Über das Gaußsche Fehlergesetz.

Von

Felix Bernstein in Halle a. S.

Teil I.

Das lineare Fehlergesetz.

1.

Einleitung.

Die Funktionalgleichung

$$f(x+y) = f(x) + f(y) \tag{1}$$

tritt in verschiedenen Gebieten der Mathematik auf und ist schon seit langer Zeit Gegenstand der Untersuchung gewesen. Zuerst hat man wohl ihre Bedeutung für den Beweis des Satzes vom *Parallelogramm der Kräfte* erkannt. Später zeigte sich, daß der Beweis des *Fundamentalsatzes der projektiven Geometrie* die Behandlung der Gleichung (1) erfordert. Endlich besitzt sie eine unmittelbare Bedeutung für den *axiomatischen Aufbau der Arithmetik.*

Cauchy*) hat zuerst bewiesen, daß die lineare Funktion Ax die einzige *stetige* Lösung der Gleichung (1) ist. Die Methode des Beweises stammt schon aus dem Altertum, und ich glaube, es besteht eine interessante historische Kontinuität. Bereits Archimedes**) hat in seinem berühmten Beweise des Hebelgesetzes eine der unsrigen verwandte Funktionalgleichung exakt gelöst.

In der Tat kommt seine Methode, das Gleichgewicht des Hebels mit ungleichen Armen auf das Gleichgewicht des gleicharmigen Hebels zurückzuführen, darauf hinaus, eine Funktionalgleichung

$$f(x+y) = f(x)\left(1 + f\left(\frac{y}{x}\right)\right) \tag{2}$$

*) Analyse algébrique.

**) Wir wollen nicht verfehlen, auf die beredte und glänzende Schilderung zu verweisen, welche Duhem (Les origines de la statique) von der geistigen Eigenart des großen Gelehrten gegeben hat, der vor allem ein Meister der Methode war.

unter der Bedingung $f(1)=1$ durch die Funktion $f(x)=x$ als die stetige Lösung derselben zu befriedigen.

G. Darboux*) hat ferner gezeigt, daß die Annahme einer endlichen *oberen Grenze* g für die Funktion $f(x)$ der Gleichung (1) in irgend einem Intervall ausreicht, die Stetigkeit derselben zu erschließen und damit den linearen Verlauf nachzuweisen.

Endlich hat G. Hamel**) neuerdings festgestellt, daß auch unstetige Lösungen der Funktionalgleichung (1) existieren, falls man annimmt, daß die Wohlordnung des Kontinuums möglich sei. Solche Funktionen würden in der Nähe jeder Stelle jedem Werte beliebig nahe kommen, so daß diejenigen Punkte der Ebene, welche die Funktion darstellen, die ganze Ebene überalldicht ausfüllen müssen. Irgend eine, diese letztere Eigenschaft ausschließende Annahme würde also hinreichen, die stetige Lösung zur *einzigen* zu machen. Die hierin liegende Erweiterung der Darbouxschen Bedingung läßt sich, worauf es wesentlich ankommt, leicht völlig elementar zeigen. Da nämlich, unter α und β *rationale* Zahlen verstanden, aus der Funktionalgleichung (1) stets

$$f(a\alpha)=\alpha f(a), \quad f(b\beta)=\beta f(b)$$

für irgend welche Werte a und b folgt, so würde die Annahme einer nichtlinearen Lösung zur Folge haben, daß wir $\frac{f(b)}{b}$ als verschieden von $\frac{f(a)}{a}$ voraussetzen können. Dann können wir α und β jedoch so wählen, daß die Größen

$$\alpha a+\beta b$$

$$\alpha f(a)+\beta f(b)$$

vorgeschriebenen Werten beliebig nahe kommen.

Die Funktionalgleichung (1) tritt nun auch in der *Wahrscheinlichkeitsrechnung* auf. Es bedient sich C. F. Gauß***) derselben bei Gelegenheit der ersten Begründung seines Fehlergesetzes. Indessen führt die Annahme, auf welcher die Ableitung des Fehlergesetzes beruht, nämlich die berühmte Hypothese „des arithmetischen Mittels", nicht unmittelbar auf jene Funktionalgleichung. In der Tat ist ja die zu bestimmende Funktion hier eine Exponentialfunktion, deren Exponent eine quadratische Funktion ist. Erst die Ableitung des negativ genommenen Logarithmus genügt der Funk-

*) G. Darboux, 1) Sur la composition des forces en statique. Bull. des Scienc. Math. 9 (1875); 2) „Sur le théorème fondamental de la géométrie projective". Math. Ann. 17 (1886), p. 58.

**) G. Hamel, Eine Basis aller Zahlen und die unstetigen Lösungen der Funktionalgleichung $f(x+y)=f(x)+f(y)$. Math. Ann. 60, p. 459. — Siehe auch die Bemerkung am Schluß dieser Arbeit (S. 447).

***) C. F. Gauß, Theoria motus etc. 1809, art. 175 u. 179 (Werke 7).

tionalgleichung (1). Der negativ genommene Logarithmus der Funktion, auf dessen Bestimmung es ankommt, genügt der Funktionalungleichung

$$(3)\quad \varphi(x_1)+\varphi(x_2)+\cdots+\varphi(x_n)<\varphi(x_1+\varepsilon)+\varphi(x_2+\varepsilon)+\cdots+\varphi(x_n+\varepsilon).$$

Hierin bedeute ε *irgend* eine von Null verschiedene Größe und es seien $x_1, x_2, \cdots, x_n$ *irgend* welche Argumente, welche durch die Gleichung

$$(4)\qquad x_1+x_2+\cdots+x_n=0$$

verbunden sind.

Die zu bestimmende Funktion $\varphi(x)$ werde als eine für alle reellen Werte von x eindeutig definierte, endliche Funktion angenommen.

Es würde, wie bemerkt sei, genügen, statt n lediglich drei Argumente in Betracht zu ziehen.

Die Funktionalbedingung (3) wird von Gauß unter der stillschweigenden Annahme gelöst, daß die Funktion $\varphi(x)$ *stetig und überdies zweimal differenzierbar sei.*

So bereitwillig man für die Naturgesetze den analytischen Charakter derselben zuzugestehen geneigt ist, so wird man zugeben, daß dies in keinem Falle für die Gesetze des Zufalls angängig ist. Hier wird man wünschen, ausschließlich mit denjenigen Annahmen auszukommen, welche sich aus der Natur der Wahrscheinlichkeitsgrößen ergeben.

Die Prüfung des Sachverhalts ergibt nun, daß dies in der Tat möglich ist.

Die Untersuchung der Funktionalbedingung (3) lehrt zwar, daß auch hier neben der stetigen Lösung durch die *quadratische* Funktion unstetige Lösungen existieren können. Indessen genügt auch hier schon die Darbouxsche Bedingung, um die unstetigen Lösungen auszuschließen.

In einer Note: *Über eine Funktionalgleichung und eine erweiterte Begründung des Gaußschen Fehlergesetzes* (Ber. d. Kgl. Sächs. Ges. d. Wiss., Math.-Phys. Kl., Leipzig, LVIII, 1906, 23. Juli) habe ich dieses Resultat bewiesen. Die dort gegebene Darstellung bedarf der Berichtigung, indem die Darbouxsche Bedingung, von welcher der Beweis Gebrauch macht, den Voraussetzungen ausdrücklich hinzugefügt werden muß, was dort nicht geschehen ist.

Für die Wahrscheinlichkeitsrechnung ist nun aber, ähnlich wie es beim Fundamentalsatz der projektiven Geometrie liegt, eine andere Zusatzbedingung von vornherein gegeben, aus der gleichfalls der vollständige Beweis sich ergibt. Es ist die Forderung der *Integrierbarkeit* der Fehlerfunktion $\overline{\varphi}(x)=e^{-\varphi(x)}$.

In der Tat bezieht sich ja die Frage der *geometrischen Wahrscheinlichkeit* auf das Vorkommen eines Fehlers in einem bestimmten *Intervall*, und es wird dieselbe durch das über dies Intervall erstreckte Integral definiert. Und es gehört dann weiter zur vollständigen Bestimmung der Fehlerfunktion noch die Annahme, *daß die Summe der Wahrscheinlichkeiten gleich Eins sei.* Ich erinnere daran, daß auch der Satz von der zusammengesetzten Wahrscheinlichkeit zu den bei der Ableitung des Fehlergesetzes benutzten Postulaten gehört.

Zu diesen allgemeinen Wahrscheinlichkeitspostulaten kommt dann nur noch ausschließlich die Hypothese des *arithmetischen Mittels* hinzu, um die Form des Fehlergesetzes zu begründen. Wir formulieren das hier festgestellte abschließende Ergebnis folgendermaßen:

(A) *Gemäß den Postulaten der geometrischen Wahrscheinlichkeit ist die Fehlerfunktion eine integrierbare Funktion derart, daß*

$$\int_{-\infty}^{+\infty} \overline{\varphi}(x)\, dx = 1$$

ist.

Setzen wir weiter

(B) *die Hypothese des arithmetischen Mittels*

voraus, so ist die Form des Fehlergesetzes vollständig bestimmt.

Da sich die Funktionalbedingung (3) in jeder Hinsicht als die nächsthöhere Stufe der Betrachtung im Verhältnis zur Funktionalgleichung (1) erweist, so war noch weiter zu untersuchen, welche Rolle hier die Zusatzbedingung spielt, insbesondere ob unter Weglassung derselben unstetige Lösungen existieren.

Dies ist in der Tat der Fall, aber die Funktionswerte füllen hier nicht die ganze Ebene aus, sondern sie erfüllen denjenigen Teil der positiven Halbebene, der oberhalb einer *stetigen* Lösung der Funktionalbedingung (3) liegt, überalldicht. Es genügt hier also schon als Zusatzbedingung die geringere Forderung, daß irgend ein Wert an irgend einer Stelle, der größer als die untere Grenze der Funktion in der Umgebung dieser Stelle ist, nicht beliebig angenähert werden könne.

Dieses Resultat wird begründet, nachdem vorher im zweiten Teile eine Verallgemeinerung auf n Variable vorgenommen worden ist. Dieselbe wird an sich durch die Bedürfnisse der Wahrscheinlichkeitsrechnung ebenfalls gefordert.

Die Funktionalbedingung für mehrere Variable besitzt als einzige Lösung eine *definite quadratische Form,* wenn noch eine entsprechende Zusatzbedingung als erfüllt angenommen wird.

Es scheint mir diese Einführung der definiten quadratischen Form durch ein Funktionaltheorem infolge ihrer Allgemeinheit besonderes Interesse zu beanspruchen, und ich möchte dasselbe schlechthin als das *Funk-*

tionaltheorem der definiten quadratischen Formen bezeichnen. Es bildet das wesentlichste Ergebnis der Untersuchung.

Die Allgemeinheit der durch dasselbe ausgedrückten Minimaleigenschaft läßt sich an dem Umstand ermessen, daß die Funktionen, welche hier zum Vergleich mit der Minimalfunktion herangezogen werden, eine Menge von der Mächtigkeit $2^{\mathfrak{c}}$ bilden, während man es in der Variationsrechnung stets mit einer Menge von Vergleichsfunktionen der Mächtigkeit $\mathfrak{c}$ zu tun hat.

2.

Beweis des Fehlergesetzes unter der Annahme, daß die Funktion $\varphi(x)$ unterhalb einer endlichen Grenze g liegt.

