I.

Theorie
der den kleinsten Fehlern unterworfenen
Combination der Beobachtungen.

Erster Theil.

(Der Königlichen Societät der Wissenschaften zu Göttingen überreicht 1821, Februar 15.)

1.

Beobachtungen, welche sich auf Grössenbestimmungen aus der Sinnenwelt beziehen, werden immer, so sorgfältig man auch verfahren mag, grösseren oder kleineren Fehlern unterworfen bleiben. Die Fehler der Beobachtungen sind im allgemeinen nicht einfache, sondern entspringen gleichzeitig mehreren Quellen, bei denen zwei Arten genau unterschieden werden müssen. Gewisse Fehlerursachen sind nämlich so beschaffen, dass ihr Einfluss auf jede Beobachtung von veränderlichen Umständen abhängt, die unter sich und mit der Beobachtung selbst in keinem wesentlichen Zusammenhang stehen; die so entstehenden Fehler werden unregelmässige oder zufällige genannt; und insoweit jene Umstände der Rechnung nicht unterworfen werden können, gilt dieses auch von den Fehlern selbst. Dahin gehören die von der Unvollkommenheit unserer Sinne herrührenden Fehler und solche, die von unregelmässigen äusseren Ursachen abhängen, z. B. von der durch das Wallen der Luft bewirkten Unsicherheit beim Sehen; auch rechnen wir hierher manche, selbst den besten Instrumenten anhaftende Unvollkommenheiten, z. B. Ungleichförmigkeiten der inneren Wandungen der Libellen, Mangel an absoluter Festigkeit u. s. w. Dagegen haben andere Fehlerursachen bei sämmtlichen Beobachtungen derselben Art ihrer Natur nach entweder einen vollkommen constanten Einfluss, oder doch einen solchen, dessen Grösse in gesetzmässig bestimmter Weise allein von Umständen abhängt, welche mit der Beobachtung wesentlich verknüpft sind. Fehler dieser Art werden constante oder regelmässige genannt.

Uebrigens ist es klar, dass diese Unterscheidung gewissermaassen nur relativ ist und von dem weiteren oder engeren Sinne abhängt, in welchem man den Begriff von Beobachtungen derselben Art fassen will. So bringen z. B. unregelmässige Fehler der Theilung der Instrumente bei Winkelmessungen einen constanten Fehler hervor, wenn es sich nur um eine beliebig oft zu wiederholende Beobachtung desselben Winkels handelt, und wenn dabei immer dieselben fehlerhaften Theilstriche benutzt werden; während der aus derselben Quelle stammende Fehler als ein zufälliger angesehen werden kann, wenn man irgendwie Winkel von beliebiger Grösse zu messen hat, und eine Tafel, die für jeden Theilstrich den zugehörigen Fehler angiebt, nicht zu Gebote steht.

2.

Die Betrachtung der regelmässigen Fehler soll von unseren Untersuchungen ausdrücklich ausgeschlossen bleiben. Es ist nämlich Sache des Beobachters, alle Ursachen, welche constante Fehler hervorzubringen vermögen, sorgfältig aufzusuchen und dieselben entweder abzustellen, oder wenigstens ihrer Wirkung und Grösse nach auf das genaueste zu erforschen, um ihren Einfluss auf jede einzelne Beobachtung bestimmen und diese von jenem befreien zu können, so dass ein Ergebniss erzielt wird, als ob der Fehler überhaupt nicht vorhanden gewesen wäre. Ganz verschieden hiervon ist aber das Wesen der unregelmässigen Fehler, welche ihrer Natur nach der Rechnung nicht unterworfen werden können. Diese wird man daher in den Beobachtungen zwar dulden, ihren Einfluss aber auf die aus den Beobachtungen abzuleitenden Grössen durch eine geschickte Combination der ersteren möglichst abschwächen müssen. Dieser wichtigen Aufgabe ist die folgende Untersuchung gewidmet.

3.

Die Fehler in den Beobachtungen gleicher Art, welche einer bestimmten einfachen Ursache entspringen, sind der Natur der Sache nach in bestimmte *Grenzen* eingeschlossen, welche man zweifelsohne genau angeben könnte, wenn die Natur dieser Ursache selbst *vollständig* erkannt wäre. Die meisten Ursachen zufälliger Fehler sind so beschaffen, dass nach dem Gesetz der Stetigkeit alle zwischen jenen Grenzen enthaltenen Fehler für möglich gehalten werden müssen, und dass die vollständige Erkenntniss einer solchen Ursache zugleich lehren würde, ob alle diese Fehler mit gleicher oder ungleicher Leichtigkeit begangen werden können, und, in letzterem Falle, eine wie grosse relative Wahrscheinlichkeit jedem Fehler beizulegen sei. Dasselbe gilt auch in Bezug auf den totalen Fehler, der sich aus mehreren einfachen Fehlern zusammensetzt, dass er nämlich zwischen bestimmten Grenzen eingeschlossen sein wird (von denen die eine der Summe aller oberen, die andere der Summe aller unteren Theilgrenzen gleich ist); alle Fehler zwischen diesen Grenzen werden zwar möglich sein, da sich indess jeder auf unendlich viele verschiedene Weisen durch Zusammensetzung der Theilfehler, welche selbst wieder mehr oder weniger wahrscheinlich sind, ergeben kann, so werden wir für den einen eine grössere, für den andern eine geringere Häufigkeit annehmen müssen, und es könnte unter der Voraussetzung, dass man die Gesetze der einfachen Fehler kennt, ein Gesetz der relativen Wahrscheinlichkeit aufgestellt werden, abgesehen von den analytischen Schwierigkeiten beim Zusammenfassen aller Combinationen.

Freilich giebt es auch gewisse Fehlerursachen, welche nicht nach dem Gesetz der Stetigkeit fortschreitende, sondern nur unstetige Fehler hervorbringen können, wie z. B. die Theilungsfehler der Instrumente (wenn man diese überhaupt zu den zufälligen Fehlern rechnen will); denn die Anzahl der Theilstriche an jedem bestimmten Instrument ist endlich. Dessen ungeachtet wird aber offenbar, wenn nur nicht alle Fehlerursachen unstetige Fehler erzeugen, die Gesammtheit aller möglichen Totalfehler eine nach dem Gesetz der Stetigkeit fortschreitende Reihe bilden, oder auch mehrere derartige getrennte Reihen, wenn es sich nämlich bei Anordnung aller möglichen unstetigen Fehler nach ihrer Grösse ergeben sollte, dass zufällig eine oder die andere Differenz zwischen zwei aufeinander folgenden Gliedern dieser Reihe grösser ist, als die Differenz zwischen den Grenzen derjenigen Totalfehler, welche den stetigen Fehlern allein entstammen. In der Praxis wird aber der letztere Fall kaum jemals eintreten, wenn nicht etwa die Theilung an groben Fehlern leidet.

4.

Bezeichnet man mit $\varphi(x)$ die relative Häufigkeit des Totalfehlers x bei einer bestimmten Gattung von Beobachtungen, so wird wegen der Stetigkeit der Fehler die Wahrscheinlichkeit eines zwischen den unendlich nahen Grenzen x und $x + dx$ liegenden Fehlers $= \varphi(x)\,dx$ zu setzen sein. Es wird in der Praxis wohl immer so gut wie unmöglich sein, diese Funktion a priori anzugeben; nichtsdestoweniger lassen sich mehrere allgemeine Eigenschaften derselben feststellen, welche hier folgen sollen. Offenbar ist die Funktion $\varphi(x)$ insofern zu den unstetigen Funktionen zu rechnen, als sie für alle Werthe des x, welche ausserhalb der Grenzen der möglichen Fehler liegen, $= 0$ sein muss; innerhalb dieser Grenzen wird sie aber überall einen positiven Werth annehmen (abgesehen von dem Fall, über den wir am Ende des vorigen Art. gesprochen haben). In den meisten Fällen wird man positive und negative Fehler von derselben Grösse als gleich häufig voraussetzen dürfen, so dass $\varphi(-x) = \varphi(x)$ sein wird. Da ferner kleinere Fehler leichter als grössere begangen werden, so wird im allgemeinen $\varphi(x)$ für $x = 0$ seinen grössten Werth erhalten und beständig abnehmen, wenn x wächst.

Allgemein giebt aber der Werth des von $x = a$ bis $x = b$ genommenen Integrals $\int\varphi(x)\,dx$ die Wahrscheinlichkeit dafür an, dass irgend ein noch unbekannter Fehler zwischen den Grenzen a und b liegt. Der Werth dieses Integrals von der unteren Grenze aller möglichen Fehler bis zu ihrer oberen Grenze wird daher immer $= 1$ sein. Und da $\varphi(x)$ für alle ausserhalb dieser Grenzen liegenden Werthe des x immer $= 0$ ist, so ist offenbar auch

der Werth des von $x = -\infty$ bis $x = +\infty$ genommenen Integrals $\int\varphi(x)\,dx$ immer $= 1$.

5.

