

Det Kgl. Danske Videnskabernes Selskab.  
Mathematisk-fysiske Meddelelser. **X**, 6.

---

ÜBER  
MONOTONE FUNKTIONEN

VON

BØRGE JESSEN



KØBENHAVN

HOVEDKOMMISSIONÆR: ANDR. FRED. HØST & SØN, KGL. HOF-BOGHANDEL  
BIANCO LUNOS BOGTRYKKERI

1929



In der Theorie der reellen Funktionen einer Veränderlichen spielen die monotonen Funktionen eine wichtige Rolle. Durch stetige, stets wachsende (oder stets abnehmende) Funktionen werden die allgemeinsten eineindeutigen und stetigen Abbildungen von Intervallen auf Intervalle vermittelt; die allgemeinste Funktion von beschränkter Variation wird als Differenz einer monoton wachsenden und einer monoton abnehmenden Funktion dargestellt. Die entscheidenden Eigenschaften der Funktionen von beschränkter Variation: die Differentiierbarkeit fast überall, sowie die Zerlegung in einen totalstetigen Teil (das unbestimmte Integral der Ableitung) und einen Teil von konstanter  $\lambda$ -Variation, sind unmittelbare Folgerungen der entsprechenden Tatsachen für monotone Funktionen. Im folgenden sollen nun diese Ergebnisse von Neuem mit Hilfe eines einfachen Kunstgriffs hergeleitet werden. Hierbei sind die Grundlagen die »Vorlesungen über reelle Funktionen« (Leipzig und Berlin, 1918 (2. Aufl. 1927)) des Herrn CARATHÉODORY, auf die sich die Paragraphenhinweise beziehen. Während in der üblichen Theorie (§§ 500—513) die Zerlegung einer monotonen Funktion in eine totalstetige Funktion und eine Funktion von konstanter  $\lambda$ -Variation direkt vorgenommen wird, und dann die übrigen Eigenschaften (die Differentiierbarkeit und die Summierbarkeit der Ableitung) aus den Eigenschaften dieser Funktionen gefolgert werden, kommen

wir im folgenden allein mit Betrachtungen über totalstetige Funktionen aus (deren Theorie als bekannt vorausgesetzt wird). Ausser den schon erwähnten Sätzen ergibt sich u. a.: Die zu einer monotonen Funktion von konstanter  $\lambda$ -Variation inverse Funktion ist immer wieder von konstanter  $\lambda$ -Variation (vgl. § 512).

Es sei  $f(x)$  eine im (endlichen) Intervalle

$$a \leq x \leq b$$

monoton wachsende Funktion, d. h. eine endliche (aber nicht notwendig stetige) Funktion für die stets

$$f(x_1) \leq f(x_2) \quad \text{wenn} \quad x_1 < x_2.$$

Wir wollen zunächst die Differentiierbarkeit dieser Funktion fast überall, d. h. für alle Werte von  $x$  bis auf die Werte einer Nullmenge, beweisen. Hierzu betrachten wir (Fig. 1) den Graph der Funktion, d. h. die ebene Kurve  $G$  die

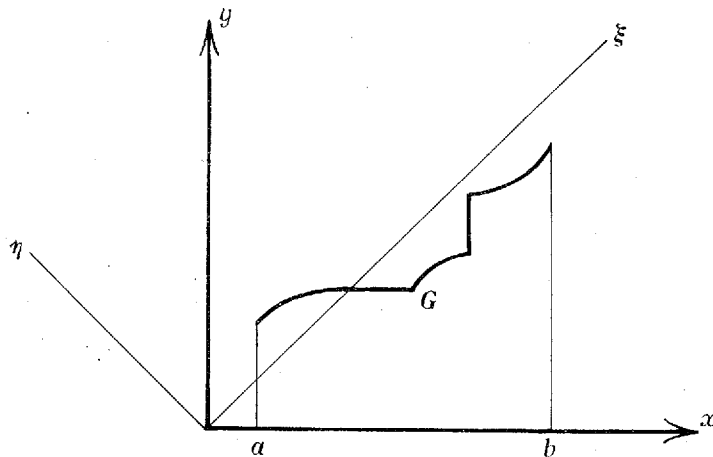


Fig. 1.

