SUB | NIEDERSÄCHSISCHE STAATS- UND UNIVERSITÄTSBIBLIOTHEK GÖTTINGEN

Werk

Titel: Mathematische Annalen
Jahr: 1919
Kollektion: Mathematica
Digitalisiert: Niedersächsische Staats- und Universitätsbibliothek Göttingen
Werk Id: PPN235181684_0079
PURL: http://resolver.sub.uni-goettingen.de/purl?PPN235181684_0079

LOG Id: LOG_0039
LOG Titel: Über partielle und totale Differenzierbarkeit von Funktionen mehrerer Variabeln und über die Transformation der Doppelintegrale
LOG Typ: article

Übergeordnetes Werk

Werk Id: PPN235181684
PURL: http://resolver.sub.uni-goettingen.de/purl?PPN235181684
OPAC: http://opac.sub.uni-goettingen.de/DB=1/PPN?PPN=235181684

Terms and Conditions

The Goettingen State and University Library provides access to digitized documents strictly for noncommercial educational, research and private purposes and makes no warranty with regard to their use for other purposes. Some of our collections are protected by copyright. Publication and/or broadcast in any form (including electronic) requires prior written permission from the Goettingen State- and University Library.
Each copy of any part of this document must contain there Terms and Conditions. With the usage of the library's online system to access or download a digitized document you accept the Terms and Conditions.
Reproductions of material on the web site may not be made for or donated to other repositories, nor may be further reproduced without written permission from the Goettingen State- and University Library.

For reproduction requests and permissions, please contact us. If citing materials, please give proper attribution of the source.

Contact

Niedersächsische Staats- und Universitätsbibliothek Göttingen
Georg-August-Universität Göttingen
Platz der Göttinger Sieben 1
37073 Göttingen
Germany
Email: gdz@sub.uni-goettingen.de

Über partielle und totale Differenzierbarkeit von Funktionen mehrerer Variabeln und über die Transformation der Doppelintegrale.

Von

HANS RADEMACHER in Wickersdorf (Thüringen)

Ist $f(x)$ eine in einem Gebiet G der xy-Ebene definierte stetige Funktion mit *stetigen* partiellen Ableitungen, so ist diese Funktion bekanntlich im Stolzschen Sinne differenzierbar, d. h. es ist

$$f(x+h, y+k) = f(x, y) + h\frac{\partial f(x, y)}{\partial x} + k\frac{\partial f(x, y)}{\partial y} + \sqrt{h^2+k^2}\cdot R(h, k),$$

wenn $R(h, k) \leqq S(\sqrt{h^2+k^2})$ und $S(\sqrt{h^2+k^2}) \to 0$ mit $h^2+k^2 \to 0$.*)

Aus dieser Gleichung schließt man, daß eine durch zwei stetig partiell differenzierbare Funktionen

$$\xi = \varphi(x, y), \quad \eta = \psi(x, y)$$

vermittelte eineindeutige Abbildung von G auf ein Gebiet Γ der $\xi\eta$-Ebene im Unendlichkleinen affin ist. Durch Vergleich dieser Abbildung mit einer affinen folgert man daraus, daß der Betrag der Funktionaldeterminante

$$\left|\frac{D(\varphi, \psi)}{D(x, y)}\right|$$

in jedem Punkte (x, y) das Vergrößerungsverhältnis der Abbildung darstellt. Von hier aus gelangt man dann zu der Transformationsformel für Doppelintegrale

$$\iint_\Gamma F(\xi, \eta)\, d\xi\, d\eta = \iint_G F(\varphi(x, y), \psi(x, y))\left|\frac{D(\varphi, \psi)}{D(x, y)}\right| dx\, dy.$$

Diese Schlußweisen sollen nun in der vorliegenden Arbeit auf Funktionen übertragen werden, die nur einer erweiterten Lipschitzschen Bedingung

*) Bei jeder *totalstetigen* Funktion zweier Variabeln findet übrigens Differenzierbarkeit im Stolzschen Sinne wenigstens in einem maßgleichen Kern von G statt; siehe C. Carathéodory, Reelle Funktionen (Leipzig 1918), S. 661, Satz 7. Dieses Buch wird im Folgenden mit R. F. zitiert.

unterworfen sind. Dabei werden wir auf dem Wege noch einige an sich bemerkenswerte Sätze über die Funktionaldeterminante selbst beweisen. Anstelle von „Differenzierbarkeit im Stolzschen Sinne" werden wir im folgenden auch kurz von „totaler Differenzierbarkeit" sprechen. Wir bemerken noch, daß wir alle Sätze nur der Einfachheit halber für *zwei* Variable formulieren; die Beweise lassen sich ohne weiteres auch auf mehr Variable übertragen.

I. Partielle und totale Differenzierbarkeit.

1. In dem beschränkten Gebiete G der xy-Ebene sei die Funktion $f(x, y)$ definiert. Wir führen dann ein

$$(1)\quad w_f(x, y; \varrho) = \text{obere Grenze von } \frac{|f(x+h, y+k)|}{\sqrt{h^2+k^2}} \text{ für } 0 < h^2+k^2 \leqq \varrho^2.$$

Hiernach ist $w_f(x, y; \varrho)$ bei festem x, y eine mit abnehmendem ϱ nicht wachsende Funktion. Es existiert also

$$(2)\quad L_f(x, y) = \lim_{\varrho \to 0} w_f(x, y; \varrho).$$

Über $f(x, y)$ setzen wir nun voraus, daß $L_f(x, y)$ in G *endlich* und *summierbar* sei. Die Endlichkeit von L_f impliziert offenbar die Stetigkeit von f.*)

Die $f(x,y)$ auferlegte Bedingung kann als eine Erweiterung der Lipschitzschen Bedingung aufgefaßt werden. Denn diese wird ausgedrückt durch

$$(3)\quad \left|\frac{f(x+h, y)-f(x, y)}{h}\right| < M, \quad \left|\frac{f(x, y+k)-f(x, y)}{k}\right| < M \qquad h, k \neq 0,$$

wo M eine in G konstante endliche Zahl ist. Aus (3) folgt aber leicht

$$(4)\quad \frac{|f(x+h, y+k)-f(x, y)|}{\sqrt{h^2+k^2}} < 2M$$

und also auch

$$L_f(x, y) \leqq 2M,$$

so daß L_f dann endlich und in dem beschränkten Gebiet G summierbar ist. Für die Erfüllung von (3) ist bekanntlich wieder notwendig und hinreichend, daß die partiellen Derivierten von $f(x, y)$ beschränkt sind.**)

Wir setzen jedoch zunächst nur Endlichkeit und Summierbarkeit, nicht Beschränktheit von L_f voraus. Nun ist

*) Die *Meßbarkeit* von $L_f(x, y)$ ist eine Folge ihrer Endlichkeit, da L_f durch die Stetigkeit von f eine Funktion der zweiten Baireschen Klasse auf G wird.

**) R. F. S. 551, Satz 9.

$$\text{obere Grenze von } \left|\frac{f(x+h, y)-f(x, y)}{h}\right|$$

$$\leqq \text{obere Grenze von } \left|\frac{f(x+h, y+k)-f(x, y)}{\sqrt{h^2+k^2}}\right|$$

$$0 < h^2 + k^2 \leqq \varrho^2, \quad h \neq 0, \qquad \text{also auch}$$

$$|D_x f(x, y)| \leqq L_f(x, y), \tag{5a}$$

worin $D_x f(x, y)$ irgend eine der vier partiellen Hauptderivierten nach x bedeute. Ebenso gilt

$$|D_y f(x, y)| \leqq L_f(x, y). \tag{5b}$$

Die partiellen Hauptderivierten von $f(x, y)$ sind also in G endlich und summierbar. Ist dann E_{y_0} die lineare offene Punktmenge, die die Gerade $y = y_0$ mit G gemeinsam hat, so existiert nach einem bekannten Satz von G. Fubini*) außer für eine lineare Nullmenge der y_0 das Integral

$$\int_{E_{y_0}} D_x f(x, y_0)\, dx. \tag{6}$$

Dies hat aber zur Folge, daß in jedem E_{y_0}, in dem (6) existiert, auch außer höchstens in einer linearen Nullmenge die partielle Ableitung $\frac{\partial f(x, y)}{\partial x}$ existiert.**) Daraus ergibt sich, daß $\frac{\partial f(x, y)}{\partial x}$ in einem maßgleichen Kern K_1 von G existieren muß. Denn die vier partiellen Hauptderivierten nach x sind meßbare Funktionen und können daher nur in einer meßbaren Menge voneinander abweichen. Da diese Menge aber von allen Geraden $y = y_0$ mit Ausnahme einer linearen Nullmenge der y_0 in linearen Nullmengen geschnitten wird, so ist sie, wie eine Anwendung des Fubinischen Satzes zeigt, selbst eine zweidimensionale Nullmenge.***)

Nach analogen Überlegungen existiert $\frac{\partial f(x, y)}{\partial y}$ ebenfalls in einem maßgleichen Kern K_2 von G; beide gemeinsam also in dem maßgleichen Kern $K = K_1 \cdot K_2$.