Es sei $\varphi(x)$ eine eindeutige Funktion der reellen Veränderlichen x, welche für *alle* x die Bedingung

$$\varphi(x) < g \tag{5}$$

erfülle, wo g ein irgendwie beschaffener endlicher Wert sei. Es sei ferner ε eine von Null verschiedene beliebige Größe. Es gilt dann das folgende Theorem:

Theorem I. *Die Funktionalbeziehung*

$$\varphi(x_1) + \varphi(x_2) + \cdots + \varphi(x_n) < \varphi(x_1 + \varepsilon) + \varphi(x_2 + \varepsilon) + \cdots + \varphi(x_n + \varepsilon), \tag{6}$$

die unter der Nebenbedingung

$$x_1 + x_2 + \cdots + x_n = 0 \tag{7}$$

für alle definierten Werte der Argumentgrößen erfüllt sein soll, besitzt als einzige Lösung die Funktion

$$\varphi(x) = h^2x^2 + \varphi(0), \tag{8}$$

wenn die Existenz der oberen Grenze g vorausgesetzt wird.

Es würde genügen, $n = 3$ anzunehmen, und es dürfte auch möglich sein, die Variabilität von $x_1, x_2, \cdots, x_n$ und ε noch weiter einzuschränken, als es hier für unseren Zusammenhang geboten ist.

Wir bedürfen für den Beweis folgender Definitionen und Sätze.

Unter einer *konvexen Funktion des Intervalles* (gh) verstehen wir eine eindeutige Funktion $f(x)$, welche für irgend zwei Werte x und y des Intervalles die Bedingung

$$f\left(\frac{x+y}{2}\right) \leqq \frac{f(x)+f(y)}{2} \tag{9}$$

erfüllt. Wir wollen das Intervall (gh) unter Einschluß seiner Endpunkte g und h mit Osgood als *abgeschlossenes* Intervall bezeichnen und uns auf ein Intervall dieser Art durchweg beziehen, ohne stets die vollständige Bezeichnung zu gebrauchen. Umgekehrt gebrauchen wir die Ausdrücke *rechts offenes, links offenes* und schlechthin *offenes* Intervall, um in leicht verständlicher Weise die anderen Fälle zu bezeichnen.

Es besteht zunächst der folgende Hilfssatz:

Hilfssatz 1. *Eine konvexe Funktion $f(x)$ des Intervalles (gh) ist stetig und besitzt an jeder Stelle des Intervalles einen endlichen vorderen und hinteren Differentialquotienten $f'_-(x)$ und $f'_+(x)$ derart, daß*

$$f'_-(x) \leqq f'_+(x)$$

ist, sobald $f(x)$ in dem Intervall eine obere Grenze g besitzt.

Dieser Satz ist von J. L. W. V. Jensen*) aufgestellt und bewiesen worden. Wir geben den Beweis des Satzes, der für das Folgende grundlegend ist, ausführlich wieder.

Aus (9) folgt für irgend welche Argumente x_1, x_2, x_3, x_4 des Intervalles

$$f\left(\frac{x_1+x_2+x_3+x_4}{4}\right) \leqq \frac{1}{2} f\left(\frac{x_1+x_2}{2}\right) + \frac{1}{2} f\left(\frac{x_3+x_4}{2}\right)$$
$$\leqq \frac{1}{4} f(x_1) + \frac{1}{4} f(x_2) + \frac{1}{4} f(x_3) + \frac{1}{4} f(x_4)$$

und allgemein, mittels Schlusses von $m-1$ auf m,

$$2^m f\left(2^{-m} \sum_{\nu=1}^{2^m} x_\nu\right) \leqq \sum_{\nu=1}^{2^m} f(x_\nu).$$

Wir setzen, unter n eine positive ganze Zahl verstanden, $2^m > n$ und wählen

$$x_{n+1} = x_{n+2} = \cdots = x_{2^m} = \frac{x_1+x_2+\cdots+x_n}{n},$$

so daß wir die Beziehung

$$2^m f\left(\frac{1}{n} \sum_{\nu=1}^{n} x_\nu\right) \leqq \sum_{\nu=1}^{n} f(x_\nu) + (2^m - n) f\left(\frac{1}{n} \sum_{\nu=1}^{n} f(x_\nu)\right)$$

erhalten. Dieselbe geht über in die *verallgemeinerte* Relation

$$(10) \qquad f\left(\frac{1}{n} \sum_{\nu=1}^{n} x_\nu\right) \leqq \frac{1}{n} \sum_{\nu=1}^{n} f(x_\nu).$$

Dieser Übergang von der mit Potenzen von 2 behafteten Beziehung zu einer allgemeineren findet sich in einem entsprechenden Gedankengang in dem angeführten Beweis des Archimedes.

Wir betrachten jetzt zwei Größen x und $x + n\delta$ des Intervalles und verstehen unter m eine natürliche Zahl, die kleiner als n ist. Wählen wir nun

*) J. L. W. V. Jensen, Sur les fonctions convexes et les inégalités entre leurs valeurs moyennes. Acta Math. 30, p. 175—201.

$$x_1 = x_2 = \cdots = x_m = x + n\delta;\; x_{m+1} = \cdots = x_n = x,$$

so haben wir

$$f(x+m\delta) \leqq \frac{m}{n} f(x+n\delta) + \frac{n-m}{n} f(x)$$

oder

$$\frac{f(x+m\delta) - f(x)}{m} \leqq \frac{f(x+n\delta) - f(x)}{n}.$$

Hieraus folgt, indem $-\delta$ statt δ gewählt wird,

$$\frac{f(x-m\delta) - f(x)}{m} \leqq \frac{f(x-n\delta) - f(x)}{n},$$

falls auch $x - n\delta$ im Intervall gelegen ist.

Bedenken wir, daß infolge der Definition (19)

$$f(x) - f(x-m\delta) \leqq f(x+m\delta) - f(x)$$

ist, so können wir die fortlaufende Beziehung schreiben:

$$(11) \quad \frac{f(x) - f(x-n\delta)}{n} \leqq \frac{f(x) - f(x-m\delta)}{m} \leqq \frac{f(x+m\delta) - f(x)}{m} \leqq \frac{f(x+n\delta) - f(x)}{n},$$

wo also

$$m < n$$

angenommen ist.

Aus (11) schließen wir jetzt, daß $f(x)$ stetig ist, wenn im ganzen Intervall $f(x) < g$ bleibt. Wir setzen hierzu $m = 1$ und ersetzen $f(x-n\delta)$ und $f(x+n\delta)$ durch g. Es folgt

$$\frac{f(x) - g}{n} \leqq f(x) - f(x-\delta) \leqq f(x+\delta) - f(x) \leqq \frac{g - f(x)}{n}.$$

Diese Formel gilt, sobald $x \pm n\delta$ im gegebenen Intervall liegt. Lassen wir jetzt δ gegen Null sinken und n gegen Unendlich so wachsen, daß diese Bedingung erfüllt bleibt, so folgt offenbar, daß

$$\lim_{\delta = 0} [f(x+\delta) - f(x)] = 0$$

ist.

Um weitere Schlüsse zu ziehen, setzen wir $\frac{\delta}{n}$ statt δ in (11) ein, und nehmen δ *positiv*, es folgt dann, falls $x \pm n \cdot \frac{\delta}{n}$ im Intervall liegt,

$$(12) \quad \frac{f(x) - f(x-\delta)}{\delta} \leqq \frac{f(x) - f\left(x - \frac{m}{n}\delta\right)}{\frac{m}{n}\delta} \leqq \frac{f\left(x + \frac{m}{n}\delta\right) - f(x)}{\frac{m}{n}\delta} \leqq \frac{f(x+\delta) - f(x)}{\delta}.$$

Bedenkt man, daß $\frac{m}{n}$ ein beliebiger echter Bruch und daß $f(x)$ *stetig* ist, so folgt, daß $\frac{f(x) - f(x-\delta)}{\delta}$ bei abnehmendem δ ständig wächst, oder konstant bleibt; andererseits bleibt der Quotient ständig unterhalb $\frac{f(x+\delta')-f(x)}{\delta'}$, wenn δ' irgend ein positiver Wert ist. Es existiert daher

$$\lim_{+\delta=0} \frac{f(x)-f(x-\delta)}{\delta} = f'_-(x)$$

und aus gleichen Gründen

$$\lim_{+\delta=0} \frac{f(x+\delta)-f(x)}{\delta} = f'_+(x)$$

und es ist zugleich

$$f'_-(x) \leqq f'_+(x). \tag{13}$$

Wir beweisen ferner den

Hilfssatz 2. *Für eine stetige nichtlineare konvexe Funktion ist bei allen Vorzeichenkombinationen*

$$f'_\pm(x) < f'_\pm(y), \tag{14}$$

wenn

$$x < y$$

ist.

Zunächst bemerken wir, daß das Gleichheitszeichen in (9), sobald es sich um eine stetige Funktion handelt, *nur* für eine lineare Funktion gilt. Wir haben daher in allen Gleichungen, die keinen Grenzübergang enthalten, das Gleichheitszeichen zu streichen. Aus (12) folgt, indem wir $\frac{m}{n}$ gegen eine zwischen Null und Eins enthaltene Zahl α konvergieren lassen, zunächst

$$\frac{f(x)-f(x-\delta)}{\delta} \leqq \frac{f(x)-f(x-\alpha\delta)}{\alpha\delta}.$$

Wir bemerken aber, daß das Gleichheitszeichen fortfallen muß, indem wir $\frac{m}{\delta}\delta$ statt δ und $\frac{n}{m}\alpha$ statt α setzen und (12) noch einmal berücksichtigen. Hierbei nehmen wir $1 > \frac{m}{n} > \alpha$ an. Analog erhalten wir schließlich

$$\frac{f(x)-f(x-\delta)}{\delta} < \frac{f(x)-f(x-\delta')}{\delta'} < \frac{f(x+\delta')-f(x)}{\delta'} < \frac{f(x+\delta)-f(x)}{\delta},$$

oder allgemein

$$\frac{f(x)-f(x-\delta_1)}{\delta_1} < \frac{f(x+\delta_2)-f(x)}{\delta_2}$$

für irgend drei Werte $x-\delta_1 < x < x+\delta_2$ des Intervalles. Es ist daher, wenn $x_1 < x_2 < x_3 < x_4$ ist,

$$\frac{f(x_2)-f(x_1)}{x_2-x_1} < \frac{f(x_3)-f(x_2)}{x_3-x_2} < \frac{f(x_4)-f(x_3)}{x_4-x_3}$$

Infolgedessen ist, für die Werte

$$x < x + \delta < y - \delta < y,$$

$$\frac{f(x+\delta) - f(x)}{\delta} < \frac{f(y) - f(y-\delta)}{\delta}.$$

Da nun für $0 < \delta' < \delta$

$$\frac{f(x+\delta') - f(x)}{\delta'} < \frac{f(x+\delta) - f(x)}{\delta}$$

infolge (12) ist, und ebenso

$$\frac{f(y) - f(y-\delta)}{\delta} < \frac{f(y) - f(y-\delta')}{\delta'}$$

wird, so folgt für $\lim \delta' = 0$

$$f'_+(x) < f'_-(y),$$

und hieraus in Verbindung mit (13) die Beziehung (14).

Der geometrische Sinn der Sätze 1 und 2 ist aber der, daß die konvexe Kurve, falls sie im Intervall unterhalb einer Geraden verläuft, stetig ist, eine Tangente besitzt, und daß sich diese beim Fortschreiten auf der Kurve im Sinne des wachsenden x selbst in positivem Sinne dreht. Dabei kann sich die Drehung an einer gewissen Zahl von Ecken sprungweise vollziehen. Diese Ecken können, wie Jensen durch ein Beispiel zeigt, überalldicht liegen.*)

Wir zeigen des Ferneren, daß die Menge der Ecken abzählbar ist. Es gehört die höchstens zweiwertige Funktion $f'(x)$ zu den punktweise unstetigen. Wir formulieren den

Hilfssatz 3. *Eine konvexe Funktion $f(x)$ besitzt höchstens eine abzählbare Anzahl von Ecken, d. h. solchen Stellen, wo*

$$f'_-(x) < f'_+(x)$$

ist.

