IV. Ueber die Bestimmung der Präcision einer Beobachtungsmethode nach der Methode der kleinsten Quadrate.

In seiner ersten Begründung der Methode der kleinsten Quadrate ging Gauss (Theoria motus corp. coel.) von der Voraussetzung aus, dass der wahrscheinlichste Werth einer beliebig oft auf dieselbe Weise direct gemessenen Grösse das arithmetische Mittel aus den durch diese Messungen erhaltenen Werthen ist, und kam auf diese Weise zu dem Ausdruck

$$\frac{h}{\sqrt{\pi}}\, e^{-h^2 t^2} dt$$

für die Wahrscheinlichkeit, dass ein Beobachtungsfehler seinem Werth nach in dem unendlich kleinen Intervall zwischen t und $t + dt$ liegt; in diesem Ausdruck bedeutet h eine positive Constante, welche für verschiedene Beobachtungsmethoden im Allgemeinen auch verschiedene Werthe hat, und zwar leuchtet ein, dass eine Beobachtungsmethode desto zuverlässiger ist, je grösser der Werth der ihr zugehörigen Constante h ist; denn die Wahrscheinlichkeit

$$\frac{h}{\sqrt{\pi}}\int_{-a}^{+a} e^{-h^2 t^2}\, dt = \frac{1}{\sqrt{\pi}}\int_{-ha}^{+ha} e^{-u^2}\, du$$

dafür, dass ein Fehler seinem absoluten Werthe nach die positive Grösse a nicht überschreitet, ist desto grösser, je grösser h ist. Aus diesem Grunde hat Gauss die Grösse h die Präcision der Beobachtungsmethode genannt; in einer spätern Abhandlung (Zeitschrift für Astronomie etc. von Lindenau und Bohnenberger, Bd. I, 1816) hat er ferner gezeigt, wie man den wahrscheinlichsten Werth der Präcision einer Beobachtungsmethode bestimmen kann, wenn eine Reihe wirklich gemachter Beobachtungsfehler bekannt ist. Es wird für das Folgende nützlich sein, hier den von Gauss zu diesem Zweck eingeschlagenen Weg wieder in Erinnerung zu bringen, welcher auf dem Satz über die Wahrscheinlichkeit a posteriori beruht.

Ist h die wahre Präcision der Beobachtungsmethode, so ist die Wahrscheinlichkeit, dass bei m aufeinanderfolgenden Beobachtungen die Fehler

$$t_1, t_2, \cdot \cdot t_m$$

gemacht werden, gleich

$$\alpha = \left(\frac{h}{\sqrt{\pi}}\right)^m e^{-h^2 S} dt_1 dt_2 \cdot \cdot dt_m,$$

worin zur Abkürzung

$$S = t_1^2 + t_2^2 + \cdot \cdot \cdot + t_m^2,$$

gesetzt ist. A priori, d. h. ehe irgend eine Messung vorgenommen ist, haben wir keinen Grund, der Präcision einer uns unbekannten Beobachtungsmethode einen Werth h eher beizulegen als einen andern; folglich ist a posteriori, d. h. nachdem wirklich die Beobachtungsfehler t_1, t_2, . . t_m gemacht sind, die Wahrscheinlichkeit der Hypothese, dass h der wahre Werth der Präcision ist, proportional dem α, also proportional dem Ausdruck

$$h^m e^{-h^2 S},$$

welcher für

$$\frac{1}{2h^2} = \frac{S}{m}, \text{ also } h = \sqrt{\frac{m}{2S}}$$

ein Maximum wird; es ist also dies der wahrscheinlichste Werth der Präcision der Beobachtungsmethode.

In allen wirklichen Fällen liegt aber die Sache ganz anders. Die Objecte der Beobachtungen sind lineare Functionen

$$v_1, v_2, \ldots v_m$$

von gewissen unbekannten Grössen $x, y, z, \ldots$, deren Anzahl n höchstens gleich der Anzahl m der Beobachtungen und deren Werthbestimmung gerade der Zweck dieser Beobachtungen ist. Sind nun

$$k_1, k_2, \ldots k_m$$

die durch die Beobachtungen gelieferten Werthe von $v_1, v_2, \ldots v_m$, so bestimmt die aus dem obigen Wahrscheinlichkeitsgesetz eines beliebigen Fehlers t gefolgerte Methode der kleinsten Quadrate die Werthe der Unbekannten $x, y, z, \ldots$ durch die Forderung, dass die Quadratsumme

$$(k_1 - v_1)^2, + (k_2 - v_2)^2 + \cdots + (k_m - v_m)^2 = \Omega$$

ein Minimum werden soll. Wären nun diese wirklich die wahren Werthe der Unbekannten, so wären die entsprechenden Werthe der Differenzen

$$k_1 - v_1, k_2 - v_2, \cdots k_m - v_m$$

auch die wahren Beobachtungsfehler, und man könnte versucht sein, den wahrscheinlichsten Werth der Präcision h nach der frühern Regel zu bestimmen, indem man statt S nur das Minimum Ω_0 der Function Ω zu substituiren brauchte, so dass also

$$\sqrt{\frac{m}{2\Omega_0}}$$

als wahrscheinlichster Werth von h anzusehen wäre. Dass diese Formel aber nicht richtig sein kann, bemerkt man am deutlichsten in dem Fall, wo $n = m$ ist; dann können nämlich die gemachten Beobachtungen sämmtlich durch ein und dasselbe Werthsystem