2. Nun sei λ eine beliebige positive Zahl. Dann gibt es wegen der Meßbarkeit und Endlichkeit von L_f eine meßbare Teilmenge $H < K$, so daß

$$mH > mK - \lambda = mG - \lambda$$

ist und für jeden Punkt (x, y) in H

$$L_f(x, y) < M(\lambda) \tag{7}$$

ist, wo $M(\lambda)$ nicht von (x, y) abhängt.

*) G. Fubini, Sugli integrali multipli (Rend. della R. Accademia dei Lincei 1907), auch R. F. S. 632, Satz 4.

**) R. F. S. 597, Satz 4 zusammen mit S. 545, Satz 2.

***) R. F. S. 628, Satz 3.

Nach einem Konvergenzsatze aus der Lebesgueschen Theorie der Meßbarkeit*) gibt es nun in H eine meßbare Teilmenge S mit

$$mS > mH - \lambda > mG - 2\lambda,$$

so daß in S die *Konvergenz* von

$$w_f(x, y; \varrho) \to L_f, \quad \frac{f(x+h, y) - f(x, y)}{h} \to \frac{\partial f}{\partial x} \text{ und } \frac{f(x, y+h) - f(x, y)}{h} \to \frac{\partial f}{\partial y}$$

gleichmäßig stattfindet. Es gibt also ein $\varrho_0(\lambda)$, so daß für jedes (x, y) in S

$$w_f(x, y; \varrho) < L_f + M(\lambda) < 2M(\lambda) \quad \text{für} \quad 0 < \varrho \leqq \varrho_0(\lambda); \tag{8}$$

und es gibt zu jedem beliebig vorgeschriebenen positiven ε ein $h_0(\lambda, \varepsilon)$, so daß in S

$$\left.\begin{aligned} \left|\frac{f(x+h, y) - f(x, y)}{h} - \frac{\partial f(x, y)}{\partial x}\right| &< \varepsilon \\ \left|\frac{f(x, y+h) - f(x, y)}{h} - \frac{\partial f(x, y)}{\partial y}\right| &< \varepsilon \end{aligned}\right\} \quad \text{für} \quad 0 < |h| \leqq h_0(\lambda, \varepsilon). \tag{9}$$

In S als einer meßbaren Menge von positivem Maße gibt es nun eine *perfekte* Teilmenge T, deren Maß um weniger als λ unter dem von S bleibt**):

$$mT > mS - \lambda > mG - 3\lambda. \tag{10}$$

In T sind nun $\frac{\partial f}{\partial x}$ und $\frac{\partial f}{\partial y}$ *stetig*. Denn es sei $h_1, h_2, \cdots, h_n, \cdots$ eine gegen Null konvergierende Folge positiver Zahlen, so konvergieren die Folgen der in T *stetigen* Funktionen von (x, y)

$$\frac{f(x+h_n, y) - f(x, y)}{h_n}, \quad \frac{f(x, y+h_n) - f(x, y)}{h_n} \quad (n = 1, 2, 3, \cdots)$$

in T gleichmäßig gegen $\frac{\partial f}{\partial x}$ bzw. $\frac{\partial f}{\partial y}$, so daß also auch diese Limites auf T stetig sind. Auf der *perfekten* Menge T sind sie dann aber auch gleichmäßig stetig***), d. h. zu jeder Zahl $\varepsilon > 0$ gibt es ein $\bar{h}(\varepsilon)$, so daß

$$\left|\frac{\partial f(x_1, y_1)}{\partial x} - \frac{\partial f(x_2, y_2)}{\partial x}\right| < \varepsilon, \quad \left|\frac{\partial f(x_1, y_1)}{\partial y} - \frac{\partial f(x_2, y_2)}{\partial y}\right| < \varepsilon, \tag{11}$$

wenn nur $|x_1 - x_2| + |y_1 - y_2| < \bar{h}(\varepsilon)$. [Bemerkung zur Schreibweise: $\frac{\partial f(x_1, y_1)}{\partial x}$ soll bedeuten, daß zunächst $\frac{\partial f(x, y)}{\partial x}$ gebildet und darin dann $x = x_1$, $y = y_1$ gesetzt wird.]

Ferner hat T als meßbare Menge von positivem Maß fast überall die

*) R. F. S. 382, Satz 12, doch bedarf für unsere Anwendung dieser Satz samt seinem Beweis einer leicht anzubringenden Erweiterung für den Fall des Grenzüberganges in einem *stetigen* Parameter.

**) R. F. S. 288, Satz 12.

***) R. F. S. 205, Satz 2.

Dichte 1.*) Das soll heißen: Ist $K(P; r)$ der Kreis mit dem Radius r um den Punkt $P = (x, y)$ von T, und ist $T(P; r)$ der Durchschnitt von T und $K(P; r)$, so ist diejenige Teilmenge E von T, in der

$$\lim_{r \to 0} \frac{mT(P; r)}{mK(P; r)} = 1$$

gilt, ein maßgleicher Kern von T,

$$E < T, \quad mE = mT. \tag{12}$$

Ist daher $N > 1$ eine positive ganze Zahl, so gibt es zu dem Punkte $P_0 = (x_0, y_0)$ von E, in dem wir die Funktion auf ihre totale Differenzierbarkeit hin untersuchen wollen, eine positive Zahl $r_N(P_0)$, so daß für $r < r_N(P_0)$

$$\frac{mT(P_0; r)}{mK(P_0; r)} > 1 - \frac{1}{N^2}$$

stattfindet. Bedeutet dann $E(P_0; r)$ den Durchschnitt von E und $K(P_0; r)$, so muß wegen (12) bei $r < r_N(P_0)$ zugleich

$$\frac{mE(P_0; r)}{mK(P_0; r)} > 1 - \frac{1}{N^2} \tag{13}$$

sein.

Zu der beliebigen positiven Zahl λ haben wir die Teilmenge E von G konstruiert, so daß

$$mE > mG - 3\lambda \tag{14}$$

ist. Zu λ haben wir die positiven Zahlen $M(\lambda)$ und $\varrho_0(\lambda)$ angegeben und zu der beliebigen positiven Zahl ε ferner die Zahlen $h_0(\varepsilon, \lambda)$ und $\bar{h}(\varepsilon)$ bestimmt, welche Zahlen sämtlich für ganz E Gültigkeit besitzen. Zu der beliebigen natürlichen Zahl N haben wir endlich für den Punkt P_0 in E ein $r_N(P_0)$ angegeben. Dann sei r_0 eine positive Zahl, die

$$r_0 < h_0(\lambda, \varepsilon), \quad r_0 < \bar{h}(\varepsilon), \quad r_0 < r_N(P_0) \tag{15}$$

genügt. Um P_0 denken wir uns den Kreis $K(P_0; r_0)$ mit dem Radius r_0 geschlagen.

3. In $K(P_0; r_0)$ sei P_1 mit den Koordinaten

$$\begin{cases} x_1 = x_0 + h_1 \cos \alpha_1 \\ y_1 = y_0 + h_1 \sin \alpha_1 \end{cases} \tag{16}$$

irgend ein von P_0 verschiedener Punkt, für den nur noch

$$0 < h_1 < r_0 \frac{N-1}{N}. \tag{17a}$$

$$h_1 < \varrho_0(\lambda)(N-1) \tag{17b}$$

*) Lebesgue, Sur l'intégration des fonctions discontinues, Annales de l'Ecole Normale Supérieure 1910 (3), Bd. 27, S. 399—401. Auch R. F. S. 497, Satz 3. Die Dichte ist aufzufassen als eine verallgemeinerte Derivierte des unbestimmten Integrals derjenigen summierbaren Funktion, die auf T gleich 1 und auf $G - T$ gleich 0 ist.

vorausgesetzt wird. In dem Kreis um P_0 mit dem Radius $h_1 \frac{N}{N-1}$ muß wegen (13), (17a) und der dritten Ungleichung (15) das Maß der zu E komplementären Punktmenge weniger als $\frac{h_1^2 \pi}{(N-1)^2}$ betragen. Ein Kreis um P_1 mit dem Radius $\frac{h_1}{N-1}$, der wegen (16) und (17a) ganz in $K(P_0; r_0)$ liegt, muß also Punkte von E in seinem Innern enthalten; es sei P_2 ein solcher Punkt mit den Koordinaten x_2, y_2, so daß mithin

$$\sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} < \frac{h_1}{N-1} \tag{18}$$

ist (P_2 darf übrigens mit P_1 zusammenfallen).