aus allen Punkten  $(x, y) = (x, f(x))$ ,  $a \leq x \leq b$ , besteht und die ausserdem noch, wenn  $f(x)$  unstetig ist, für jede Unstetigkeitsstelle  $x_0$  das Geradenstück  $x = x_0$ ,  $f(x_0 - 0) \leq y \leq f(x_0 + 0)$  enthält. Die Aufgabe,  $f(x)$  zu differenzieren, ist mit der Aufgabe äquivalent, Tangenten an  $G$  zu legen. Führen wir auf  $G$  als Parameter die Grösse  $t = x + y$  ein, so wollen wir zeigen, dass  $G$  in fast allen Punkten (d. h. für fast alle Werte von  $t$ ) eine bestimmte Tangente hat. Hiermit sind wir fertig, denn eine Nullmenge auf  $G$  projiziert sich offenbar auf die  $x$ -Achse wieder in eine Nullmenge. Dass  $G$  fast überall eine bestimmte Tangente hat, ist aber eine unmittelbare Folge der Tatsache, dass wenn wir die Gerade  $x = y$  als Achse einer neuen Variablen  $\xi$  einführen, die Kurve  $G$  das Bild einer in  $\xi$  totalstetigen Funktion wird, nämlich das Bild einer Funktion, deren Differenzenquotienten zwischen den Grenzen  $-1$  und  $+1$  liegen. — Aus der angewendeten Betrachtung folgt (vgl. § 478), dass noch für monotone Funktionen alle Sätze über Limesfunktionen der Derivierten wörtlich wie für totalstetige Funktionen gelten.

Fassen wir, was auch im folgenden bequem sein wird, die Funktion  $f(x)$  als mehrdeutig auf, indem wir ihr in jeder Unstetigkeitsstelle  $x_0$  alle Werte des Intervalles  $f(x - 0) \leq y \leq f(x_0 + 0)$  zuschreiben, so können wir die neue unabhängige Variable  $\xi$  in der Form schreiben

$$\xi = \frac{\sqrt{2}}{2}(x + f(x));$$

die neue abhängige Variable ist

$$\eta = \frac{\sqrt{2}}{2}(-x + f(x)).$$

Die Auffassung von  $G$  als Bild einer Funktion  $\eta = \eta(\xi)$  ist also mit der Betrachtung von  $-x + f(x)$  als Funktion von  $x + f(x)$  gleichbedeutend. Für die folgenden Rechnungen wird es bequemer sein, statt  $-x + f(x)$  nur  $f(x)$  selber als Funktion von  $x + f(x)$  zu betrachten. Indem wir diese Transformation einführen, leiten wir zunächst wieder die Differentierbarkeit ab und zeigen dann die Summierbarkeit der Ableitung.

Gleichzeitig mit  $f(x)$  ist auch

$$t = x + f(x)$$

eine monotone Funktion von  $x$ ; sie transformiert das Intervall  $a \leq x \leq b$  in das Intervall  $t_0 \leq t \leq T$ , wobei  $t_0 = a + f(a)$ ,  $T = b + f(b)$  ist. Wir betrachten (Fig. 2) die Funktion

$$F(t) = f(x),$$

d. h. die Funktion  $f(x)$  als Funktion der neuen unabhängigen Variablen  $x + f(x)$  aufgefasst. Sie ist stetig und mono-

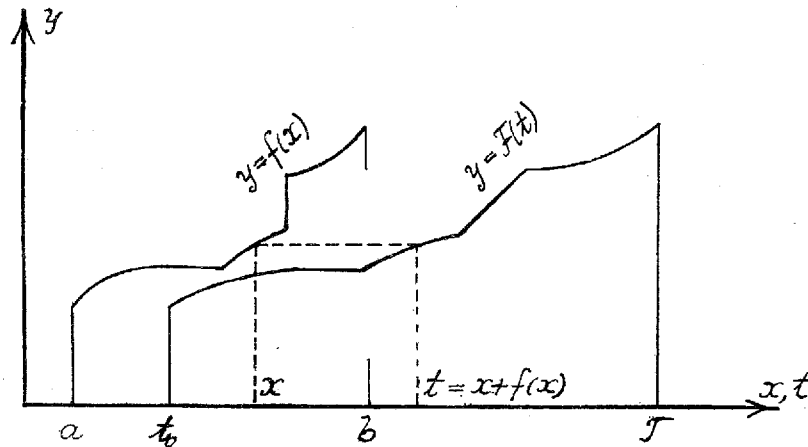


Fig. 2.