$x, y, z, \ldots$ befriedigt werden, Ω_0 ist $= 0$, und man würde $h = \infty$, also das Resultat erhalten, dass die Beobachtungsmethode höchst wahrscheinlich absolut genau ist, während doch erst dann ein Urtheil über die Präcision gestattet ist, wenn ein Ueberschuss von Beobachtungen vorliegt.

In einer spätern Abhandlung (Theoria combinationis etc. art. 39), in welcher das Princip des arithmetischen Mittels und damit zugleich das obige Wahrscheinlichkeitsgesetz eines Fehlers t ganz verlassen ist, hat Gauss für eine ähnliche Frage (die nach dem wahrscheinlichsten Werthe des sogenannten mittlern Fehlers) die richtige Antwort gegeben, welche, auf die frühere Darstellungsweise übertragen, den Ausdruck

$$\sqrt{\frac{m-n}{2\Omega_0}}$$

als wahrscheinlichsten Werth der Präcision h liefert, so dass also das Minimum Ω_0 als eine Summe von nur $(m - n)$ Fehlerquadraten zu behandeln ist. Man sieht, dass diese Formel in dem Fall $n = m$ unter die ganz unbestimmte Form $\frac{0}{0}$ tritt, und in der That ist in diesem Fall gar kein Schluss auf die Präcision gestattet.

Es erscheint nun wünschenswerth, einen Beweis dieses Satzes auch aus dem obigen Wahrscheinlichkeitsgesetze abzuleiten, da dies meines Wissens in befriedigender Weise noch nicht geschehen ist. *) Dazu führt folgender einfache Weg.

*) So z. B. geht Wittstein (Anhang zu der Uebersetzung von Navier's Differentialrechnung) von dem unrichtigen Satz aus, dass, wenn h die wahre Präcision, der wahrscheinlichste Werth eines Fehlerquadrates $= \frac{1}{2h^2}$, statt $= 0$ ist.

In der Hypothese B, dass $h, x, y, z, \ldots$ die wahren Werthe der Präcision, der ersten, zweiten, dritten u. s. w. Unbekannten sind, ist die Wahrscheinlichkeit, dass für die Functionen

$$v_1,\ v_2, \cdot\ \cdot v_m$$

die Werthe

$$k_1,\ k_2, \cdot\ \cdot k_m$$

durch Beobachtung geliefert, dass also die Beobachtungsfehler

$$k_1 - v_1,\ k_2 - v_2, \cdot\ \cdot k_m - v_m$$

gemacht werden, proportional dem Ausdruck

$$h^m e^{-h^2\Omega};$$

da nun alle denkbaren Hypothesen B a priori gleich wahrscheinlich sind, so ist a posteriori, d. h. nachdem wirklich die Werthe $k_1, k_2, \ldots k_m$ beobachtet sind, die Wahrscheinlichkeit der Hypothese B proportional demselben Ausdruck; dieselbe ist daher

$$= Ch^m e^{-h^2\Omega}\, dh\, dx\, dy\, dz \cdot \cdot \cdot,$$

worin

$$\frac{1}{C} = \int_0^\infty dh \int_{-\infty}^{+\infty} dx \int_{-\infty}^{+\infty} dy \int_{-\infty}^{+\infty} dz \cdot\ \cdot h^m e^{-h^2\Omega}$$

Fragt man nun nach dem wahrscheinlichsten Werthsystem von $h, x, y, z, \ldots$, so würde man untersuchen müssen, für welche Werthe $h, x, y, z, \ldots$ der Ausdruck

$$h^m e^{-h^2\Omega}$$

ein Maximum wird. Allein wir fragen nach dem wahrscheinlichsten Werth der Präcision allein; wir haben daher zunächst den Ausdruck der Wahrscheinlichkeit herzustellen, dass der Werth der Präcision zwischen h und $h + dh$ liegt. Diesen erhält man aus dem vorhergehenden durch Integration über alle reellen Werthe von x, y, z, . . Es ist aber nach bekannten Sätzen

$$\int_{-\infty}^{+\infty} dx \int_{-\infty}^{+\infty} dy \int_{-\infty}^{+\infty} dz \cdot \cdot e^{-h^2 \Omega} = K \frac{1}{h^n} e^{-h^2 \Omega_0}$$