Wir bilden nun den Differenzenquotienten

$$\frac{f(x_1, y_1) - f(x_0, y_0)}{h_1} = \frac{f(x_1, y_1) - f(x_2, y_2)}{h_1} + \frac{f(x_2, y_2) - f(x_0, y_2)}{h_1} + \frac{f(x_0, y_2) - f(x_0, y_0)}{h_1}. \tag{19}$$

Die Brüche der rechten Seite schätzen wir nacheinander ab. Ist P_2 mit P_1 identisch, so verschwindet der erste Bruch. Sonst ist wegen (1) und (18)

$$\frac{|f(x_1, y_1) - f(x_2, y_2)|}{\sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}} \leqq w_f\left(x_2, y_2; \frac{h_1}{N-1}\right),$$

und folglich unter Berücksichtigung von (8) und (17b)

$$\left|\frac{f(x_1, y_1) - f(x_2, y_2)}{h_1}\right| < 2M(\lambda) \cdot \frac{\sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}}{h_1} < \frac{2M(\lambda)}{N-1}, \tag{20}$$

was auch für zusammenfallende P_1 und P_2 gültig bleibt.

Führt man noch

$$x_2 = x_0 + h_2 \cos \alpha_2$$
$$y_2 = y_0 + h_2 \sin \alpha_2$$

ein, so kann man für den zweiten Bruch schreiben

$$\frac{f(x_2, y_2) - f(x_0, y_2)}{h_1} = \frac{f(x_2, y_2) - f(x_0, y_2)}{x_2 - x_0} \frac{h_2}{h_1} \cos \alpha_2. \tag{21}$$

Nun ist, da P_2 in E liegt und $|x_2 - x_0| < h_0(\lambda, \varepsilon)$ ist, wegen (9)

$$\left|\frac{f(x_2, y_2) - f(x_0, y_2)}{x_2 - x_0} - \frac{\partial f(x_0, y_2)}{\partial x}\right| < \varepsilon,$$

ferner auch wegen $|y_2 - y_0| < \bar{h}(\varepsilon)$ nach (11)

$$\left|\frac{\partial f(x_0, y_2)}{\partial x} - \frac{\partial f(x_0, y_0)}{\partial x}\right| < \varepsilon,$$

so daß zusammengenommen

$$\frac{f(x_2, y_2) - f(x_0, y_2)}{x_2 - x_0} = \frac{\partial f(x_0, y_0)}{\partial x} + \eta_1 \tag{22}$$

sich ergibt, wo $|\eta_1| < 2\varepsilon$ ist.

Ferner ist nach dem Satz über die Differenz zweier Seiten im Dreieck unter Berücksichtigung von (18)

$$|h_2 - h_1| < \frac{h_1}{N-1} \qquad \text{oder}$$

$$\frac{h_2}{h_1} = 1 + \eta_2, \quad |\eta_2| < \frac{1}{N-1}. \tag{23}$$

Endlich ist $\cos\alpha_2 - \cos\alpha_1 = 2\sin\frac{\alpha_1-\alpha_2}{2}\sin\frac{\alpha_1+\alpha_2}{2}$,

$$|\cos\alpha_2 - \cos\alpha_1| \leqq 2\left|\sin\frac{\alpha_1-\alpha_2}{2}\right|.$$

Nun ersieht man aus dem Dreieck $P_0P_1P_2$ ohne weiteres

$$\sin|\alpha_2 - \alpha_1| < \frac{h_1}{N-1} : h_1 = \frac{1}{N-1},$$

woraus $$|\cos\alpha_2 - \cos\alpha_1| < \frac{2}{N-1}$$ und somit

$$\cos\alpha_2 = \cos\alpha_1 + \eta_3, \quad |\eta_3| < \frac{2}{N-1} \tag{24}$$

folgt. Für (21) erhält man also mit Hilfe von (22), (23), (24)

$$\begin{aligned}\frac{f(x_2,y_2)-f(x_0,y_2)}{h_1} &= \left(\frac{\partial f(x_0,y_0)}{\partial x}+\eta_1\right)(1+\eta_2)(\cos\alpha_1+\eta_3) \\ &= \frac{\partial f(x_0,y_0)}{\partial x}\cos\alpha_1 + \Theta(\eta_1,\eta_2,\eta_3),\end{aligned} \tag{25}$$

wo $\Theta(\eta_1,\eta_2,\eta_3)$ mit η_1, η_2, η_3 gegen Null strebt. Es lassen sich aber η_1, η_2, η_3 beliebig klein machen, indem man nur ε hinreichend klein und N hinreichend groß bestimmt, was man durch hinreichend kleines r_0 und damit hinreichend kleines h_1 erreichen kann, also $\Theta \to 0$ mit $h_1 \to 0$.

Endlich schätzt man den dritten Bruch rechts in (19) nach demselben Verfahren ab, was noch einfacher gelingt als bei dem zweiten Bruch, weil man nicht die durch (11) ausgesprochene gleichmäßige Stetigkeit von $\frac{\partial f}{\partial y}$ auf E heranzuziehen braucht. Man erhält dann

$$\frac{f(x_0,y_2)-f(x_0,y_0)}{h_1} = \frac{\partial f(x_0,y_0)}{\partial y}\sin\alpha_1 + \Theta',$$

wo wieder Θ' mit h_1 gegen Null strebt. Zusammen mit (20) und (25) ergibt sich für (19), wenn man noch die Bezeichnungen (16) wieder aufnimmt,

$$\begin{aligned}\frac{f(x_0+h_1\cos\alpha_1, y_0+h_1\sin\alpha_1)-f(x_0,y_0)}{h_1} &= \frac{\partial f(x_0,y_0)}{\partial x}\cos\alpha_1 \\ &+ \frac{\partial f(x_0,y_0)}{\partial y}\sin\alpha_1 + R(h_1;\alpha_1),\end{aligned} \tag{26}$$

wo $\lim R(h_1; \alpha_1) = 0$ für $h_1 \to 0$, was, wie unsere Abschätzungen zeigen, gleichmäßig in α_1 gilt. Setzt man

$$h_1 \cos \alpha_1 = h, \quad h_1 \sin \alpha_1 = k,$$

so kann man für (26) schreiben

$$(27) \quad f(x_0+h, y_0+k) = f(x_0, y_0) + h\frac{\partial f(x_0, y_0)}{\partial x} + k\frac{\partial f(x_0, y_0)}{\partial y} + \sqrt{h^2+k^2}\,R(h, k).$$

$$R(h, k) < S(\sqrt{h^2+k^2}),$$

$$\text{wo } \lim S(\sqrt{h^2+k^2}) = 0 \quad \text{mit } \sqrt{h^2+k^2} \to 0.$$

Im Punkte $P_0 = (x_0, y_0)$ ist $f(x, y)$ also *total (nach Stolz) differenzierbar.*

Nun war P_0 ein beliebiger Punkt von E; in ganz E ist also $f(x, y)$ von derselben Eigenschaft. Wir lassen nun λ der Reihe nach die Werte $1, \frac{1}{2}, \frac{1}{3}, \cdots, \frac{1}{n}, \cdots$ durchlaufen. Zu jedem dieser Werte bestimmen wir eine Menge E, die für $\lambda = \frac{1}{n}$ mit E_n bezeichnet werden möge, so daß nach (14)

$$(28) \quad mE_n > mG - \frac{3}{n}$$

ist. Es sei E^* die Vereinigungsmenge aller E_n. Dann ist in E^* überall $f(x, y)$ total differenzierbar, da jeder Punkt von E^* zu mindestens einem E_n gehört, in dem diese Eigenschaft der Funktion nachgewiesen ist. Anderseits folgt aus (28) $mE^* = mG$,

so daß wir folgenden Satz aussprechen können:

Satz I. *Ist $f(x, y)$ eine in dem beschränkten Gebiete G der xy-Ebene definierte Funktion, für welche die aus*

$$(1) \quad w_f(x, y; \varrho) = \text{obere Grenze von } \frac{f(x+h, y+k) - f(x, y)}{\sqrt{h^2+k^2}} \text{ für } 0 < h^2+k^2 \leqq \varrho^2,$$

$$(2) \quad L_f(x, y) = \lim_{\varrho \to 0} w_f(x, y; \varrho)$$

entspringende Funktion $L_f(x, y)$ in G endlich und summierbar ist, so ist $f(x, y)$ in einem maßgleichen Kern von G total (im Stolzschen Sinne) differenzierbar.

