = \tilde{y}(x_1),$$

bestimmt im allgemeinen die Abszisse x_1 des Punktes P_1 und die beiden Integrationskonstanten α, β der allgemeinen Lösung der Euler'schen Differentialgleichung.

Beispiel I (siehe pp. 1, 33, 41):

$$f = y\sqrt{1 + y'^2}.$$

Hier lautet die Transversalitätsbedingung

$$\left.\frac{y(1 + y'\tilde{y}')}{\sqrt{1 + y'^2}}\right|^1 = 0,$$

welche besagt, daß die Kettenlinie

---

[1]) Dabei ist stillschweigend vorausgesetzt, daß die Funktion $f(x, y, y')$ nicht von x_1, y_1 abhängt; vgl. p. 50 und § 34, c).

[2]) Im Gebrauch des Wortes „transversal" folge ich Osgood, *Sufficient conditions etc.*, p. 112. Kneser, von dem der Ausdruck herrührt, sagt umgekehrt, $\tilde{\mathfrak{C}}$ werde von $\mathfrak{C}$ transversal geschnitten.

$$y = \alpha \operatorname{Ch} \frac{x - \beta}{\alpha}$$

die gegebene Kurve $\tilde{\mathfrak{C}}$ im Punkt P_1 *orthogonal* schneiden muß; wir erhalten also dieselbe Bedingung wie im speziellen Fall a).[1] Dasselbe Resultat gilt allgemein[2] für

$$f = G(x, y) \sqrt{1 + y'^2}.$$

Wir bemerken noch für späteren Gebrauch, daß sich aus (58) der folgende Ausdruck für die totale Variation ergibt:

$$\Delta J = -\varepsilon [f(x_1, y_1, y_1') + (\tilde{y}_1' - y_1') f_{y'}(x_1, y_1, y_1')] + \varepsilon(\varepsilon). \quad (60)$$

Wenn statt des Punktes P_1 der Punkt P_2 auf einer gegebenen Kurve $\tilde{\mathfrak{C}}$ beweglich, dagegen P_1 fest ist, so führt eine ganz analoge Betrachtung zu der entsprechenden Bedingung

$$f(x_2, y_2, y_2') + (\tilde{y}_2' - y_2') f_{y'}(x_2, y_2, y_2') = 0, \quad (59\text{a})$$

und dem Ausdruck

$$\Delta J = \varepsilon [f(x_2, y_2, y_2') + (\tilde{y}_2' - y_2') f_{y'}(x_2, y_2, y_2')] + \varepsilon(\varepsilon). \quad (60\text{a})$$

Sind beide Endpunkte beweglich, P_1 auf einer Kurve $\tilde{\mathfrak{C}}_1$, P_2 auf einer Kurve $\tilde{\mathfrak{C}}_2$, so muß die gesuchte Kurve eine Extremale sein, und es müssen gleichzeitig die beiden Transversalitätsbedingungen (59) und (59a) erfüllt sein.

§ 8. Der allgemeine δ-Prozeß.[3]

Wir knüpfen an die Entwicklungen des letzten Paragraphen eine Besprechung des allgemeinen δ-Prozesses, der eine so hervorragende Rolle in der älteren Variationsrechnung gespielt hat.[4] In § 4 haben wir eine vorläufige Definition desselben gegeben für spezielle Variationen der Form

$$\Delta y = \varepsilon \eta$$

---

[1] Die Bedingung (56) kann als Grenzfall von (59) aufgefaßt werden, für $\tilde{y}_1' = \infty$; vgl. übrigens die Behandlung des Problems in Parameterdarstellung, §§ 34, 35.

[2] Vgl. dazu *Übungsaufgabe* Nr. 17 am Ende von Kap. III.

[3] Wir empfehlen dem Leser, diesen Paragraphen vorläufig zu überschlagen und erst bei Bedarf darauf zurückzugreifen.