Wir betrachten ferner das Integral $\int x\,\varphi(x)\,dx$ zwischen denselben Grenzen, und setzen seinen Werth $= k$. Sind alle einfachen Fehlerursachen nun so beschaffen, dass kein Anlass vorhanden ist, zwei gleichen, aber mit entgegengesetzten Vorzeichen versehenen Fehlern verschiedene Häufigkeit beizulegen, so wird dasselbe auch für den totalen Fehler gelten, es ist also $\varphi(-x) = \varphi(x)$, und deshalb nothwendig $k = 0$. Wir schliessen hieraus, dass, jedesmal wenn k nicht verschwindet, sondern etwa eine positive Grösse ist, nothwendig eine oder die andere Fehlerursache vorhanden sein müsse, welche entweder nur positive Fehler, oder wenigstens häufiger positive als negative zu erzeugen vermöge. Diese Grösse k, welche in der That das Mittel aller möglichen Fehler oder der mittlere Werth der Grösse x ist, kann passend der constante Theil des Fehlers genannt werden. Uebrigens ist leicht zu beweisen, dass der constante Theil des totalen Fehlers gleich der Summe der constanten Theile derjenigen Fehler ist, welche aus den einzelnen einfachen Ursachen hervorgehen. Setzt man jetzt die Grösse k als bekannt voraus, subtrahirt dieselbe von jeder Beobachtung und bezeichnet den Fehler der so verbesserten Beobachtung mit x', die entsprechende Wahrscheinlichkeit aber mit $\varphi'(x')$, so wird $x' = x - k$, $\varphi'(x') = \varphi(x)$ und folglich

$$\int x'\varphi'(x')\,dx' = \int x\varphi(x)\,dx - \int k\varphi(x)\,dx = k - k = 0,$$

d. h. die Fehler der verbesserten Beobachtungen werden keinen constanten Theil haben, was auch an sich klar ist.

6.

Wie das Integral $\int x\varphi(x)\,dx$, oder der mittlere Werth von x, das Fehlen oder Vorhandensein und die Grösse eines constanten Fehlers anzeigt, ebenso erscheint das von $x = -\infty$ bis $x = +\infty$ ausgedehnte Integral

$$\int x^2\varphi(x)\,dx$$

(oder der mittlere Werth des Quadrates x^2) am geeignetsten, die Unsicherheit von Beobachtungen allgemein zu definiren und zu messen, so dass bei zwei Beobachtungsgruppen, die sich hinsichtlich der Häufigkeit der Fehler unterscheiden, diejenigen Beobachtungen für die genaueren zu halten sind, für welche das Integral $\int x^2\varphi(x)\,dx$ den kleineren Werth erhält. Wenn nun jemand einwenden würde, diese Festsetzung sei ohne zwingende Nothwendigkeit willkürlich getroffen, so stimmen wir gern zu. Enthält doch diese Frage der Natur der Sache nach etwas Unbestimmtes, welches nur durch ein in gewisser Hinsicht willkürliches Princip bestimmt begrenzt werden kann. Die Bestimmung einer Grösse durch eine einem grösseren oder kleineren Fehler unterworfene Beobachtung wird nicht unpassend mit einem Glücksspiel verglichen, in welchem man nur verlieren, aber nicht gewinnen kann, wobei also jeder zu befürchtende Fehler einem Verluste entspricht. Das Risiko eines solchen Spieles wird nach dem wahrscheinlichen Verlust geschätzt, d. h. nach der Summe der Produkte der einzelnen möglichen Verluste in die zugehörigen Wahrscheinlickeiten. Welchem Verluste man aber jeden einzelnen Beobachtungsfehler gleichsetzen soll, ist keineswegs an sich klar; hängt doch vielmehr diese Bestimmung zum Theil von unserem Ermessen ab. Den Verlust dem Fehler selbst gleichzusetzen, ist offenbar nicht erlaubt; würden nämlich positive Fehler wie Verluste behandelt, so müssten negative als Gewinne gelten. Die Grösse des Verlustes muss vielmehr durch eine solche Funktion des Fehlers ausgedrückt werden, die ihrer Natur nach immer positiv ist. Bei der unendlichen Mannigfaltigkeit derartiger Funktionen scheint die einfachste, welche diese Eigenschaft besitzt, vor den übrigen den Vorzug zu verdienen, und diese ist unstreitig das Quadrat. Somit ergiebt sich das oben aufgestellte Princip.

Laplace hat die Sache zwar auf eine ähnliche Weise betrachtet, er hat aber den immer positiv genommenen Fehler selbst als Maass des Verlustes gewählt. Wenn wir jedoch nicht irren, so ist diese Festsetzung sicherlich nicht weniger willkürlich, als die unsrige: ob nämlich der doppelte Fehler für ebenso erträglich zu halten ist, wie der einfache, zweimal wiederholte, oder für schlimmer, und ob es daher angemessener ist, dem doppelten Fehler nur das doppelte Moment, oder ein grösseres beizulegen, ist eine Frage, die weder an sich klar, noch durch mathematische Beweise zu entscheiden, sondern allein dem freien Ermessen zu überlassen ist. Ausserdem kann man nicht leugnen, dass die in Rede stehende Festsetzung gegen die Stetigkeit verstösst: und gerade deshalb widerstrebt dieses Verfahren in höherem Grade der analytischen Behandlung, während die Resultate, zu welchen unser Princip führt, sich sowohl durch Einfachheit als auch durch Allgemeinheit ganz besonders auszeichnen.

7.

Wir setzen den Werth des von $x = -\infty$ bis $x = +\infty$ genommenen Integrals $\int x^2 \varphi(x)\, dx = m^2$, und nennen die Grösse m *den mittleren zu befürchtenden Fehler*, oder einfach *den mittleren Fehler* der Beobachtungen, deren unbestimmte Fehler x die relative Wahrscheinlichkeit $\varphi(x)$ haben. Jene Bezeichnung werden wir nicht auf unmittelbare Beobachtungen beschränken, sondern auch auf alle aus Beobachtungen abgeleiteten Bestimmungen ausdehnen. Man muss sich indess sehr wohl davor hüten, den mittleren Fehler mit dem arithmetischen Mittel aller Fehler, von welchem im Art. 5. die Rede war, zu verwechseln.

Wo mehrere Gattungen von Beobachtungen oder mehrere aus Beobachtungen erhaltene Bestimmungen, denen nicht dieselbe Genauigkeit zukommt, zu vergleichen sind, verstehen wir unter dem relativen *Gewicht* derselben eine Grösse, die dem m^2 umgekehrt proportional ist, während die *Genauigkeit* einfach dem m umgekehrt proportional genommen wird. Um demnach das Gewicht durch eine Zahl ausdrücken zu können, muss man das Gewicht einer gewissen Gattung von Beobachtungen als Einheit annehmen.

8.

Enthalten die Beobachtungsfehler einen constanten Theil, so wird durch seine Elimination der mittlere Fehler verringert, das Gewicht und die Genauigkeit vermehrt. Unter Beibehaltung der Bezeichnungen des Art. 5. erhält man, wenn m' den mittleren Fehler der verbesserten Beobachtungen bedeutet,

$$m'^2 = \int x'^2\,\varphi'(x')\,dx' = \int (x - k)^2\,\varphi(x)\,dx = \int x^2\,\varphi(x)\,dx$$
$$-2k\int x\,\varphi(x)\,dx + k^2\int \varphi(x)\,dx = m^2 - 2k^2 + k^2 = m^2 - k^2.$$

Wenn man aber an Stelle des wahren constanten Theiles k eine andere Grösse l von den Beobachtungen abgezogen hätte, so würde das Quadrat des neuen mittleren Fehlers $= m^2 - 2kl + l^2 = m'^2 + (l - k)^2$ werden.

9.

Bezeichnet man mit λ einen bestimmten Coefficienten und mit μ den Werth des Integrals $\int \varphi(x)\,dx$ von $x = -\lambda m$ bis $x = +\lambda m$, so wird μ die Wahrscheinlichkeit dafür sein, dass der Fehler irgend einer Beobachtung (dem absoluten Werthe nach) kleiner als λm sei, dagegen $1 - \mu$ die Wahrscheinlichkeit dafür, dass der Fehler grösser als λm sei. Wenn also der Werth $\mu = \frac{1}{2}$ dem Werth $\lambda m = \varrho$ entspricht, so wird der Fehler ebenso leicht unterhalb ϱ als oberhalb ϱ liegen können, so dass ϱ passend der *wahrscheinliche* Fehler genannt werden kann. Die Beziehung zwischen den Grössen λ und μ hängt offenbar von der Natur der Funktion $\varphi(x)$ ab, welche im allgemeinen unbekannt ist. Es wird deshalb die Mühe lohnen, jene Beziehung für einige besondere Fälle näher zu betrachten.

I. Sind die Grenzen aller möglichen Fehler $-a$ und $+a$, und sind alle Fehler in diesen Grenzen gleich wahrscheinlich, so wird $\varphi(x)$ in den Grenzen $x = -a$ und $x = +a$ constant und folglich $= \frac{1}{2a}$ sein. Hieraus ergiebt sich $\underline{m = a\sqrt{\frac{1}{3}}}$ und $\mu = \lambda\sqrt{\frac{1}{3}}$, so lange λ nicht grösser als $\sqrt{3}$ ist; endlich wird $\varrho = m\sqrt{\frac{3}{4}} = 0{,}8660254\,m$,

$m^2 = \int x^2(\varphi(x))\,dx = \frac{1}{2a}\int_{-a}^{+a} x^2\,dx = \frac{1}{3}a^2$; $m = a\sqrt{\frac{1}{3}}$; $k\int_{-a}^{+a} dx = 1$; $k = \varphi(x)$

$\mu = \int_{-\lambda m}^{+\lambda m} \varphi(x)\,dx = \frac{1}{2a}\int_{-\lambda m}^{+\lambda m} dx = \frac{1}{2a}\,2\lambda m = \lambda\left(\frac{m}{a}\right) = \lambda\sqrt{\frac{1}{3}}$; $\varrho = \lambda m = m, \mu\sqrt{3} - m\sqrt{\frac{3}{4}}$ [illegible] und die Wahrscheinlichkeit, dass ein Fehler nicht grösser als der mittlere Fehler werde, $= \sqrt{\frac{1}{3}} = 0{,}5773503$.

