II.

Aus der
Theorie der Bewegung der Himmelskörper,
welche in Kegelschnitten die Sonne umlaufen.

(1809).

Zweites Buch. Dritter Abschnitt.
Bestimmung der Bahn, die beliebig viele Beobachtungen möglichst genau erfüllt.

172.

Wenn die astronomischen Beobachtungen und die übrigen Zahlenwerthe, auf denen die Berechnung der Bahnen beruht, sich einer absoluten Genauigkeit erfreuten, so würden sich auch die Elemente, mag man sie nun aus drei oder aus vier Beobachtungen abgeleitet haben, sofort völlig genau ergeben (wenigstens insoweit man voraussetzt, dass die Bewegungen genau nach den *Kepler*'schen Gesetzen erfolgen), so dass man durch Hinzunahme immer neuer Beobachtungen nur eine Bestätigung und keine Verbesserung erhalten könnte. Da nun aber in der That alle unsere Messungen und Beobachtungen nur Annäherungen an die Wahrheit sind, und dasselbe von allen auf ihnen beruhenden Rechnungen gelten muss, so wird das höchste Ziel aller, auf bestimmte Erscheinungen sich beziehenden Rechnungen darin bestehen müssen, der Wahrheit möglichst nahe zu kommen. Dies kann aber nur durch eine zweckmässige Combination von *mehr* Beobachtungen geschehen, als deren zur Bestimmung der unbekannten Grössen unbedingt erforderlich sind. Dies Geschäft wird man jedoch dann erst unternehmen können, wenn man bereits eine angenäherte Kenntniss der Bahn besitzt, welche alsdann so verbessert werden muss, dass sie alle Beobachtungen *möglichst genau* erfüllt. Wenn nun auch dieser Ausdruck etwas Unbestimmtes zu enthalten scheint, so werden doch im Folgenden Principien vorgetragen werden, vermittelst deren die Aufgabe einer gesetzmässigen und methodischen Lösung unterworfen wird.

Die höchste Genauigkeit zu erstreben, kann nur dann die Mühe lohnen, wenn an die zu bestimmende Bahn gleichsam die letzte Hand anzulegen ist. So lange man dagegen die Hoffnung hegt, dass bald neue Beobachtungen Anlass zu neuen Verbesserungen bieten werden, so empfiehlt es sich, je nach der Sachlage mehr oder weniger auf die äusserste Genauigkeit zu verzichten, wenn man hierdurch die Weitschichtigkeit der Rechnungen wesentlich verringern kann. Wir werden uns bemühen, für jeden der beiden Fälle Rath zu schaffen.

173.

Vor allem ist es von grösster Bedeutung, die einzelnen geocentrischen Positionen des Himmelskörpers, welche man der bebetreffenden Bahn zu Grunde legen will, nicht aus vereinzelten Beobachtungen abzuleiten, sondern, wenn möglich, aus mehreren so combinirten, dass die zufällig begangenen Fehler sich gegenseitig, so weit es geht, vernichtet haben. Solche Beobachtungen nämlich, welche nur um den Zwischenraum weniger Tage auseinander liegen — oder, nach Lage der Sache, sogar um einen Zwischenraum von 15 oder 20 Tagen — werden bei der Rechnung nicht als eben so viele, verschiedene Positionen zu verwenden sein, sondern es wird vielmehr aus ihnen eine einzige Position abgeleitet werden, welche unter allen gleichsam die mittlere ist und deshalb eine bei weitem grössere Genauigkeit gewährt, als die einzelnen, gesondert betrachteten Beobachtungen. Dies Verfahren stützt sich auf folgende Principien.