4. Die Überlegungen im Anfang von § 1 haben dargetan, daß für die Endlichkeit und Summierbarkeit von $L_f(x, y)$ die Beschränktheit der partiellen Derivierten von $f(x, y)$ in G hinreichend ist. Wir haben somit das

Korollar: *Eine in einem Gebiete G der xy-Ebene definierte Funktion $f(x, y)$ mit beschränkten partiellen Derivierten ist in einem maßgleichen Kern von G total (im Stolzschen Sinne) differenzierbar.**)

*) Eine solche Funktion zweier Variabeln braucht nach einem Beispiel von C. Carathéodory (R. F. S. 655) nicht totalstetig zu sein. Vgl. S. 340 Anm. *).

Selbst wenn aber, was in diesem Korollar nicht einmal vorausgesetzt ist, die Funktion $f(x, y)$ *überall* in G partielle Ableitungen besitzt, die beschränkt sind, braucht sie darum doch nicht überall total differenzierbar zu sein. Das zeigt das folgende einfache Beispiel

$$f(x, y) = \frac{xy^2}{x^2+y^2} \quad \text{für} \quad x^2+y^2>0,$$
$$f(x, y) = 0 \qquad \text{für} \quad x^2+y^2=0.$$

Leichte Rechnungen zeigen, daß überall

$$\left|\frac{\partial f}{\partial x}\right| \leqq 2, \qquad \left|\frac{\partial f}{\partial y}\right| \leqq 2 \qquad \text{und daß}$$

$$\left|\frac{\partial f}{\partial x}\right|_{\substack{x_0=0\\y_0=0}} = 0, \qquad \left|\frac{\partial f}{\partial y}\right|_{\substack{x_0=0\\y_0=0}} = 0 \tag{29}$$

ist. Dagegen haben wir

$$f(h\cos\alpha,\ h\sin\alpha) = h\sin^2\alpha\cos\alpha$$

und folglich

$$\frac{df(h\cos\alpha,\ h\sin\alpha)}{dh} = \sin^2\alpha\cos\alpha,$$

was mit (29) zusammen zeigt, daß für $x_0=0$, $y_0=0$ in diesem Falle (26) nicht erfüllt ist.

II. Eigenschaften der Funktionaldeterminante.

5. Nach dem Beweis unseres Satzes I schlagen wir den in der Einleitung skizzierten Weg ein. Aus (27) erhält man zunächst durch die Substitution $h = at, \quad k = bt$

$$\frac{f(x_0+at,\ y_0+bt)-f(x_0,\ y_0)}{t} = a\frac{\partial f(x_0,\ y_0)}{\partial x} + b\frac{\partial f(x_0,\ y_0)}{\partial y} + \sqrt{a^2+b^2}\,R(at, bt). \tag{30}$$

Nun seien

$$x = x(s), \qquad y = y(s) \tag{31}$$

zwei differenzierbare Funktionen (deren Ableitungen übrigens nicht beschränkt zu sein brauchen), deren Werte x und y für ein gewisses Intervall $s_1 < s < s_2$ in das Definitionsgebiet G von $f(x, y)$ fallen mögen. Wir betrachten einen Punkt $x_0 = x(s_0)$, $y_0 = y(s_0)$, der zur Menge E^* gehört, falls es einen solchen gibt. In diesem Punkte seien $x'(s_0)$ und $y'(s_0)$ die Ableitungen der Funktionen (31). Dann gilt in einer linearen Umgebung um s_0:

$$x(s) = x_0 + (x'(s_0)+\delta_1)(s-s_0),$$
$$y(s) = y_0 + (y'(s_0)+\delta_2)(s-s_0),$$

worin $\delta_1 \to 0$, $\delta_2 \to 0$ mit $s \to s_0$. Für hinreichend kleines $(s-s_0)$ schließt man daher aus (30) auf

$$\frac{f(x(s), y(s)) - f(x_0, y_0)}{s - s_0} = (x'(s) + \delta_1)\frac{\partial f(x_0, y_0)}{\partial x} + (y'(s_0) + \delta_2)\frac{\partial f(x_0, y_0)}{\partial y}$$

$$+ \sqrt{(x'(s_0) + \delta_1)^2 + (y'(s_0) + \delta_2)^2}\, R((x' + \delta_1)(s - s_0), (y' + \delta_2)(s - s_0)),$$

woraus im Limes $s \to s_0$

$$\left.\frac{df(x(s), y(s))}{ds}\right|_{s = s_0} = x'(s_0)\frac{\partial f(x_0, y_0)}{\partial x} + y'(s_0)\frac{\partial f(x_0, y_0)}{\partial y} \tag{32}$$

hervorgeht. Diese Gleichung gilt für ganz beliebige differenzierbare Funktionen $x(s)$, $y(s)$ (deren Werte in G liegen) außer in einer durch die Funktion $f(x, y)$ völlig bestimmten Nullmenge $(G - E^*)$ der (x, y).*) Die Funktionen $x(s)$ und $y(s)$ könnten also, ohne den Gültigkeitsbereich von (32) zu ändern, noch einen Parameter enthalten. Es seien

$$x = x(\xi, \eta), \quad y = y(\xi, \eta)$$

zwei in einem Gebiet Γ der $\xi\eta$-Ebene definierte Funktionen, deren Funktionswerte x, y in G hineinfallen und die in Γ überall partielle Ableitungen nach ξ und η besitzen, über deren Beschränktheit wir nichts voraussetzen. Dann gelten analog zu (32)

$$\begin{cases} \dfrac{\partial f(x(\xi, \eta), y(\xi, \eta))}{\partial \xi} = \dfrac{\partial f(x, y)}{\partial x}\dfrac{\partial x(\xi, \eta)}{\partial \xi} + \dfrac{\partial f(x, y)}{\partial y}\dfrac{\partial y(\xi, \eta)}{\partial \xi}, \\ \dfrac{\partial f(x(\xi, \eta), y(\xi, \eta))}{\partial \eta} = \dfrac{\partial f(x, y)}{\partial x}\dfrac{\partial x(\xi, \eta)}{\partial \eta} + \dfrac{\partial f(x, y)}{\partial y}\dfrac{\partial x(\xi, \eta)}{\partial \eta} \end{cases} \tag{33}$$

außer in einer Nullmenge $G - E^*$ der (x, y). Ist $g(x, y)$ eine weitere Funktion in G, die dort gleichfalls den Voraussetzungen von Satz I genügt, so gelten (33) und die analogen Formeln für $g(x, y)$ zugleich in allen Punkten von G außer in einer durch $f(x, y)$ und $g(x, y)$ bestimmten Nullmenge, die Z heißen möge.