II. Sind die Grenzen der möglichen Fehler, wie vorher, $-a$ und $+a$, und nimmt man an, dass die Wahrscheinlichkeit dieser Fehler vom Fehler 0 ab nach beiden Seiten in arithmetischer Progression abnehme, so wird

$$\varphi(x) = \frac{a-x}{a^2}, \text{ für die Werthe von } x \text{ zwischen } 0 \text{ und } +a,$$

$$\varphi(x) = \frac{a+x}{a^2}, \text{ für die Werthe von } x \text{ zwischen } 0 \text{ und } -a,$$

sein. Es folgt hieraus $m = a\sqrt{\frac{1}{6}}$, $\mu = \lambda\sqrt{\frac{2}{3}} - \frac{1}{6}\lambda^2$, so lange λ zwischen 0 und $\sqrt{6}$ liegt, und endlich $\lambda = \sqrt{6} - \sqrt{6 - 6\mu}$, so lange μ zwischen 0 und 1 liegt, und hieraus

$$\varrho = m(\sqrt{6} - \sqrt{3}) = 0{,}7174389\, m.$$

Die Wahrscheinlichkeit eines den mittleren nicht übersteigenden Fehlers wird in diesem Falle

$$= \sqrt{\frac{2}{3}} - \frac{1}{6} = 0{,}6498299.$$

III. Nehmen wir die Funktion $\varphi(x)$ proportional zu $e^{-\frac{x^2}{h^2}}$ (was zwar in der Wirklichkeit nur sehr nahe richtig sein kann), so wird

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

sein müssen, wobei π den halben Kreisumfang für den Radius 1 bezeichnet, woraus wir ferner ableiten

$$m = h\sqrt{\frac{1}{2}}$$

(Siehe: Disquisitiones generales circa seriem infinitam etc., art. 28.). Bezeichnet man ferner den Werth des von $z = 0$ an genommenen Integrals

$$\frac{2}{\sqrt{\pi}} \int e^{-z^2} dz$$

mit $\Theta(z)$, so wird

$$\mu = \Theta\left(\lambda\sqrt{\frac{1}{2}}\right).$$

Die folgende Tafel giebt einige Werthe dieser Grösse:

λ	μ
0,6744897	0,5
0,8416213	0,6
1,0000000	0,6826895
1,0364334	0,7
1,2815517	0,8
1,6448537	0,9
2,5758293	0,99
3,2918301	0,999
3,8905940	0,9999
∞	1

10.

Obgleich die Beziehung zwischen λ und μ von der Natur der Funktion $\varphi(x)$ abhängt, so kann man doch einige allgemeine Eigenschaften derselben feststellen. Wie nämlich diese Funktion auch beschaffen sei, so wird sicher, wenn sie nur die Eigenschaft hat, dass ihr Werth bei wachsendem absoluten Werthe von x immer abnimmt oder wenigstens nicht wächst,

λ kleiner oder wenigstens nicht grösser als $\mu\sqrt{3}$ sein, wenn μ kleiner als $\frac{2}{3}$,

λ nicht grösser als $\frac{2}{3\sqrt{1-\mu}}$ sein, wenn μ grösser als $\frac{2}{3}$ ist.

Für $\mu = \frac{2}{3}$ fallen beide Grenzen zusammen, und λ kann alsdann nicht grösser als $\sqrt{\frac{4}{3}}$ sein.

Um diesen merkwürdigen Lehrsatz zu beweisen, bezeichnen wir mit y den Werth des von $z = -x$ bis $z = +x$ genommenen Integrals $\int \varphi(z)\,dz$; y wird alsdann die Wahrscheinlichkeit dafür sein, dass irgend ein Fehler in den Grenzen $-x$ und $+x$ enthalten ist. Ferner setzen wir

$$x = \psi(y), \qquad d\,\psi(y) = \psi'(y)\,dy, \qquad d\psi'(y) = \psi''(y)\,dy.$$

Es wird demnach $\psi(0) = 0$ sein, und

$$\psi'(y) = \frac{1}{\varphi(x) + \varphi(-x)},$$

woraus mit Rücksicht auf die Voraussetzung folgt, dass $\psi'(y)$ von $y = 0$ bis $y = 1$ beständig wächst oder wenigstens nirgends abnimmt; oder, was dasselbe ist, dass der Werth von $\psi''(y)$ immer positiv oder wenigstens nicht negativ ist. Ferner haben wir $d(y\,\psi'(y)) = \psi'(y)\,dy + y\,\psi''(y)\,dy$, und folglich

$$y\,\psi'(y) - \psi(y) = \int y\,\psi''(y)\,dy\,,$$

wenn man die Integration mit $y = 0$ beginnen lässt. Der Werth des Ausdrucks $y\,\psi'(y) - \psi(y)$ wird deshalb immer eine positive oder wenigstens keine negative Grösse, und folglich

$$1 - \frac{\psi(y)}{y\,\psi'(y)}$$

eine positive Grösse kleiner als 1 sein. Es sei f ihr Werth für $y = \mu$, also es sei, da man $\psi(\mu) = \lambda m$ hat,

$$f = 1 - \frac{\lambda m}{\mu\,\psi'(\mu)} \quad \text{oder} \quad \psi'(\mu) = \frac{\lambda m}{(1-f)\mu}\,.$$

Nach diesen Vorbereitungen betrachten wir folgende Funktion des y

$$\frac{\lambda m}{(1-f)\,\mu}(y - \mu f)\,,$$

welche wir $=$ F(y) setzen, wobei $d\,\mathrm{F}(y) = \mathrm{F}'(y)\,dy$. Offenbar wird dann

$$\mathrm{F}(\mu) = \lambda m = \psi(\mu)$$
$$\mathrm{F}'(\mu) = \frac{\lambda m}{(1-f)\,\mu} = \psi'(\mu)\,.$$

Da nun $\psi'(y)$ mit wachsendem y immer wächst (oder wenigstens nicht abnimmt, was stets hinzuzudenken ist), und da $\mathrm{F}'(y)$ andererseits constant ist, so wird die Differenz $\psi'(y) - \mathrm{F}'(y) = \frac{d(\psi(y) - \mathrm{F}(y))}{dy}$ für Werthe von y, welche grösser als μ sind, positiv, für kleinere negativ sein. Hieraus ist leicht zu ersehen, dass $\psi(y) - \mathrm{F}(y)$ immer eine positive Grösse, und dass ferner $\psi(y)$ immer absolut grösser oder wenigstens nicht kleiner als F(y) ist, wenigstens so lange als der Werth von F(y) positiv ist, d. h. von $y = \mu f$ bis $y = 1$. Deshalb wird der Werth des Integrals $\int [\mathrm{F}(y)]^2\,dy$ von $y = \mu f$ bis $y = 1$ kleiner sein als der Werth des Integrals $\int [\psi(y)]^2\,dy$ in denselben Grenzen, und um so mehr auch kleiner als der Werth dieses Integrals von $y = 0$ bis $y = 1$, welches $= m^2$ ist. Der Werth des ersteren Integrals ergiebt sich aber

$$= \frac{\lambda^2 m^2 (1-\mu f)^3}{3\mu^2 (1-f)^2},$$

woraus man entnimmt, dass λ^2 kleiner sei als $\frac{3\mu^2(1-f)^2}{(1-\mu f)^3}$, wo die Grösse f zwischen 0 und 1 liegt. Der Werth des Bruches $\frac{3\mu^2(1-f)^2}{(1-\mu f)^3}$, dessen Differential, wenn f als Variable betrachtet wird,

$$= -\frac{3\mu^2(1-f)}{(1-\mu f)^4}(2-3\mu+\mu f)\,df$$

ist, nimmt ferner beständig ab, wenn f vom Werthe 0 bis zum Werthe 1 steigt, sobald μ kleiner ist als $\frac{2}{3}$; der grösstmögliche Werth wird deshalb dem Werthe $f=0$ entsprechen und folglich $=3\mu^2$ werden, so dass in diesem Fall λ sicher kleiner oder doch nicht grösser als $\mu\sqrt{3}$ wird. W. z. b. w. Wenn hingegen μ grösser als $\frac{2}{3}$ ist, so wird der grösste Werth jenes Bruches für $2-3\mu+\mu f=0$ eintreten, d. h. für $f=3-\frac{2}{\mu}$, und zwar wird derselbe $=\frac{4}{9(1-\mu)}$, und λ kann also in diesem Falle nicht grösser als $\frac{2}{3\sqrt{1-\mu}}$ sein. W. z. b. w.

So kann z. B. für $\mu=\frac{1}{2}$ sicher λ nicht grösser als $\sqrt{\frac{3}{4}}$ werden, d. h. der wahrscheinliche Fehler kann die Grenze $0{,}8660254\,m$ nicht übersteigen, welcher Werth für ihn im ersten Beispiel des Art. 9. gefunden wurde. Ferner schliesst man aus unserem Satze leicht, dass μ nicht kleiner als $\lambda\sqrt{\frac{1}{3}}$ sei, so lange λ kleiner als $\sqrt{\frac{4}{3}}$ ist, dass hingegen μ nicht kleiner als $1-\frac{4}{9\lambda^2}$ sein könne, wenn der Werth von λ grösser als $\sqrt{\frac{4}{3}}$ ist.