worin K von h unabhängig ist; folglich ist das aus den gemachten Beobachtungen resultirende Wahrscheinlichkeitsgesetz für die Präcision von der Form

$$H \cdot h^{m-n} e^{-h^2 \Omega_0} dh,$$

worin

$$\frac{1}{H} = \int_0^\infty h^{m-n} e^{-h^2 \Omega_0} dh$$

ist. Vergleicht man diese Form mit der frühern

$$H' h^m e^{-h^2 S} dh, \text{ wo } \frac{1}{H'} = \int_0^\infty h^m e^{-h^2 S} dh,$$

welche sich ergab, wenn m wahre Beobachtungsfehler vorlagen, deren Quadratsumme $= S$ war, so findet man in der That vollständige Uebereinstimmung, wenn man das Minimum Ω_0 der Summe von m Fehlerquadraten wie eine Summe von $m - n$ wirklichen Fehlerquadraten ansieht. Der wahrscheinlichste Werth zu der Präcision ist daher wirklich

$$= \sqrt{\frac{m-n}{2\Omega_0}}.$$

Hiermit ist der eigentliche Gegenstand dieser Mittheilung beendigt; zum Schluss mag noch folgende Bemerkung gemacht werden. Wir haben als wahrscheinlichsten Werth h einen andern gefunden, als denjenigen, welcher dem h in dem wahrscheinlichsten System von Werthen h, x, y, z, . . zukommt. Man könnte nun befürchten, dass auch die Bestimmung der wahrscheinlichsten Werthe von $x, y, z, \ldots$, wenn sie nach demselben Princip ausgeführt, wenn also für jede einzelne Unbekannte besonders der wahrscheinlichste Werth aufgesucht würde, von der durch die Methode der kleinsten Quadrate geforderten Regel abweichen könnte. Allein man überzeugt sich leicht, dass diese Befürchtung ungegründet ist, und dass das System der wahrscheinlichsten Werthe von $x, y, z, \ldots$ übereinstimmt mit dem wahrscheinlichsten Werthsystem dieser Unbekannten.

Das letztere ist offenbar dasjenige, für welches die Quadratsumme Ω ein Minimum wird, und darin besteht ja gerade der Hauptsatz der Methode der kleinsten Quadrate; die entsprechenden Werthe der n Unbekannten $x, y, z, \ldots$ findet man bekanntlich dadurch, dass man, was immer möglich ist, die Function Ω auf die Form

$$\Omega = Y^2 + Z^2 + \cdot \cdot + X^2 + \Omega_0$$

bringt, worin Y eine lineare Function aller n Unbekannten ist, die dadurch bestimmt wird, dass $\Omega - Y^2$ unabhängig von y wird; ähnlich ist Z eine lineare Function der übrigen $(n - 1)$ Unbekannten, und dadurch bestimmt, dass $\Omega - Y^2 - Z^2$ unabhängig von y, z wird, u. s. f., so dass endlich X eine lineare Function von der n^{ten} Unbekannten x allein ist. Die Werthe, welche Ω zu einem Minimum machen, sind diejenigen, welche die n Gleichungen

$$X = 0, \cdot \cdot \cdot Z = 0, \; Y = 0$$

befriedigen, und das letzte Glied Ω_0 in dieser Form stellt offenbar den Minimumwerth von Ω dar.

Fragt man nun aber nach dem wahrscheinlichsten Werth der Unbekannten x allein, so hat man zunächst den Ausdruck der Wahrscheinlickeit abzuleiten, dass der Werth dieser Unbekannten zwischen den Grenzen x und $x + dx$ enthalten ist. Diesen erhält man durch Integration des obigen Werthes

$$Ch^m e^{-h^2 \Omega} dh\, dx\, dy\, dz \ldots$$

in Bezug auf alle zulässige Werthe der Unbekannten $h, y, z, \ldots$ Bringt man die Summe Ω auf die oben erwähnte Form, so giebt die successive Integration in Bezug auf die $(n - 1)$ Unbekannten $y, z, \ldots$ ein Resultat

$$C' h^{m-n+1} e^{-h^2 (X^2 + \Omega_0)} dh\, dx,$$

worin C' unabhängig von h und x ist; integrirt man endlich noch in Bezug auf h, so erhält man für die gesuchte Wahrscheinlichkeit den Ausdruck

$$\frac{c\,dx}{(X^2 + \Omega_0)^{\frac{m-n+2}{2}}},$$

worin

$$\frac{1}{c} = \int_{-\infty}^{+\infty} \frac{dx}{(X^2 + \Omega_0)^{\frac{m-n+2}{2}}};$$

und hieraus folgt, dass derjenige Werth von x, für welchen $X = 0$ wird, unter allen der wahrscheinlichste ist. Dieser Werth stimmt daher wirklich mit dem durch die Methode der kleinsten Quadrate erhaltenen überein.