Aus (33) und den analogen Gleichungen für $g(x, y)$ erhält man den Multiplikationssatz der Funktionaldeterminanten

$$\begin{vmatrix} \dfrac{\partial f}{\partial \xi} & \dfrac{\partial f}{\partial \eta} \\ \dfrac{\partial g}{\partial \xi} & \dfrac{\partial g}{\partial \eta} \end{vmatrix} = \begin{vmatrix} \dfrac{\partial f}{\partial x}\dfrac{\partial x}{\partial \xi} + \dfrac{\partial f}{\partial y}\dfrac{\partial y}{\partial \xi} & \dfrac{\partial f}{\partial x}\dfrac{\partial x}{\partial \eta} + \dfrac{\partial f}{\partial y}\dfrac{\partial y}{\partial \eta} \\ \dfrac{\partial g}{\partial x}\dfrac{\partial x}{\partial \xi} + \dfrac{\partial g}{\partial y}\dfrac{\partial y}{\partial \xi} & \dfrac{\partial g}{\partial x}\dfrac{\partial x}{\partial \eta} + \dfrac{\partial g}{\partial y}\dfrac{\partial y}{\partial \eta} \end{vmatrix} = \begin{vmatrix} \dfrac{\partial f}{\partial x} & \dfrac{\partial f}{\partial y} \\ \dfrac{\partial g}{\partial x} & \dfrac{\partial g}{\partial y} \end{vmatrix} \cdot \begin{vmatrix} \dfrac{\partial x}{\partial \xi} & \dfrac{\partial x}{\partial \eta} \\ \dfrac{\partial y}{\partial \xi} & \dfrac{\partial y}{\partial \eta} \end{vmatrix}. \tag{34}$$

Satz II. *Sind $f(x, y)$ und $g(x, y)$ zwei in einem Gebiet G der xy-Ebene den Voraussetzungen von Satz I genügende Funktionen und $x(\xi, \eta)$, $y(\xi, \eta)$ zwei in einem Gebiet Γ der $\xi\eta$-Ebene definierte partiell differenzierbare Funktionen, deren Funktionswerte (x, y) in das Gebiet G fallen, so gilt die Multiplikationsformel (34) der Funktionaldeterminanten außer in*

*) Es wäre denkbar, daß diese Nullmenge die ganze Kurve (31) enthielte, so daß die Formel (32) für keinen einzigen Wert von s zu gelten brauchte.

einer Nullmenge Z der (x, y) in G, die nur von $f(x, y)$ und $g(x, y)$, nicht von der Wahl der Funktionen $x(\xi, \eta)$ und $y(\xi, \eta)$ abhängt.

Über das Maß der Punktmenge der (ξ, η) in Γ, in der (34) nicht gilt, werden wir unter gewissen Einschränkungen nachher noch eine Aussage machen können.

6. Durch die Funktionen

$$u = f(x, y), \quad v = g(x, y),$$

die wieder die im vorigen Absatz diskutierten Eigenschaften haben sollen, wird das Gebiet G der xy-Ebene auf einen Bereich U der uv-Ebene eindeutig abgebildet. Da diese Abbildung jedoch keine eindeutige Umkehrung zu besitzen braucht, so braucht U kein Gebiet zu sein.

Die Gleichung (27) liefert in der Umgebung eines Punktes P_0, in dem sowohl $f(x, y)$ als auch $g(x, y)$ total differenzierbar ist,

$$(35)\quad \begin{cases} u = f(x_0 + h, y_0 + k) = f(x_0, y_0) + h\dfrac{\partial f(x_0, y_0)}{\partial x} + k\dfrac{\partial f(x_0, y_0)}{\partial y} \\ \qquad\qquad + \sqrt{h^2 + k^2}\, R(h, k) \\ v = g(x_0 + h, y_0 + k) = g(x_0, y_0) + h\dfrac{\partial g(x_0, y_0)}{\partial x} + k\dfrac{\partial g(x_0, y_0)}{\partial y} \\ \qquad\qquad + \sqrt{h^2 + k^2}\, R'(h, k). \end{cases}$$

Die Abbildung ist also im Unendlichkleinen um P_0 affin. Würden die Restglieder in (35) fehlen, so würde durch die dann hervorgehende affine Abbildung

$$(36)\quad \begin{cases} \bar{u} = f_0 + h\dfrac{\partial f}{\partial x} + k\dfrac{\partial f}{\partial y}, \\ \bar{v} = g_0 + h\dfrac{\partial g}{\partial x} + k\dfrac{\partial g}{\partial y} \end{cases}$$

das Quadrat Q in G

$$(37)\qquad x_0 - l \leqq x \leqq x_0 + l, \qquad y_0 - l \leqq y \leqq y_0 + l$$

auf ein Parallelogramm $\bar{V}$ abgebildet werden.

Wir erledigen zunächst den Fall

$$(38)\qquad D = \begin{vmatrix} \dfrac{\partial f}{\partial x} & \dfrac{\partial f}{\partial y} \\ \dfrac{\partial g}{\partial x} & \dfrac{\partial g}{\partial y} \end{vmatrix} = 0.$$

Dann artet das Parallelogramm $\bar{V}$ in eine Strecke $\bar{A}$ aus, deren Mittelpunkt das Bild von P_0 ist; verschwinden jedoch sämtliche Ableitungen, so wird Q durch die affine Transformation nur auf einen Punkt abgebildet. Die Länge a der Bildstrecke $\bar{A}$ ist der größere der beiden Abstände der Bilder gegenüberliegender Ecken von Q. Diese Bildabstände sind bzw.

$$\cdot\ 2l\sqrt{\left(\frac{\partial f}{\partial x}+\frac{\partial f}{\partial y}\right)^2+\left(\frac{\partial g}{\partial x}+\frac{\partial g}{\partial y}\right)^2} \quad \text{und} \quad 2l\sqrt{\left(\frac{\partial f}{\partial x}-\frac{\partial f}{\partial y}\right)^2+\left(\frac{\partial g}{\partial x}-\frac{\partial g}{\partial y}\right)^2},$$

und wegen (5a) und (5b) ist folglich

$$a \leqq 2l\sqrt{8L_f(x_0, y_0)^2} < 6lL_f, \tag{39}$$

worunter auch der Fall, daß $\overline{V}$ in einen Punkt ausartet, mit einbegriffen ist.

Der Vergleich von (35) und (36) zeigt, daß zwei zu demselben Wertepaar (h, k) gehörende Punkte (u, v) und $(\bar{u}, \bar{v})$ einen kleineren Abstand als $\sqrt{h^2+k^2}\,R^*\sqrt{2}$ voneinander haben, wo R^* der größere der beiden absoluten Beträge $|R|$ und $|R'|$ ist und daher $R^* \longrightarrow 0$ mit $h^2+k^2 \longrightarrow 0$. Also liegt das durch (35) von Q entworfene Bild V ganz in der Menge der Punkte, die von $\bar{A}$ höchstens den Abstand $2lR_l^*$ haben, wo R_l^* der Maximalwert von R^* in dem Quadrat Q und den Seitenlängen $2l$ sei, so daß also auch $R_l^* \longrightarrow 0$ mit $l \longrightarrow 0$ gilt. Diese Punktmenge (sie besteht aus einem Rechteck mit zwei angesetzten Halbkreisen) hat aber wegen (39) einen kleineren Inhalt als

$$24l^2L_fR_l^* + 4l^2R_l^{*2}\pi,$$

also ist auch das Maß von V kleiner als diese Zahl. Folglich

$$\frac{mV}{mQ} < (6L_f + \pi R_l^*)R_l^*,$$

woraus für $\operatorname{Lim} l = 0$ das *Vergrößerungsverhältnis* $\Delta(P_0)$ der Abbildung im Punkte P_0 folgt: $\Delta(P_0) = 0$.

Im Falle (38) ist also in einem Punkt von totaler Differenzierbarkeit der f und g das Vergrößerungsverhältnis gleich dem absoluten Betrag der Funktionaldeterminante.

Nehmen wir nun

$$D = \begin{vmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \\ \frac{\partial g}{\partial x} & \frac{\partial g}{\partial y} \end{vmatrix} \neq 0 \tag{40}$$

an, so ist $\overline{V}$ ein Parallelogramm mit nicht zusammenfallenden Seiten und mit dem Bild von P_0 als Mittelpunkt. Die Parallelogrammseiten werden einen kleinsten mit l proportionalen Abstand μl vom Mittelpunkte haben. Ich wähle nun l von vornherein so klein, daß $4R_l^* < \mu$ ist. Das durch (35) vom Rande des Quadrates Q entworfene Bild sei die Kurve C. Da jeder Punkt von C vom Rande von $\overline{V}$ höchstens einen Abstand $d = 2lR_l^*$ haben kann, so liegt C ganz in einem ringartigen Streifen, der die Breite $2d$ hat und den Mittelpunkt von $\overline{V}$ umschließt, aber nicht enthält. Ein Parallelogramm $\overline{V}_{-d}$ also, das ganz in $\overline{V}$ liegt und dessen Seiten parallel zu denen von $\overline{V}$ im Abstand d verlaufen, enthält keinen Punkt von C. Ein Parallelogramm $\overline{V}_{+d}$, dessen Seiten außerhalb von $\overline{V}$ parallel zu den Seiten von $\overline{V}$ im Abstande d liegen, enthält anderseits C ganz im Innern.