11.

Da mehrere der später zu behandelnden Aufgaben auch mit dem Werth des Integrals $\int x^4\varphi(x)\,dx$ im Zusammenhang stehen, so wird es die Mühe lohnen, denselben für einige specielle Fälle zu ermitteln. Wir bezeichnen den Werth dieses von $x=-\infty$ bis $x=+\infty$ genommenen Integrals mit n^4.

I. Für $\varphi(x) = \frac{1}{2a}$ wird, wenn x zwischen $-a$ und $+a$ eingeschlossen ist, $n^4 = \frac{1}{5} a^4 = \frac{9}{5} m^4$.

II. Für den zweiten Fall des Art. 9., wenn $\varphi(x) = \frac{a \mp x}{a^2}$ für Werthe von x zwischen 0 und $\pm a$ ist, hat man $n^4 = \frac{1}{15} a^4 = \frac{12}{5} m^4$.

III. Im dritten Fall, wenn

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

findet man, nach den in der oben angeführten Abhandlung erhaltenen Resultaten, $n^4 = \frac{3}{4} h^4 = 3m^4$.

Ausserdem lässt sich zeigen, dass der Werth von $\frac{n^4}{m^4}$ nicht kleiner als $\frac{9}{5}$ sein kann, wenn nur die Voraussetzung des vorigen Art. erfüllt ist.

12.

Bezeichnen x, x', x'' etc. allgemein Fehler von Beobachtungen derselben Art, die unabhängig von einander seien, und drückt ein vorgesetztes Zeichen φ ihre relativen Wahrscheinlichkeiten aus, ist ferner y eine gegebene rationale Funktion der Variabeln x, x', x'' etc., dann wird das vielfache Integral

$$\int \varphi(x)\, \varphi(x')\, \varphi(x'') \ldots .\, dx\, dx'\, dx'' \ldots ., \tag{I}$$

erstreckt über alle Werthe der Variabeln x, x', x'' etc., für welche der Werth des y zwischen die gegebenen Grenzen 0 und η fällt, die Wahrscheinlichkeit dafür ausdrücken, dass der Werth des y irgendwo zwischen 0 und η liegt. Offenbar wird dieses Integral eine Funktion von η sein, deren Differential wir $= \psi(\eta)\, d\eta$ setzen, so dass das Integral selbst dem Integral $\int \psi(\eta)\, d\eta$, von $\eta = 0$ angefangen, gleich ist. Alsdann wird das Zeichen $\psi(\eta)$ die relative Wahrscheinlichkeit eines jeden Werthes von y ausdrücken müssen. Da man x nun als eine Funktion der Variabeln y, x', x'' etc. ansehen kann, die mit $f(y, x', x'' \ldots)$ bezeichnet werden möge, so verwandelt sich das Integral (I) in

$$\int \varphi[f(y, x', x'' \ldots)] \frac{d f(y, x', x'' \ldots)}{dy}\, \varphi(x')\, \varphi(x'') \ldots dy\, dx'\, dx'' \ldots,$$

wo y von $y = 0$ bis $y = \eta$ genommen werden muss, die übrigen Variablen aber über alle Werthe zu erstrecken sind, denen ein reeller Werth von $f(y, x', x'' \ldots)$ entspricht. Hieraus schliesst man, dass

$$\psi(y) = \int \varphi\,[f(y, x', x'' \ldots)]\,\frac{df(y, x', x'' \ldots)}{dy}\,\varphi(x')\,\varphi(x'')\ldots dx'\,dx''\ldots,$$

wo die Integration, bei welcher y als constant betrachtet werden muss, über alle Werthe der Variabeln x', x'' etc. zu erstrecken ist, die für $f(y, x', x'' \ldots)$ einen reellen Werth ergeben.

13.

Um obige Integration wirklich auszuführen, müsste man die Funktion φ kennen, welche im allgemeinen unbekannt ist. Selbst wenn aber auch die Funktion bekannt wäre, würde die Integration meistens die Kräfte der Analysis übersteigen. Deshalb werden wir zwar die Wahrscheinlichkeit der einzelnen Werthe des y nicht angeben können; anders aber wird es sich verhalten, wenn man nur den mittleren Werth des y verlangt; derselbe ergiebt sich nämlich durch Integration von $\int y\,\psi(y)\,dy$ über alle möglichen Werthe des y. Und da man offenbar für alle Werthe, welche y nicht annehmen kann — sei es der Natur der mit y bezeichneten Funktion wegen (z. B. bei $y = x^2 + x'^2 + x''^2 +$ etc. für die negativen Werthe), sei es um der den Fehlern x, x', x'' etc. gesetzten bestimmten Grenzen willen —, $\psi(y) = 0$ setzen muss, so darf man mit demselben Rechte offenbar jene Integration über alle reellen Werthe von y erstrecken, also von $y = -\infty$ bis $y = +\infty$. Nun ist aber das in den bestimmten Grenzen von $y = \eta$ bis $y = \eta'$ genommene Integral $\int y\,\psi(y)\,dy$ gleich dem Integral

$$\int y\,\varphi\,[f(y, x', x'' \ldots)]\,\frac{df(y, x', x'' \ldots)}{dy}\,\varphi(x')\,\varphi(x'')\ldots dy\,dx'\,dx''\ldots,$$

welches gleichfalls von $y = \eta$ bis $y = \eta'$ und über alle Werthe der Variabeln x', x'' etc., denen ein reeller Werth von $f(y, x', x'' \ldots)$ entspricht, zu erstrecken ist; oder, was dasselbe ist, auch gleich dem Werthe des Integrals

$$\int y\,\varphi(x)\,\varphi(x')\,\varphi(x'')\ldots dx\,dx'\,dx''\ldots,$$

wenn bei dieser Integration für y sein Werth als Funktion von x, x', x'' etc. eingesetzt, und dieselbe über alle Werthe dieser Variabeln, welchen ein zwischen η und η' liegender Werth des y entspricht, ausgedehnt wird. Hieraus folgern wir, dass das über alle Werthe des y, von $y = -\infty$ bis $y = +\infty$ ausgedehnte Integral $\int y\,\psi(y)\,dy$ aus der Integration von

$$\int y\,\varphi(x)\,\varphi(x')\,\varphi(x'')\ldots dx\,dx'\,dx''\ldots$$

erhalten wird, wenn man dieselbe über alle reellen Werthe von x, x', x'' etc. erstreckt, demnach von $x = -\infty$ bis $x = +\infty$, von $x' = -\infty$ bis $x' = +\infty$ etc.

14.

Besteht daher die Funktion y nur aus einer Summe von Gliedern von der Form

$$\mathrm{A}\, x^{\alpha}\, x'^{\beta}\, x''^{\gamma}\ldots,$$

so wird der Werth des über alle Werthe von y erstreckten Integrals $\int y\,\psi(y)\,dy$, oder der mittlere Werth von y, einer Summe von Gliedern

$$\mathrm{A} \times \int x^{\alpha}\,\varphi(x)\,dx \times \int x'^{\beta}\,\varphi(x')\,dx' \times \int x''^{\gamma}\,\varphi(x'')\,dx''\ldots$$

gleich sein, bei welchen die Integrationen von $x = -\infty$ bis $x = +\infty$, von $x' = -\infty$ bis $x' = +\infty$ etc. zu nehmen sind; oder, was dasselbe ist, gleich einer Summe von Gliedern, welche entstehen, wenn man für die einzelnen Potenzen x^{α}, x'^{β}, x''^{γ} etc. ihre mittleren Werthe einsetzt. Die Richtigkeit dieses so wichtigen Lehrsatzes hätte auch leicht aus anderen Ueberlegungen gefolgert werden können.

15.

Wir wollen den im vorigen Art. aufgestellten Lehrsatz auf den speciellen Fall anwenden, dass

$$y = \frac{x^2 + x'^2 + x''^2 + \text{etc.}}{\sigma},$$

wo σ die Anzahl der Glieder im Zähler bezeichnet. Den mittleren Werth des y finden wir hier ohne Weiteres $= m^2$, indem wir dem Buchstaben m dieselbe Bedeutung wie oben geben. Der wahre Werth des y kann sich zwar in einem bestimmten Fall grösser oder kleiner als dieser mittlere ergeben, ebenso wie der wahre Werth eines einzelnen Gliedes x^2; die Wahrscheinlichkeit aber, dass ein gelegentlicher Werth des y von dem mittleren m^2 nicht wesentlich abweiche, wird sich stetig um so mehr der Gewissheit nähern, je mehr die Zahl σ wächst. Um dieses noch klarer zu zeigen, werden wir, da wir die Wahrscheinlichkeit selbst nicht genau zu bestimmen im Stande sind, den mittleren bei der Annahme $y = m^2$ zu befürchtenden Fehler suchen. Nach den im Art. 6. aufgestellten Principien wird dieser Fehler offenbar gleich der Quadratwurzel aus dem mittleren Werthe der Funktion

$$\left(\frac{x^2 + x'^2 + x''^2 + \text{etc.}}{\sigma} - m^2\right)^2$$

sein; zur Auffindung desselben genügt die Bemerkung, dass der mittlere Werth eines Gliedes von der Form $\frac{x^4}{\sigma^2}$ gleich $\frac{n^4}{\sigma^2}$ ist (wo der Buchstabe n dieselbe Bedeutung wie im Art. 11. hat), dass dagegen der mittlere Werth eines Gliedes von der Form $\frac{2\,x^2\,x'^2}{\sigma^2}$ gleich $\frac{2m^4}{\sigma^2}$ ist; woraus unmittelbar der mittlere Werth jener Funktion

$$= \frac{n^4 - m^4}{\sigma}$$

folgt.