[4] Bei Lagrange und bei allen älteren Autoren ist „calcul des variations" geradezu identisch mit der Theorie des δ-Prozesses, und die Theorie der Extrema bestimmter Integrale wird als Anwendung dieses Variationskalküls aufgefaßt.

und für den Fall fester Endpunkte. Wie jedoch schon die in § 7 gegebenen Entwickelungen zeigen, kommt man nicht immer mit Variationen dieser einfachsten Art aus, und es wird häufig nötig, Variationen von dem allgemeinen Typus

$$\Delta y = \omega(x, \varepsilon)$$

heranzuziehen. Es soll jetzt gezeigt werden, wie dabei der δ-Prozeß zu modifizieren ist.

a) **Die erste Variation:**

Es sei

$$\mathfrak{C}: \qquad y = y(x), \qquad x_1 \overline{\gtrless} x \overline{\gtrless} x_2;$$

ein Kurvenbogen, von dem wir voraussetzen, daß die Funktion $y(x)$ von der Klasse C' ist in einem Intervall $[X_1 X_2]$, das nach beiden Seiten über $[x_1 x_2]$ hinausreicht, so daß $X_1 < x_1$, $X_2 > x_2$.

Die Endpunkte des Bogens $\mathfrak{C}$ seien P_1 und P_2. Wir betrachten jetzt eine Schar[1]) von Bogen mit einem Parameter ε, welcher den Bogen $\mathfrak{C}$ als Individuum enthält und zwar für $\varepsilon = 0$. Eine solche Schar können wir auf unendlich viele Weisen konstruieren, indem wir einfach setzen

$$y = y(x) + \omega(x, \varepsilon) \equiv \bar{y}(x, \varepsilon), \qquad \bar{x}_1(\varepsilon) \overline{\gtrless} x \overline{\gtrless} \bar{x}_2(\varepsilon), \tag{61}$$

wobei die Funktionen $\omega(x, \varepsilon)$, $\bar{x}_1(\varepsilon)$, $\bar{x}_2(\varepsilon)$ den Anfangsbedingungen genügen:

$$\omega(x, 0) = 0 \tag{62}$$

$$\bar{x}_1(0) = x_1, \qquad \bar{x}_2(0) = x_2. \tag{63}$$

Über die Funtion $\omega(x, \varepsilon)$ machen wir noch die weitere Annahme, daß sie selbst, sowie ihre partiellen Ableitungen ω_x, ω_ε, $\omega_{x\varepsilon}$ existieren und stetig sind in dem Bereich

$$X_1 \overline{\gtrless} x \overline{\gtrless} X_2 \qquad |\varepsilon| \overline{\gtrless} k,$$

[1]) Der Gedanke, die zu variierende Kurve als Individuum einer einparametrigen Kurvenschar aufzufassen, rührt von Euler her, (*Methodus nova et facilis calculum variationum tractandi,* Novi Commentarii Acad. Imp. Sc. Petropolitanae, Bd. 16 (1772), auch *Instit. Calc. Integr.*, Bd. 4, Suppl. XI). Derselbe ist von der größten Wichtigkeit für die Variationsrechnung gewesen, da er den Lagrange'schen „Variationskalkül" auf einen Differentiationsprozeß reduzierte und dadurch überhaupt erst auf eine feste Grundlage stellte. Zugleich ist er aber auch von nachteiligem Einfluß gewesen, insofern er zu dem naheliegenden Irrtum führte, daß man glaubte, auf diese Weise den allgemeinsten Ausdruck einer Variation zu erhalten (vgl. § 15, a).

wofern die positiven Größen k, $x_1 - X_1$, $X_2 - x_2$ hinreichend klein gewählt werden. Daraus folgt dann, daß auch $\omega_{\varepsilon x}$ in demselben Bereich existiert und gleich $\omega_{x\varepsilon}$ ist.[1])

Von den beiden Funktionen $\bar{x}_1(\varepsilon)$, $\bar{x}_2(\varepsilon)$ setzen wir außer der Gleichung (63) noch voraus, daß sie in dem Bereich $|\,\varepsilon\,| \overline{\gtrless} k$ von der Klasse C' sein sollen.

Wir variieren nun den Bogen $\mathfrak{C}$, indem wir ihn durch einen Bogen $\bar{\mathfrak{C}}$ der Schar (61) ersetzen. Eine solche Variation, welche alle soeben angeführten Bedingungen erfüllt, wollen wir eine „*Normal-Variation*“ nennen.[2])

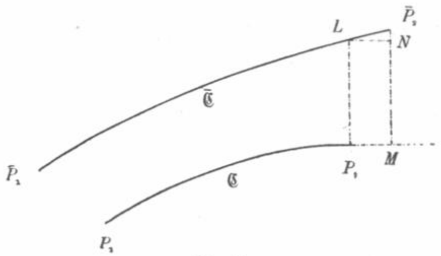

Fig. 7.