ton wachsend im Intervalle  $t_0 \leq t \leq T$ . Für beliebige Werte  $t_1 = x_1 + f(x_1)$  und  $t_2 = x_2 + f(x_2)$  ( $\neq t_1$ ) ist

$$(1) \quad \frac{F(t_2) - F(t_1)}{t_2 - t_1} = \frac{f(x_2) - f(x_1)}{x_2 - x_1 + f(x_2) - f(x_1)} = \frac{1}{\frac{1}{\frac{f(x_2) - f(x_1)}{x_2 - x_1}} + 1}$$

also

$$0 \leq \frac{F(t_2) - F(t_1)}{t_2 - t_1} \leq 1.$$

Die Funktion  $F(t)$  hat also beschränkte Differenzenquotienten; sie ist totalstetig. Dasselbe gilt für

$$\varphi(t) = t - F(t),$$

deren Differenzenquotienten ebenfalls zwischen 0 und 1 liegen. Die letzte Funktion transformiert das Intervall  $t_0 \leq t \leq T$  in das Intervall  $a \leq x \leq b$  und man hat

$$f(\varphi(t)) = F(t).$$

Bei der Abbildung  $x = \varphi(t)$  geht ein beliebiges Teilintervall  $x_1 + f(x_1) \leq t \leq x_2 + f(x_2)$  von  $t_0 \leq t \leq T$  in ein nicht grösseres Intervall, nämlich in das Intervall  $x_1 \leq x \leq x_2$  über; hieraus folgt, dass auch jede beliebige messbare Teilmenge von  $t_0 \leq t \leq T$  bei der Abbildung in eine messbare Teilmenge von  $a \leq x \leq b$  übergehen muss und zwar in eine Menge mit höchstens demselben Inhalt; speziell gehen also Nullmengen in Nullmengen über.<sup>1</sup>

<sup>1</sup> Analog kann man, wie LEBESGUE in der neuen Ausgabe der »Leçons sur l'intégration« (Paris, 1928), eine beliebige Funktion  $f(x)$  von beschränkter Variation als Funktion der neuen Variablen  $x + V(x)$  betrachten, wobei  $V(x)$  die totale Variation von  $f(x)$  im Intervalle  $a \leq \xi \leq x$  bedeutet, und hierdurch u. a. zu einer sehr übersichtlichen Theorie der

Wenden wir auf eine beliebige monoton wachsende totalstetige Funktion  $F(t)$ , deren Differenzenquotienten zwischen 0 und 1 liegen, die Transformation

$$(2) \quad x = t - F(t), \quad f(x) = F(t)$$

an, so bekommen wir eine monotone Funktion  $f(x)$ . Der gewonnene Zusammenhang zwischen monotonen Funktionen  $f(x)$  und totalstetigen Funktionen  $F(t)$  ist also ein gegenseitiger. Statt (2) können wir auch schreiben

$$(3) \quad x = a + \int_{t_0}^t (1 - \Phi(t)) dt, \quad f(x) = f(a) + \int_{t_0}^t \Phi(t) dt,$$

wobei  $\Phi(t)$  eine summierbare Funktion zwischen 0 und 1 bedeutet. (3) gibt eine Parameterdarstellung der allgemeinsten monoton wachsenden Funktion.

Aus (1) geht unmittelbar hervor:

$$(4) \quad DF(t) = \frac{1}{\frac{1}{Df(x)} + 1} \quad \text{oder} \quad Df(x) = \frac{DF(t)}{1 - DF(t)}$$

für zusammengehörige Derivierte von  $f(x)$  und  $F(t)$ . Nun ist  $DF(t)$  eindeutig bestimmt (gleich  $\Phi(t)$ ) für fast alle Werte von  $t$ ; es ist demnach  $Df(x)$  eindeutig bestimmt für fast alle Werte von  $x$ :

STIELTJÉSintegrale gelangen. Der Ausgangspunkt des Verfassers war übrigens von dem LEBESGUE'schen ganz verschieden. Bedeutet  $f(x)$  die Totalkrümmung eines konvexen Bogens, dessen Länge von einem festen Punkt gerechnet mit  $x$  bezeichnet wird, so bedeutet  $x + f(x)$  die Länge des äusseren Parallelbogens im Abstand 1. Die Krümmung des Bogens ist die Ableitung der Funktion. Durch Übergang zum äusseren Parallelbogen erreicht man aber eben, mit einem Bogen beschränkter Krümmung rechnen zu können.



Satz 1. Jede monotone Funktion ist differenzierbar in einem massgleichen Kern ihres Definitionsbereiches.