Wir behaupten nun, daß sämtliche Punkte von $\overline{V}_{-d}$ Bilder von Punkten von Q bei der Abbildung (35) sind. Durchläuft nämlich der Punkt (h, k) einmal den Rand von Q, so tut $(\overline{u}, \overline{v})$ das Gleiche auf dem Rande von $\overline{V}$, und (u, v) durchläuft C. Dabei umkreist $(\overline{u}, \overline{v})$ einmal einen beliebigen Punkt P^* in V_{-d} und (u, v) gleichfalls, da (u, v) von $(\overline{u}, \overline{v})$ innerhalb eines Kreises vom Radius d und dem Mittelpunkt $(\overline{u}, \overline{v})$ mitherumgeschleppt wird. Deformiert sich C stetig, so kann sich die Umlaufungszahl nur ändern, wenn C durch P^* geht. Wir ziehen nun den Quadratrand von Q stetig auf den Punkt P_0 zusammen, wobei C stetig deformiert wird. Geht dabei einmal C durch P^*, so ist P^* gewiß Bild eines Punktes von Q. Tritt dies nie ein, so umläuft C stets P^*. Da sich aber C zugleich mit dem Quadratrand auf einen Punkt zusammenzieht, so muß P^* der Grenzpunkt von C und also das Bild von P_0 sein. In jedem Falle ist folglich ein beliebiger Punkt P^* von $\overline{V}_{-d}$ Bild eines Punktes von Q.*)

*) Der hier für diese Behauptung gegebene Beweis ist der einzige der vorliegenden Arbeit, der an der Dimensionszahl 2 hängt. Für n Dimensionen zieht man am besten den Begriff der „Ordnung" eines Punktes in Bezug auf eine geschlossene Fläche und den Kroneckerschen „Index" eines Funktionensystems auf einer geschlossenen Fläche heran. (Siehe hierzu die Note von J. Hadamard „Sur quelques applications de l'indice de Kronecker" in J. Tannery, Introduction à la théorie des fonctions d'une variable, t. 2e, éd. 2e, Paris 1910). Ich skizziere den allgemeinen Beweis. Durch

$$(A)\quad u_i = f_i(x_1 + h_1, x_2 + h_2, \cdots x_n + h_n) = f_i(x_1, x_2, \cdots x_n) + \sum_{j=1}^{n} h_j \frac{\partial f_i}{\partial x_j} + r \cdot R_i,$$

$$\text{wo } r^2 = \sum_j h_j^2 \text{ und } R_i \longrightarrow 0 \text{ mit } r \longrightarrow 0, \qquad i = 1, 2 \ldots n$$

wird der Würfel W

$$h_i \leq l$$

mit dem Mittelpunkt $(x_1, x_2 \ldots x_n)$ auf einen Bereich V abgebildet. (Vorausgesetzt ist nicht-verschwindende Funktionaldeterminante). Die Transformation

$$\overline{u}_i = f_i + \sum_j h_j \frac{\partial f_i}{\partial x_j}$$

bildet dagegen W auf ein Parallelepiped $\overline{V}$ mit der Oberfläche $\overline{S}$ ab. Man bestimmt V_{-d} wie im Text. Jeder Punkt von $\overline{V}_{-d}$ hat in bezug auf $\overline{S}$ die Ordnung 1 (l. c. S. 460, § 28). Die Oberfläche von W wird auf eine (ganz außerhalb $\overline{V}_{-d}$ liegende) Fläche S abgebildet, in bezug auf die jeder Punkt von $\overline{V}_{-d}$ dieselbe Ordnung hat wie in bezug auf $\overline{S}$ (l. c. S. 462 ff., § 30). Die Ordnung des Punktes $(u_1^*, u_2^*, \ldots u_n^*)$ in $\overline{V}_{-d}$ in Bezug auf S ist aber nichts anderes als der Kroneckersche Index des Funktionensystems

$$f_i(x_1 + h_1, x_2 + h_2, \ldots x_n + h_n) - u_i^* \qquad i = 1, 2, \ldots n$$

Somit überdeckt V ganz $\overline{V}_{-d}$ und wird selbst von $\overline{V}_{+d}$ überdeckt, so daß man hat:

$$m\overline{V}_{-d} \leqq mV \leqq m\overline{V}_{+d}. \tag{41}$$

Durch eine bekannte elementare Rechnung erschließt man aus (36) und (37)

$$m\overline{V} = 4l^2|D|.$$

Die offenbar mit l proportionalen und wegen (40) nicht verschwindenden Abstände von je zwei Gegenseiten von $\overline{V}$ seien $2\mu l$ und $2\nu l$ (wo μ und ν Ausdrücke in den Koeffizienten der Transformation (36) sind). Das Parallelogramm $\overline{V}_{+d}$, dessen Seitenabstände je um $2d = 4lR_l^*$ größer sind als die von $\overline{V}$, hat mithin den Inhalt

$$m\overline{V}_{+d} = 4l^2|D|\left(1 + \frac{2R_l^*}{\mu}\right)\left(1 + \frac{2R_l^*}{\nu}\right),$$

und analog gilt für $\overline{V}_{-d}$

$$m\overline{V}_{-d} = 4l^2|D|\left(1 - \frac{2R_l^*}{\mu}\right)\left(1 - \frac{2R_l^*}{\nu}\right).$$

Hieraus und aus (41) ergibt sich dann

$$|D|\left(1 - \frac{2R_l^*}{\mu}\right)\left(1 - \frac{2R_l^*}{\nu}\right) \leqq \frac{mV}{mQ} \leqq |D|\left(1 + \frac{2R_l^*}{\mu}\right)\left(1 + \frac{2R_l^*}{\nu}\right),$$

und folglich ist für Limes $l = 0$ auch in diesem Fall das Vergrößerungsverhältnis $\Delta(P_0)$ der Abbildung gleich dem absoluten Betrag der Funktionaldeterminante

$$\Delta(P_0) = |D| = \frac{D(f, g)}{D(x, y)}.$$

Dies gilt für alle Punkte P_0 in G außer denen der Nullmenge Z, in denen mindestens eine der beiden Funktionen nicht total differenzierbar ist. Wir haben damit den folgenden Satz gewonnen:

Satz III. *Die im Gebiete G definierten, den Bedingungen des Satzes I genügenden Funktionen $u = f(x, y)$ und $v = g(x, y)$ vermitteln eine eindeutige Abbildung des Gebietes G der xy-Ebene auf einen Bereich U der uv-Ebene. Dann ist in dem maßgleichen Kern von G in dem $f(x, y)$ und $g(x, y)$ zugleich total differenzierbar sind, das Vergrößerungsverhältnis der Abbildung gleich dem absoluten Betrag der Funktionaldeterminante.*

Daß übrigens selbst bei *eineindeutigen* Abbildungen die Funktionaldeterminante, auch wenn sie überall in G existiert, doch nicht überall

---

auf der Oberfläche von W. Dieses Funktionensystem ist in W und auf seiner Oberfläche *stetig* definiert. Wegen des nicht verschwindenden Index müssen seine Funktionen eine gemeinsame Nullstelle in W haben (l. c. S. 466, § 33). Also ist jeder Punkt von $\overline{V}_{-d}$ durch (*A*) Bild mindestens eines Punktes von W.

das Vergrößerungsverhältnis der Abbildung zu sein braucht, habe ich früher durch ein Beispiel gezeigt.*)

Bei einer eineindeutigen stetigen Abbildung eines Gebietes wird entweder der Umlaufungssinn in jedem Punkte aufrecht erhalten oder in jedem Punkte umgekehrt. Unsere Überlegungen zeigen nun zugleich, daß das *Vorzeichen* der Funktionaldeterminante dort, wo sie nicht verschwindet, fast überall angibt, ob die Abbildung den Umlaufungssinn erhält oder umkehrt. Denn bekanntlich kann man dies bei einer affinen Transformation aus dem Vorzeichen der Determinante entnehmen, und mit affinen Transformationen haben wir fast überall unsere Abbildung vergleichen können.