Jetzt sei $\varphi(x, y, y')$ irgend eine Funktion von x, y, y' welche von der Klasse C' ist in einer gewissen Umgebung $\mathfrak{A}'$ der Kurve

$$\mathfrak{C}': \quad y = y(x), \quad y' = y'(x),$$
$$x_1 \overline{\gtrless} x \overline{\gtrless} x_2$$

im Raum der Variabeln x, y, y'. Wir substituieren darin $y + \omega$, $y' + \omega'$ für y, y' und bezeichnen[3])

$$\overline{\varphi} = \varphi(x, y + \omega, y' + \omega').$$

Alsdann definieren[4]) wir

$$\delta\varphi = \varepsilon\left(\frac{\partial\overline{\varphi}}{\partial\varepsilon}\right)_{\varepsilon=0}. \tag{64}$$

---

[1]) Nach A IV 7.

[2]) Abweichend von Kneser, *Lehrbuch*, p. 91.

[3]) Akzente sollen Differentiation nach x auch dann bezeichnen, wenn neben x andere Variable vorkommen, also hier: $\omega' = \omega_x$, $\omega'_\varepsilon = \omega_{\varepsilon x}$ usw.

[4]) Definition und Bezeichnung sind schwankend. Die hier gegebene Definition von $\delta\varphi$ schließt sich an die von Jordan, *Cours d'Analyse*, III, Nr. 348 für Variationen der speziellen Form $\varepsilon\eta$ gegebene Definition an. Euler und nach ihm Stegemann (*Lehrbuch*) und Erdmann (Schlömilch's Zeitschrift, Bd. XXVI (1881) p. 76) definieren: $\delta\varphi = \left(\frac{\partial\overline{\varphi}}{\partial\varepsilon}\right)_0 d\varepsilon$. Dagegen definieren Ohm, Strauch und Moigno: $\delta\varphi = \left(\frac{\partial\overline{\varphi}}{\partial\varepsilon}\right)_0$ wofür Stegemann den Ausdruck „Variationsquotient“ und das Zeichen δ' gebraucht.

Falls $\overline{\varphi}$ Entwicklung nach Potenzen von ε zuläßt:

$$\overline{\varphi} = \varphi + \varepsilon \varphi_1 + \frac{\varepsilon^2}{2!} \varphi_2 + \cdots$$

so ist $\delta \varphi$ identisch mit dem Glied erster Ordnung $\varepsilon \varphi_1$ in dieser Entwicklung.

Wenden wir dies insbesondere auf die Funktionen $\varphi = y$ und $\varphi = y'$ an so kommt

$$\delta y = \varepsilon \omega_\varepsilon (x, 0), \qquad \delta y' = \varepsilon \omega_{x\varepsilon}(x, 0) \tag{65}$$

woraus wegen $\omega_{x\varepsilon} = \omega_{\varepsilon x}$ folgt:

$$\frac{d}{dx} \delta y = \delta \left(\frac{dy}{dx}\right). \tag{66}$$

Die Operationen der Variation und Differentiation sind also vertauschbar. Unter Benutzung von (65) wird jetzt

$$\delta \varphi = \varphi_y \, \delta y + \varphi_{y'} \delta y'. \tag{67}$$

Wir betrachten weiter das Integral

$$J = \int f(x, y, y') \, dx$$

unter der Voraussetzung, daß die Funktion $f(x, y, y')$ in einer gewissen Umgebung der Kurve $\mathfrak{C}'$ von der Klasse C' ist. Dieses Integral, genommen entlang einer Kurve $\overline{\mathfrak{C}}$ der Schar (61) ist eine Funktion von ε, die wir wieder mit $J(\varepsilon)$ bezeichnen:

$$J(\varepsilon) = \int_{\overline{x}_1(\varepsilon)}^{\overline{x}_2(\varepsilon)} f(x, y(x) + \omega(x, \varepsilon), y'(x) + \omega'(x, \varepsilon)) \, dx.$$

Dann definieren wir analog:

$$\delta J = \varepsilon \left(\frac{dJ(\varepsilon)}{d\varepsilon}\right)_{\varepsilon = 0}. \tag{68}$$