Hieraus ersehen wir, dass, wenn nur eine hinlänglich grosse Anzahl von einander unabhängiger, zufälliger Fehler x, x', x'' etc. vorhanden ist, aus ihnen ein angenäherter Werth des m mit grosser Sicherheit mittelst der Formel

$$m = \sqrt{\frac{x^2 + x'^2 + x''^2 + \text{etc.}}{\sigma}}$$

gefunden werden könne, und dass der mittlere bei dieser Bestimmung zu befürchtende Fehler des Quadrates m^2

$$= \sqrt{\frac{n^4 - m^4}{\sigma}}$$

sei. Da indess diese letzte Formel die Grösse n enthält, so wird es genügen, falls es sich nur um die Erlangung einer ungefähren Vorstellung von dem Genauigkeitsgrade jener Bestimmung handelt, irgend eine specielle Form der Funktion φ anzunehmen. Z. B. wird bei der dritten Annahme der Art. 9. und 11. dieser Fehler $= m^2\sqrt{\frac{2}{\sigma}}$. Wenn dies weniger befriedigt, so kann ein angenäherter Werth von n^4 aus den Fehlern selbst mit Hülfe der Formel

$$\frac{x^4 + x'^4 + x''^4 + \text{etc.}}{\sigma}$$

abgeleitet werden. Im allgemeinen können wir aber versichern, dass für eine zweimal grössere Genauigkeit jener Bestimmung eine vierfache Anzahl von Fehlern erforderlich ist, oder dass das Gewicht der Bestimmung der Anzahl σ selbst proportional ist.

Auf ähnliche Weise wird man ferner, wenn die Beobachtungsfehler einen constanten Theil besitzen, einen angenäherten Werth dieses Theils um so sicherer aus dem arithmetischen Mittel vieler Fehler ableiten können, je grösser deren Anzahl war. Und zwar wird der mittlere zu befürchtende Fehler dieser Bestimmung durch

$$\sqrt{\frac{m^2 - k^2}{\sigma}}$$

ausgedrückt, wenn k den constanten Theil selbst und m den mittleren Fehler der von dem constanten Theil noch nicht befreiten Beobachtungen ausdrückt; oder einfach durch $\frac{m}{\sqrt{\sigma}}$, wenn m den mittleren Fehler der von dem constanten Theil freien Beobachtungen bezeichnet. (Siehe Art. 8.)

16.

In den Art. 12. bis 15. haben wir vorausgesetzt, dass die Fehler x, x', x'' etc. sich auf dieselbe Gattung von Beobachtungen beziehen, so dass die Wahrscheinlichkeit jedes einzelnen durch dieselbe Funktion ausgedrückt werde. Augenscheinlich kann aber die allgemeine Untersuchung der Art. 12. bis 14. ebenso leicht auf den allgemeineren Fall ausgedehnt werden, wo die Wahrscheinlichkeiten der Fehler x, x', x'' etc. durch verschiedene Funktionen $\varphi(x)$, $\varphi'(x')$, $\varphi''(x'')$ etc. ausgedrückt werden, d. h. wo sich jene Fehler auf Beobachtungen verschiedener Schärfe oder Unsicherheit beziehen. Nehmen wir an, x sei der Fehler einer Beobachtung, deren mittlerer zu befürchtender Fehler $= m$ ist; ebenso seien x', x'' etc. die Fehler anderer Beobachtungen, deren mittlere zu befürchtende Fehler bezw. m', m'' etc. sind. Dann wird der mittlere Werth der Summe $x^2 + x'^2 + x''^2 +$ etc. gleich $m^2 + m'^2 + m''^2 +$ etc. sein. Ist nun anderweit schon bekannt, dass die Grössen m, m', m'' etc. in gegebenem Verhältniss zueinander stehen, also den Zahlen 1, μ', μ'' etc. bezw. proportional sind, so wird der mittlere Werth des Ausdrucks

$$\frac{x^2 + x'^2 + x''^2 + \text{etc.}}{1 + \mu'^2 + \mu''^2 + \text{etc.}}$$

$= m^2$ sein. Setzen wir aber einen bestimmten Werth dieses Ausdrucks, jenachdem der Zufall Fehler x, x', x'' etc. liefert, dem m^2 gleich, so wird der mittlere Fehler, welcher dieser Bestimmung noch anhaftet, auf ähnliche Weise wie im vorhergehenden Art.

$$= \frac{\sqrt{n^4 + n'^4 + n''^4 + \text{etc.} - m^4 - m'^4 - m''^4 - \text{etc.}}}{1 + \mu'^2 + \mu''^2 + \text{etc.}}$$

gefunden, wo n', n'' etc. in Bezug auf die Beobachtungen, zu welchen die Fehler x', x'' etc. gehören, dieselbe Bedeutung haben sollen, wie n in Bezug auf die erste Beobachtung. Wenn man nun die Zahlen n, n', n'' etc. den m, m', m'' etc. proportional annehmen darf, so wird jener mittlere zu befürchtende Fehler

$$= \frac{\sqrt{n^4 - m^4}\ \sqrt{1 + \mu'^4 + \mu''^4 + \text{etc.}}}{1 + \mu'^2 + \mu''^2 + \text{etc.}}.$$

Diese Methode, einen angenäherten Werth von m zu bestimmen, ist aber nicht die zweckmässigste. Um dies desto deutlicher zu machen, betrachten wir den allgemeineren Ausdruck

$$y = \frac{x^2 + \alpha' x'^2 + \alpha'' x''^2 + \text{etc.}}{1 + \alpha'\mu'^2 + \alpha''\mu''^2 + \text{etc.}},$$

dessen mittlerer Werth ebenfalls $= m^2$ wird, wie man auch die Coefficienten α', α'' etc. wählen möge. Der mittlere zu befürchtende Fehler aber wird, wenn man einen bestimmten Werth von y, jenachdem der Zufall Fehler x, x', x'' etc. liefert, gleich m^2 annimmt, mit Hülfe der oben vorgetragenen Principien

$$= \frac{\sqrt{(n^4 - m^4) + \alpha'^2 (n'^4 - m'^4) + \alpha''^2 (n''^4 - m''^4) + \text{etc.}}}{1 + \alpha'\mu'^2 + \alpha''\mu''^2 + \text{etc.}}$$

gefunden. Damit dieser mittlere Fehler so klein als möglich wird, ist zu setzen:

$$\alpha' = \frac{n^4 - m^4}{n'^4 - m'^4}\mu'^2$$

$$\alpha'' = \frac{n^4 - m^4}{n''^4 - m''^4}\mu''^2 \text{ etc.}$$

Offenbar können diese Werthe nur dann berechnet werden, wenn überdies die Beziehung der Grössen n, n', n'' etc. zu m, m', m'' etc. anderweit bekannt ist; fehlt aber diese genaue Kenntniss, so erscheint es wenigstens am sichersten*), sie zu einander proportional anzunehmen (siehe Art. 11.), woraus man die Werthe erhält

$$\alpha' = \frac{1}{\mu'^2}, \quad \alpha'' = \frac{1}{\mu''^2} \text{ etc.},$$

d. h. die Coefficienten α', α'' etc. müssen den relativen Gewichten der Beobachtungen, zu welchen die Fehler x', x'' etc. gehören, gleich gesetzt werden, nachdem man das Gewicht der Beobachtung, zu welcher der Fehler x gehört, als Einheit angenommen hat. Wenn hiernach, wie oben, σ die Anzahl der vorhandenen Fehler bezeichnet, so wird der mittlere Werth des Ausdrucks

$$\frac{x^2 + \alpha' x'^2 + \alpha'' x''^2 + \text{etc.}}{\sigma}$$

$= m^2$, und der mittlere zu befürchtende Fehler, wenn wir einen zufällig bestimmten Werth dieses Ausdrucks als wahren Werth von m^2 annehmen, ergiebt sich

$$= \frac{\sqrt{n^4 + \alpha'^2 n'^4 + \alpha''^2 n''^4 + \text{etc.} - \sigma m^4}}{\sigma},$$

und folglich, wenn wir nur die n, n', n'' etc. den m, m', m'' etc. proportional annehmen dürfen,

$$= \sqrt{\frac{n^4 - m^4}{\sigma}},$$

welche Formel mit der oben für den Fall von Beobachtungen derselben Art gefundenen übereinstimmt.

17.