Die aus den angenäherten Elementen berechneten geocentrischen Oerter des Himmelskörpers dürfen von den wahren Oertern nur wenig abweichen, und die Differenzen zwischen beiden dürfen nur sehr langsamen Aenderungen unterliegen, so dass sie im Verlaufe von wenigen Tagen nahezu wie Constanten behandelt, oder dass wenigstens ihre Aenderungen als den Zeiten proportional angesehen werden können. Wären demnach die Beobachtungen völlig fehlerfrei, so würden die den Zeiten t, t', t'', t''' etc. entsprechenden Differenzen zwischen den beobachteten und den aus den Elementen berechneten Oertern, d. h. die Differenzen der beobachteten und berechneten Längen und Breiten, oder Rectascensionen und Declinationen, entweder hinlänglich gleiche oder wenigstens gleichmässig und sehr langsam wachsende oder abnehmende Grössen sein. Es mögen z. B. jenen Zeiten die beobachteten Rectascensionen α, α', α'', α''' etc. entsprechen, berechnet aber seien $\alpha + \delta$, $\alpha' + \delta'$, $\alpha'' + \delta''$, $\alpha''' + \delta'''$ etc.; alsdann werden die Unterschiede δ, δ', δ'', δ''' etc. von den wahren Abweichungen der Elemente nur in so weit verschieden sein, als die Beobachtungen selbst fehlerhaft sind; darf man also jene Abweichungen für alle diese Beobachtungen als constant ansehen, so werden die Werthe δ, δ', δ'', δ''' etc. eben so viele verschiedene Bestimmungen derselben Grösse darstellen, als deren verbesserten Werth man deshalb das arithmetische Mittel aus jenen Bestimmungen annehmen darf, so lange wenigstens kein Grund vorliegt, eine oder die andere zu bevorzugen. Glaubt man aber, den einzelnen Beobachtungen nicht denselben Genauigkeitsgrad beilegen zu dürfen, so nehme man an, dass der Genauigkeitsgrad bei den einzelnen bezw. den Zahlen e, e', e'', e''' etc. proportional zu schätzen sei, d. h. dass Fehler, welche diesen Zahlen umgekehrt proportional sind, bei den Beobachtungen gleich leicht hätten begangen werden können; dann wird nach den unten anzugebenden Principien der mittlere wahrscheinlichste Werth nicht mehr das einfache arithmetische Mittel, sondern

$$= \frac{e^2\delta + e'^2\delta' + e''^2\delta'' + e'''^2\delta''' + \text{etc.}}{e^2 + e'^2 + e''^2 + e'''^2 + \text{etc.}}$$

sein. Setzt man nun diesen mittleren Werth $= \Delta$, so wird man für die wahren Rectascensionen bezw. $\alpha + \delta - \Delta$, $\alpha' + \delta' - \Delta$, $\alpha'' + \delta'' - \Delta$, $\alpha''' + \delta''' - \Delta$ etc. annehmen dürfen, und dann wird es gleichgültig sein, welche man in der Rechnung benutzt. Wenn aber entweder die Beobachtungen durch einen allzugrossen Zeitraum von einander getrennt sind, oder wenn hinlänglich angenäherte Elemente der Bahn noch nicht vorliegen, so dass man ihre Abweichungen nicht für alle Beobachtungen als constant ansehen darf, so ist leicht ersichtlich, dass hierdurch kein anderer Unterschied entsteht, als dass die so gefundene mittlere Abweichung nicht sowohl für alle Beobachtungen gemeinsam zu gelten hat, sondern dass sie vielmehr auf eine gewisse mittlere Zeit zu beziehen ist, welche man aus den einzelnen Zeitmomenten ebenso ableiten muss, wie Δ aus den einzelnen Abweichungen, im allgemeinen also auf die Zeit

$$\frac{e^2 t + e'^2 t' + e''^2 t'' + e'''^2 t''' + \text{etc.}}{e^2 + e'^2 + e''^2 + e'''^2 + \text{etc.}}.$$