Wir bezeichnen $f'_+(x) - f'_-(x)$ mit δ_x. Wie klein wir auch die positive Größe ε wählen, es gibt nur eine *endliche* Zahl n von Stellen, an denen $\delta_x > \varepsilon$ im Intervall $(g'h')$ ist, wo g' und h' irgend ein Teilintervall von (gh) bestimmen. In der Tat ist ja $f'(x)$ eine ständig wachsende Funktion und also höchstens in g negativ und höchstens in h positiv unendlich, sonst aber durchweg endlich, so daß die Differenz

$$f'_+(h') - f'_-(g') = J$$

eine bestimmte positive Größe ist. Dann aber ist offenbar n durch die Bedingung

$$n\varepsilon < n\,\delta_x < J \quad \text{oder} \quad n < \frac{J}{\varepsilon}$$

endlich beschränkt.

*) In einer Note, die im Archiv d. Math. u. Physik erscheinen wird, werde ich ein allgemeines Verfahren angeben, solche Kurven zu bilden.

In Ausdehnung auf das ganze Intervall (gh) sehen wir daher, daß die Anzahl der Stellen, an denen $\delta_x > \varepsilon$ ist, höchstens abzählbar unendlich ist.

Bedeuten jetzt

$$\delta_1,\ \delta_2,\ \cdots,\ \delta_n,\ \cdots$$

eine unendliche Reihe positiver, gegen Null abnehmender Größen, so ist, wenn $\varepsilon = \delta_n$ genommen wird, die Menge S_n der Stellen $\delta_x > \varepsilon$ abzählbar, also ist die Menge S aller Sprungstellen, welche das kleinste gemeinschaftliche Vielfache dieser Mengen ist, ebenfalls abzählbar. Die Grenzen g und h können auch $-\infty$ und $+\infty$ sein.

Nach diesen Vorbereitungen treten wir in den Beweis des Theorems I ein. Um zu einer *konvexen* Funktion zu gelangen, bemerken wir, daß die Ungleichheitsbedingung (6)

$$(6)\qquad \varphi(x_1)+\varphi(x_2)+\cdots+\varphi(x_n) < \varphi(x_1+\varepsilon)+\varphi(x_2+\varepsilon)+\cdots+\varphi(x_n+\varepsilon)$$

und die Nebenbedingung (7)

$$(7)\qquad x_1+x_2+\cdots+x_n = 0$$

ebenfalls für die *gerade* Funktion

$$(15)\qquad \psi(x) = \varphi(x)+\varphi(-x)$$

erfüllt sind, so daß also für diese

$$(16)\qquad \psi(x_1)+\psi(x_2)+\cdots+\psi(x_n) < \psi(x_1+\varepsilon)+\psi(x_2+\varepsilon)+\cdots+\psi(x_n+\varepsilon)$$

ist. Wir setzen $n=2$ und $x_1=-x_2$ und haben infolge von $\psi(x)=\psi(-x)$

$$(17)\qquad 2\psi(x_1) < \psi(x_1+\varepsilon)+\psi(x_1-\varepsilon).$$

Setzen wir jetzt

$$x_1 = \frac{x+y}{2},\qquad \varepsilon = \frac{x-y}{2}\ (>0)$$

so folgt

$$(18)\qquad 2\psi\left(\frac{x+y}{2}\right) < \psi(x)+\psi(y)$$

für

$$x \neq y.$$

Ferner folgt aus

$$\varphi(x) < g \quad \text{und} \quad \varphi(-x) < g$$

$$(19)\qquad \psi(x) < 2g.$$

Es ist daher $\psi(x)$ infolge (18) und (19) eine konvexe Funktion, welche für alle reellen x existiert und unterhalb einer festen Grenze liegt. Wir folgern daher aus Hilfssatz 1, daß $\psi(x)$ stetig ist und daß $\psi(x)$ eine ständig wachsende Ableitung $\psi'(x)$ hat, welche an höchstens einer abzählbaren Menge S von Stellen einen Sprung besitzt. Wir beweisen jetzt folgenden Satz:

Hilfssatz 4. *Es gibt eine Stelle a, welche mitsamt allen rationalen Vielfachen $\pm r \cdot a$ nicht Sprungstelle von $\psi'(x)$ ist.*

Wir wollen, um Ausnahmen zu vermeiden, a und r verschieden von Null annehmen.

In der Tat ist die Menge $\bar{S}$ aller rationalen Vielfachen der hypothetischen Sprungstellen ebenfalls abzählbar. *Es gibt daher nach dem Theorem von G. Cantor, welches besagt, daß das Kontinuum nicht abzählbar ist, wenigstens eine Stelle $a(\neq 0)$, welche nicht zu $\bar{S}$ gehört.* Das gleiche gilt dann offenbar auch für $\pm ra(r \neq 0)$.

Sind jetzt $x_1, x_2, \cdots, x_n$ solche Stellen, an denen die Funktion $\psi(x)$ einen *einzigen* Differentialquotienten besitzt, so besitzt die Funktion

$$\Psi(\varepsilon) = \psi(x_1+\varepsilon) + \psi(x_2+\varepsilon) + \cdots + \psi(x_n+\varepsilon) \tag{20}$$

als Funktion von ε an der Stelle $\varepsilon = 0$ den einzigen Differentialquotienten

$$[\Psi'(\varepsilon)]_{\varepsilon=0} = \psi'(x_1) + \psi'(x_2) + \cdots + \psi'(x_n). \tag{21}$$

Besitzt aber eine eindeutige stetige Funktion an einer Stelle ein Minimum und ist ein einziger Differentialquotient der rechts und links differenzierbaren Funktion vorhanden, so muß derselbe verschwinden. Infolgedessen ist, wenn die Gleichung

$$x_1 + x_2 + \cdots + x_n = 0$$

für die genannten Stellen gilt, auch

$$0 = \psi'(x_1) + \psi'(x_2) + \cdots + \psi'(x_n). \tag{22}$$

Zufolge Hilfssatz 4 erhalten wir aber für

$$x_1 = x_2 = \cdots = x_{n-1} = a, \quad x_n = -(n-1)a, \tag{23}$$

indem wir statt $n-1$ wieder n schreiben, für jedes positive ganze n

$$n\psi'(a) + \psi'(-na) = 0. \tag{24}$$

Offenbar gilt (24) ebenso für alle von Null verschiedenen rationalen Vielfachen von a. Setzen wir $n = 1$, so erhalten wir

$$\psi'(a) = -\psi'(-a)$$

und es geht daher (24) in die für alle ganzzahligen n geltende Formel

$$n\psi'(a) = \psi'(na) \quad \text{oder} \quad \psi'(a) = \frac{1}{n}\psi'(na)$$

über.

Ersetzen wir a durch $\frac{m}{n}a$, so folgt (für m und n verschieden von Null)

$$\psi'\left(\frac{m}{n}a\right) = \frac{1}{n}\psi'(ma) = \frac{m}{n}\psi'(a)$$

oder

$$\psi'(ra) = r\psi'(a), \tag{25}$$

wenn $r \neq 0$ irgend eine rationale Zahl ist.

Wir wissen jetzt andrerseits, daß gemäß Hilfssatz 3 $\psi'(x)$ eine *ständig wachsende* Funktion ist.

Infolgedessen ist für irgend ein irrationales x oder $x=0$, das zwischen den rationalen Zahlen r_1 und r_2 liegt,

$$\psi(r_1 a) < \psi(xa) < \psi(r_2 a)$$

oder

$$r_1 \psi(a) < \psi(xa) < r_2 \psi(a).$$

Es ist daher $\psi(xa)$ mit dem *Schnitt* $x\psi(a)$ der Wertmengen $(r_1\psi(a))$ und $(r_2\psi(a))$ identisch und wir haben für *alle* reellen x

$$\psi'(xa) = x\,\psi'(a)$$

oder

$$\psi'(x) = x\psi'(1). \tag{26}$$

Außerdem ist, da die Funktion beständig wächst, $\psi'(1)$ positiv. Wir setzen

$$\psi'(x) = 4h^2 x$$

und haben also

$$\psi(x) = 2h^2x^2 + \psi^*(x),$$

wo $\psi^*(x)$ eine Funktion ist, deren Differentialquotient verschwindet. Da $\psi^*(x)$ stetig ist und die Ableitung Null besitzt, so ist $\psi^*(x)$ eine Konstante, und wir setzen also

$$\psi(x) = 2h^2x^2 + \psi(0). \tag{27}$$

Um die Schreibweise zu vereinfachen, setzen wir zunächst $h=1$, $\psi(0)=0$ und haben

$$\psi(x) = 2x^2. \tag{28}$$

Um jetzt zu zeigen, daß $\varphi(x) = x^2$ ist, beweisen wir den Satz:

Hilfssatz 5. *Es ist $\varphi(x)$ stetig und $2\varphi'(x) = \psi'(x)$.*

Beweis. Aus

$$\varphi(x) + \varphi(-x) < \varphi(x+\varepsilon) + \varphi(-x+\varepsilon)$$

folgt durch Addition und Subtraktion von $\varphi(x-\varepsilon)$ auf der rechten Seite

$$\psi(x) < \psi(x-\varepsilon) + \varphi(x+\varepsilon) - \varphi(x-\varepsilon).$$

Indem wir das Vorzeichen von ε vertauschen, erhalten wir

$$\psi(x) < \psi(x+\varepsilon) + \varphi(x-\varepsilon) - \varphi(x+\varepsilon)$$

oder

$$\psi(x) - \psi(x-\varepsilon) < \varphi(x+\varepsilon) - \varphi(x-\varepsilon) < \psi(x+\varepsilon) - \psi(x). \tag{29}$$

Hieraus folgt sofort, daß $\varphi(x)$ *stetig* ist. Um die Ableitung herzustellen, dividieren wir durch ε und erhalten

$$\frac{\psi(x) - \psi(x-\varepsilon)}{\varepsilon} < \frac{\varphi(x+\varepsilon) - \varphi(x-\varepsilon)}{\varepsilon} < \frac{\psi(x+\varepsilon) - \psi(x)}{\varepsilon}.$$

Hieraus folgt, indem wir $x+\varepsilon$ oder $x-\varepsilon$ gleich y setzen und berücksichtigen, daß $\pm 2\varepsilon = \eta$ der Zuwachs von y wird,

$$\psi'(y) = 2\varphi'_{\pm}(y). \tag{30}$$

Also ist in der Tat mit Rücksicht auf $\psi'(y) = 2y$

$$\varphi(y) = y^2.$$

Haben wir allgemeiner

$$\psi(x) = 2h^2x^2 + \psi(0),$$

so folgt

$$\varphi(x) = h^2x^2 + \varphi(0) \qquad \text{q. e. d.} \tag{31}$$

3.

Ein Hilfssatz über konvexe Funktionen.