Wenn der Werth einer Grösse, die von einer anderen unbekannten Grösse abhängt, durch eine nicht völlig genaue Beobachtung bestimmt ist, so wird der hieraus berechnete Werth der Unbekannten auch einem Fehler unterworfen sein; es bleibt aber bei dieser Bestimmungsweise nichts der Willkür überlassen. Wenn aber *mehrere* von derselben Unbekannten abhängige Grössen durch nicht völlig genaue Beobachtungen bestimmt sind, so kann man den

*) Wir können uns nämlich die Kenntniss der Grössen μ', μ'' etc. nur in dem einen Falle erlangt denken, wo der Natur der Sache nach Fehler x, x', x'' etc., welche zu 1, μ', μ'' etc. proportional sind, als gleich wahrscheinlich anzunehmen sind, oder vielmehr, wo

$$\varphi(x) = \mu' \varphi'(\mu' x) = \mu'' \varphi''(\mu'' x) \text{ etc.}$$

Werth der Unbekannten entweder aus irgend einer dieser Beobachtungen ableiten, oder auch aus irgend einer Combination mehrerer Beobachtungen, was auf unendlich verschiedene Weisen geschehen kann. Wenn nun auch der auf eine solche Weise erhaltene Werth der Unbekannten immer einem Fehler unterworfen bleibt, so wird doch bei der einen Combination ein grösserer, bei einer anderen ein kleinerer Fehler zu befürchten sein. Aehnlich wird es sich verhalten, wenn mehrere Grössen, die von mehreren Unbekannten zugleich abhängen, beobachtet sind: jenachdem die Anzahl der Beobachtungen entweder der Anzahl der Unbekannten gleich, oder kleiner oder grösser als diese ist, wird die Aufgabe entweder bestimmt oder unbestimmt oder überbestimmt sein (wenigstens im allgemeinen), und im dritten Fall wird man zur Bestimmung der Unbekannten die Beobachtungen auf unendlich verschiedene Weisen combiniren können. Aus dieser Mannigfaltigkeit der Combinationen diejenigen auszuwählen, welche der Sache am besten dienen, d. h. welche die mit den kleinsten Fehlern behafteten Werthe der Unbekannten liefern, ist unstreitig bei der Anwendung der Mathematik auf die Naturwissenschaften eine der wichtigsten Aufgaben.

In der „Theorie der Bewegung der Himmelskörper“ haben wir gezeigt, wie die *wahrscheinlichsten* Werthe der Unbekannten abzuleiten sind, wenn das Gesetz für die Wahrscheinlichkeit der Beobachtungsfehler bekannt ist; und da dieses Gesetz seiner Natur nach in beinahe allen Fällen hypothetisch bleibt, so haben wir jene Theorie auf das plausibelste Gesetz angewendet, wobei die Wahrscheinlichkeit eines Fehlers x der Exponentialgrösse $e^{-h^2x^2}$ proportional genommen wird; und hieraus ist das Verfahren entstanden, welches von uns schon lange und zwar besonders bei astronomischen Rechnungen gebraucht wurde, und jetzt unter dem Namen der Methode der kleinsten Quadrate von den meisten Rechnern angewandt wird.

Später zeigte *Laplace*, indem er die Sache anders angriff, dass gerade dieses Princip, wie auch das Gesetz der Wahrscheinlichkeit der Fehler beschaffen sei, allen anderen immer noch vorzuziehen sei, wenn nur die Anzahl der Beobachtungen eine sehr grosse ist. Ist jedoch die Anzahl der Beobachtungen eine mässige, so bleibt die Frage unentschieden, so dass bei Verwerfung unseres hypothetischen Gesetzes die Methode der kleinsten Quadrate nur deshalb vor anderen empfohlen zu werden verdiente, weil sie zur Vereinfachung der Rechnungen am besten geeignet ist.

Wir hoffen deshalb, den Mathematikern einen Dienst zu erweisen, indem wir bei dieser neuen Behandlung des Gegenstandes zeigten, dass die Methode der kleinsten Quadrate die beste von allen Combinationen liefere, und zwar nicht angenähert, sondern unbedingt, welches auch das Wahrscheinlichkeitsgesetz für die Fehler, und welches auch die Anzahl der Beobachtungen sei, wenn man nur die Definition des mittleren Fehlers nicht im Sinne von *Laplace*, sondern so, wie es von uns in den Art. 5. und 6. geschehen ist, feststellt.

Uebrigens muss hier ausdrücklich hervorgehoben werden, dass es sich in allen folgenden Untersuchungen nur um die unregelmässigen und vom constanten Theil freien Fehler handelt, da es im Grunde zu einer vollkommenen Beobachtungskunst gehört, alle Ursachen constanter Fehler möglichst fernzuhalten. Was für Vortheile aber ein Rechner, welcher solche Beobachtungen zu discutiren unternimmt, von denen man mit Recht argwöhnt, dass sie von constanten Fehlern nicht frei seien, aus der Wahrscheinlichkeitsrechnung selbst erlangen kann, darüber behalten wir uns vor, eine besondere Untersuchung bei einer anderen Gelegenheit zu veröffentlichen.

18.

Aufgabe. Es bezeichne U eine gegebene Funktion der unbekannten Grössen V, V′, V″ etc.; man sucht den mittleren bei der Bestimmung des Werthes von U zu befürchtenden Fehler M, wenn für V, V′, V″ etc. nicht ihre wahren Werthe, sondern diejenigen genommen werden, welche aus von einander unabhängigen und bezw. mit den mittleren Fehlern m, m', m'' etc. behafteten Beobachtungen hervorgehen.

Lösung. Es seien e, e', e'' etc. die Fehler der beobachteten Werthe von V, V′, V″ etc.; alsdann kann der aus ihnen folgende Fehler des Werthes von U durch die lineare Funktion

$$\lambda e + \lambda' e' + \lambda'' e'' + \text{etc.} = \mathrm{E}$$

ausgedrückt werden, wo λ, λ', λ'' etc. die Werthe der Differentialquotienten $\frac{d\mathrm{U}}{d\mathrm{V}}$, $\frac{d\mathrm{U}'}{d\mathrm{V}'}$, $\frac{d\mathrm{U}''}{d\mathrm{V}''}$ etc. für die wahren Werthe der V, V′, V″ etc. sind, wenn nur die Beobachtungen hinlänglich genau sind, um die Quadrate und Produkte der Fehler vernachlässigen zu dürfen. Hieraus folgt erstens, da ja die Beobachtungsfehler als von constanten Theilen frei angenommen werden, dass der mittlere Werth von E gleich 0 sein müsse. Ferner wird der mittlere zu befürchtende Fehler des Werthes von U gleich der Quadratwurzel aus dem mittleren Werthe von E^2 sein, oder M^2 wird der mittlere Werth der Summe

$$\lambda^2 e^2 + \lambda'^2 e'^2 + \lambda''^2 e''^2 + \text{etc.} + 2\lambda\lambda' ee' + 2\lambda\lambda'' ee'' + 2\lambda'\lambda'' e'e'' + \text{etc.}$$

sein. Der mittlere Werth von $\lambda^2 e^2$ ist aber $\lambda^2 m^2$, der mittlere Werth von $\lambda'^2 e'^2$ ist $= \lambda'^2 m'^2$ etc., endlich sind die mittleren Werthe der Produkte $2\lambda\lambda' ee'$ etc. sämmtlich $= 0$. Hieraus schliessen wir also:

$$M = \sqrt{\lambda^2 m^2 + \lambda'^2 m'^2 + \lambda''^2 m''^2 + \text{etc.}}$$

Dieser Lösung wollen wir einige Anmerkungen beifügen.

I. Insoweit man die Beobachtungsfehler als Grössen erster Ordnung ansieht und Grössen höherer Ordnung vernachlässigt, darf man in unserer Formel für λ, λ', λ'' etc. auch diejenigen Werthe der Quotienten $\frac{dU}{dV}$ etc. nehmen, welche aus den beobachteten Werthen der Grössen V, V′, V″ etc. hervorgehen. Wenn U eine lineare Funktion ist, so ist hierbei offenbar kein Unterschied vorhanden.

II. Will man an Stelle der mittleren Fehler der Beobachtungen lieber deren Gewichte einführen, so seien diese, auf eine willkürliche Einheit bezogen, bezw. p, p', p'' etc., und P sei das Gewicht der Bestimmung des sich aus den beobachteten Werthen der Grössen V, V′, V″ etc. ergebenden Werthes von U. Wir erhalten dann

$$P = \frac{1}{\frac{\lambda^2}{p} + \frac{\lambda'^2}{p'} + \frac{\lambda''^2}{p''} + \text{etc.}}.$$

III. Ist T eine andere gegebene Funktion der Grössen V, V′, V″ etc., und ist für deren wahre Werthe

$$\frac{dT}{dV} = \varkappa, \quad \frac{dT}{dV'} = \varkappa', \quad \frac{dT}{dV''} = \varkappa'' \text{ etc.},$$

so wird der Fehler in der aus den beobachteten Werthen von V, V′, V″ etc. erhaltenen Bestimmung des Werthes von T

$$= \varkappa e + \varkappa' e' + \varkappa'' e'' + \text{etc.} = E',$$

und der mittlere bei jener Bestimmung zu befürchtende Fehler

$$= \sqrt{\varkappa^2 m^2 + \varkappa'^2 m'^2 + \varkappa''^2 m''^2 + \text{etc.}}$$

sein. Die Fehler E, E′ werden aber offenbar nicht mehr von einander unabhängig sein, und der mittlere Werth des Produktes EE′ wird, im Gegensatz zum mittleren Werthe des Produktes ee', nicht $= 0$, sondern $= \varkappa\lambda m^2 + \varkappa'\lambda' m'^2 + \varkappa''\lambda'' m''^2 +$ etc. sein.

