MATH. SCAND. 1 (1953)

BEMERKUNG ZUR FUNKTIONALANALYSIS

ROLF NEVANLINNA

§ 1. Einleitung.

1. Im Folgenden soll ein Beweis des Hauptsatzes über die lokale Umkehrbarkeit einer eindeutigen differenzierbaren Abbildung gegeben werden. Da die Beweismethode unabhängig von Koordinatendarstellungen und von Dimensionszahlen der betrachteten Räume sein wird, wollen wir uns auf den Standpunkt allgemeiner Hilbertscher Räume stellen. Der von Fréchet und seinen Nachfolgern eingeführte Begriff der Ableitung eines Funktionals wird für die Untersuchung wesentlich sein, und wir wollen vorbereitenderweise die hier zur Anwendung kommenden Grundbegriffe und Bezeichnungen zusammenstellen.

2. Es seien H_x und H_y zwei Hilbertsche (oder Banachsche) Räume, deren Elemente (Punkte) mit x bzw. y bezeichnet werden sollen; der Körper der skalaren Multiplikatoren sei z. B. derjenige der reellen Zahlen $(\lambda, \mu, \ldots)$. Wir betrachten eine eindeutige Abbildung $y = y(x)$ der Kugel $|x| \leqq R$ in den Raum H_y.

Die *Ableitung* $dy(x)/dx$ von $y(x)$ im Punkte x erklären wir, falls eine solche existiert, als diejenige lineare Transformation $D = dy/dx$ des Raumes H_x in den Raum H_y, für welche der Zuwachs $\Delta y = y(x+\Delta x)-y(x)$ eine Entwicklung

$$\Delta y = D\Delta x+|\Delta x|(\Delta x)$$

zulässt, wobei $(\Delta x) \to 0$ für $\Delta x \to 0$ (Differenzierbarkeit von $y = y(x)$).

Falls die Räume H_x und H_y von endlicher Dimension n_x bzw. n_y sind, so ist dy/dx gleich der durch die Funktionalmatrix

$$\frac{\partial y_i}{\partial x_k} \quad (i = 1,\ldots,n_y;\ k = 1,\ldots,n_x)$$

bestimmten linearen Transformation, wobei y_i, x_k Koordinaten in zwei affinen Systemen sind.

Eingegangen am 28. April 1953.

3. Die Existenz der Ableitung D impliziert die Stetigkeit von $y(x)$ im Punkte x, sofern D eine stetige (beschränkte) Transformation ist, was wir von allen vorkommenden Ableitungen voraussetzen wollen. Dann gilt auch die Kettenregel

$$\frac{dz}{dx} = \frac{dz}{dy}\frac{dy}{dx},$$

wo rechts das Produkt der Transformationen dy/dx und dz/dy steht.

4. Unter dem Differential dy verstehen wir das Element

$$dy = \frac{dy}{dx}\,dx\,,$$

wobei dx einfach gleich Δx ist. Für den totalen Zuwachs Δy hat man also

$$\Delta y = dy + |dx|(dx)\,.$$

§ 2. Der Mittelwertsatz.

5. Wir stellen diesen fundamentalen Satz in einer allgemeinen Form auf, in der er als eine direkte Erweiterung des elementaren Mittelwertsatzes erscheint und diesen als einfachsten Spezialfall enthält (H_x und H_y eindimensional). Der Beweis stützt sich hier, wie in der elementaren Analysis, auf den Satz von Rolle.

Für das Folgende ist es wesentlich, dass die Räume H Hilbertsch, also mit einer bilinearen, symmetrischen, positiv definiten metrischen Form ausgerüstet sind. Dieses innere Produkt von zwei Elementen x_1, x_2 wird, wie üblich, mit (x_1, x_2) bezeichnet; für die Norm $+(x, x)^{1/2}$ von x verwenden wir das Zeichen $|x|$. Im Folgenden wird Vollständigkeit (nicht aber Separabilität) der Hilbertschen Räume H_x, H_y vorausgesetzt.