Die Bedingung, daß $\varphi(x)$ unter einer endlichen Grenze liegt, besitzt vom Standpunkte der Wahrscheinlichkeitsrechnung aus keine natürliche Motivierung. Wir haben, wie auseinandergesetzt, dort als natürliche Bedingung die gewonnen, daß $e^{-\varphi(x)}$ *integrierbar* ist. Ich lasse dahingestellt, durch welchen Zusatz man von der ersten Bedingung zur zweiten übergehen kann. Als natürlicher Weg zur Erledigung der Frage ergibt sich hier der, die Analogie zu den im Falle der Funktionalgleichung

$$f(x+y) = f(x) + f(y) \qquad (f(x) \text{ integrierbar})$$

auftretenden Verhältnissen aufzusuchen. In dem einfacheren Falle würde man sich an die für alle rationalen Vielfachen rx von x bestimmte *Teillösung* $rf(x) = f(rx)$ halten. Die Lösung ist *gleichmäßig* stetig und infolgedessen definiert sie eine stetige lineare Funktion, deren Integral mit dem Integral von $f(x)$ übereinstimmt. Hieraus schließt man dann leicht, daß $f(x)$ selbst linear ist. Genau so werden wir zeigen, daß stets für eine gewisse abzählbare überalldichte Argumentmenge eine *Teillösung* existiert, welche in dieser Menge *gleichmäßig* stetig ist, und also eine stetige Funktion definiert, deren Integral mit dem Integral der Funktion verglichen wird. Es ergibt sich dann leicht der Schluß auf die Form der Fehlerfunktion.

Die Existenz der Teillösung zeigen wir mittels eines Hilfssatzes über konvexe Funktionen, der zunächst unter der Voraussetzung, daß die Funktion *oberhalb* einer festen Grenze liegt, bewiesen wird.

Wir transformieren ein beliebiges Intervall $(a\,b)$ auf das Intervall $(0\;1)$, indem wir $f(a+x(b-a))$ statt $f(x)$ betrachten. Wir setzen

$$g = \text{Max.}\,(f(0),\, f(1)). \tag{32}$$

Bedeutet dann $r = \frac{m}{n}$ eine rationale Zahl zwischen Null und Eins, so gilt der Satz:

Hilfssatz 6a. *Es ist, wenn $f(x)$ eine konvexe Funktion ist,*

$$f(r) \leqq g \qquad \left(r = \frac{m}{n},\ 0 \leqq m \leqq n\right). \tag{33}$$

Beweis. Aus der Formel (10)

$$f\left(\frac{1}{n}\sum_{\nu=1}^{n} x_\nu\right) \leqq \frac{1}{n}\sum_{\nu=1}^{n} f(x_\nu)$$

folgt, wenn wir

$$x_1 = x_2 = \cdots = x_{n-m} = 0, \quad x_{n-m+1} = \cdots = x_n = 1$$

setzen,

$$f\left(\frac{m}{n}\right) \leqq \frac{1}{n} \cdot ng \leqq g,$$

wenn wir mit g das Maximum von $f(0)$ und $f(1)$ bezeichnen.

Es besteht nun der folgende

Hilfssatz 6b. *Es sei $f(x)$ eine positive konvexe Funktion des Intervalles $(0\,1)$, so ist $f(r)$ für alle rationalen r des Teilintervalles $(\eta \leqq r \leqq 1-\eta)$ eine gleichmäßig stetige Funktion und zwar bei noch so kleinem positivem η, d. h. es läßt sich bei vorgegebenem ε eine rationale Größe $\delta > 0$ durch die Bedingung $\left[\frac{g}{\varepsilon} + 1\right]\delta < \eta$ so bestimmen, daß*

$$|f(r+\delta) - f(r)| < \varepsilon \tag{34}$$

für jedes r des Teilintervalles ist.

Beweis. Wir bezeichnen die ganze Zahl $\left[\frac{g}{\varepsilon} + 1\right]$ mit n und haben also, da g *positiv* ist,

$$\left[\frac{g}{\varepsilon} + 1\right] = n, \qquad n > \frac{g}{\varepsilon}, \tag{35}$$

und andrerseits, infolge

$$\left[\frac{g}{\varepsilon} + 1\right]\delta < \eta, \tag{36}$$

$$n\delta < \eta. \tag{37}$$

Infolgedessen ist sowohl $r - n\delta$ (da $\eta < r$ ist) größer als Null, als auch $r + n\delta$ kleiner als 1. Wir können daher die Formel (11), in der wir $m = 1$ setzen, zur Anwendung bringen und erhalten

$$\frac{f(r) - f(r - n\delta)}{n} \leqq f(r+\delta) - f(r) \leqq \frac{f(r+n\delta) - f(r)}{n}. \tag{38}$$

Sämtliche Argumente liegen zwischen Null und Eins; da also gemäß dem Hilfssatz 6a für jedes r' zwischen Null und Eins

$$0 \leqq f(r) \leqq g$$

ist, so folgt aus (38)

$$-\frac{g}{n} \leqq f(r+\delta) - f(r) \leqq +\frac{g}{n}, \tag{39}$$

also, da infolge von (35) $\frac{g}{n} < \varepsilon$ ist

$$-\varepsilon \leqq f(r+\delta) - f(r) \leqq +\varepsilon \tag{40}$$

für jedes r und δ, die den Bedingungen

$$\eta \leqq r \leqq 1-\eta, \quad \left[\frac{g}{n}+1\right]\delta < \eta \tag{41}$$

genügen, womit der Satz bewiesen ist.

Zu einer solchen gleichmäßig stetigen Funktion einer überalldichten Punktmenge gehört eine für *alle* Werte $\eta \leqq x \leqq 1-\eta$ *stetige* Funktion hinzu. Dieselbe kann, da η beliebig klein ist, auf das ganze Intervall ausgedehnt werden. Wir formulieren den

Satz 7. *Es gibt zu einer positiven konvexen Funktion eine zugehörige stetige Funktion, welche mit derselben in den Endpunkten a und b eines Intervalls, sowie in sämtlichen rationalen Teilpunkten desselben übereinstimmt. Diese stetige Funktion ist konvex und durch die auferlegte Bedingung eindeutig bestimmt.*

Hierbei verstehen wir unter den rationalen Teilpunkten des Intervalls diejenigen, welche das Intervall in rationalem Verhältnis teilen. Es ist ohne weiteres klar, daß die Teillösung, sowie die zugehörige stetige Funktion auch in die äußeren rationalen Teilpunkte fortgesetzt werden kann. Denn, wenn wir a und b als äußere Teilpunkte so annehmen, daß das ursprüngliche Intervall jetzt eingeschlossen wird, so wird die zugehörige Teillösung die frühere Teillösung infolge der Eindeutigkeit der Bestimmung derselben enthalten. Jede Teillösung läßt sich für das ganze Definitionsintervall der konvexen Funktion aufstellen und kann nur an den Grenzen desselben unendlich werden.

Zusatz. Um jetzt die Einschränkung zu beseitigen, welche in der Annahme liegt, daß $f(x)$ positiv sein soll, teilen wir das Intervall $(0\,1)$ in n gleiche Teile von der Länge $\delta = \frac{1}{n}$. Setzen wir

$$\Delta_m = f(m\delta) - f((m-1)\delta),$$

so ist infolge der Definitionsgleichung

$$\Delta_1 \leqq \Delta_2 \leqq \cdots \leqq \Delta_m \leqq \cdots \leqq \Delta_n. \tag{42}$$

Es ist also

$$m\Delta_1 \leqq \Delta_1 + \Delta_2 + \cdots + \Delta_m$$

oder, infolge von

$$f(m\delta) = f(0) + \Delta_1 + \Delta_2 + \cdots + \Delta_m,$$

$$f(0) + \frac{m}{n}\,\frac{f(\delta)-f(0)}{\delta} \leqq f\left(\frac{m}{n}\right). \tag{43}$$

In der Reihe (42) kann nur an einer Stelle ein Übergang von negativen zu positiven Werten erfolgen. Indem wir n gegen Unendlich wachsen lassen, erhalten wir höchstens *einen* Schnitt x im Intervall Null bis Eins derart, daß $f(r)$ für alle $0 \leqq r < x$ ständig abnimmt, und für alle $x < r \leqq 1$ ständig zunimmt. Wenn es keinen Schnitt x gibt, so muß $f(0)$ oder $f(1)$ der kleinste Wert von $f(r)$ sein, so daß wir für $f(r)$ die gesuchte untere Grenze bereits haben. Existiert aber der Schnitt, so sei $\frac{m'}{n'}$ irgend ein Wert kleiner als x. Dann ist im Intervall 0 bis $\frac{m'}{n'}$ stets $f(r)$ größer als $f\left(\frac{m'}{n'}\right)$. Es existiert daher in diesem Intervall eine Teillösung, deren zugehörige stetige Lösung einen ständig wachsenden rechten Differentialquotienten besitzt. Derselbe besitzt also in einem rationalen Punkte p zwischen 0 und $\frac{m'}{n'}$ einen negativen endlichen Wert k. Wir denken uns, um die Teillösung fortzusetzen, p als Endpunkt des Intervalles $(p\ 1)$, wollen aber, um die Schreibweise zu vereinfachen, wieder $p = 0$ setzen, und also annehmen, es existiere die Teillösung in einem in Null beginnenden Intervall, und es habe dort die zugehörige stetige Lösung einen rechten Differentialquotienten $f'_+(0) = k$. Es ist dann

$$k \leqq \frac{f(\delta) - f(0)}{\delta}. \tag{44}$$

Aus (43) und (44) folgt

$$f(0) + \frac{m}{n} k \leqq f\left(\frac{m}{n}\right). \tag{45}$$

Da die linke Seite für $0 \leqq \frac{m}{n} \leqq 1$ offenbar eine untere Grenze besitzt, so ist also festgestellt, daß auch $f\left(\frac{m}{n}\right)$ stets oberhalb einer endlichen Grenze h liegt. Die Funktion $f(r) - h$ ist also eine positive konvexe Funktion. Wir erhalten daher den

Satz 7b. *Irgend zwei Punkte einer konvexen Funktion bestimmen eindeutig eine Teillösung, deren Punkte überalldicht auf einer stetigen konvexen Kurve liegen.*

4.

Beweis des Gaußschen Fehlergesetzes.

Wir beweisen jetzt das Gaußsche Fehlergesetz unter den Annahmen:

(A) *Das arithmetische Mittel ist der wahrscheinlichste Wert.*

(B) *Die Summe der Fehlerwahrscheinlichkeiten ist* 1.

Aus (B) entnehmen wir, daß $e^{-\varphi(x)}$ integrierbar ist, so daß

$$1 = \int_{-\infty}^{+\infty} e^{-\varphi(x)} dx$$

wird.

Es seien jetzt u und v irgend zwei Argumentwerte. Die Ungleichheitsbedingung (3) gibt für $x_1 = 0$ und jedes ε

$$\psi(\varepsilon) > \psi(0).$$

Es ist also $\psi(x)$ eine *positive* konvexe Funktion. Es folgt daher nach Hilfssatz 6a und 6b, daß es eine stetige konvexe Funktion $\psi_{u,v}(x)$ gibt, die an den Stellen u und v gleich $\psi(x)$ wird und für eine überalldichte Argumentmenge mit $\psi(x)$ übereinstimmt.