Mit  $Df(x)$  und  $DF(t)$  bezeichnen wir zwei zusammengehörige Derivierte von  $f(x)$  und  $F(t)$  d. h. zwei eindeutige, durch die Relationen (4) verbundene, Funktionen von  $x$  bzw.  $t$  die in jedem Punkt eine der Derivierten von  $f(x)$  bzw.  $F(t)$  darstellen. Es soll gezeigt werden, dass  $Df(x)$  über  $a \leq x \leq b$  summierbar ist. Hierzu sei zuerst bemerkt, dass  $Df(x)$  infolge seiner Darstellung (4) jedenfalls eine messbare Funktion sein muss; denn für jede beliebige Zahl  $A$  geht die Menge  $M\{Df(x) < A\}$  aus der Menge  $M\left\{DF(t) < \frac{1}{\frac{1}{A} + 1}\right\}$  durch die Transformation  $x = t - F(t)$

hervor. Für jede natürliche Zahl  $n$  ist also die abgeschnittene Funktion

$$[Df(x)]_n = \begin{cases} Df(x) & \text{wenn } Df(x) < n \\ n & \text{wenn } Df(x) \geq n \end{cases}$$

notwendig summierbar. Substituieren wir aber (§ 497) in das Integral dieser Funktion für  $x$  die monotone totalstetige Funktion  $\varphi(t) = t - F(t)$ , so kommt

$$\int_a^b [Df(x)]_n dx = \int_{t_0}^T \left[ \frac{DF(t)}{1 - DF(t)} \right]_n (1 - DF(t)) dt \leq \int_{t_0}^T DF(x) dt$$

und aus der Beschränktheit von  $\int_a^b [Df(x)]_n dx$  folgt die Summierbarkeit von  $Df(x)$ . Hieraus folgt speziell

Satz 2. Eine monotone Funktion hat höchstens in einer Nullmenge eine unendliche Derivierte.

Es bedeute  $\dot{f}(x)$  die Ableitung von  $f(x)$ , d. h. die eindeutige Funktion von  $x$ , die in allen Punkten, wo  $f(x)$  differenzierbar ist und eine endliche Derivierte  $Df(x)$  besitzt, gleich dieser Derivierten ist, aber sonst gleich Null ist. Diese Funktion ist mit allen Derivierten von  $f(x)$  äquivalent. Mit  $\dot{F}(t)$  bezeichnen wir die Ableitung von  $F(t)$ . Wir setzen

$$\dot{F}(t) = \dot{F}_0(t) + \dot{F}_1(t)$$

wobei

$$\dot{F}_0(t) = \begin{cases} 0 & \text{wenn } \dot{F}(t) = 1 \\ \dot{F}(t) & \text{wenn } \dot{F}(t) < 1 \end{cases}; \quad \dot{F}_1(t) = \begin{cases} 1 & \text{wenn } \dot{F}(t) = 1 \\ 0 & \text{wenn } \dot{F}(t) < 1. \end{cases}$$

Dann ist

$$\dot{f}(x) = \frac{\dot{F}_0(t)}{1 - \dot{F}_0(t)}.$$

Für beliebige Werte von  $x$  im Intervalle  $a \leq x \leq b$  ist also (§ 499, Satz 4) mit  $t = x + f(x)$

$$\begin{aligned} \int_a^x \dot{f}(x) dx &= \int_{t_0}^t \frac{\dot{F}_0(t)}{1 - \dot{F}_0(t)} (1 - \dot{F}(t)) dt \\ &= \int_{t_0}^t \dot{F}_0(t) dt = \int_{t_0}^t \dot{F}(t) dt - \int_{t_0}^t \dot{F}_1(t) dt \\ &= f(x) - f(a) - \int_{t_0}^t \dot{F}_1(t) dt \end{aligned}$$

und demnach

$$(5) \quad f(x) = f(a) + \int_a^x \dot{f}(x) dx + \int_{t_0}^t \dot{F}_1(t) dt.$$

Diese Formel gibt eine Zerlegung von  $f(x)$  in zwei monotone Funktionen

$$f(x) = \tau(x) + \nu(x),$$

wo

$$\nu(x) = f(a) + \int_a^x f(x) dx$$

totalstetig ist und fast überall dieselbe Ableitung wie  $f(x)$  hat, während

$$\nu(x) = \int_{t_0}^x \dot{F}_1(t) dt$$

fast überall die Ableitung Null hat. Diese letzte Funktion wollen wir nun näher untersuchen.