6. Satz von Rolle. *Wir nehmen an:*

$1°$ *$y = y(x)$ ist stetig auf der Strecke $x = a + \tau(b-a)$, $0 \leqq \tau \leqq 1$.*

$2°$ *Die Ableitung $\dfrac{dy}{dx}$ existiert für $0 < \tau < 1$.*

$3°$ *$y(a) = y(b)$.*

Dann gibt es für jeden Vektor $e \in H_y$ einen Wert τ mit $0 < \tau < 1$, so dass

$$(dy(\xi), e) = 0, \quad \xi = a + \tau(b-a)\,. \tag{1}$$

Beweis. Die Projektion des Vektors $y(x)$ auf einen beliebig gewählten Einheitsvektor e ist

$$p(x) = (y(x), e)e\,.$$

Hier genügt das innere Produkt $(y(x), e)$ als Funktion von τ den Bedingungen des gewöhnlichen Rolleschen Satzes. Wegen der Linearität von (y, e) in y ist

$$d(y, e) = (dy, e) ,$$

und es wird für ein $\xi = a+\tau(b-a)$, $0 < \tau < 1$,

$$(dy(\xi), e) = 0 ,$$

w.z.b.w.

7. Wir behalten die Voraussetzungen 1° und 2° des Rolleschen Satzes bei, während die Punkte $y(a)$ und $y(b)$ jetzt nicht mehr zusammenzufallen brauchen. Die Anwendung des Rolleschen Satzes auf

$$y(x)-A(x) ,$$

wo A diejenige auf $x = a+\tau(b-a)$ wohlbestimmte lineare Transformation ist, welche die Strecke $\Delta x = b-a$ in die Strecke $\Delta y = y(b)-y(a)$ überführt, ergibt für einen beliebigen Vektor e und einen zugeordneten Wert τ des Intervalles $0 < \tau < 1$ mit $\xi = a+\tau(b-a)$

$$0 = (dy(\xi)-dA(b-a), e) = (dy(\xi), e)-(\Delta y, e) .$$

Dies ist der

Mittelwertsatz. *Unter den Voraussetzungen* 1° *und* 2° *gibt es zu jedem Vektor e einen Wert* $\xi = a+\tau(b-a)$, $0 < \tau < 1$, *so dass*

$$(\Delta y, e) = (dy(\xi), e) . \tag{2}$$

8. Der Mittelwertsatz lässt sich zur Abschätzung des Zuwachses Δy verwenden. Vorausgesetzt, dass $\Delta y \neq 0$, wähle man $e = \Delta y/|\Delta y|$; dann ergibt die Beziehung (2) mit Hilfe der Schwarzschen Ungleichung

$$|\Delta y| \leqq |dy(\xi)| = \left|\frac{dy(\xi)}{dx}\Delta x\right| . \tag{3}$$

Führt man die Bezeichnung

$$M_x = \sup\left|\frac{dy(x)}{dx}\right| \equiv \sup_{|h|=1}\left|\frac{dy(x)}{dx}h\right|$$

ein, und ist

$$M_x \leqq M \text{ für } |x| \leqq R ,$$

so gilt also in dieser Kugel die Lipschitzbedingung

$$|\Delta y| \leqq M|\Delta x| .$$

Im Folgenden werden wir den Mittelwertsatz verwenden, um eine untere Abschätzung von $|\Delta y|$ herzuleiten.

§ 3. Lokale Umkehrbarkeit einer differenzierbaren Abbildung.

9. SATZ 1. *Falls die für $|x| \leqq R$ eindeutige und differenzierbare Abbildung $y = y(x)$ den Bedingungen*

$$\text{(A)} \qquad \inf \left| \frac{dy(0)}{dx} \right| \equiv \inf_{|h|=1} \left| \frac{dy(0)}{dx} h \right| = m > 0$$

und

$$\text{(B)} \qquad \sup \left| \frac{dy(x)}{dx} - \frac{dy(0)}{dx} \right| < m \text{ für } |x| \leqq \varrho \leqq R$$

genügt, so ist die Abbildung für $|x| \leqq \varrho$ eineindeutig.