Nach den über die Funktionen y, ω, $\overline{x}_1$, $\overline{x}_2$ und f gemachten Annahmen dürfen wir bei der Differentiation von $J(\varepsilon)$ die gewöhnlichen Regeln[1]) über die Differentiation eines bestimmten Integrals nach einem Parameter anwenden, wobei insbesondere zu beachten ist, daß jetzt auch die Grenzen von ε abhängen. Definieren wir noch

$$\delta x_1 = \varepsilon \left(\frac{d\overline{x}_1}{d\varepsilon}\right)_{\varepsilon = 0}, \qquad \delta x_2 = \varepsilon \left(\frac{d\overline{x}_2}{d\varepsilon}\right)_{\varepsilon = 0}, \tag{69}$$

so ergibt die Ausführung der Rechnung:

[1]) Vgl. A V 7.

$$\delta J = \int_{x_1}^{x_2} (f_y \delta y + f_{y'} \delta y')\, dx + \Big[f \delta x\Big]_1^2,$$

oder, wenn man das bestimmte Integral in der üblichen Weise durch partielle Integration umformt (die Zulässigkeit der letzteren vorausgesetzt)

$$\begin{aligned} \delta J = \int_{x_1}^{x_2} \delta y \left(f_y - \frac{d}{dx} f_{y'}\right) dx + f(x_2, y_2, y_2')\, \delta x_2 + f_{y'}(x_2, y_2, y_2')\, (\delta y)_2 \\ - f(x_1, y_1, y_1')\, \delta x_1 - f_{y'}(x_1, y_1, y_1')\, (\delta y)_1; \end{aligned} \tag{70}$$

dabei bedeutet $(\delta y)_i$ den Wert von δy für $x = x_i$, also

$$(\delta y)_1 = \varepsilon \omega_\varepsilon(x_1, 0), \qquad (\delta y)_2 = \varepsilon \omega_\varepsilon(x_2, 0). \tag{71}$$

Dagegen bezeichnen wir

$$\delta(y_1) = \varepsilon \left(\frac{d \bar{y}_1}{d \varepsilon}\right)_0, \qquad \delta(y_2) = \varepsilon \left(\frac{d \bar{y}_2}{d \varepsilon}\right)_0. \tag{72}$$

Da

$$\frac{d \bar{y}_i}{d \varepsilon} = \frac{d \bar{y}(\bar{x}_i, \varepsilon)}{d \varepsilon} = \bar{y}'(\bar{x}_i, \varepsilon) \frac{d \bar{x}_i}{d \varepsilon} + \bar{y}_\varepsilon(\bar{x}_i, \varepsilon),$$

so ist

$$\delta(y_1) = y_1' \delta x_1 + (\delta y)_1, \qquad \delta(y_2) = y_2' \delta x_2 + (\delta y)_2. \tag{73}$$

Die Variationen $\delta(y_i)$ pflegt man *„gemischte[1]) Variationen"* zu nennen, im Gegensatz zu den sogenannten *„reinen Variationen"* $(\delta y)_i$.

Führt man mittels der Relationen (73) in der Formel (70) die gemischten Variationen statt der reinen ein, so erhält man schließlich:

$$\left.\begin{aligned} \delta J = \int_{x_1}^{x_2} \delta y \left(f_y - \frac{d}{dx} f_{y'}\right) dx + [f(x_2, y_2, y_2') - y_2' f_{y'}(x_2, y_2, y_2')]\, \delta x_2 + \\ + f_{y'}(x_2, y_2, y_2')\, \delta(y_2) - [f(x_1, y_1, y_1') - y_1' f_{y'}(x_1, y_1, y_1')]\, \delta x_1 - \\ - f_{y'}(x_1, y_1, y_1')\, \delta(y_1). \end{aligned}\right\} \tag{74}$$

---

[1]) Über reine und gemischte Variationen vgl. Kneser, *Encyclopädie*, II A p. 576, und Stegemann, *Lehrbuch*, § 56 und wegen der verschiedenen dafür gebrauchten Bezeichnungen die Vorrede zu letzterem, p. VII. Noch anders bezeichnet Erdmann, Schlömilch's Zeitschrift, Bd. XXII (1878), p. 363. Er schreibt $[\delta y_i]$, δy_i für $\delta(y_i)$, $(\delta y)_i$ resp. In Fig. 7 ist annähernd: $(\delta y)_2 = P_2 L$; $\delta(y_2) = M \bar{P}_2$.