IV. Man kann unsere Aufgabe auch auf den Fall ausdehnen, wo die Werthe der Grössen V, V′, V″ etc. nicht unmittelbar aus den Beobachtungen gefunden, sondern irgendwie aus Combinationen der Beobachtungen abgeleitet werden, wenn nur die Bestimmungen der einzelnen von einander unabhängig sind, d. h. auf verschiedenen Beobachtungen beruhen: sobald aber diese Bedingung nicht erfüllt ist, würde die Formel für M falsch werden. Wäre z. B. eine oder die andere zur Bestimmung des Werthes von V verwendete Beobachtung auch zur Bestimmung des Werthes von V′ benutzt worden, so würden die Fehler e und e' nicht mehr von einander unabhängig, und der mittlere Werth des Produkts ee' deshalb auch nicht mehr $= 0$ sein. Wenn aber in einem solchen Fall der Zusammenhang der Grössen V und V′ mit den einfachen Beobachtungen, aus denen sie abgeleitet sind, genau bekannt ist, so wird man den mittleren Werth des Produktes ee' nach der Anmerkung III. bestimmen, und so die Formel für M vervollständigen können.

19.

Es seien V, V′, V″ etc. Funktionen der Unbekannten x, y, z etc.; die Anzahl jener sei $= \pi$, die Anzahl der Unbekannten $= \varrho$; wir nehmen an, durch Beobachtungen seien unmittelbar oder mittelbar die Werthe der Funktionen $V = L$, $V' = L'$, $V'' = L''$ etc. gefunden, jedoch so, dass diese Bestimmungen unabhängig von einander sind. Ist ϱ grösser als π, so ist die Aufsuchung der Unbekannten offenbar eine unbestimmte Aufgabe; ist ϱ gleich π, so können die einzelnen x, y, z etc. als Funktionen von V, V′, V″ etc. entweder dargestellt oder in dieser Form gedacht werden, so dass aus den beobachteten Werthen von diesen die Werthe von jenen gefunden werden können, worauf man mit Hülfe des vorigen Art. die diesen einzelnen Bestimmungen zukommende relative Genauigkeit berechnen kann; ist endlich ϱ kleiner als π, so lassen sich die einzelnen x, y, z etc. auf unendlich verschiedene Weisen als Funktionen von V, V′, V″ etc. darstellen, und man kann deshalb für jene auf unendlich verschiedene Weisen Werthe ableiten. Diese Bestimmungen müssten nun völlig identisch sein, wenn den Beobachtungen absolute Genauigkeit zukäme; da dies indess nicht der Fall ist, so werden andere Weisen andere Werthe ergeben, und ebenso werden die aus verschiedenen Combinationen erhaltenen Bestimmungen mit verschiedener Genauigkeit begabt sein.

Wenn übrigens im zweiten oder dritten Fall die Funktionen V, V′, V″ etc. so beschaffen wären, dass $\pi - \varrho + 1$ oder mehrere unter ihnen als Funktionen der übrigen betrachtet werden könnten, so würde die Aufgabe in Bezug auf die letzteren Funktionen immer noch überbestimmt sein, in Bezug auf die Unbekannten x, y, z etc. aber unbestimmt; und man könnte die Werthe der letzteren selbst dann nicht einmal bestimmen, wenn die Werthe der Funktionen V, V′, V″ etc. völlig genau gegeben wären; diesen Fall werden wir aber von unseren Untersuchungen ausschliessen.

Sobald V, V′, V″ etc. nicht von vorn herein *lineare* Funktionen ihrer Variabeln sind, so kann man ihnen diese Form geben, indem man an Stelle der ursprünglichen Unbekannten deren Unterschiede gegen angenäherte Werthe, welche man als anderweit bekannt voraussetzen darf, einsetzt. Die mittleren in den Bestimmungen $V = L$, $V' = L'$, $V'' = L''$ etc. zu befürchtenden Fehler bezeichnen wir bezw. mit m, m', m'' etc., und die Gewichte der Bestimmungen mit p, p', p'' etc., so dass $pm^2 = p'm'^2 = p''m''^2 =$ etc. ist. Wir setzen das Verhältniss der mittleren Fehler zu einander als bekannt voraus, so dass die Gewichte, von denen man eines beliebig annehmen kann, ebenfalls bekannt sind. Endlich setzen wir

$$(V - L)\sqrt{p} = v, \quad (V' - L')\sqrt{p'} = v', \quad (V'' - L'')\sqrt{p''} = v'' \text{ etc.}$$

Dann wird sich die Sache offenbar ebenso verhalten, als wenn unmittelbare Beobachtungen von gleicher Genauigkeit, deren mittlerer Fehler also $= m\sqrt{p} = m'\sqrt{p'} = m''\sqrt{p''}$ etc. ist, oder denen das Gewicht $= 1$ beigelegt wird, auf

$$v = 0, \quad v' = 0, \quad v'' = 0 \text{ etc.}$$

geführt hätten.

20.

Aufgabe. Wir bezeichnen mit v, v', v'' etc. die folgenden linearen Funktionen der Variabeln x, y, z etc.:

$$\left.\begin{aligned} v &= ax + by + cz + \text{etc.} + l \\ v' &= a'x + b'y + c'z + \text{etc.} + l' \\ v'' &= a''x + b''y + c''z + \text{etc.} + l'' \text{ etc.} \end{aligned}\right\} \quad (1)$$

Es soll aus allen Systemen von Coefficienten $\varkappa$, $\varkappa'$, $\varkappa''$ etc., welche allgemein

$$\varkappa v + \varkappa' v' + \varkappa'' v'' + \text{etc.} = x - k$$

geben, wo k eine bestimmte, d. h. von x, y, z etc. unabhängige Grösse ist, das System ermittelt werden, für welches $\varkappa^2 + \varkappa'^2 + \varkappa''^2 + \text{etc.}$ den kleinsten Werth erhält.

Lösung. Wir setzen

$$\left.\begin{aligned} av + a'v' + a''v'' + \text{etc.} &= \xi \\ bv + b'v' + b''v'' + \text{etc.} &= \eta \\ cv + c'v' + c''v'' + \text{etc.} &= \zeta \end{aligned}\right\} \qquad (2)$$

etc. Dann sind auch ξ, η, ζ etc. lineare Funktionen von x, y, z etc., nämlich

$$\left.\begin{aligned} \xi &= x\Sigma a^2 + y\Sigma ab + z\Sigma ac + \text{etc.} + \Sigma al \\ \eta &= x\Sigma ab + y\Sigma b^2 + z\Sigma bc + \text{etc.} + \Sigma bl \\ \zeta &= x\Sigma ac + y\Sigma bc + z\Sigma c^2 + \text{etc.} + \Sigma cl \text{ etc.} \end{aligned}\right\} \qquad (3)$$

(wo Σa^2 die Summe $a^2 + a'^2 + a''^2 + \text{etc.}$ bezeichnet, und analog bei den übrigen); die Anzahl der ξ, η, ζ etc. ist hierbei der Anzahl der Variabeln x, y, z etc. gleich, nämlich $= \varrho$. Man kann deshalb durch Elimination eine Gleichung folgender Art ableiten*):

$$x = \mathrm{A} + [\alpha\alpha]\,\xi + [\alpha\beta]\,\eta + [\alpha\gamma]\,\zeta + \text{etc.},$$

aus welcher durch Substitution der Werthe von ξ, η, ζ etc. nach (3) eine identische Gleichung hervorgehen muss. Wenn man folglich

$$\left.\begin{aligned} a\,[\alpha\alpha] + b\,[\alpha\beta] + c\,[\alpha\gamma] + \text{etc.} &= \alpha \\ a'\,[\alpha\alpha] + b'\,[\alpha\beta] + c'\,[\alpha\gamma] + \text{etc.} &= \alpha' \\ a''\,[\alpha\alpha] + b''\,[\alpha\beta] + c''\,[\alpha\gamma] + \text{etc.} &= \alpha'' \text{ etc.} \end{aligned}\right\} \qquad (4)$$

setzt, so wird nothwendig allgemein

$$\alpha v + \alpha' v' + \alpha'' v'' + \text{etc.} = x - \mathrm{A}. \qquad (5)$$

Diese Gleichung zeigt, dass unter die Werthsysteme der Coefficienten $\varkappa$, $\varkappa'$, $\varkappa''$ etc. sicher auch dieses: $\varkappa = \alpha$, $\varkappa' = \alpha'$, $\varkappa'' = \alpha''$ etc. zu rechnen ist, ebenso dass für ein beliebiges System allgemein

$$(\varkappa - \alpha)\,v + (\varkappa' - \alpha')\,v' + (\varkappa'' - \alpha'')\,v'' + \text{etc.} = \mathrm{A} - k$$

werden muss; eine Gleichung, welche die folgenden einschliesst:

*) Der Grund, weshalb wir für die aus einer solchen Elimination hervorgehenden Coefficienten gerade diese Bezeichnung ausgewählt haben, wird später einleuchten.