BEMERKUNG. Die Bedingung (B) ist für hinreichend kleines ϱ sicher dann erfüllt, wenn die Ableitung dy/dx für $x = 0$ *stetig* ist, d.h. wenn

$$\sup \left| \frac{dy(x)}{dx} - \frac{dy(0)}{dx} \right| \to 0 \text{ für } x \to 0 .$$

Falls die Dimensionszahlen n_x und n_y der Räume H_x bzw. H_y endlich sind, so bedeutet die Bedingung (A), dass $n_y \leqq n_x$ und dass die Funktionalmatrix von $y(x)$ den Rang n_y hat.

BEWEIS VON SATZ 1. Man setze im Mittelwertsatze: $|a| \leqq \varrho$, $|b| \leqq \varrho$. Wenn e ein beliebiger Einheitsvektor ist, findet man dann für $\Delta x = b - a \neq 0$ mittels der Schwarzschen Ungleichung

$$|\Delta y| \geqq |(\Delta y, e)| = \left|(dy(\xi), e)\right| = \left|(dy(0), e) + (dy(\xi) - dy(0), e)\right|$$
$$\geqq \left|(dy(0), e)\right| - \sup \left| \frac{dy(\xi)}{dx} - \frac{dy(0)}{dx} \right| |\Delta x| .$$

Wählt man

$$e = \frac{dy(0)}{dx} \Delta x \Big/ \left| \frac{dy(0)}{dx} \Delta x \right| ,$$

so wird also mit Rücksicht auf (A) und (B)

$$\text{(4)} \qquad |\Delta y| \geqq \left(m - \sup \left| \frac{dy(\xi)}{dx} - \frac{dy(0)}{dx} \right| \right) |\Delta x| \neq 0 ,$$

womit der Satz bewiesen ist.

10. Wir nehmen jetzt an, dass die Ableitung dy/dx für $x = 0$ nicht nur stetig ist, sondern einer Lipschitzbedingung genügt:

$$\text{(B)}' \qquad \sup \left| \frac{dy(x)}{dx} - \frac{dy(0)}{dx} \right| \leqq M|x| \text{ für } |x| \leqq R .$$

Dann ist die Bedingung (B) sicher in der Kugel $|x| \leqq \varrho = \min(mM^{-1}, R)$ erfüllt, und man hat den

Satz 2. *Unter den Bedingungen* (A) *und* (B)′ *ist die Abbildung* $y = y(x)$ *für*

(5) $$|x| \leqq \varrho \equiv \min(mM^{-1}, R)$$

eineindeutig.

Es sei bemerkt, dass der durch (5) bestimmte »Schlichtheitsradius« unter den angenommenen Voraussetzungen (A) und (B)′ nicht vergrössert werden kann. Dies zeigt das elementare Beispiel der Wurfparabel $y = mx + \frac{1}{2}Mx^2$. Hier ist

$$\left|\frac{dy}{dx}\right| = \inf\left|\frac{dy}{dx}\right| = m, \qquad \sup\left|\frac{dy(x)}{dx} - \frac{dy(0)}{dx}\right| = M;$$

die Abbildung ist in dem Intervall $|x| \leqq \varrho$ für $\varrho = mM^{-1}$ eineindeutig, hingegen für $\varrho > mM^{-1}$ bereits zweiwertig.

§ 4. Die Bildwertmenge.

11. Es gilt jetzt zu entscheiden, unter welchen Bedingungen die Bildpunktmenge $y = y(x)$ eine *volle Umgebung* des Punktes $y = y(0)$ enthält.

Wir werden zeigen, dass dies unter den obigen Bedingungen (A) und (B)′ gilt, sofern die Ableitung dy/dx im Punkte $x = 0$ die entsprechende Eigenschaft hat:

(C) *Der Wertvorrat der linearen Transformation* $dy(0)/dx$ *ist der ganze Raum* H_y.

Diese Eigenschaft ist für endlich dimensionale Räume H_x, H_y eine Folgerung aus der Bedingung (A), falls die Dimensionen n_x und n_y gleich sind: jene Bedingung spricht nämlich dann das Nichtverschwinden der Funktionaldeterminante $\|\partial y/\partial x\|$ aus. Für unendlichdimensionale Räume ($H_x = H_y$) ist die Bedingung (C) z. B. sicher dann erfüllt, wenn neben $dy(0)/dx$ auch die adjungierte lineare Transformation $dy^*(0)/dx$ eine *positive* untere Grenze hat.