Will man daher die grösste Genauigkeit erlangen, so wird der geocentrische Ort für dieselbe Zeit aus den Elementen zu berechnen und alsdann von dem mittleren Fehler Δ zu befreien sein, so dass eine möglichst genaue Position hervorgeht; zumeist wird es jedoch weitaus genügen, wenn man den mittleren Fehler auf die der mittleren Zeit nächste Beobachtung bezieht. Was wir hier von den Rectascensionen gesagt haben, gilt ebenso von den Declinationen, oder, wenn man lieber will, von den Längen und Breiten; es wird aber immer vortheilhafter sein, unmittelbar die aus den Elementen berechneten Rectascensionen und Declinationen mit den beobachteten zu vergleichen; so gewinnen wir nämlich nicht nur eine raschere Rechnung, besonders wenn wir die in den Art. 53. bis 60. auseinandergesetzten Methoden benutzen, sondern jene Methode empfiehlt sich ausserdem aus dem Grunde, weil wir auch unvollständige Beobachtungen gebrauchen können, und weil ausserdem, wenn wir Alles auf Längen und Breiten beziehen würden, zu befürchten wäre, dass eine Beobachtung, welche in Bezug auf Rectascension richtig, in Bezug auf Declination schlecht angestellt ist (oder umgekehrt), in beiden Beziehungen verschlechtert und somit ganz unbrauchbar würde. — Uebrigens wird nach den bald zu entwickelnden Principien der dem so gefundenen Mittel beizulegende Grad der Genauigkeit $= \sqrt{e^2 + e'^2 + e''^2 + e'''^2 + \text{etc.}}$ sein, so dass vier oder neun gleich genaue Beobachtungen erforderlich sind, wenn sich das Mittel der doppelten oder dreifachen Genauigkeit erfreuen soll, und so weiter.

174.

Wenn man die Bahn eines Himmelskörpers nach den in den vorhergehenden Abschnitten entwickelten Methoden aus drei oder vier derartigen geocentrischen Positionen bestimmt hat, welche selbst einzeln nach der Regel des vorhergehenden Art. aus mehreren Beobachtungen gebildet waren, so wird diese Bahn zwischen allen diesen Beobachtungen gleichsam die Mitte halten, und in den Differenzen zwischen den beobachteten und berechneten Oertern wird keine Spur einer Regelmässigkeit übrig bleiben, welche sich durch Verbesserung der Elemente beseitigen oder merklich abschwächen liesse. So lange nun der ganze Vorrath der Beobachtungen keinen allzugrossen Zeitraum umfasst, wird man auf diese Weise das erwünschteste Zusammenstimmen der Elemente mit allen Beobachtungen erreichen können, wenn man nur drei oder vier gleichsam normale Positionen geschickt auswählt. Bei der Bestimmung der Bahnen neuer Kometen oder Planeten, deren Beobachtungen ein Jahr noch nicht überschreiten, werden wir durch diese Methode meist so viel erreichen, als die Natur der Sache selbst erlaubt. Ist daher die zu bestimmende Bahn um einen beträchtlichen Winkel gegen die Ekliptik geneigt, so leite man sie im allgemeinen aus drei Beobachtungen ab, welche wir möglichst weit entfernt von einander wählen; wenn wir dabei aber zufällig auf einen der oben ausgeschlossenen Fälle (Art. 160. bis 162.) gerathen, oder wenn die Neigung der Bahn allzuklein erscheint, so werden wir die Bestimmung aus vier Positionen vorziehen, welche wir ebenfalls so weit wie möglich von einander entfernt annehmen.