Aus 2) folgt: $x - A = (a[\alpha\alpha] + b[\alpha\beta] + c[\alpha\gamma])\,v + (a'[\alpha\alpha] + b'[\alpha\beta] + c'[\alpha\gamma])\,v' + \ldots = \alpha v + \alpha' v' \ldots$

$$(\varkappa-\alpha)\,a+(\varkappa'-\alpha')\,a'+(\varkappa''-\alpha'')\,a''+\text{etc.}=0$$
$$(\varkappa-\alpha)\,b+(\varkappa'-\alpha')\,b'+(\varkappa''-\alpha'')\,b''+\text{etc.}=0$$
$$(\varkappa-\alpha)\,c+(\varkappa'-\alpha')\,c'+(\varkappa''-\alpha'')\,c''+\text{etc.}=0\ \text{etc.}$$

Multipliciren wir diese Gleichungen bezw. mit $[\alpha\alpha]$, $[\alpha\beta]$, $[\alpha\gamma]$ etc. und addiren, so erhalten wir wegen (4):

$$(\varkappa-\alpha)\,\alpha+(\varkappa'-\alpha')\,\alpha'+(\varkappa''-\alpha'')\,\alpha''+\text{etc.}=0$$

oder, was dasselbe ist,

$$\varkappa^2+\varkappa'^2+\varkappa''^2+\text{etc.}$$
$$=\alpha^2+\alpha'^2+\alpha''^2+\text{etc.}+(\varkappa-\alpha)^2+(\varkappa'-\alpha')^2+(\varkappa''-\alpha'')^2+\text{etc.},$$

woraus folgt, dass die Summe $\varkappa^2+\varkappa'^2+\varkappa''^2+$ etc. den kleinsten Werth erhält, wenn man $\varkappa=\alpha$, $\varkappa'=\alpha'$, $\varkappa''=\alpha''$ etc. setzt. Was zu finden war.

Dieser kleinste Werth selbst wird übrigens auf folgende Weise ermittelt. Die Gleichung (5) zeigt, dass

$$\alpha a+\alpha'a'+\alpha''a''+\text{etc.}=1$$
$$\alpha b+\alpha'b'+\alpha''b''+\text{etc.}=0$$
$$\alpha c+\alpha'c'+\alpha''c''+\text{etc.}=0\ \text{etc.}$$

ist. Multiplicirt man diese Gleichungen bezw. mit $[\alpha\alpha]$, $[\alpha\beta]$, $[\alpha\gamma]$ etc. und addirt, so erhält man unter Berücksichtigung der Gleichungen (4) sofort

$$\alpha^2+\alpha'^2+\alpha''^2+\text{etc.}=[\alpha\alpha].$$

21.

Wenn die Beobachtungen die (der Wahrheit sehr nahe kommenden) Gleichungen $v=0$, $v'=0$, $v''=0$ etc. geliefert haben, so muss man, um aus ihnen den Werth der Unbekannten x zu finden, eine solche Combination

$$\varkappa v+\varkappa'v'+\varkappa''v''+\text{etc.}=0$$

dieser Gleichungen aufsuchen, dass der Coefficient von x gleich 1 wird, und die übrigen Unbekannten y, z etc. eliminirt werden; dieser Bestimmung wird nach Art. 18. das Gewicht

$$=\frac{1}{\varkappa^2+\varkappa'^2+\varkappa''^2+\text{etc.}}$$

zu geben sein. Aus dem vorigen Art. folgt daher, die zweckmässigste Bestimmung werde die sein, wenn man $\varkappa=\alpha$, $\varkappa'=\alpha'$, $\varkappa''=\alpha''$ etc. setzt. Alsdann erhält x den Werth A; offenbar kann man denselben Werth (ohne Kenntniss der Multiplicatoren α, α', α'' etc.) auch direkt durch Elimination aus den Gleichungen $\xi = 0$, $\eta = 0$, $\zeta = 0$ etc. ableiten. Das dieser Bestimmung zu ertheilende Gewicht wird $= \frac{1}{[\alpha\alpha]}$, oder der mittlere bei ihr zu befürchtende Fehler wird

$$= m\sqrt{p\,[\alpha\alpha]} = m'\sqrt{p'\,[\alpha\alpha]} = m''\sqrt{p''\,[\alpha\alpha]} \text{ etc.}$$

sein.

Auf analoge Weise wird ferner die zweckmässigste Bestimmung der übrigen Unbekannten y, z etc. für sie dieselben Werthe ergeben, welche durch Elimination aus den nämlichen Gleichungen $\xi = 0$, $\eta = 0$, $\zeta = 0$ etc. hervorgehen.

Bezeichnen wir die allgemeine Summe $v^2 + v'^2 + v''^2 +$ etc. oder, was dasselbe ist,

$$p\,(\mathrm{V}-\mathrm{L})^2 + p'\,(\mathrm{V}'-\mathrm{L}')^2 + p''\,(\mathrm{V}''-\mathrm{L}'')^2 + \text{etc.}$$

mit Ω, so sind offenbar 2ξ, 2η, 2ζ etc. die partiellen Differentialquotienten der Funktion Ω, nämlich

$$2\xi = \frac{d\Omega}{dx},\quad 2\eta = \frac{d\Omega}{dy},\quad 2\zeta = \frac{d\Omega}{dz} \text{ etc.}$$

Demnach werden die Werthe der Unbekannten, welche aus der zweckmässigsten Combination der Beobachtungen hervorgehen, und welche man passend die *plausibelsten Werthe* nennen kann, mit denen identisch sein, die Ω zu einem Minimum machen. Nun drückt $\mathrm{V}-\mathrm{L}$ allgemein die Differenz des berechneten und des beobachteten Werthes aus. Die plausibelsten Werthe der Unbekannten werden deshalb dieselben sein, welche die Summe der mit den Gewichten der Beobachtungen multiplicirten Quadrate der Differenzen zwischen den beobachteten und berechneten Werthen der Grössen V, V′, V″ etc. zu einem Minimum machen, ein Princip, welches wir in der „Theoria Motus Corporum Coelestium“ von einem ganz anderen Gesichtspunkte aus festgestellt hatten. Und wenn ausserdem die relative Genauigkeit der einzelnen Bestimmungen angegeben werden soll, so muss man die x, y, z etc. durch unbestimmte Elimination aus den Gleichungen (3) in folgender Form ableiten:

$$\left.\begin{aligned} x &= \mathrm{A} + [\alpha\alpha]\,\xi + [\alpha\beta]\,\eta + [\alpha\gamma]\,\zeta + \text{etc.}\\ y &= \mathrm{B} + [\beta\alpha]\,\xi + [\beta\beta]\,\eta + [\beta\gamma]\,\zeta + \text{etc.}\\ z &= \mathrm{C} + [\gamma\alpha]\,\xi + [\gamma\beta]\,\eta + [\gamma\gamma]\,\zeta + \text{etc.}\end{aligned}\right\} \qquad (7)$$

etc.,

wonach die plausibelsten Werthe der Unbekannten x, y, z etc.

bezw. A, B, C etc., und die diesen Bestimmungen zukommenden Gewichte $\frac{1}{[\alpha\alpha]}$, $\frac{1}{[\beta\beta]}$, $\frac{1}{[\gamma\gamma]}$ etc., oder die mittleren bei denselben zu befürchtenden Fehler

$$\begin{aligned}
&\text{für } x \ldots\ldots m\sqrt{p\,[\alpha\alpha]} = m'\sqrt{p'\,[\alpha\alpha]} = m''\sqrt{p''\,[\alpha\alpha]} \text{ etc.}\\
&\text{für } y \ldots\ldots m\sqrt{p\,[\beta\beta]} = m'\sqrt{p'\,[\beta\beta]} = m''\sqrt{p''\,[\beta\beta]} \text{ etc.}\\
&\text{für } z \ldots\ldots m\sqrt{p\,[\gamma\gamma]} = m'\sqrt{p'\,[\gamma\gamma]} = m''\sqrt{p''\,[\gamma\gamma]} \text{ etc.}\\
&\text{etc.}
\end{aligned}$$

sein werden, ein Resultat, welches mit dem in der „Theoria Motus Corporum Coelestium" abgeleiteten übereinstimmt.

22.

Wir wollen den allereinfachsten, zugleich aber auch häufigsten Fall, dass nur eine einzige Unbekannte vorhanden ist, und $V = x$, $V' = x$, $V'' = x$ etc. wird, in Kürze besonders behandeln. Es wird nämlich $a = \sqrt{p}$, $a' = \sqrt{p'}$, $a'' = \sqrt{p''}$ etc., $l = -L\sqrt{p}$, $l' = -L'\sqrt{p'}$, $l'' = -L''\sqrt{p''}$ etc., und folglich

$$\xi = (p + p' + p'' + \text{ etc.})\,x - (pL + p'L' + p''L'' + \text{etc.})\,.$$

Hieraus weiter

$$[\alpha\alpha] = \frac{1}{p + p' + p'' + \text{etc.}}$$

$$A = \frac{pL + p'L' + p''L'' + \text{etc.}}{p + p' + p'' + \text{etc.}}\,.$$

Wenn man demnach aus mehreren Beobachtungen von ungleicher Genauigkeit, deren Gewichte bezw. p, p', p'' etc. sind, den Werth einer und derselben Grösse ermittelt hat, und zwar aus der ersten $= L$, aus der zweiten $= L'$, aus der dritten $= L''$ etc., so wird der plausibelste Werth derselben

$$= \frac{pL + p'L' + p''L'' + \text{etc.}}{p + p' + p'' + \text{etc.}}\,,$$

und das Gewicht dieser Bestimmung $= p + p' + p'' +$ etc. sein. Sind alle Beobachtungen von gleicher Genauigkeit, so wird der plausibelste Werth

$$= \frac{L + L' + L'' + \text{etc.}}{\pi}$$

sein, d. h. gleich dem arithmetischen Mittel der beobachteten Werthe, und das Gewicht dieser Bestimmung $= \pi$, wenn man das Gewicht der Beobachtungen als Einheit annimmt.