Andrerseits ergibt die für jedes x und ε gültige Relation

$$\psi(x) - \psi(x-\varepsilon) < \varphi(x+\varepsilon) - \varphi(x-\varepsilon) < \psi(x+\varepsilon) - \psi(x)$$

für ein der Menge der rationalen Teilpunkte des Intervalles (uv) angehöriges x und ε

$$\psi_{u,v}(x) - \psi_{u,v}(x-\varepsilon) < \varphi(x+\varepsilon) - \varphi(x-\varepsilon) < \psi_{u,v}(x+\varepsilon) - \psi_{u,v}(x).$$

In jedem Intervall, in dem $\psi_{u,v}(x)$ *gleichmäßig* stetig ist, ist daher auch $\varphi(x)$ für die entsprechende Wertmenge *gleichmäßig* stetig. Also existiert eine stetige Funktion, welche in u und v mit $\varphi(x)$ übereinstimmt und so bestimmt ist, daß sie in allen rationalen Teilpunkten der Strecke (uv) mit $\varphi(x)$ gleichen Wert hat. Wir bezeichnen dieselbe mit $\varphi_{u,v}(x)$. Die *stetige* Funktion $\varphi_{u,v}(x)$ genügt aber den Bedingungen (3) und (4). Also ist sie von der Form

$$\varphi_{u,v}(x) = h_{u,v}^2 x^2 + \varphi_{u,v}(0) \tag{46}$$

oder

$$\varphi_{u,v}(x) = \frac{\varphi(u) - \varphi(v)}{u^2 - v^2}\, x^2 + \frac{u^2\varphi(v) - v^2\varphi(u)}{u^2 - v^2}. \tag{47}$$

Wir formulieren den folgenden Satz:

Satz 8. *Die Teillösungen der Funktionalbedingung* (3) *des Gaußschen Fehlergesetzes liegen auf Parabeln der Form* $y = h^2x^2 + y(0)$.

Es seien jetzt u und v irgend zwei von Null verschiedene Werte, wir bezeichnen mit $\varphi_{0,u}(x)$, kürzer $\varphi_u(x)$, und $\varphi_{0,v}(x)$, kürzer $\varphi_v(x)$, die zu den Teillösungen der Strecken $(0u)$ und $(0v)$ zugehörigen stetigen Funktionen. Es ist also

$$\begin{aligned} \varphi_u(u) &= \varphi(u), \qquad & \varphi_u(0) &= \varphi(0), \\ \varphi_v(v) &= \varphi(v), \qquad & \varphi_v(0) &= \varphi(0). \end{aligned} \tag{48}$$

Da $\varphi_u(x)$ und $\varphi_v(x)$ mit $\varphi(x)$ an einer *überalldichten* Menge von Stellen übereinstimmen, so gilt das gleiche für $e^{-\varphi_u(x)}$, $e^{-\varphi_v(x)}$ und $e^{-\varphi(x)}$; infolgedessen ist

$$1 = \int_{-\infty}^{+\infty} e^{-\varphi_u(x)}\, dx = \int_{-\infty}^{+\infty} e^{-\varphi_v(x)}\, dx = \int_{-\infty}^{+\infty} e^{-\varphi(x)}\, dx. \tag{49}$$

Ist daher

$$\begin{aligned}\varphi_u(x) &= h_u^2 x^2 + \varphi(0), \\ \varphi_v(x) &= h_v^2 x^2 + \varphi(0),\end{aligned} \tag{50}$$

so folgt aus (49)

$$h_u^2 = h_v^2,$$

für jedes u und v. Wir bezeichnen diese Größe mit h^2.

[Denn es ist bekanntlich

$$\int_{-\infty}^{+\infty} e^{-h^2 x^2}\, dx = \frac{\sqrt{\pi}}{|h|}$$

und also

$$\frac{\sqrt{\pi}}{|h_u|} = \frac{\sqrt{\pi}}{|h_v|}.\Big]$$

Da nun andrerseits für jedes u identisch

$$\varphi(u) = \varphi_u(u)$$

ist, so folgt für jedes u

$$\varphi(u) = h^2 u^2 + \varphi(0).$$

Die Konstante $e^{-\varphi(0)}$ wird dann aus (49) als $\frac{|h|}{\sqrt{\pi}}$ gefunden, so daß wir definitiv für das Fehlergesetz die Form

$$e^{-\varphi(x)} = e^{-h^2 x^2} \frac{|h|}{\sqrt{\pi}} \tag{51}$$

erhalten. Es besteht also das

Theorem II. *Aus den Annahmen*

(A) $\varphi(x_1) + \varphi(x_2) + \cdots + \varphi(x_n) < \varphi(x_1+\varepsilon) + \varphi(x_2+\varepsilon) + \cdots + \varphi(x_n+\varepsilon)$

unter der Nebenbedingung

$$x_1 + x_2 + \cdots + x_n = 0$$

und

$$\int_{-\infty}^{+\infty} e^{-\varphi(x)}\, dx = 1 \tag{B}$$

folgt

$$e^{-\varphi(x)} = e^{-h^2 x^2} \frac{|h|}{\sqrt{\pi}}.$$

5.

Existenzbeweis der unstetigen Lösungen.

Die Bedeutung der Zusatzbedingungen, die wir hinsichtlich der Funktion $\varphi(x)$ gemacht haben, erhellt aus dem Umstande, daß sich in analoger Weise, wie für die Gleichung $f(x+y) = f(x) + f(y)$, unstetige Lösungen der Funktionalbedingung (3) konstruieren lassen.

Hierbei bedienen wir uns des von G. Hamel eingeführten Hilfsmittels einer *Basis aller Zahlen*. Es sei $a, b, c, \cdots$ eine Menge reeller Zahlen derart, daß sich jede reelle Zahl x auf eine und nur auf eine Weise in der Form

$$x = \alpha a + \beta b + \gamma c + \cdots \tag{52}$$

darstellen läßt, wo die Zahlen $\alpha, \beta, \gamma, \cdots$ rational sind, aber in jedem Fall nur eine *endliche* Anzahl derselben von Null verschieden ist. Wir nennen $a, b, c, \cdots$ eine Basis aller reellen Zahlen. Ist $x = 0$, so sind also alle $\alpha, \beta, \gamma, \cdots$ gleich Null.

Die Existenz dieser Teilmenge des Kontinuums ist unter der Voraussetzung der Möglichkeit einer *Wohlordnung des Kontinuums* von G. Hamel erbracht worden. Wir diskutieren diese Frage hier nicht näher.

Es seien jetzt

$$f(a) > 0, \; f(b) > 0, \; f(c) > 0, \cdots \tag{53}$$

irgend welche positive endliche Werte, die größer als Null sind.

Die Funktion

$$f(x) = \alpha^2 f(a) + \beta^2 f(b) + \gamma^2 f(c) + \cdots \tag{54}$$

besitzt die verlangte Eigenschaft. In der Tat es ist, wenn

$$\begin{aligned} x_1 &= \alpha_1 a + \beta_1 b + \gamma_1 c + \cdots, \\ x_2 &= \alpha_2 a + \beta_2 b + \gamma_2 c + \cdots, \\ &\cdots\cdots\cdots\cdots \\ x_n &= \alpha_n a + \beta_n b + \gamma_n c + \cdots \end{aligned} \tag{55}$$

gesetzt wird, mit Rücksicht auf die Eindeutigkeit dieser Darstellung infolge der Gleichung

$$0 = x_1 + x_2 + \cdots + x_n \tag{4}$$

notwendig

$$\begin{aligned} \alpha_1 + \alpha_2 + \cdots + \alpha_n &= 0, \\ \beta_1 + \beta_2 + \cdots + \beta_n &= 0, \\ \gamma_1 + \gamma_2 + \cdots + \gamma_n &= 0, \\ &\cdots\cdots\cdots \end{aligned} \tag{56}$$

Es sei jetzt

$$\varepsilon = \mathrm{A} a + \mathrm{B} b + \Gamma c + \cdots. \tag{57}$$

Wir haben dann

$$\begin{aligned} f(x+\varepsilon) &= (\alpha + \mathrm{A})^2 a + (\beta + \mathrm{B})^2 b + (\gamma + \Gamma)^2 + \cdots \\ &= f(x) + f(\varepsilon) + 2\mathrm{A}\alpha a + 2\mathrm{B}\beta b + 2\Gamma\gamma c + \cdots. \end{aligned}$$

Es folgt daher mit Rücksicht auf (56)

$$f(x_1+\varepsilon) + f(x_2+\varepsilon) + \cdots + f(x_n+\varepsilon) = f(x_1) + f(x_2) + \cdots + f(x_n) + n f(\varepsilon).$$

Nun ist aber für ein von Null verschiedenes ε irgend eine der Größen $\mathsf{A}, \mathsf{B}, \Gamma, \cdots$ von Null verschieden, es ist also mit Rücksicht auf (53) $f(\varepsilon)$ größer als Null. Es ist daher für ein von Null verschiedenes ε

$$(58)\quad f(x_1+\varepsilon)+f(x_2+\varepsilon)+\cdots+f(x_n+\varepsilon)>f(x_1)+f(x_2)+\cdots+f(x_n),$$

was zu beweisen war.

Wir wollen noch ein zweites Beispiel bilden, aus dessen Betrachtung hervorgeht, daß eine solche Lösung nicht notwendig die ganze positive Halbebene ausfüllt. Es liegen hier die Verhältnisse nicht so einfach, wie bei der linearen Funktionalgleichung. Eine erschöpfende Diskussion werden wir im zweiten Teile durchführen.

Wir setzen, indem wir irgend eine Basiszahl a herausgreifen,

$$(59)\qquad x=\alpha a+y$$

und bilden

$$(60)\qquad f(x)=a^2\alpha^2+y^2.$$

Aus

$$x_1+x_2+\cdots+x_n=0$$

folgt offenbar

$$\alpha_1+\alpha_2+\cdots+\alpha_n=0$$

und

$$y_1+y_2+\cdots+y_n=0.$$

Ferner wird, wenn

$$(61)\qquad \varepsilon=\mathsf{A}a+\mathsf{E}$$

gesetzt wird,

$$\begin{aligned} f(x+\varepsilon)&=a^2(\alpha+\mathsf{A})^2+(y+\varepsilon)^2\\ &=f(x)+f(\varepsilon)+2\mathsf{A}\alpha a^2+2\mathsf{E}y. \end{aligned}$$

Da nun für ein von Null verschiedenes ε

$$f(\varepsilon)>0$$

ist, so ist wiederum (58) erfüllt. Andrerseits ist aber

$$(62)\qquad \begin{aligned} f(x)=a^2\alpha^2+y^2&\geqq 2\cdot\left(\frac{a\alpha+y}{2}\right)^2\\ &\geqq\frac{x^2}{2}. \end{aligned}$$

Man erkennt übrigens aus dem Umstande, daß stets

$$a\alpha \neq y$$

ist, den Wegfall des Gleichheitszeichens.

Andrerseits kann man

$$y=b\beta+c\gamma+\cdots$$

und $a\alpha$ so nahe an $\frac{x}{2}$, als man will, annehmen. Infolgedessen wird der Wert der Funktion in der Umgebung einer beliebigen Stelle x so nahe, als man will, an $\frac{x^2}{2}$ liegen. Die Funktion $y = \frac{x^2}{2}$ ist in der Umgebung jeder Stelle die untere Grenze unserer Funktion. Da wir an Stelle von a jede andere Basiszahl setzen können, so sehen wir dieselbe untere Grenze für eine unendliche Menge von solchen Funktionen eintreten. Wir werden zeigen, daß jede unstetige Lösung ähnliche Verhältnisse zeigt, und daß die Funktion jeden Wert oberhalb der unteren Grenze in der Umgebung jeder Stelle annimmt, so daß die Ebene oberhalb der Grenzkurve überall dicht ausgefüllt wird.

Teil II.

Das Fehlergesetz für n Variable.

6.

Das Funktionaltheorem der definiten quadratischen Formen.

