

Sitzungsberichte

der

mathematisch-naturwissenschaftlichen
Abteilung

der

Bayerischen Akademie der Wissenschaften

zu München

1926. Heft II
Mai- bis Julisitzung

München 1926

Verlag der Bayerischen Akademie der Wissenschaften
in Kommission des Verlags R. Oldenbourg München

Über Maxima und Minima und eine Modifikation des Begriffs der höheren Ableitungen.

Von O. Perron.

Vorgelegt in der Sitzung am 12. Juni 1926.

§ 1. Einleitung.

In der Differentialrechnung beweist man gewöhnlich die beiden folgenden Sätze:

A. Die Funktion $f(x)$ hat, wenn $f'(x_0) = 0$, wenn ferner $f''(x_0)$ positiv (negativ) ist und wenn $f''(x)$ an der Stelle x_0 stetig ist, an dieser Stelle ein Minimum (Maximum).

B. Die Funktion $f(x_1, x_2, \dots, x_m)$ hat, wenn

$$f'_{x_\mu}(x_1^0, x_2^0, \dots, x_m^0) = 0 \quad (\mu = 1, 2, \dots, m),$$

wenn ferner die quadratische Form

$$\sum_{\lambda=1}^m \sum_{\mu=1}^m f''_{x_\lambda x_\mu}(x_1^0, x_2^0, \dots, x_m^0) X_\lambda X_\mu$$

positiv (negativ) definit ist und wenn die zweiten Ableitungen an der Stelle $x_1^0, \dots, x_m^0$ stetig sind, an dieser Stelle ein Minimum (Maximum).

In dem Satz A kann die Voraussetzung, daß $f''(x)$ stetig ist, glatt gestrichen werden; $f''(x)$ braucht außerhalb der Stelle x_0 überhaupt nicht zu existieren¹⁾. Dagegen darf in Satz B die Voraussetzung, daß die zweiten Ableitungen stetig sind, nicht

¹⁾ Ohne diese Voraussetzung ist der Satz z. B. bei H. von Mangoldt bewiesen: Einführung in die höhere Mathematik, Bd. 2.

gestrichen werden, selbst dann nicht, wenn durchweg $f''_{x_\lambda x_\mu} = f''_{x_\mu x_\lambda}$ ist, wie folgendes Beispiel zeigt:

$$f(x, y) = \begin{cases} x^2 + y^2 - \frac{5x^2y^2}{x^2 + y^2} & \text{für } x^2 + y^2 > 0, \\ 0 & \text{für } x = y = 0. \end{cases}$$

Hier ist nämlich am Nullpunkt

$$\begin{aligned} f'_x &= 0, & f'_y &= 0, \\ f''_{xx} &= 2, & f''_{xy} &= f''_{yx} = 0, & f''_{yy} &= 2. \end{aligned}$$

Die quadratische Form $2X^2 + 2Y^2$ ist positiv definit, aber trotzdem hat die Funktion im Nullpunkt kein Minimum; denn es ist z. B.

$$f(x, x) = 2x^2 - \frac{5x^2}{2} < 0.$$

Dieser Unterschied im Verhalten der Funktionen einer Variablen und mehrerer Variablen kann beseitigt werden, wenn man den Begriff „Ableitung“ durch einen etwas anderen ersetzt, den ich „Derivat“ nenne. Die Derivate werden bei Funktionen einer Variablen und bei Funktionen mehrerer Variablen völlig einheitlich definiert. Trotzdem greift der Begriff „Derivat“ bei Funktionen einer Variablen über den Begriff „Ableitung“ hinaus, insofern die Existenz der n^{ten} Ableitung auch die Existenz des n^{ten} Derivats nach sich zieht, aber nicht umgekehrt. Bei Funktionen von mehreren Variablen greift er aber nur teilweise über den Begriff der Ableitung hinaus und bleibt in anderer Richtung dahinter zurück.

§ 2. Die Derivate einer Funktion von einer Variablen.