Die Funktion  $\dot{F}_1(t)$  nimmt nur die Werte 0 und 1 an. Die Menge  $M\{\dot{F}_1(t) = 1\}$  hat den Inhalt  $\nu(b) - \nu(a)$  ( $= \nu(b)$ ) und wird durch die Funktion  $x = \varphi(t)$  in die Menge der Punkte von  $a \leq x \leq b$  transformiert, für die  $f(x)$  mit unendlicher Derivierten differentierbar ist, also in eine Menge mit dem Inhalt Null. Diese Menge bezeichnen wir mit  $A$ . Es sei  $U = \delta_1 + \delta_2 + \dots$  eine beliebige Punktmenge, die aus offenen Intervallen

$$\delta_k: a_k < x < b_k,$$

ohne gemeinsame Punkte besteht und die  $A$  überdeckt. Indem wir (wenn es nötig ist)  $f(x) = f(a)$  für  $x < a$  und  $f(x) = f(b)$  für  $x > b$  setzen, entspricht der Menge  $U$  bei der Abbildung  $t = x + f(x)$  eine Menge von Intervallen

$$a_k + f(a_k) < t < b_k + f(b_k),$$

die die Menge  $M\{\dot{F}_1(t) = 1\}$  überdeckt. Es wird also

$$(6) \quad \nu(b) - \nu(a) = \int_{t_0}^T \dot{F}_1(t) dt = \sum_k (\nu(b_k) - \nu(a_k)).$$

Da wir den Inhalt von  $U$  beliebig klein wählen können, wird also die Nullvariation von  $\nu(x)$  im Intervalle  $a \leq x \leq b$  gleich der totalen Variation  $\nu(b) - \nu(a)$ : Die Funktion  $\nu(x)$  ist von konstanter  $\lambda$ -Variation (§ 501).

Satz 3. Jede im Intervalle  $a \leq x \leq b$  definierte monoton wachsende Funktion  $f(x)$  kann als Summe von zwei monoton wachsenden Funktionen

$$(7) \quad f(x) = \tau(x) + \nu(x)$$

angesehen werden, von denen die erste ein unbestimmtes Integral der Ableitung von  $f(x)$  ist und die zweite von konstanter  $\lambda$ -Variation ist; für die letzte gilt noch die Gleichung (6).

Dass diese Zerlegung in eine totalstetige Funktion und eine Funktion von konstanter  $\lambda$ -Variation nur in einer Weise möglich ist, wenn wir  $\nu(a) = 0$  verlangen, folgt daraus, dass bei jeder Zerlegung (7)  $\nu(x) - \nu(a)$  die Nullvariation von  $f(x)$  im Intervalle  $a \leq \xi \leq x$  bedeuten muss.

Aus der Darstellung

$$f(x) = f(a) + \int_{t_0}^x \dot{F}_0(t) dt + \int_{t_0}^x \dot{F}_1(t) dt$$

folgt, dass  $f(x)$  dann und nur dann totalstetig bzw. von konstanter  $\lambda$ -Variation ist, wenn  $\dot{F}_1(t)$  bzw.  $\dot{F}_0(t) \infty 0$  ist. Unsere Parameterdarstellung (3) einer beliebigen monoton wachsenden Funktion stellt also dann und nur dann eine totalstetige Funktion bzw. eine Funktion von konstanter  $\lambda$ -Variation dar, wenn fast überall in  $t_0 \leq t \leq T$

$$0 \leq \Phi(t) < 1$$

bzw.

$$(8) \quad \Phi(t) = 0 \text{ oder } 1$$

ist. Hieraus folgt sofort

Satz 4. Die inverse Funktion einer monotonen Funktion von konstanter  $\lambda$ -Variation ist immer wieder von konstanter  $\lambda$ -Variation; die inverse Funktion einer totalstetigen Funktion ist dann und nur dann wieder totalstetig, wenn die Ableitung der gegebenen Funktion fast überall von Null verschieden ist.

Es kann dies auch zu einer leichten Verschärfung der Abbildungssätze in § 512 benutzt werden.

Wir bemerken noch, dass man in der Definition der allgemeinsten monoton wachsenden Funktion von konstanter  $\lambda$ -Variation die Verwendung des Integralbegriffs (oder des gleichwertigen Begriffs des zweidimensionalen Inhalts) vermeiden kann. Wie aus (3) und (8) hervorgeht, hat man, wenn  $A$  und  $B$  zwei komplementäre messbare Teilmengen von  $t_0 \leqq t \leqq T$  bedeuten und wenn das Intervall von  $t_0$  bis  $t$  mit  $I_t$  bezeichnet wird, in

$$x = a + mAI_t, \quad f(x) = f(a) + mBI_t$$

eine Parameterdarstellung dieser Funktion. Hieraus folgt unmittelbar die Existenz sogar von stetigen und stets wachsenden Funktionen von konstanter  $\lambda$ -Variation.