Es besteht der folgende merkwürdige Satz:

Theorem III. *Es sei $\varphi(x_1, x_2, \cdots, x_n)$ eine Funktion der reellen, unbeschränkt veränderlichen Argumente $x_1, x_2, \cdots, x_n$. Es genüge $\varphi(x_1, x_2, \cdots, x_n)$ für ein System von Argumenten*

$$(63)\qquad \begin{aligned} 0 &= x_1' + x_1'' + x_1''' + \cdots + x_1^{(k)}, \\ 0 &= x_2' + x_2'' + x_2''' + \cdots + x_2^{(k)}, \\ &\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot \\ 0 &= x_n' + x_n'' + x_n''' + \cdots + x_n^{(k)} \end{aligned}$$

der Ungleichheitsrelation

$$\text{(A)}\ (64)\qquad \sum_{\varkappa=1}^{\varkappa=k} \varphi(x_1^{(\varkappa)} + \varepsilon_1, x_2^{(\varkappa)} + \varepsilon_2, \cdots, x_n^{(\varkappa)} + \varepsilon_n) > \sum_{\varkappa=1}^{\varkappa=k} \varphi(x_1^{(\varkappa)}, \cdots, x_n^{(\varkappa)})$$

für jedes Wertsystem $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$, für welches nicht gleichzeitig alle Größen verschwinden. Bleibt dann noch für alle Argumentwerte

$$\text{(B)}\qquad \varphi(x_1, x_2, \cdots, x_n)$$

unterhalb einer Grenze g, so ist $\varphi(x_1, x_2, \cdots, x_n) - \varphi(0, 0, \cdots, 0)$ eine homogene positiv definite quadratische Form der n Veränderlichen $x_1, x_2, \cdots, x_n$.

Die Bedingung (B) ist wiederum nicht entbehrlich, da wir dies bereits im Falle $n = 1$ gezeigt haben. Sie kann aber durch andere Bedingungen ersetzt werden, z. B. durch die Bedingung, daß $\varphi(x_1, x_2, \cdots, x_n)$ *integrierbar* sein soll, oder daß die Funktion *stetig* sein soll. Wir wollen diese Bedingung als die *Zusatzbedingung* fernerhin bezeichnen.

Das Theorem III heiße das *Funktionaltheorem der definiten quadratischen Formen.*

Wir suchen wieder zu einer konvexen Funktion zu gelangen. Die Funktion

$$(65)\qquad \psi(x_1, x_2, \cdots, x_n) = \varphi(x_1, x_2, \cdots, x_n) + \varphi(-x_1, -x_2, \cdots, -x_n)$$

besitzt diese Eigenschaft. Zunächst liegt ψ unterhalb der Grenze $2g$. Setzen wir ferner

$$(66)\qquad x_1' = -x_1'', \cdots, x_n' = -x_n''$$

und

$$0 = x_1''' = \cdots = x_n''' = \cdots = x_1^{(k)} = \cdots = x_n^{(k)},$$

so erhalten wir, wie früher aus (64), indem wir die Akzente fortlassen, wenn nicht alle ε Null sind,

$$(67)\qquad 2\psi(x_1, \cdots, x_n) < \psi(x_1 + \varepsilon_1, \cdots, x_n + \varepsilon_n) + \psi(x_1 - \varepsilon_1, \cdots, x_n - \varepsilon_n).$$

Setzen wir jetzt wieder

$$(68)\qquad x_1 + \varepsilon_1 = u, \qquad x_1 - \varepsilon_1 = v,$$

so folgt, wenn alle andern ε gleich Null genommen werden,

$$(69)\qquad 2\psi\left(\frac{u+v}{2}, x_2, \cdots, x_n\right) < \psi(u, x_2, \cdots, x_n) + \psi(v, x_2, \cdots, x_n),$$

sobald $u \neq v$ ist. Die Funktion $\psi(x_1, x_2, \cdots, x_n)$ ist daher, was auch $x_2, \cdots, x_n$ für spezielle Werte annehmen mögen, eine *konvexe* Funktion von x_1.

Da $\psi(x_1, x_2, \cdots, x_n)$ überdies unterhalb $2g$ liegt, so ist die Funktion in x_1 *stetig* und besitzt an jeder Stelle, mit eventueller Ausnahme einer abzählbaren Menge S, einen ständig wachsenden partiellen Differentialquotienten $\psi_1(x_1, \cdots, x_n) = \frac{\partial}{\partial x_1}\psi(x_1, \cdots, x_n)$. Die Funktion $\psi_1(x_1, \cdots, x_n)$ besitzt an den Stellen S einen positiven Sprung. Wir bezeichnen mit $\bar{S}$ die Menge der rationalen Vielfachen dieser eventuellen Ausnahmestellen.

Diese Menge ist für ein spezielles Wertesystem $x_2, \cdots, x_n$ gebildet und genauer mit $\bar{S}(x_2, \cdots, x_n)$ zu bezeichnen. Gemäß der Gleichung (65) ist ψ eine gerade Funktion, und es ist $\frac{\partial}{\partial x_1}\psi(+x_1, -x_2, \cdots, -x_n)$ mit $\frac{\partial}{\partial x_1}\psi(-x_1, x_2, \cdots, x_n)$ identisch und also $\bar{S}(x_2, x_3, \cdots, x_n)$ mit $\bar{S}(-x_2, \cdots, -x_n)$ zusammenfallend. Diese Bemerkung ist für das Folgende wichtig.

Wir wenden jetzt wieder das Theorem von G. Cantor an, und schließen aus dem Umstande, daß das Kontinuum nicht abzählbar ist, daß es eine Stelle $x_1 = a$ geben muß, welche mitsamt ihren von Null verschiedenen rationalen Vielfachen keine Ausnahmestelle ist.

Ist jetzt

$$(70)\qquad r_1 + r_2 + r_3 + r_4 = 0 \qquad (r_i \neq 0)$$

eine Gleichung zwischen vier rationalen, von Null verschiedenen Zahlen, so besitzt mit Rücksicht darauf, daß jetzt $r_1 a$, $r_2 a$, $r_3 a$, $r_4 a$ nicht zu $\bar{S}(\pm x_2, \cdots, \pm x_n)$ gehören, die stetige Funktion

$$(71) \qquad \begin{aligned} \Psi(\varepsilon) = \psi(\varepsilon + r_1 a, x_2, \cdots, x_n) + \psi(\varepsilon + r_2 a, -x_2, \cdots, -x_n) \\ + \psi(\varepsilon + r_3 a, x_2, \cdots, x_n) + \psi(\varepsilon + r_1 a, -x_2, \cdots, -x_n) \end{aligned}$$

an der Stelle $\varepsilon = 0$ einen einzigen Differentialquotienten nach x_1 und also nach ε. Außerdem hat sie an derselben Stelle ein Minimum. Denn einerseits genügt $\psi(x_1, x_2, \cdots, x_n)$ der Ungleichheit (64), andererseits ist infolge von (70) gerade die Nebenbedingung (63) erfüllt. Der eigentümliche Kunstgriff, welcher hier benutzt ist, um diese Nebenbedingung für alle Argumente *gleichzeitig* zu befriedigen, setzt voraus, daß (63) und (64) für $k = 4$ gelten, während wir uns früher mit $k = 3$ begnügen konnten. Es muß also $\frac{\partial \psi(\varepsilon)}{\partial \varepsilon}$ an der Stelle $\varepsilon = 0$ verschwinden. Da nun

$$(72) \qquad \begin{aligned} \Psi(\varepsilon) = \psi(\varepsilon + r_1 a, x_2, \cdots, x_n) + \psi(-\varepsilon - r_2 a, x_2, \cdots, x_n) \\ + \psi(\varepsilon + r_3 a, x_2, \cdots, x_n) + \psi(-\varepsilon - r_4 a, x_2, \cdots, x_n) \end{aligned}$$

ist, so erhalten wir, indem wir $x_2, \cdots, x_n$ nicht mitschreiben,

$$(73) \qquad 0 = \left[\frac{\partial}{\partial \varepsilon} \Psi(\varepsilon)\right]_{\varepsilon=0} = \psi_1(r_1 a) - \psi_1(-r_2 a) + \psi_1(r_3 a) - \psi_1(-r_4 a).$$

Wir setzen jetzt

$$(74) \qquad \begin{aligned} r_1 &= s + d \qquad & r_2 &= -s \\ r_3 &= t \qquad & r_4 &= -t - d \end{aligned} \qquad (r_i \neq 0)$$

und erhalten, indem wir ax statt x_1 als Variable einführen,

$$(75) \qquad \psi_1(s + d) - \psi_1(s) = \psi_1(t + d) - \psi_1(t).$$

Die Zahlen s, t, d sind nur dadurch beschränkt, daß s und t von Null sowohl als $-d$ verschieden sein müssen; wir schreiben dies in der Form:

$$(76) \qquad s, t \neq {0 \atop -d}.$$

Von der Funktion $\psi_1(x_1, x_2, \cdots, x_n)$ wissen wir überdies, daß sie, infolge des konvexen Charakters von $\psi(x_1, x_2, \cdots, x_n)$ hinsichtlich x_1, *ständig wachsend* ist. Das gleiche gilt also von $\psi_1(x)$, wenn wir, was keine Einschränkung bedeutet, a als positiv gewählt denken. Wir beweisen den

Hilfssatz. *Eine ständig wachsende Funktion, welche den Bedingungen* (75) *und* (76) *genügt, ist eine lineare Funktion.*

Wir betrachten zuerst positive Argumente, dann negative Argumente und zeigen schließlich den stetigen Anschluß im Nullpunkt.

a) Es seien s, t und d positiv. Wir setzen s der Reihe nach gleich

$$t + d, \; t + 2d, \; \cdots, \; t + md,$$

dann wird, indem wir die entstehenden Gleichungen addieren, für alle ganzzahligen positiven m und für $m = 0$

$$\psi_1(t+md) - \psi_1(t) = m[\psi_1(t+d) - \psi_1(t)] \tag{77}$$

oder auch, indem wir d statt md setzen und n statt m schreiben,

$$\psi_1\left(t+\frac{d}{n}\right) - \psi_1(t) = \frac{1}{n}[\psi_1(t+d) - \psi_1(t)]. \tag{78}$$

Ersetzen wir in (77) d durch $\frac{d}{n}$, so erhalten wir für jede rationale positive Zahl $\frac{m}{n} = r \geqq 0$

$$\psi_1(t+rd) - \psi_1(t) = r[\psi_1(t+d) - \psi_1(t)]. \tag{79}$$

Die Funktion $\psi_1(x)$ ist eine ständig wachsende Funktion von x; infolgedessen ist auch $\psi_1(t+xd) - \psi_1(t)$ bei irgend welchen positiven, festgewählten t und d eine ständig wachsende Funktion von x. Dieselbe nimmt gemäß der Gleichung (79), falls r alle rationalen Werte von Null bis $+\infty$ durchläuft, eine lückenlose, *wachsende* Folge von Werten an, die mit Null beginnt und also positiv ansteigt und über jede Grenze wächst.

Wie früher gezeigt wurde, muß also $\psi_1(t+xd) - \psi_1(t)$ stetig sein, und es findet daher für jedes $x \geqq 0$ die Gleichung

$$\psi_1(t+xd) - \psi_1(t) = x[\psi_1(t+d) - \psi_1(t)] \tag{80}$$

oder einfacher

$$\psi_1(t+x) - \psi_1(t) = x[\psi_1(t+1) - \psi_1(t)]$$

statt.

b) Wir betrachten jetzt an Stelle von $\psi_1(x)$ die ständig abnehmende Funktion $\psi_1(-x) = \overline{\psi}_1(x)$. Für diese gilt dieselbe Betrachtung, da sie ebenfalls den Bedingungen (75) und (76) genügt. Es ist also für positive s, t, d

$$\overline{\psi}_1(t+xd) - \overline{\psi}_1(t) = x[\overline{\psi}_1(t+d) - \overline{\psi}_1(t)],$$

also

$$\psi_1(-t+x\cdot(-d)) - \psi_1(-t) = x[\psi_1(-t-d) - \psi_1(-t)] \tag{81}$$

für jedes $x \geqq 0$ eine ständig abnehmende Funktion, oder, indem wir wieder $d = 1$ wählen,

$$\psi_1(-t-x) - \psi_1(-t) = x[\psi_1(-t) - \psi_1(-t-1)] \tag{82}$$

für $x \leqq 0$ eine ständig zunehmende Funktion.