Sei $f(x)$ eine reelle Funktion der reellen Variablen x . Das erste Derivat von $f(x)$ definieren wir durch die Formel

$$(1) \quad \mathfrak{D}^1 f(x) = \mathfrak{D} f(x) = \lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h}.$$

Das erste Derivat deckt sich also mit der ersten Ableitung; seine Existenz (d. h. auch Endlichkeit) an einer Stelle x zieht die Stetigkeit der Funktion an der Stelle x nach sich. Dagegen definieren wir die höheren Derivate rekurrent durch die Formel

$$(2) \quad \mathfrak{D}^n f(x) = n! \lim_{h \rightarrow 0} \frac{f(x+h) - f(x) - h \mathfrak{D} f(x) - \dots - \frac{h^{n-1}}{(n-1)!} \mathfrak{D}^{n-1} f(x)}{h^n}.$$

Es wird also bei der Definition des n^{ten} Derivates an einer Stelle x auch die Existenz der vorausgehenden Derivate nur an der Stelle x verlangt, während die Definition der n^{ten} Ableitung erfordert, daß die früheren Ableitungen nicht nur an der betreffenden Stelle, sondern auch in der Umgebung existieren. Nun beweisen wir den

Satz 1. Wenn an einer Stelle x die n^{te} Ableitung $f^{(n)}(x)$ existiert, so existiert auch das n^{te} Derivat $\mathfrak{D}^n f(x)$ und es ist $\mathfrak{D}^n f(x) = f^{(n)}(x)$.

Der Satz ist zunächst richtig für $n = 1$. Seine Allgemeingültigkeit beweisen wir durch den Schluß von $n - 1$ auf n . Wir nehmen daher als bereits bewiesen an, daß

$$(3) \quad \mathfrak{D}^v f(x) = f^{(v)}(x) \quad (v = 1, 2, \dots, n-1)$$

ist. Dann muß gezeigt werden, daß, wenn $f^{(n)}(x)$ existiert, der Bruch in (2) für $h \rightarrow 0$ dem Grenzwert $\frac{1}{n!} f^{(n)}(x)$ zustrebt. Dieser Bruch ist aber mit Rücksicht auf (3) gleich

$$\frac{f(x+h) - f(x) - hf'(x) - \dots - \frac{h^{n-1}}{(n-1)!} f^{(n-1)}(x)}{h^n}.$$

Hier streben nun Zähler und Nenner nach 0; nach einer bekannten Regel wird man also die Berechnung des Grenzwerts dadurch versuchen, daß man Zähler und Nenner nach h differenziert. In dem neuen Bruch streben dann, weil $f'(x)$ wegen der Existenz von $f^{(n)}(x)$ stetig ist, Zähler und Nenner wieder nach 0 und man kann das Verfahren wiederholen usw. Nach dem $(n-1)^{\text{ten}}$ und n^{ten} Schritt ergibt sich schließlich:

$$\begin{aligned} \lim_{h \rightarrow 0} \frac{f(x+h) - f(x) - hf'(x) - \dots - \frac{h^{n-1}}{(n-1)!} f^{(n-1)}(x)}{h^n} \\ = \lim_{h \rightarrow 0} \frac{f^{(n-1)}(x+h) - f^{(n-1)}(x)}{n! h} = \frac{1}{n!} f^{(n)}(x), \end{aligned}$$

womit der Satz bewiesen ist.

Daß anderseits der Begriff des Derivats allgemeiner als der der Ableitung ist, mögen folgende Beispiele zeigen:

Erstes Beispiel:

$$f'(x) = \begin{cases} x^3 \sin \frac{1}{x} & \text{für } x \neq 0 \\ 0 & \text{für } x = 0. \end{cases}$$

Hier ist die erste Ableitung, also auch das erste Derivat am Nullpunkt gleich 0: $\mathfrak{D}f(0) = f'(0) = 0$.

Die zweite Ableitung ist am Nullpunkt nicht vorhanden, wohl aber das zweite Derivat, und zwar ist

$$\mathfrak{D}^2 f(0) = 2 \lim_{h \rightarrow 0} \frac{f(h) - f(0) - h \mathfrak{D}f(0)}{h^2} = 2 \lim_{h \rightarrow 0} h \sin \frac{1}{h} = 0.$$

Zweites Beispiel:

$$f(x) = a_0 + a_1 x + \frac{a_2}{2!} x^2 + \dots + \frac{a_n}{n!} x^n + x^{n+1} w(x),$$

wo $w(x)$ eine stetige, aber nirgends differenzierbare Funktion ist.