Ist aber bereits eine längere, mehrere Jahre umfassende Beobachtungsreihe vorhanden, so wird man aus ihr mehrere Normalörter ableiten können, und man würde daher der grössten Genauigkeit wenig Rechnung tragen, wollte man zur Bestimmung der Bahn nur drei oder vier Positionen aussuchen, und alle übrigen gänzlich vernachlässigen. Man wird sich vielmehr in einem solchen Falle, wenn man die höchste Genauigkeit erlangen will, Mühe geben, so viele ausgesuchte Positionen wie möglich zusammenzustellen und zu benutzen. Dann werden also mehr Daten vorhanden sein, als zur Bestimmung der Unbekannten erforderlich sind; alle diese Daten werden aber Fehlern, wenn auch nur kleinen, unterworfen sein, so dass es im allgemeinen unmöglich ist, allen völlig zu genügen. Da nun kein Grund vorhanden ist, weshalb man aus diesen Daten diese oder jene sechs als absolut genau annehmen soll, sondern da man vielmehr nach den Principien der Wahrscheinlichkeit bei allen ohne Unterschied grössere oder kleinere Fehler als gleich möglich voraussetzen muss, und da ferner im allgemeinen geringere Fehler häufiger begangen werden als gröbere, so ist es offenbar, dass eine solche Bahn, welche zwar sechs Daten vollkommen befriedigt, von den übrigen aber mehr oder weniger abweicht, für eine mit den Principien der Wahrscheinlichkeitsrechnung weniger übereinstimmende zu halten ist, als eine andere, welche zwar auch von jenen sechs Daten um ein Geringes unterschieden ist, desto besser aber mit den übrigen zusammenstimmt. Die Aufsuchung der Bahn, welche im strengen Sinn die *grösste* Wahrscheinlichkeit für sich hat, wird von der Kenntniss des Gesetzes abhängen, nach welchem die Wahrscheinlichkeit wachsender Fehler abnimmt; dieses Gesetz hängt aber von so vielen unbestimmten oder zweifelhaften — auch physiologischen — Erwägungen ab, welche der Rechnung nicht unterworfen werden können, dass man ein solches wohl kaum jemals in irgend einem Falle der praktischen Astronomie richtig anzugeben vermag. Nichtsdestoweniger wird die Aufsuchung des Zusammenhangs zwischen diesem Gesetz und der wahrscheinlichsten Bahn, welche wir nunmehr in grösster Allgemeinheit unternehmen wollen, keineswegs für eine unfruchtbare Speculation zu halten sein.

175.

Zu diesem Zweck steigen wir von unserer besonderen Aufgabe zu einer ganz allgemeinen Untersuchung auf, welche sich bei jeder Anwendung der Mathematik auf naturwissenschaftliche Fragen sehr fruchtbar erweist. Es seien V, V′, V″ etc. Funktionen der Unbekannten p, q, r, s etc., μ die Anzahl dieser Funktionen, ν die Anzahl der Unbekannten; wir setzen voraus, als Werthe der Funktionen seien durch unmittelbare Beobachtungen $V = M$, $V' = M'$, $V'' = M''$ etc. gefunden worden. Im allgemeinen wird daher die Entwickelung der Werthe der Unbekannten sich als unbestimmte, bestimmte oder überbestimmte Aufgabe darstellen, jenachdem $\mu < \nu$, $\mu = \nu$ oder $\mu > \nu$ ist*). Hier wird nur von dem letzten Fall die Rede sein, in welchem offenbar eine genaue Darstellung sämmtlicher Beobachtungen nur dann möglich wäre, wenn letztere alle absolut fehlerfrei wären. Da dies aber in Wirklichkeit nicht stattfindet, so wird jedes Werthsystem der Unbekannten p, q, r, s etc. für möglich zu halten sein, aus welchem sich Werthe der Funktionen $M - V$, $M' - V'$, $M'' - V''$ etc. ergeben, welche nicht grösser sind als die Grenzen der Fehler, die bei jenen Beobachtungen begangen werden konnten, was jedoch keineswegs so zu verstehen ist, als ob diese einzelnen möglichen Systeme einen gleichen Grad von Wahrscheinlichkeit besässen.

Wir nehmen zuerst an, es sei bei allen Beobachtungen die Sachlage derartig gewesen, dass kein Grund vorhanden ist, die eine

*) Wenn im dritten Falle die Funktionen V, V′, V″ etc. so beschaffen wären, dass man $\mu + 1 - \nu$ von ihnen oder mehrere als Funktionen der übrigen ansehen könnte, so würde die Aufgabe in Bezug auf diese Funktionen immer noch überbestimmt, in Bezug auf die Grössen p, q, r, s etc. aber unbestimmt sein; man würde nämlich ihre Werthe nicht einmal dann bestimmen können, wenn die Werthe der Funktionen V, V′, V″ etc. absolut genau gegeben wären. Diesen Fall schliessen wir aber von unserer Untersuchung aus.