12. Wir führen daher die Untersuchung unter Annahme der zusätzlichen Bedingung (C) weiter und beweisen den

Satz 3. *Falls die Abbildung* $y = y(x)$ *den drei Bedingungen* (A), (B)′ *und* (C) *genügt, so enthält die Bildpunktmenge* $y = y(x)$ *für* $|x| \leqq \frac{1}{2}\varrho \equiv \frac{1}{2}\min(mM^{-1}, R)$ *die volle Kugel* $|y - y(0)| < \frac{1}{8}m\varrho$.

13. Zum Beweise normieren wir $y(0) = 0$ und zeigen, dass die Antithese:

(6) $$y(x) \neq c \quad (|c| < \tfrac{1}{8}m\varrho)$$

für $|x| \leqq \frac{1}{2}\varrho$, zu einem Widerspruch führt, wobei ϱ wieder der im Satz 3 erwähnte Schlichtheitsradius ist.

Hierzu bemerke man zunächst, dass die nichtnegative untere Grenze

(7) $$r \equiv \inf_{|x| \leqq \frac{1}{2}\varrho} |y(x)-c|$$

positiv sein muss. Ist nämlich $|a| \leqq \frac{1}{2}\varrho$, $|b| \leqq \frac{1}{2}\varrho$, $\Delta x = b-a$, $\Delta y = y(b)-y(a)$, so wird nach (4) und (B)'

(8) $$|\Delta y| \geqq (m-\tfrac{1}{2}M\varrho)|\Delta x| \geqq \tfrac{1}{2}m|\Delta x| \, .$$

Ist nun $x_1, x_2, \ldots$ eine Folge von Punkten in $|x| \leqq \frac{1}{2}\varrho$, derart dass

(9) $$|y(x_i)-c| \to r \text{ für } i \to \infty \, ,$$

und wäre $r = 0$, so hätte man $y(x_i)-y(x_j) \to 0$ für $i, j \to \infty$ und nach (8) auch $x_i - x_j \to 0$. Für den Grenzpunkt $x_0 = \lim x_i$ ($|x_0| \leqq \frac{1}{2}\varrho$) hätte man dann wegen der Stetigkeit von $y(x)$ die Gleichheit $y(x_0) = c$, im Widerspruch zur Antithese.

Es ist also $r > 0$ und andererseits, da $y(0) = 0$,

(10) $$0 < r \leqq |y(0)-c| = |c| < \tfrac{1}{8}m\varrho \, .$$

14. Wir wählen nun N so gross, dass der Betrag von

$$k_i = c - y(x_i)$$

der Bedingung

(11) $$|k_i| < \tfrac{1}{8}m\varrho \text{ für } i \geqq N$$

genügt, was gemäss (9) und (10) möglich ist[1].

Mit dieser Festsetzung gilt nach (8) (wo $a = 0$, $b = x_i$ zu setzen ist)

(12) $$\begin{aligned} |x_i| &\leqq 2m^{-1}|y(x_i)| \leqq 2m^{-1}(|y(x_i)-c|+|c|) \\ &\leqq 2m^{-1}(\tfrac{1}{8}m\varrho+|c|) = \tfrac{1}{2}\varrho - \tfrac{1}{4}\theta\varrho \, , \end{aligned}$$

wo

(12)' $$\theta = 1-8|c|m^{-1}\varrho^{-1} \quad (0 < \theta < 1) \, .$$

Der Punkt x_i liegt also für $i \geqq N$ in der Kugel $|x| < \frac{1}{2}\varrho$.

---

[1] Für *kompakte* Räume würden sich die nachfolgenden Überlegungen etwas vereinfachen, da die Schranke r dann *erreicht* wird.