Die Funktion $\psi_1(x)$ selbst ist bis jetzt in dem Intervalle $(+t \cdots +\infty)$ durch (80) und in dem Intervall $(-\infty \cdots -t)$ durch die Gleichung (82) definiert. Man übersieht sofort, daß die geradlinigen Stücke dieselbe Neigung haben, da infolge von (75) und (76)

$$\psi_1(t+1) - \psi_1(t) = \psi_1((-t-1)+1) - \psi_1(-t-1)$$

ist. Es ist also nur nötig zu zeigen, daß sich die Stücke von linearen

Funktionen für $x = 0$ *stetig* aneinanderschließen. Hieraus würde übrigens auch die unbeschränkte Gültigkeit von (80) und damit zugleich wieder die Gleichheit der Neigungen folgen.

c) Wählen wir jetzt, unter ϱ eine positive Größe verstanden, um in das Intervall $(-\varrho + \cdots + \varrho)$ einzudringen, t und d gemäß den Gleichungen

(83) $$\text{a)} \quad -\varrho < t < 0 \qquad 0 < t + d < +\varrho,$$

also sicher

$$\text{b)} \quad 0 < d < 2\varrho,$$

so können wir die Differenz

$$\psi_1(t+d) - \psi_1(t)$$

mittels der Gleichung (75) auf

$$\psi_1(s+d) - \psi_1(s)$$

zurückführen, indem wir für s irgend eine positive Größe annehmen, da ja die Argumente $t, t+d, s, s+d$ von Null verschieden sind. Es gilt also, wenn t und d die Bedingungen (83) erfüllen, stets infolge (80) für positive s und d

(84) $$\psi_1(t+d) - \psi_1(t) = d[\psi_1(s+1) - \psi_1(s)].$$

Infolge (83) b) ist also

$$|\psi_1(t+d) - \psi_1(t)| < 2\varrho\, |\psi_1(s+1) - \psi_1(s)|.$$

Für hinreichend kleines ϱ ist diese Differenz also beliebig klein und da $t + d$ stets positiv, t stets negativ ist, so folgt, daß die Funktion $\psi_1(x)$ an der Stelle $x = 0$ keinen Sprung erleidet. *Es ist also $\psi_1(x)$ eine lineare Funktion von x*, und wir setzen für jedes x

(85) $$\psi_1(x) = x[\psi_1(1) - \psi_1(0)] + \psi_1(0).$$

Um jetzt zu zeigen, daß der Koeffizient von x unabhängig von

$$x_2, x_3, \cdots, x_n$$

ist, betrachten wir folgendes Wertsystem:

(86) $$\begin{array}{llll} x_1' = x_1, & x_2' = x_2, & x_3' = x_3, & \cdots, \; x_n' = x_n \\ x_1'' = 0, & x_2'' = -x_2, & x_3'' = -x_3, & \cdots, \; x_n'' = -x_n \\ x_1''' = -x_1, & x_2''' = 0, & x_3''' = 0, & \cdots, \; x_n''' = 0. \end{array}$$

Wir haben bewiesen, daß $\psi(x_1, x_2, \cdots, x_n)$, wie auch $x_2, \cdots, x_n$ und wie auch schließlich x_1 beschaffen ist, eine einzige partielle Ableitung $\psi_1(x_1, \cdots, x_n)$ nach x_1 besitzt. Es ist also, damit die Minimalbedingung (64) erfüllt wird, notwendig in allen Werten der Argumente

(87) $$0 = \psi_1(x_1, x_2, \cdots, x_n) + \psi_1(0, -x_2, \cdots, -x_n, x_n) + \psi_1(-x_1, 0, \cdots, 0),$$

oder, wenn wir vor dem Differenzieren nach ε die Zeichen verändern, in allen Argumenten

(88) $$\psi_1(x_1, x_2, \cdots, x_n) = \psi_1(x_1, 0, \cdots, 0) + \psi_1(0, x_2, \cdots, x_n).$$

An der Stelle $x_2 = \cdots = x_n = 0$ ist aber die Funktion $\psi(x_1, x_2, \cdots, x_n)$, wie wir wissen, quadratisch von der Form $2h^2 x_1{}^2 + \psi(0)$, also ist $\psi_1(x_1, 0, \cdots, 0)$ von der Form $x_1 \psi_1(1, 0, \cdots, 0)$ und wir haben die fundamentale Formel gewonnen:

Es ist identisch in allen Argumenten

$$\psi_1(x_1, x_2, \cdots, x_n) = x_1 \psi_1(1, 0, \cdots, 0) + \psi_1(0, x_2, \cdots, x_n), \tag{89}$$

wo überdies

$$\psi_1(1, 0, \cdots, 0) > 0$$

ist.

Wir vollenden jetzt den Beweis für ψ mittels des Schlusses von $n-1$ auf n.

Zunächst folgt aus (89), daß

$$\psi(x_1, x_2, \cdots, x_n) = \frac{1}{2}\psi_1(1, 0, \cdots, 0)\, x_1{}^2 + \psi_1(0, x_2, \cdots, x_n) x_1 + \psi(0, x_2, \cdots, x_n) \tag{89}$$

Denn es besitzt die Differenz

$$\psi(x_1, x_2, \cdots, x_n) - \frac{1}{2}\psi_1(1, 0, \cdots, 0)\, x_1{}^2 - \psi_1(0, x_2, \cdots, x_n)\, x_1$$

den Charakter einer in x_1 stetigen Funktion, welche in bezug auf x_1 den Differentialquotienten Null hat. Sie ist also von x_1 unabhängig und kann daher gleich $\psi(0, x_2, \cdots, x_n)$ gesetzt werden, indem wir $x_1 = 0$ annehmen.

Andererseits ist $\psi(0, x_2, \cdots, x_n)$ für die $n-1$ Variabeln $x_2, \cdots, x_n$ von genau dem gleichen Charakter wie $\psi(x_1, \cdots, x_n)$ für die n Variabeln. Nun ist unser Satz für $n = 1$ bewiesen. Wir nehmen jetzt an, er sei für $n-1$ bewiesen und es sei also

$$\chi(x_2, \cdots, x_n) = \psi(0, x_2, \cdots, x_n) - \psi(0, 0, \cdots, 0) \tag{90}$$

eine homogene positive quadratische Form, so ist also identisch, wenn wir $\frac{1}{2}\psi_1(0, 0, \cdots, 0)$ gleich A_1 setzen und $\psi(x_1, x_2, \cdots, x_n)$ statt

$$\psi(x_1, x_2, \cdots, x_n) - \psi(0, 0, \cdots, 0)$$

schreiben,

$$\psi(x_1, x_2, \cdots, x_n) = A_1 x_1{}^2 + \psi_1(0, x_2, \cdots, x_n)\, x_1 + \chi(x_2, \cdots, x_n). \tag{91}$$

Wir setzen $x_1 = x_2$ und haben

$$x_2 \psi_1(0, x_2, \cdots, x_n) = \psi(x_2, x_2, \cdots, x_n) - A_1 x_2{}^2 - \chi(x_2, \cdots, x_n). \tag{92}$$

Nun ist aber auch $\psi(x_2, x_2, \cdots, x_n)$ eine Funktion von $n-1$ Variabeln, welche also eine homogene quadratische Form sein muß. Infolgedessen ist auch die linke Seite eine homogene quadratische Form der $n-1$ Variabeln und also $\psi_1(0, x_2, \cdots, x_n)$ eine homogene Form der Variabeln $x_2, \cdots, x_n$. Hiermit ist in Rücksicht auf (91) bewiesen, daß $\psi(x_1, x_2, \cdots, x_n)$ eine homogene quadratische Form der n Variabeln ist. Aus der Ungleichheitsrelation (64) folgt überdies, indem wir $k = 1$ und alle x gleich Null nehmen, daß $\varphi(\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n) - \varphi(0, \cdots, 0)$ und also auch ψ für alle Werte der Argumente positiv sind.

Hiermit ist bewiesen, daß $\psi(x_1, \cdots, x_n) - \psi(0, 0, \cdots, 0)$ *eine positiv definite quadratische Form ist.*

Setzt man übrigens

$$x_1 = x_1 t, \; x_2 = x_2 t, \; \cdots, \; x_n = x_n t,$$

so bemerkt man sofort a priori, daß die Gleichung

$$(93) \quad \varphi(x_1 t, x_2 t, \cdots, x_n t) = t^2[\varphi(x_1, x_2, \cdots, x_n) - \varphi(0, \cdots, 0)] + \varphi(0, \cdots, 0)$$

besteht und zwar identisch in $t, x_1, \cdots, x_n$.

Hier ist der homogene Charakter sofort erkennbar und ferner wissen wir, daß der Koeffizient von t^2 *positiv* ist.

Den Übergang von ψ zu φ vollziehen wir durch den folgenden dem Hilfssatz 5 (Teil I) entsprechenden

Hilfssatz 9. *Es ist* $\varphi(x_1, x_2, \cdots, x_n)$ *stetig und besitzt einen Differentialquotienten nach* x_ν, *für welchen*

$$(94) \qquad 2 \frac{\partial}{\partial x_\nu} \varphi(\cdots, x_\nu, \cdots) = \frac{\partial}{\partial x_\nu} \psi(\cdots, x_\nu, \cdots) \qquad (\nu = 1, \cdots, n)$$

gilt.

Beweis. Wir beweisen diese Behauptung für $\nu = 1$. Wählen wir die Argumente x nach (66) und nehmen alle ε außer ε_1 gleich Null, so haben wir gemäß (64):

$$(95) \quad \psi(x_1, x_2, \cdots, x_n) < \varphi(x_1 + \varepsilon_1, x_2, \cdots, x_n) + \psi(-x_1 + \varepsilon_1, x_2, \cdots, x_n).$$

Durch Addition und Subtraktion von $\varphi(x_1 - \varepsilon_1, x_2, \cdots, x_n)$ folgt

$$(96) \qquad \psi(x_1, x_2, \cdots, x_n) < \psi(x_1 + \varepsilon_1, x_2, \cdots, x_n) + \varphi(x_1 - \varepsilon_1, x_2, \cdots, x_n) \\ - \varphi(x_1 + \varepsilon_1, x_2, \cdots, x_n).$$

Indem wir das Vorzeichen von ε_1 vertauschen, erhalten wir

$$(97) \quad \psi(x_1, \cdots, x_n) - \psi(x_1 - \varepsilon_1, x_2, \cdots, x_n) < \varphi(x_1 + \varepsilon_1, x_2, \cdots, x_n) - \varphi(x_1 - \varepsilon_1, x_2, \cdots, x_n) \\ < \psi(x_1 + \varepsilon_1, x_2, \cdots, x_n) - \psi(x_1, \cdots, x_n).$$

Hieraus folgt sofort, daß $\varphi(x_1)$ in x_1 stetig ist. Um die Ableitung herzustellen, dividieren wir durch ε_1 und gehen zur Grenze über, indem wir entweder $x_1 + \varepsilon_1$ oder $x_1 - \varepsilon_1$ festhalten und $\pm 2\varepsilon = \eta$ als Zuwachs nehmen. Wir erhalten für die vordere wie für die hintere Ableitung $\frac{\partial}{d x_1} \varphi(x_1, \cdots, x_n)$ denselben Wert und zwar die Hälfte der Ableitung $\frac{\partial}{\partial x_1} \psi(x_1, \cdots, x_n)$.