für weniger genau als die andere zu erachten, oder dass man gleich grosse Fehler bei den einzelnen für gleich wahrscheinlich halten muss. Die Wahrscheinlichkeit, welche irgend einem Fehler Δ beizulegen ist, wird daher durch eine Funktion von Δ ausgedrückt, welche wir mit $\varphi(\Delta)$ bezeichnen wollen. Wenn man nun auch diese Funktion nicht genau angeben kann, so kann man doch wenigstens versichern, dass ihr Werth ein Maximum für $\Delta = 0$ werden müsse, dass er im allgemeinen für gleiche und entgegengesetzte Werthe von Δ der gleiche sei, und endlich, dass er verschwinde, wenn man für Δ den grössten Fehler oder einen noch grösseren Werth annimmt. Eigentlich ist deshalb $\varphi(\Delta)$ zu der Gattung der unstetigen Funktionen zu rechnen, und wenn wir uns erlauben, zum praktischen Gebrauch an ihrer Stelle eine analytische Funktion einzuführen, so muss diese so beschaffen sein, dass sie von $\Delta = 0$ ab nach beiden Seiten gleichsam asymptotisch gegen 0 convergirt, so dass sie ausserhalb der betreffenden Grenze als wirklich verschwindend angesehen werden kann. Ferner wird die Wahrscheinlichkeit, dass der Fehler zwischen den Grenzen Δ und $\Delta + d\Delta$ liege, welche um die unendlich kleine Differenz $d\Delta$ von einander abstehen, durch $\varphi(\Delta)\,d\Delta$ auszudrücken sein; hiernach wird allgemein die Wahrscheinlichkeit, dass der Fehler zwischen D und D′ liege, durch das von $\Delta = \mathrm{D}$ bis $\Delta = \mathrm{D}'$ genommene Integral $\int\varphi(\Delta)\,d\Delta$ dargestellt. Nimmt man dieses Integral von dem grössten negativen Werthe bis zum grössten positiven Werthe von Δ, oder allgemeiner von $\Delta = -\infty$ bis $\Delta = +\infty$, so muss es nothwendig $= 1$ werden.

Nehmen wir also an, irgend ein bestimmtes Werthsystem der Grössen p, q, r, s etc. sei gegeben, so wird die Wahrscheinlichkeit, dass für V aus der Beobachtung der Werth M hervorgehen werde, durch $\varphi(\mathrm{M} - \mathrm{V})$ ausgedrückt, indem man in V für p, q, r, s etc. ihre Werthe einsetzt; ebenso drücken $\varphi(\mathrm{M}' - \mathrm{V}')$, $\varphi(\mathrm{M}'' - \mathrm{V}'')$ etc. die Wahrscheinlichkeiten aus, dass sich aus den Beobachtungen für die Funktionen V′, V″ etc. die Werthe M′, M″ etc. ergeben werden. Deshalb wird, wenn man nur alle Beobachtungen als von einander unabhängige Ergebnisse ansehen darf, das Produkt

$$\varphi(\mathrm{M} - \mathrm{V})\,\varphi(\mathrm{M}' - \mathrm{V}')\,\varphi(\mathrm{M}'' - \mathrm{V}'')\text{ etc.} = \Omega$$

die Erwartung oder die Wahrscheinlichkeit ausdrücken, dass alle diese Werthe gleichzeitig aus den Beobachtungen hervorgehen werden.

176.

So wie nun bei Annahme irgend welcher bestimmten Werthe der Unbekannten jedem System von Werthen der Funktionen V, V′, V″ etc. vor Anstellung der Beobachtung eine bestimmte Wahrscheinlichkeit zukommt, ebenso wird umgekehrt, nachdem aus den Beobachtungen bestimmte Werthe der Funktionen erhalten sind, für die einzelnen Werthsysteme der Unbekannten, aus welchen jene hervorgehen konnten, sich eine bestimmte Wahrscheinlichkeit ergeben; offenbar sind nämlich diejenigen Systeme für wahrscheinlicher zu halten, bei welchen die Erwartung des erhaltenen Ergebnisses die grössere gewesen war. Die Schätzung dieser Wahrscheinlichkeit stützt sich auf folgenden Lehrsatz:

Wenn bei irgend einer zu Grunde gelegten Hypothese H *die Wahrscheinlichkeit irgend eines bestimmten Ergebnisses* E *gleich* h *ist, bei Annahme einer anderen, die erstere ausschliessenden und an sich gleich wahrscheinlichen Hypothese* H′ *aber die Wahrscheinlichkeit desselben Ergebnisses gleich* h′ *ist: dann behaupte ich, wenn das Ergebniss* E *wirklich eingetreten ist, dass die Wahrscheinlichkeit dafür, dass* H *die richtige Hypothese gewesen, sich zur Wahrscheinlichkeit, dass* H′ *die richtige Hypothese gewesen, verhalte wie* h *zu* h′.