Satz IV. *Die Funktionaldeterminante einer eineindeutigen Abbildung, die im übrigen die Voraussetzungen von Satz III erfüllt, nimmt auf dem maßgleichen Kern von G, in dem f und g beide total differenzierbar sind, entweder nicht das negative oder nicht das positive Zeichen an, ersteres, wenn die Abbildung den Umlaufungssinn aufrechterhält, letzteres, wenn sie ihn umkehrt.***)

7. Im folgenden soll Satz II noch eine Ergänzung erfahren. Wir stellen zunächst noch zwei Eigenschaften einer solchen Abbildung fest, welche durch die der Lipschitzschen Bedingung (3) und somit auch (4) genügenden Funktionen $u = f(x, y)$, $v = g(x, y)$ vermittelt sind. Vorausgesetzt wird also

$$\frac{|f(x+h, y+h) - f(x, y)}{\sqrt{h^2+k^2}} < M', \qquad \frac{|g(x+h, y+k) - g(x, y)|}{\sqrt{h^2+k^2}} < M'$$

und folglich auch

$$(42)\quad \sqrt{(f(x+h, y+k) - f(x, y))^2 + (g(x+h, y+k) - g(x, y))^2} < M'\sqrt{2}\sqrt{h^2+k^2}.$$

Dann gilt der

Hilfssatz 1: *Ist W eine beliebige Menge in G, so gilt für das Bild W^* in U*

$$\overline{m}\, W^* \leqq 2\, M'^2\, \overline{m}\, W,$$

wo $\overline{m}$ das äußere Maß bezeichnen soll.

Man kann nämlich W mit abzählbar vielen Kreisen K restlos so überdecken, daß

$$\sum mK < \overline{m}\, W + \varepsilon$$

---

*) Eineindeutige Abbildungen und Meßbarkeit, Göttinger Diss. 1917, S. 107 f. (auch Monatshefte f. Math. u. Phys., Bd. 27, 1916, S. 288 f.) Im Folgenden zitiert als Dissertation.

**) Für mehr als zwei Dimensionen wäre hier statt des Umlaufungssinns eine allgemeiner definierte „Indikatrix" zu nehmen.

für beliebiges $\varepsilon > 0$ ist.*) Wegen (42) liegt aber das Bild K^* eines Kreises K mit dem Radius ϱ ganz in einem Kreise mit dem Radius $M'\sqrt{2}\varrho$, so daß

$$mK^* < 2M'^2 mK$$

und damit

$$\sum' mK^* < 2M'^2(\overline{m}W + \varepsilon)$$

ist. Da aber W^* von den K^* ganz überdeckt wird, so folgt weiter

$$\overline{m}W^* < 2M'^2(\overline{m}W + \varepsilon)$$

und wegen der Willkürlichkeit von ε die zu beweisende Ungleichung.

Speziell wird also eine Nullmenge stets auf eine Nullmenge abgebildet. Die Meßbarkeit von W zieht übrigens die von W^* nach sich, denn eine meßbare Menge ist stets darstellbar als Vereinigungsmenge abzählbar vieler abgeschlossener Mengen und einer Nullmenge, von denen die abgeschlossenen Mengen wegen der Stetigkeit der Abbildung wieder auf abgeschlossene, also meßbare Mengen, die Nullmenge aber auf eine Nullmenge abgebildet wird, deren Vereinigungsmenge wieder meßbar ist.

Hilfssatz 2: *Wird unter der Voraussetzung* (42) *eine meßbare Menge W von positivem Maß auf eine Nullmenge W^* abgebildet, so ist fast überall in W die Funktionaldeterminante* $\frac{D(f, g)}{D(x, y)}$ *gleich Null.***)

In W gibt es nämlich einen maßgleichen Kern W_0, in dem erstens die Dichte von W gleich 1 ist und in dem zweitens der absolute Betrag der Funktionaldeterminante das Vergrößerungsverhältnis darstellt. Es sei P_0 ein Punkt W_0. Dann gibt es nach der Definition der Dichte zu $\varepsilon > 0$ ein Quadrat Q_0 mit dem Mittelpunkt P_0, so daß

$$\frac{m(Q_0 W)}{mQ_0} > 1 - \varepsilon$$

ist.***) Nun ist das Bild von $Q_0 W$ eine Nullmenge, die in der Nullmenge W^* enthalten ist. Ferner ist

$$m(Q_0 - Q_0 W) < \varepsilon m Q_0;$$

das Bild von $(Q_0 - Q_0 W)$ hat daher nach Hilfssatz 1 ein Maß, das kleiner als

$$2M'^2 \varepsilon m Q_0,$$

*) Diss. S. 9, Satz II.

**) Herr Ch. J. de la Vallée Poussin hat den entsprechenden Satz für *eine* Variable bewiesen und seine Wichtigkeit hervorgehoben, Transactions of the American Mathematical Society 16, 1915, S. 467, Corollaire.

***) Obgleich wir im § 2 die Dichte in einem Punkte P_0 mit Hilfe von Kreisen um P_0 definiert haben, können wir hier Quadrate zu dem gleichen Zweck benutzen, da der in Anmerkung *), S. 344 zitierte Satz für die *verallgemeinerte* Derivierte gilt.

ist. Insgesamt hat also die Menge Q_0^*, die Bild des Quadrates Q_0 ist, ein Maß von weniger als $2M'^2\varepsilon\, m\,Q_0$. Also

$$\frac{m\,Q_0^*}{m\,Q_0} < 2M'^2\varepsilon,$$

woraus für das Vergrößerungsverhältnis in P_0 wegen der Willkürlichkeit von ε

$$\Delta(P_0) = 0$$

folgt. Dann muß aber wegen der Bestimmung von W_0 auch

$$\frac{D(f,g)}{D(x,y)} = 0$$

in P_0 und daher ebenso in ganz W_0 sein, womit der Hilfssatz bewiesen ist.

8. Es seien nun $f(x, y)$ und $g(x, y)$ im Gebiete G der xy-Ebene, $x(\xi, \eta)$, $y(\xi, \eta)$ im Gebiete Γ der $\xi\,\eta$-Ebene definiert. Das durch $x = x(\xi, \eta)$, $y = y(\xi, \eta)$ von Γ entworfene Bild G_1 liege ganz in G. Die vier Funktionen mögen ferner beschränkte partielle Derivierte haben. Dann sind $x(\xi, \eta)$ und $y(\xi, \eta)$ außer in einer Nullmenge Z' in Γ partiell differenzierbar. Das Bild Z' von Z' ist nach Hilfssatz 1 eine Nullmenge in G_1. Wären $x(\xi, \eta)$ und $y(\xi, \eta)$ überall in Γ partiell differenzierbar, so würden (33) und die für $g(x, y)$ analogen Formeln von § 5 in G_1 bis auf diejenige Nullmenge Z'' gelten, in der f und g nicht beide total differenzierbar sind. Unter der jetzigen Voraussetzung sind daher (33) und die analogen Formeln in G_1 gültig außer in der Nullmenge $Z'' + Z' = Z_0$ in G_1. Die Menge der (ξ, η) in Γ, in der (33) und die analogen nicht gelten, heiße $\mathfrak{Z}_0$. Dann ist Z_0 das durch $x = x(\xi, \eta)$, $y = y(\xi, \eta)$ von $\mathfrak{Z}_0$ entworfene Bild.

$\mathfrak{Z}_0$ ist meßbar, denn die in (33) auftretenden Funktionen sind sämtlich meßbar in der $\xi\eta$-Ebene. Die Funktionen $f(x(\xi, \eta), y(\xi, \eta))$ und $g(x(\xi, \eta), y(\xi, \eta))$ sind nämlich als stetige Funktionen stetiger Funktionen in (ξ, η) stetig, ihre partiellen Derivierten also meßbar. Das Gleiche gilt von den Derivierten von $x(\xi, \eta)$ und $y(\xi, \eta)$. Daß die zunächst nur in der xy-Ebene meßbaren Derivierten von $f(x, y)$ und $g(x, y)$ nach x und y auch als Funktionen von ξ, η meßbar sind, sieht man folgendermaßen ein. Es ist

$$\frac{f(x(\xi, \eta) + h,\ y(\xi, \eta)) - f(x(\xi, \eta),\ y(\xi, \eta))}{h}$$

für jedes $h > 0$ in ξ, η stetig. Der Limes superior dieser Funktion für $h \to 0$, der die obere vordere partielle Derivierte nach x von $f(x, y)$ darstellt als Funktion der ξ, η, ist also als Funktion der zweiten Baireschen Klasse in der $\xi\eta$-Ebene meßbar. Das Gleiche gilt von den andern Derivierten von f und g. Auf der rechten und linken Seite von (33) und den analogen Formeln stehen also meßbare Funktionen, die sich nur in einer meßbaren Menge $\mathfrak{Z}_0$ unterscheiden können.