Hier ist

$$\mathfrak{D}f(0) = f'(0) = \lim_{h \rightarrow 0} \frac{f(h) - a_0}{h} = a_1.$$

Für $x \neq 0$ ist dagegen die erste Ableitung (also auch das erste Derivat) gar nicht vorhanden, so daß die höheren Ableitungen auch für $x = 0$ nicht existieren. Wohl aber existieren die Derivate bis zum n^{ten} einschließlich, und zwar ist

$$\mathfrak{D}^r f(0) = a_r \quad (r = 1, 2, \dots, n).$$

Die Theorie der Derivate liefert nun folgendes Kriterium für Maxima und Minima, welches nach dem Bewiesenen allgemeiner ist als das gewöhnliche, in dem Ableitungen statt Derivate vorkommen:

Satz 2. Sei $f(x)$ eine an einer Stelle x_0 stetige reelle Funktion. Wenn dann $\mathfrak{D}f(x_0) = 0$ und wenn n die kleinste Zahl ist, für die $\mathfrak{D}^n f(x_0) \neq 0$, so hat $f(x)$ an der Stelle x_0 bei ungeradem n kein Extremum, bei geradem n aber

ein Minimum, falls $\mathfrak{D}^n f(x_0) > 0$,

ein Maximum, falls $\mathfrak{D}^n f(x_0) < 0$.

Nach der Definition des n^{ten} Derivates ist nämlich

$$\frac{f(x_0 + h) - f(x_0) - h \mathfrak{D} f(x_0) - \cdots - \frac{h^{n-1}}{(n-1)!} \mathfrak{D}^{n-1} f(x_0)}{h^n} \\ = \frac{1}{n!} [\mathfrak{D}^n f(x_0) + \varepsilon_h],$$

wo $\lim_{h \rightarrow 0} \varepsilon_h = 0$, also nach den Voraussetzungen des Satzes:

$$f(x_0 + h) - f(x_0) = \frac{h^n}{n!} [\mathfrak{D}^n f(x_0) + \varepsilon_h],$$

woraus der Satz unmittelbar folgt.

§ 3. Die Derivate einer Funktion von mehreren Variabeln.

Sei $f(x_1, \dots, x_m)$ eine reelle Funktion von m reellen Variablen. Wenn dann $x_1, \dots, x_m$ ein festes Wertsystem und wenn $x_1 + h_1, \dots, x_m + h_m$ ein variables Nachbarsystem ist, so sagt man bekanntlich, die Funktion hat an der Stelle $x_1, \dots, x_m$ ein vollständiges Differential, wenn

$$(4) \quad \begin{cases} f(x_1 + h_1, \dots, x_m + h_m) - f(x_1, \dots, x_m) \\ = a_1 h_1 + \cdots + a_m h_m + (|h_1| + \cdots + |h_m|) \varphi(h_1, \dots, h_m) \end{cases}$$

ist, wobei

$$(5) \quad \varphi(h_1, \dots, h_m) \rightarrow 0 \text{ für } |h_1| + \cdots + |h_m| \rightarrow 0.$$

Die Zahlen $a_1, \dots, a_m$ sind durch diese Forderung eindeutig bestimmt. Wir nennen sie die ersten partiellen Derivate der Funktion:

$$a_\mu = \mathfrak{D}_{x_\mu} f(x_1, \dots, x_m) \quad (\mu = 1, 2, \dots, m).$$

Dann können wir die Gleichung (4) symbolisch so schreiben:

$$(6) \quad \begin{cases} f(x_1 + h_1, \dots, x_m + h_m) - f(x_1, \dots, x_m) \\ = (h_1 \mathfrak{D}_{x_1} + \cdots + h_m \mathfrak{D}_{x_m}) f(x_1, \dots, x_m) \\ + (|h_1| + \cdots + |h_m|) \varphi(h_1, \dots, h_m). \end{cases}$$

Nach dieser Definition zieht die Existenz der ersten Derivate offenbar die Stetigkeit der Funktion für das betreffende Wertsystem nach sich. Wenn die ersten partiellen Ableitungen f'_{x_μ} vorhanden und stetig sind, so ist bekanntlich $a_\mu = f'_{x_\mu}$, also auch

$$(7) \quad \mathfrak{D}_{x_\mu} f = f'_{x_\mu} \quad (\mu = 1, 2, \dots, m).$$

Wenn dagegen die ersten partiellen Ableitungen $f''_{x_{\mu}}$ zwar vorhanden, aber nicht sämtlich stetig sind, braucht ein vollständiges Differential nicht zu existieren, also brauchen auch die ersten Derivate nicht zu existieren¹⁾. Andererseits, wenn die ersten Derivate existieren, wenn es also ein vollständiges Differential gibt, so existieren stets auch die ersten Ableitungen und es gelten die Gleichungen (7); die Ableitungen brauchen aber nicht stetig zu sein²⁾.