Dasselbe gilt für jedes x_ν, womit der Hilfssatz bewiesen ist.

Mit Rücksicht auf die bewiesene Stetigkeit der Funktion φ folgt ferner, daß identisch in $x_1, x_2, \cdots, x_n$

(98) $$2\varphi(x_1, \cdots, x_n) = \psi(x_1, \cdots, x_n)$$

ist. Und es ist daher unser Funktionaltheorem für $\varphi(x_1, \ldots, x_n)$ vollständig bewiesen.

Teil III.

Die Haupteigenschaft der unstetigen Lösungen und das verallgemeinerte Funktionaltheorem.

Wie bereits in Teil I, 5. angegeben worden ist, läßt sich auf Grund des soeben bewiesenen Satzes ein allgemeines Resultat über die Natur der unstetigen Lösungen angeben. Wir formulieren den

Satz 10. *Die Werte der Funktion $y = \varphi(x)$ liegen entweder auf einer Parabel*

$$y = h^2x^2 + \varphi(0)$$

oder sie füllen den ebenen Raum oberhalb einer solchen Parabel überalldicht aus.

Wir wollen, worin keine Einschränkung liegt, fernerhin stets $\varphi(0) = 0$ annehmen.

Methodisch sei betont, daß wir von der Annahme der Existenz einer Basis keinen Gebrauch machen und diese Hypothese nur insoweit heranziehen, als sie erläutert, daß der zweite Fall tatsächlich im Bereich der Möglichkeit liegt.

Wir beweisen zunächst den

Hilfssatz 10a. *Es seien a und b zwei reelle Zahlen derart, daß die Gleichung*

(99) $$0 = a\alpha + b\beta$$

mit rationalen Koeffizienten α und β die Gleichung

(100) $$\alpha = 0, \qquad \beta = 0$$

nach sich zieht; dann ist eine der Ungleichheitsbedingung des Fehlergesetzes genügende Funktion

(101) $$\varphi(a\alpha + b\beta)$$

für die sämtlichen rationalen Werte von α und β eine positive quadratische Form in α und β.

In der Tat ist die Darstellung einer Zahl c in der Form

(102) $$c = a\alpha + b\beta$$

eindeutig in α und β (falls c von dieser Form ist). Sind also die Zahlen

(103) $$x_\nu = a\alpha_\nu + b\beta_\nu \qquad (\nu = 1, 2, \cdots, n)$$

so gewählt, daß

$$x_1 + x_2 + \cdots + x_n = 0$$

wird, so muß dazu zugleich

$$\alpha_1 + \alpha_2 + \cdots + \alpha_n = 0 \quad (104) \quad \text{und} \quad \beta_1 + \beta_2 + \cdots + \beta_n = 0$$

sein. Wählen wir ferner ε von der Form

(105) $$\varepsilon = a\mathsf{A} + b\mathsf{B},$$

wo A und B rationale Zahlen sind, so besteht also die Beziehung

(106) $$\sum_{\nu=1}^{\nu=n} \varphi[a(\alpha_\nu + \mathsf{A}) + b(\beta_\nu + \mathsf{B})] > \sum_{\nu=1}^{\nu=n} \varphi[a\alpha_\nu + b\beta_\nu]$$

für irgend welche rationalen α_ν, β_ν, die (104) genügen, und für irgend welche rationalen A und B.

Indem wir $\psi(x) = \varphi(x) + \varphi(-x)$ bilden, stellen wir fest, daß $\psi(a\alpha + b\beta)$ infolge (106) in bezug auf jede der rationalen Variabeln eine konvexe Funktion ist. Überdies ist $\psi(a\alpha + b\beta)$ positiv.

Es ist daher $\psi(a\alpha + b\beta)$ innerhalb eines endlichen Gebietes der α,β-Ebene eine für alle rationalen α und β *gleichmäßig* stetige Funktion.*) Es existiert daher eine stetige konvexe Funktion $2 \cdot F(x_1, x_2)$, welche für $x_1 = \alpha$, $x_2 = \beta$ mit $\psi(a\alpha + b\beta)$ übereinstimmt. Dieselbe genügt außerdem der Bedingung des Fehlergesetzes für zwei Variabele für alle Argumentwerte und ist daher eine positive quadratische Form. Aus der gleichmäßigen Stetigkeit von $\psi(a\alpha + b\beta)$ folgt auch die gleichmäßige Stetigkeit von $\varphi(a\alpha + b\beta)$, so daß $\varphi(a\alpha + b\beta)$ ebenfalls mit einer positiven quadratischen Form $F(x_1, x_2)$ so zusammenhängt, daß für

(107) $$x_1 = \alpha, \qquad x_2 = \beta$$

(108) $$F(\alpha, \beta) = \varphi(a\alpha + b\beta), \quad F(x_1, x_2) = Ax_1^2 + 2Bx_1x_2 + Cx_2^2$$

ist. Um A, B, C zu bestimmen, setzen wir der Reihe nach $\alpha = \beta$, $\alpha = 0$, $\beta = 0$ und erhalten

(109) $$A = \varphi(a), \quad C = \varphi(b), \quad B = \varphi(a+b) - \varphi(a) - \varphi(b).$$

Wir untersuchen jetzt die Werteverteilung der Funktion $\varphi(a\alpha + b\beta)$ näher. Sei

$$c = a\alpha + b\beta$$

ein Wert der Argumentmenge. Der Wert der quadratischen Form $F(x_1, x_2)$ für $x_1 = \alpha$, $x_2 = \beta$ ist jedenfalls größer oder gleich dem Minimum $m(c)$ von $F(x_1, x_2)$ für ein der Gleichung:

(110) $$c = x_1 a + x_2 b$$

genügendes Wertepaar x_1, x_2. Wir haben also

*) Den Beweis hierfür entnimmt man sehr leicht den Betrachtungen des Teiles I, Satz 6a und b. Wir kommen später darauf zurück.

$$(111) \qquad \varphi(a\alpha + b\beta) \geqq m(c).$$

$m(c)$ ist eine stetige Funktion von c von der Form

$$(112) \qquad y = x^2 \frac{AC - B^2}{Ab^2 - 2Bab + Ca^2},$$

wie man durch direkte Ausrechnung feststellt.

Es seien $x_1(c)$ und $x_2(c)$ Werte von x_1 und x_2, welche für irgend ein c der Bedingung (110) genügen. Wir bestimmen dann zwei rationale Zahlen α' und β' derart, daß

$$|x_1(c) - \alpha'| < \varepsilon, \qquad |x_2(c) - \beta'| < \eta$$

ist. Dann ist für dieses c

$$|a\alpha' + b\beta' - c| < |a|\varepsilon + |b|\eta.$$

Nun ist es infolge der Stetigkeit von $F(x_1, x_2)$ möglich, bei gegebenem ζ die Größen ε und η so klein zu wählen, daß

$$|F(x_1, x_2) - F(\alpha', \beta')| < \zeta$$

ist für alle α', β', welche den angegebenen Bedingungen genügen.

Infolgedessen gibt es in der Umgebung jeder Stelle c Werte $F(\alpha', \beta') = \varphi(a\alpha' + b\beta')$, die jedem Werte der Form $F(x_1, x_2)$ für $c = x_1 a + x_2 b$ beliebig nahe kommen. Hinsichtlich der Form $F(x_1, x_2)$ gibt es aber nur zwei Fälle: entweder ist sie für $ax_1 + bx_2 = c$ konstant, oder sie nimmt auf dieser Geraden alle Werte, die größer als $m(c)$ sind, an. *Es erfüllt also auch die Wertmenge $F(\alpha', \beta')$ entweder die Parabel $y = m(c)$ oder den oberen Teil der Halbebene, welche durch sie begrenzt wird.*

Betrachten wir endlich die schon früher behandelten Teillösungen, welche durch irgend ein $\varphi(a)$ bestimmt werden und zu dem Intervall $(0a)$ gehören. Dieselben besitzen Parabelform. Es sei A die untere Grenze der zugehörigen Konstanten h_a^2.

Es sind nun zwei Fälle möglich; entweder sind alle Teillösungen von der Form $y = Ax^2$, dann ist Ax^2 die einzige stetige Lösung.

Oder es gibt zwei voneinander verschiedene Teillösungen $\varphi_{0a}(a\alpha)$ und $\varphi_{0b}(b\beta)$, von denen eine, etwa $\varphi_{0a}(x)$, entweder mit Ax^2 zusammenfällt, oder doch so nahe an Ax^2, als man will, gewählt werden kann. In jedem Falle füllt dann $\varphi(a\alpha + b\beta)$ einen Teil der Ebene aus, der sowohl $y = \varphi_{0a}(a\alpha)$ als $y = \varphi_{0b}(b\beta)$ enthält. Es wird also der Teil der Ebene oberhalb Ax^2 überalldicht ausgefüllt.

Wir sind jetzt in der Lage, die Zusatzbedingung des Theorems I in ihrer reduziertesten Form zu geben.

Wir bezeichnen zu diesem Zwecke mit $\lim\limits_{\text{inf.}} \varphi(x)$ die untere Grenze aller Werte einer Funktion $\varphi(x)$ in der Umgebung der Stelle x. Es besteht dann, wie durch die voranstehenden Betrachtungen bewiesen ist, das folgende Theorem:

Verallgemeinertes Theorem I. *Es genüge die eindeutige Funktion* $\varphi(x)$ *der Bedingung*

$$\varphi(x_1)+\varphi(x_2)+\cdots+\varphi(x_n)<\varphi(x_1+\varepsilon)+\varphi(x_2+\varepsilon)+\cdots+\varphi(x_n+\varepsilon)$$

unter der Nebenbedingung

$$x_1+x_2+\cdots+x_n=0,$$

dann besitzt die untere Grenze $\lim\limits_{\text{inf.}}\varphi(x)$ *die Form*

$$\lim_{\text{inf.}}\varphi(x)=h^2x^2+\varphi(0).$$

Ist ferner irgend ein Wertepaar $x, y > \lim\limits_{\text{inf.}}\varphi(x)$ *vorhanden, welches durch* $x, \varphi(x)$ *nicht beliebig angenähert werden kann, so ist*

$$\varphi(x)\equiv\lim_{\text{inf.}}\varphi(x).$$

Die entsprechende Verallgemeinerung des Funktionaltheorems der definiten quadratischen Formen, sowie einige andere hiermit zusammenhängende Fragen gedenke ich später zu entwickeln.

Bemerkung bei der Korrektur. Während der Korrektur des vorliegenden Aufsatzes erscheint eine Note von H. Lebesgue: Sur les transformations ponctuelles, transformant les plans en plans, qu'on peut définir par des procédés analytiques (Atti d. R. Ac. d. Sc. d. Torino XLII, 10 marzo 1907). In derselben wird zunächst mittels eines der Hamelschen Methode verwandten Verfahrens das für komplexes Gebiet geltende System

$$\varphi(z_1+z_2)=\varphi(z_1)+\varphi(z_2),\quad \varphi(z_1\cdot z_2)=\varphi(z_1)\,\varphi(z_2)$$

durch eine unstetige Lösung befriedigt. Dann aber gibt der Autor den methodisch und sachlich neuen Nachweis, daß die unstetige Lösung sich nicht in einer sehr allgemeinen analytischen Form, die er früher studiert hat, darstellen läßt. Dies Resultat gilt genau ebenso für die hier betrachteten unstetigen Lösungen.

Inhaltsverzeichnis.