Auf die *„Variation der unabhängigen Variabeln"*, die bei den älteren Autoren über Variationsrechnung eine große Rolle gespielt hat, gehen wir hier nicht ein. Sie hat viel Unklarheit in die Variationsrechnung gebracht, und man

Bei manchen Aufgaben[1]) hängt die Funktion f außer von x, y, y' auch noch von den Koordinaten der beiden Endpunkte ab. Es ist dann

$$J(\varepsilon) = \int_{\bar{x}_1}^{\bar{x}_2} f(x, \bar{y}, \bar{y}'; \bar{x}_1, \bar{y}_1; \bar{x}_2, \bar{y}_2)\, dx.$$

Daher muß man dem Ausdruck (70) oder (74) für δJ jetzt noch das Zusatzglied

$$\int_{x_1}^{x_2} \left(\frac{\partial f}{\partial x_1}\delta x_1 + \frac{\partial f}{\partial y_1}\delta(y_1) + \frac{\partial f}{\partial x_2}\delta x_2 + \frac{\partial f}{\partial y_2}\delta(y_2)\right) dx$$

hinzufügen, das sich unter Benutzung von (73) auch schreiben läßt[2])

$$\delta x_1 \int_{x_1}^{x_2} \left(\frac{\partial f}{\partial x_1} + y_1' \frac{\partial f}{\partial y_1}\right) dx + \delta x_2 \int_{x_1}^{x_2} \left(\frac{\partial f}{\partial x_2} + y_2' \frac{\partial f}{\partial y_2}\right) dx +$$
$$+ (\delta y)_1 \int_{x_1}^{x_2} \frac{\partial f}{\partial y_1}\, dx + (\delta y)_2 \int_{x_1}^{x_2} \frac{\partial f}{\partial y_2}\, dx.$$

Die bisherigen Definitionen und Resultate lassen sich unmittelbar auf den allgemeineren Fall von Variationen übertragen, welche von mehreren Parametern $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_m$ abhängen, wobei man über die Funktionen $\omega(x, \varepsilon_1, \varepsilon_2, \ldots, \varepsilon_m)$, $\bar{x}_1(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_m)$, $\bar{x}_2(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_m)$ die den obigen analogen Annahmen zu machen hat.

Für diesen Fall soll definiert werden:

$$\delta\varphi = \varepsilon_1 \left(\frac{\partial \bar{\varphi}}{\partial \varepsilon_1}\right)_0 + \varepsilon_2 \left(\frac{\partial \bar{\varphi}}{\partial \varepsilon_2}\right)_0 + \cdots + \varepsilon_m \left(\frac{\partial \bar{\varphi}}{\partial \varepsilon_m}\right)_0,$$
$$\delta J = \varepsilon_1 \left(\frac{\partial J}{\partial \varepsilon_1}\right)_0 + \varepsilon_2 \left(\frac{\partial J}{\partial \varepsilon_2}\right)_0 + \cdots + \varepsilon_m \left(\frac{\partial J}{\partial \varepsilon_m}\right)_0,$$
$$\delta x_i = \varepsilon_1 \left(\frac{\partial \bar{x}_i}{\partial \varepsilon_1}\right)_0 + \varepsilon_2 \left(\frac{\partial \bar{x}_i}{\partial \varepsilon_2}\right)_0 + \cdots + \varepsilon_m \left(\frac{\partial \bar{x}_i}{\partial \varepsilon_m}\right)_0.$$

---

kann vollständig ohne sie auskommen, um so mehr als der ihr zugrunde liegende Gedanke in der Parameterdarstellung seinen Ausdruck findet. Eine rationelle Begründung der Variation der unabhängigen Variabeln findet man bei JORDAN, *Cours d'Analyse,* III, Nr. 351.

[1]) Dies tritt bei der Aufgabe der *Brachistochrone* ein, bei welcher das Integral

$$J = \int_{x_1}^{x_2} \frac{\sqrt{1+y'^2}\, dx}{\sqrt{y - y_1 + k}}, \quad k \text{ eine Konstante},$$

zu einem Minimum zu machen ist. Vgl. § ~~34, e).~~ 36, b)

[2]) Vgl. LINDELÖF-MOIGNO, loc. cit. Nr. 63.

Mit dieser allgemeineren Bedeutung des Symbols δ bleibt dann auch die Formel (74) gültig.