15 Wir geben jetzt x_i einen Zuwachs

$$\Delta x_i = \tfrac{1}{3}\theta h_i\,, \tag{13}$$

wo h_i der durch die Gleichung

$$\frac{dy(0)}{dx} h_i = k_i \tag{13$'$}$$

eindeutig bestimmte Punkt in H_x ist. Man hat nach (A)

$$m|h_i| \leqq |k_i|\,. \tag{13$''$}$$

Nach (11) und (13)$''$ wird

$$|\Delta x_i| = \tfrac{1}{3}\theta|h_i| \leqq \tfrac{1}{3}\theta m^{-1}|k_i| < \tfrac{1}{24}\theta\varrho \tag{14}$$

und wegen (12)

$$|x_i+\Delta x_i| \leqq |x_i|+|\Delta x_i| \leqq \tfrac{1}{2}\varrho-\tfrac{1}{4}\theta\varrho+\tfrac{1}{24}\theta\varrho < \tfrac{1}{2}\varrho\,.$$

Auch der Punkt $x_i{}^* = x_i+\Delta x_i$ liegt also für $i \geqq N$ in der Kugel $|x| < \frac{1}{2}\varrho$.

16. Für den Punkt $y(x_i{}^*) = y(x_i)+\Delta y$ hat man nun

$$|y(x_i{}^*)-c|^2 = |\Delta y-k_i|^2 = |k_i|^2-2(\Delta y, k_i)+|\Delta y|^2\,. \tag{15}$$

Der Mittelwertsatz ergibt hier, $\xi = x_i+\tau\Delta x_i$ gesetzt, mit Rücksicht auf (13)$'$, (B)$'$ und (13)$''$

$$\begin{aligned}(\Delta y, k_i) &= \big(dy(\xi), k_i\big) = \big(dy(0), k_i\big)+\big(dy(\xi)-dy(0), k_i\big)\\ &\geqq \tfrac{1}{3}\theta|k_i|^2-\tfrac{1}{2}M\varrho|\Delta x_i||k_i| \geqq \tfrac{1}{3}\theta|k_i|^2-\tfrac{1}{6}\theta|k_i|^2 = \tfrac{1}{6}\theta|k_i|^2\,.\end{aligned}$$

Ferner gilt mit $\xi' = x_i+\tau'\Delta x_i$ nach (3), (13), (13)$'$, (B)$'$ und (14)

$$\begin{aligned}|\Delta y| &\leqq \left|\frac{dy(\xi')}{dx}\Delta x_i\right| \leqq \left|\frac{dy(0)}{dx}\Delta x_i\right| + \left|\left(\frac{dy(\xi')}{dx}-\frac{dy(0)}{dx}\right)\Delta x_i\right|\\ &\leqq \tfrac{1}{3}\theta|k_i|+\tfrac{1}{2}M\varrho|\Delta x_i| \leqq \tfrac{1}{3}\theta|k_i|+\tfrac{1}{6}\theta|k_i| = \tfrac{1}{2}\theta|k_i|\,.\end{aligned}$$

Zusammenfassend erhält man aus (15)

$$|y(x_i{}^*)-c|^2 \leqq |k_i|^2-\tfrac{1}{3}\theta|k_i|^2+\tfrac{1}{4}\theta^2|k_i|^2 \leqq (1-\tfrac{1}{12}\theta)|k_i|^2$$

für jedes $i \geqq N$, wo $x_i{}^* = x_i+\Delta x_i$, $|x_i{}^*| < \frac{1}{2}\varrho$.

Wählt man nun $i \geqq N$ so gross, dass

$$|k_i|^2 < r^2(1+\tfrac{1}{12}\theta)\,,$$

so ergibt sich der Widerspruch

$$r^2 \leqq |y(x_i{}^*)-c|^2 < (1-\tfrac{1}{144}\theta^2)r^2\,,$$

und der Satz ist bewiesen.

17. Das in Abschnitt 10 betrachtete elementare Beispiel ergibt für den Radius ϱ_y der grössten Kugel, welche von der Bildmenge $y = y(x)$ für $|x| \leqq \frac{1}{2}\varrho = \frac{1}{2} \min (mM^{-1}, R)$ voll überdeckt wird, den Wert

$$\varrho_y = \tfrac{3}{8} m^2 M^{-1} = \tfrac{3}{8} m\varrho ,$$

während unser Satz allgemein $\varrho \geqq \frac{1}{8} m\varrho$ behauptet. Die gefundene Schranke von ϱ_y ist also höchstens dreimal zu klein. Durch eine genauere Analyse des obigen Beweises liesse sich die Schranke etwas erhöhen.