Ist nun Z_0 selbst eine Nullmenge, so gilt die Multiplikationsformel (34) in einem maßgleichen Kern von Γ. Ist $mZ_0 > 0$, so wollen wir zeigen, daß dennoch (34) fast überall in Γ gilt.

Zunächst sieht man, daß auch $f(x(\xi, \eta), y(\xi, \eta))$ und $g(x(\xi, \eta), y(\xi, \eta))$ in (ξ, η) die Lipschitzsche Bedingung erfüllen. Denn z. B.

$$\frac{f(x(\xi_1, \eta), y(\xi_1, y)) - f(x(\xi_2, \eta), y(\xi_2, \eta))}{\xi_1 - \xi_2}$$

$$= \frac{f(x_1, y_1) - f(x_2, y_2)}{\sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}} \cdot \frac{\sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}}{\xi_1 - \xi_2} < 2M \cdot \sqrt{2}\, M,$$

was aus (3) und (4) und den entsprechenden Lipschitzschen Ungleichungen für $x(\xi, \eta)$ und $y(\xi, \eta)$ folgt; zur Abkürzung ist dabei $x_1 = x(\xi_1, \eta)$ usw. geschrieben.

Da nun Z_0 auf die Nullmenge Z_0 abgebildet wird, so ist nach Hilfssatz 2 fast überall in Z_0

$$\frac{D(x, y)}{D(\xi, \eta)} = 0.$$

Da ferner das durch $u = f(x, y)$, $v = g(x, y)$ von Z_0 entworfene Bild Z_0^* nach Hilfssatz 1 ebenfalls eine Nullmenge sein muß, so geht Z_0 auch durch $u = f(x(\xi, \eta), y(\xi, \eta))$, $v = g(x(\xi, \eta), y(\xi, \eta))$ in eine Nullmenge über, und fast überall in Z_0 ist daher nach Hilfssatz 2

$$\frac{D(f, g)}{D(\xi, \eta)} = 0,$$

so daß also in Z_0 fast überall beide Seiten von (34) verschwinden und die Gleichung somit gewiß erfüllt ist.

Damit haben wir den

Satz V: *Sind $f(x, y)$ und $g(x, y)$ zwei in einem Gebiete G der xy-Ebene und $x(\xi, \eta)$, $y(\xi, \eta)$ zwei in einem Gebiete Γ der $\xi\eta$-Ebene definierte Funktionen, deren partielle Derivierte beschränkt sind (die Funktionen genügen mit andern Worten der Lipschitzschen Bedingung), wobei ferner die Funktionswerte von $x(\xi, \eta)$ und $y(\xi, \eta)$ in das Gebiet G hineinfallen, so gilt fast überall in Γ die Multiplikationsformel der Funktionaldeterminanten*

$$\frac{D(f, g)}{D(\xi, \eta)} = \frac{D(f, g)}{D(x, y)} \frac{D(x, y)}{D(\xi, \eta)},$$

*wobei den Funktionaldeterminanten dort, wo sie etwa nicht existieren, ein beliebiger Wert beigelegt sein mag.**)

*) Das folgende, mit dem im § 3 gegebenen verwandte Beispiel zeigt übrigens, daß die Multiplikationsformel wirklich nur *fast* überall gilt:

III. Die Transformation der Doppelintegrale.

9. Es werde endlich durch die Funktionen

$$u = \varphi(x, y), \quad v = \psi(x, y)$$

das beschränkte Gebiet G eineindeutig und stetig auf das Gebiet U der u, v-Ebene abgebildet, wobei über die nach (1) und (2) definierten Funktionen $L_\varphi(x, y)$ und $L_\psi(x, y)$ wieder vorausgesetzt sei, daß sie endlich und summierbar in G seien. Dann ist nach Satz III fast überall in G das Vergrößerungsverhältnis gleich $\left|\frac{D(\varphi, \psi)}{D(x, y)}\right|$.

Die Abbildung ist ferner meßbar, d. h. sie führt jede meßbare Teilmenge von G in eine meßbare Teilmenge von U über. Denn bilden wir:

$$w(x, y; \varrho) = \text{obere Grenze } \frac{\sqrt{(\varphi(x+h, y+k)-\varphi(x, y))^2 + (\psi(x+h, y+k)-\psi(x, y))^2}}{\sqrt{h^2+k^2}}$$

$$0 < h^2 + k^2 \leqq \varrho^2$$

und

$$L(x, y) = \lim_{\varrho \to 0} w(x, y; \varrho),$$

so ist für die Meßbarkeit der Abbildung hinreichend, daß $L(x, y)$ in G endlich ist.*) Nun ist aber

$$w(x, y; \varrho) \leqq \text{obere Grenze } \frac{|\varphi(x+h, y+k)-\varphi(x, y)| + |\psi(x+h, y+k)-\psi(x, y)|}{\sqrt{h^2+k^2}}$$

$$\leqq \text{obere Grenze } \frac{|\varphi(x+h, y+k)-\varphi(x, y)|}{\sqrt{h^2+k^2}}$$

$$+ \text{obere Grenze } \frac{|\psi(x+h, y+k)-\psi(x, y)|}{\sqrt{h^2+k^2}} = w_\varphi(x, y; \varrho) + w_\psi(x, y; \varrho);$$

also

$$L(x, y) \leqq L_\varphi(x, y) + L_\psi(x, y),$$

folglich $L(x, y)$ endlich.**) Nach einem Satze aus der Transformationstheorie der Lebesgueschen Integrale***) ergibt sich dann

---

$$\left.\begin{aligned} f(x, y) &= \frac{x^2 y}{x^2+y^2} \\ g(x, y) &= \frac{xy^2}{x^2+y^2} \end{aligned}\right\} \text{ für } x^2+y^2 > 0 \qquad \left.\begin{aligned} f(x, y) &= 0 \\ g(x, y) &= 0 \end{aligned}\right\} \text{ für } x^2+y^2 = 0, \qquad \begin{aligned} x &= \xi + \eta \\ y &= -\xi + \eta. \end{aligned}$$

Im Punkte $\xi = 0$, $\eta = 0$ ergibt sich nämlich:

$$\frac{D(f, g)}{D(\xi, \eta)} = \frac{1}{2}, \qquad \frac{D(f, g)}{D(x, y)} = 0, \qquad \frac{D(x, y)}{D(\xi, \eta)} = 2.$$

*) Diss. S. 57, Satz XXI und S. 59, Satz XXII.

**) Man erkennt leicht, daß $L_\varphi(x, y) \leqq L(x, y)$, $L_\psi(x, y) \leqq L(x, y)$ ist, also die Endlichkeit und Summierbarkeit von $L(x, y)$ mit unserer Voraussetzung äquivalent ist.

***) Diss. S. 52, Satz XIX.

Satz VI: *Wird das Gebiet G der xy-Ebene eineindeutig auf das Gebiet U der uv-Ebene abgebildet durch zwei Funktionen $u = \varphi(x, y)$, $v = \psi(x, y)$, für welche die nach (1) und (2) gebildeten Funktionen $L_\varphi(x, y)$ und $L_\psi(x, y)$ in G endlich und summierbar sind (wofür hinreicht, daß φ und ψ der Lipschitzschen Bedingung genügen), so ist für eine in U definierte endliche Funktion $F(u, v)$*

$$\iint_U F(u, v)\, du\, dv = \iint_G F(u(x, y),\ v(x, y)) \left| \frac{D(\varphi, \psi)}{D(x, y)} \right| dx\, dy,$$

wobei aus der Existenz des einen Integrals die des andern folgt und der Funktionaldeterminante auf der Nullmenge, wo sie etwa nicht existiert, ein beliebiger Wert beigelegt ist.

Nicht ganz so allgemein habe ich diesen Satz schon früher, allerdings auf eine nicht sehr durchsichtige Weise, mit Hilfe der Approximation durch Tonellische Polynome bewiesen.*)

Es soll noch einmal hervorgehoben werden, daß sämtliche Beweise und Sätze dieser Arbeit sich ohne weiteres auf mehr als zwei Variable übertragen lassen.

Wickersdorf, den 4. Oktober 1918.

---

*) Diss. S. 107, Satz XXVI.