Die im vorangehenden entwickelten Begriffsbildungen lassen sich ohne weiteres auf Variationsprobleme übertragen, in welchen *mehrere unbekannte Funktionen* von x vorkommen. Für das Integral

$$J = \int_{x_1}^{x_2} f(x, y_1, y_2, \ldots, y_n, y_1', y_2', \ldots, y_n')\, dx$$

erhält man so im Fall fester Grenzen x_1, x_2, wenn man die Kurve

$$\mathfrak{C}: \qquad y_i = y_i(x), \qquad x_1 \overline{\overline{<}} x \overline{\overline{<}} x_2, \qquad i = 1, 2, \ldots, n,$$

im Raum der Variabeln $x, y_1, \ldots, y_n$ durch die Kurve

$$\overline{\mathfrak{C}}: \qquad y_i = y_i(x) + \omega_i(x, \varepsilon), \qquad x_1 \overline{\overline{<}} x \overline{\overline{<}} x_2, \qquad i = 1, 2, \ldots, n,$$

ersetzt:

$$\delta J = \sum_{i=1}^{n} \left[\frac{\partial f}{\partial y_i'} \delta y_i\right]_{x_1}^{x_2} + \int_{x_1}^{x_2} \sum_{i=1}^{n} \delta y_i \left(\frac{\partial f}{\partial y_i} - \frac{d}{dx}\frac{\partial f}{\partial y_i'}\right) dx, \tag{75}$$

wobei

$$\delta y_i = \varepsilon \left.\frac{\partial \omega_i(x, \varepsilon)}{\partial \varepsilon}\right|^{\varepsilon = 0}. \tag{76}$$

Dabei ist vorausgesetzt, daß die Funktionen $y_i(x)$ im Intervall $[x_1 x_2]$ von der Klasse C'' sind, und daß die Funktionen $\omega_i(x, \varepsilon)$ dieselben Eigenschaften besitzen, wie oben die Funktion $\omega(x, \varepsilon)$.

Wenn insbesondere die Kurve $\mathfrak{C}$ eine „Extremale" ist, d. h. wenn die Funktionen $y_i(x)$ dem System von Differentialgleichungen

$$\frac{\partial f}{\partial y_i} - \frac{d}{dx}\frac{\partial f}{\partial y_i'} = 0, \qquad i = 1, 2, \ldots, n, \tag{77}$$

genügen, so reduziert sich der Ausdruck für die erste Variation auf

$$\delta J = \sum_{i=1}^{n} \left[\frac{\partial f}{\partial y_i'} \delta y_i\right]_{x_1}^{x_2}, \tag{78}$$

und daraus erhält man[1]) nach der Definition der Ableitung die folgende *Fundamentalformel für die Variation eines Extremalenbogens* bei festen Grenzen x_1, x_2:

$$\Delta J = \sum_{i=1}^{n} \left[\frac{\partial f}{\partial y_i'} \delta y_i\right]_{x_1}^{x_2} + \varepsilon(\varepsilon). \tag{79}$$

[1]) Vgl. dazu Gleichung (21).

b) Die höheren Variationen:

Wir schließen hieran eine kurze Besprechung der höheren Variationen. Da es sich dabei wesentlich um formale Entwicklungen handelt, wollen wir der Einfachheit halber annehmen, daß alle auftretenden Funktionen und Ableitungen in den in Betracht kommenden Bereichen existieren und stetig sind, daß alle vorkommenden Reihenentwicklungen konvergieren, und daß überhaupt die gewöhnlichen Regeln der formalen Differential- und Integralrechnung anwendbar sind.

Wir betrachten — etwas allgemeiner als bisher — eine Funktion $\varphi(x, y, y', \ldots, y^{(p)})$ von x, y und einigen Ableitungen von y. Wir ersetzen darin y durch $\overline{y}(x, \varepsilon)$ und dementsprechend $y^{(\lambda)}$ durch $\overline{y}^{(\lambda)}(x, \varepsilon)$. Die so entstehende Funktion von x und ε

$$\overline{\varphi} = \varphi(x, \overline{y}, \overline{y}', \ldots, \overline{y}^{(p)}),$$

entwickeln wir nach Potenzen von ε; es sei

$$\overline{\varphi} = \varphi + \frac{\varepsilon}{1!}\varphi_1 + \frac{\varepsilon^2}{2!}\varphi_2 + \cdots$$

Alsdann heißt[1]) die Größe $\varepsilon^n \varphi_n$ *die n^{te} Variation von φ* und wird mit $\delta^n \varphi$ bezeichnet; also

$$\delta^n \varphi = \varepsilon^n \left(\frac{d^n \overline{\varphi}}{d\varepsilon^n}\right)_{\varepsilon=0}. \tag{80}$$