§ 5. Anwendung auf endlichdimensionale Räume.

18. Wenn $H_x = H_y$ von endlicher Dimensionszahl n ist, so enthalten die obigen Ergebnisse Folgendes (wobei die grösste Genauigkeit nicht angestrebt worden ist).

In einem affinen Koordinatensystem sei

$$\left(\frac{\partial y_i}{\partial x_k}\right), \quad i, k = 1, 2, \ldots, n,$$

die Funktionalmatrix von y nach x und

$$\left\| \frac{\partial y_i}{\partial x_k} \right\| = d$$

die Funktionaldeterminante. Sei ferner

$$\max_{i,k} \left| \frac{\partial y_i}{\partial x_k} \right|_{x=0} = \alpha$$

und

$$\left| \frac{\partial y_i(x)}{\partial x_k} - \frac{\partial y_i(0)}{\partial x_k} \right| \leqq \beta\, |x| \text{ für } |x| \leqq R .$$

Dann gilt für die Ableitung $D = dy(x)/dx$ im Punkte $x = 0$

$$|Dx|^2 = (D^* Dx, x)$$

und, wenn $\lambda_1 \leqq \lambda_2 \leqq \ldots \leqq \lambda_n$ die Eigenwerte dieser positiven quadratischen Form sind,

$$d^2 = \lambda_1 \lambda_2 \ldots \lambda_n, \quad \inf \left| \frac{dy(0)}{dx} \right| = \lambda_1^{1/2} .$$

Es wird ferner

$$\lambda_n = \left(\sup \left| \frac{dy(0)}{dx} \right| \right)^2 \leqq \sum_{i,k} \left(\frac{\partial y_i}{\partial x_k} \right)^2 \leqq n^2 \alpha^2$$

und

$$\sup \left| \frac{dy(x)}{dx} - \frac{dy(0)}{dx} \right| \leqq n\beta |x| .$$

Schliesslich hat man

$$\lambda_1 = \frac{d^2}{\lambda_2 \ldots \lambda_n} \geqq \frac{d^2}{\lambda_n^{\,n-1}} \geqq \left(\frac{d}{n^{n-1}\alpha^{n-1}}\right)^2,$$

somit

$$\inf \left| \frac{dy(0)}{dx} \right| \geqq \frac{|d|}{n^{n-1}\alpha^{n-1}}.$$

19. Die obigen Sätze sind also erfüllt mit

$$m = \frac{|d|}{n^{n-1}\alpha^{n-1}}, \qquad M = n\beta\,,$$

und es gilt somit das

KOROLLAR. *Falls die Funktionen*

$$y_i = y_i(x_1, \ldots, x_n), \quad i = 1, \ldots, n\,,$$

für $|x|^2 = \Sigma x_i^2 \leqq R^2$ *partielle Ableitungen erster Ordnung besitzen, die den Bedingungen*

$$\left| \frac{\partial y_i(0)}{\partial x_k} \right| \leqq \alpha, \quad i, k = 1, \ldots, n\,,$$

und

$$\left| \frac{\partial y_i(x)}{\partial x_k} - \frac{\partial y_i(0)}{\partial x_k} \right| \leqq \beta |x|$$

genügen, so ist die Abbildung für

$$|x| < \varrho = \min \left(R, \frac{|d|}{n^n \alpha^{n-1} \beta} \right)$$

eineindeutig, und die Bildpunktmenge enthält (*bei der Normierung* $y_i(0) = 0$) *die volle Kugel*

$$|y| < \frac{|d|\varrho}{8n^{n-1}\alpha^{n-1}}.$$

Hierbei ist $|d|$ *der Betrag der Funktionaldeterminante*

$$\left\| \frac{\partial y_i(0)}{\partial x_k} \right\|.$$

FINNISCHE AKADEMIE, HELSINKI