Die n^{ten} partiellen Derivate

$$\mathfrak{D}_{x_1 \lambda_1 \dots x_m \lambda_m}^n f(x_1, \dots, x_m) \quad (\lambda_1 + \lambda_2 + \dots + \lambda_m = n)$$

werden, wenn die 1^{ten} , 2^{ten} , $\dots$, $(n-1)^{\text{ten}}$ bereits definiert sind, durch die symbolische Formel definiert:

$$(8) \quad \left\{ \begin{aligned} & f(x_1 + h_1, \dots, x_m + h_m) - f(x_1, \dots, x_m) = (h_1 \mathfrak{D}_{x_1} + \dots \\ & + h_m \mathfrak{D}_{x_m}) f + \dots + \frac{1}{n!} (h_1 \mathfrak{D}_{x_1} + \dots + h_m \mathfrak{D}_{x_m})^n f \\ & + ([h_1] + \dots + [h_m])^n q_n(h_1, \dots, h_m), \end{aligned} \right.$$

wobei wieder

$$(9) \quad q_n(h_1, \dots, h_m) \rightarrow 0 \text{ für } |h_1| + \dots + |h_m| \rightarrow 0.$$

Man sieht leicht, daß sie durch diese Formel eindeutig bestimmt sind. Speziell für $m = 1$ läuft die Formel, wenn man sie durch h_1^n dividiert, genau auf die Definitionsformel (2) hinaus.

Wenn für ein Wertsystem $x_1, \dots, x_m$ alle n^{ten} partiellen Ableitungen existieren und stetig sind, so existieren auch die n^{ten} Derivate und sie sind den entsprechenden Ableitungen gleich. Wenn aber die Ableitungen existieren und nicht alle stetig sind, brauchen die Derivate nicht zu existieren; der Derivatbegriff bleibt also hinter dem Ableitungsbegriff in dieser Richtung etwas zu-

¹⁾ Beispiel:

$$f(x, y) = \begin{cases} \frac{x y^2}{x^2 + y^2} & \text{für } x^2 + y^2 > 0 \\ 0 & \text{für } x = y = 0. \end{cases}$$

²⁾ Beispiel:

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

rück; das gilt sogar schon für die ersten Derivate. In anderer Richtung bleibt aber auch der Ableitungsbegriff hinter dem Derivatbegriff etwas zurück. Denn wenn für ein Wertsystem $x_1, \dots, x_n$ die n^{ten} Derivate existieren, so brauchen die r^{ten} Derivate für $r < n$ ebenfalls nur für das betreffende Wertsystem zu existieren, in welchem Fall die n^{ten} Ableitungen nicht vorhanden sind. Als Beispiel betrachten wir die Funktion

$$f(x_1, \dots, x_m) = (x_1 + \dots + x_m)^{n+1} w(x_1 + \dots + x_m),$$

wo w eine stetige, aber nirgends differenzierbare Funktion ist. Hier sind für das Wertsystem $x_1 = \dots = x_m = 0$ alle Derivate bis zu den n^{ten} einschließlich vorhanden und gleich 0, während die zweiten und höheren Ableitungen gar nicht existieren.

Die Derivate liefern nun sofort das folgende Kriterium für Maxima und Minima, welches dem für Funktionen einer Variablen völlig analog ist.

Satz 3. Sei $f(x_1, \dots, x_m)$ eine für das Wertsystem $x_1^0, \dots, x_m^0$ stetige reelle Funktion. Wenn dann die ersten Derivate $\mathfrak{D}_{x_\mu} f(x_1^0, \dots, x_m^0) \quad (\mu = 1, 2, \dots, m)$

sämtlich verschwinden und wenn n die kleinste Zahl ist derart, daß die n^{ten} Derivate nicht sämtlich verschwinden, so hat die Funktion für das Wertsystem $x_1^0, \dots, x_m^0$ ein Minimum (Maximum), wenn die symbolische Potenz

$$(\mathfrak{D}_{x_1} X_1 + \dots + \mathfrak{D}_{x_m} X_m)^n f(x_1^0, \dots, x_m^0)$$

eine positiv (negativ) definite Form ist. Dagegen liegt kein Extremum vor, wenn die symbolische Potenz eine indefinite Form ist, also insbesondere, wenn n ungerade ist.