Insbesondere ist demnach

$$\delta^n y = \varepsilon^n \left(\frac{\partial^n \omega(x, \varepsilon)}{\partial \varepsilon^n}\right)_{\varepsilon=0},$$

$$\delta^n y^{(\lambda)} = \varepsilon^n \left(\frac{\partial^{n+\lambda} \omega(x, \varepsilon)}{\partial x^\lambda \partial \varepsilon^n}\right)_{\varepsilon=0} = \varepsilon^n \frac{d^\lambda}{dx^\lambda}\left(\frac{\partial^n \omega(x, \varepsilon)}{\partial \varepsilon^n}\right)_{\varepsilon=0},$$

also

$$\delta^n \left(\frac{d^\lambda y}{dx^\lambda}\right) = \frac{d}{dx^\lambda}\, \delta^n y. \tag{81}$$

Für die Variationen niedrigster Ordnung[2]) von φ erhält man nach den Regeln für die Differentiation zusammengesetzter Funktionen

$$\delta\varphi = \varphi_y \delta y + \varphi_{y'} \delta y' + \cdots + \varphi_{y^{(p)}} \delta y^{(p)}, \tag{82}$$

wofür wir auch in der üblichen Symbolik[3]) schreiben

---

[1]) Vgl. Fußnote [4]) auf p. 47.

[2]) Eine *independente* Darstellung für $\delta^n \varphi$ findet sich für $p = 2$ bei ERDMANN, Schlömilch's Zeitschrift, Bd. XXVI (1881) p. 76.

[3]) Vgl. z. B. JORDAN, *Cours d'Analyse,* I, Nr. 129 und 253.

$$\delta\varphi = \left(\delta y \frac{\partial}{\partial y} + \delta y' \frac{\partial}{\partial y'} + \cdots + \delta y^{(p)} \frac{\partial}{\partial y^{(p)}}\right) \varphi;$$

ferner

$$\left.\begin{aligned} \delta^2\varphi = &\left(\delta y \frac{\partial}{\partial y} + \delta y' \frac{\partial}{\partial y'} + \cdots + \delta y^{(p)} \frac{\partial}{\partial y^{(p)}}\right)^2 \varphi + \\ &+ \varphi_y \delta^2 y + \varphi_{y'} \delta^2 y' + \cdots + \varphi_{y^{(p)}} \delta^2 y^{(p)} \end{aligned}\right\}. \tag{83}$$

In ganz analoger Weise wird $\delta^n J$ definiert; wir entwickeln das Integral:

$$J(\varepsilon) = \int_{\bar{x}_1}^{\bar{x}_2} f(x, \bar{y}, \bar{y}')\, dx$$

nach Potenzen von ε:

$$J(\varepsilon) = J_0 + \frac{\varepsilon}{1!} J_1 + \frac{\varepsilon^2}{2!} J_2 \cdots;$$

dann heißt $\varepsilon^n J_n$ *die* n^{te} *Variation des Integrals* J und wird mit $\delta^n J$ bezeichnet. Es ist also wieder

$$\delta^n J = \varepsilon^n \left(\frac{d^n J(\varepsilon)}{d\varepsilon^n}\right)_{\varepsilon=0}; \tag{84}$$

daraus ergibt sich der explizite Ausdruck von $\delta^n J$ nach den gewöhnlichen Regeln für die Differentiation eines Integrals nach einem Parameter; z. B. findet man[1]), wenn man noch definiert

$$\delta^n x_1 = \varepsilon^n \left(\frac{d^n \bar{x}_1}{d\varepsilon^n}\right)_0, \qquad \delta^n x_2 = \varepsilon^n \left(\frac{d^n \bar{x}_2}{d\varepsilon^n}\right)_0, \tag{85}$$

$$\delta^2 J = \int_{x_1}^{x_2} \delta^2 f\, dx + \left[\frac{df}{dx} \delta x^2 + 2 f_y\, \delta x\, \delta y + 2 f_{y'}\, \delta x\, \delta y' + f \delta^2 x\right]_1^2, \tag{86}$$

wobei nach (83)

$$\delta^2 f = f_{yy}\, \delta y^2 + 2 f_{yy'}\, \delta y\, \delta y' + f_{y'y'}\, \delta y'^2 + f_y \delta^2 y + f_{y'} \delta^2 y'.$$

---

[1]) Zuerst von Erdmann gegeben, Schlömilch's Zeitschrift, Bd. XXIII (1878) p. 363.