Um dies zu beweisen, nehmen wir an, durch Unterscheidung aller Umstände, von denen es abhängt, ob H oder H′ oder eine andere Hypothese Platz greift, um das Ergebniss E oder ein anderes hervorzubringen, sei ein gewisses System der verschiedenen Fälle aufgestellt, welche einzeln für sich (d. h. so lange es ungewiss ist, ob das Ergebniss E oder ein anderes eintreten werde) als gleich wahrscheinlich betrachtet werden müssen, und diese Fälle seien so eingetheilt,

dass unter ihnen gefunden werden	bei denen die Hypothese statthaben muss	mit solchen Modificationen, dass das Ergebniss eintreten muss
m	H	E
n	H	von E verschieden
m'	H′	E
n'	H′	von E verschieden
m''	von H und H′ verschieden	E
n''	von H und H′ verschieden	von E verschieden.

Dann wird $h = \frac{m}{m+n}$, $h' = \frac{m'}{m'+n'}$ sein; ferner war vor Bekanntwerden des Ergebnisses die Wahrscheinlichkeit der Hypothese H $= \frac{m+n}{m+n+m'+n'+m''+n''}$, nach dem Bekanntwerden des Ergebnisses aber, wo die Fälle n, n', n'' aus der Anzahl der möglichen ausscheiden, wird die Wahrscheinlichkeit derselben Hypothese $= \frac{m}{m+m'+m''}$ sein; ebenso wird die Wahrscheinlichkeit der Hypothese H′ vor und nach dem Ergebniss bezw. durch $\frac{m'+n'}{m+n+m'+n'+m''+n''}$ und $\frac{m'}{m+m'+m''}$ ausgedrückt; da nun den Hypothesen H und H′ vor dem Bekanntwerden des Ergebnisses dieselbe Wahrscheinlichkeit beigelegt ist, so wird also

$$m+n = m'+n'$$

sein, woraus sich die Richtigkeit des Lehrsatzes von selbst ergiebt.

So lange wir nun annehmen, dass ausser den Beobachtungen $V = M$, $V' = M'$, $V'' = M''$ etc. keine anderen Daten zur Bestimmung der Unbekannten vorhanden seien, und dass deshalb alle Werthsysteme dieser Unbekannten vor jenen Beobachtungen gleich wahrscheinlich gewesen seien, so wird offenbar die Wahrscheinlichkeit eines jeden bestimmten Systems nach jenen Beobachtungen dem Ω proportional sein. Dies ist so zu verstehen, dass die Wahrscheinlichkeit dafür, dass die Werthe der Unbekannten bezw. zwischen den unendlich nahen Grenzen p und $p+dp$, q und $q+dq$, r und $r+dr$, s und $s+ds$ etc. liegen, durch $\lambda\,\Omega\,dp\,dq\,dr\,ds\ldots$ ausgedrückt werde, wo λ eine von p, q, r, s etc. unabhängige, constante Grösse sein wird. Und zwar wird offenbar $\frac{1}{\lambda}$ der Werth des Integrals ν^{ter} Ordnung $\int^{(\nu)} \Omega\,dp\,dq\,dr\,ds\ldots$ sein, wenn man die einzelnen Variabeln p, q, r, s etc. von dem Werthe $-\infty$ bis zum Werthe $+\infty$ ausdehnt.

177.

Hieraus folgt schon von selbst, dass das wahrscheinlichste Werthsystem der Grössen p, q, r, s etc. dasjenige sein wird, bei welchem Ω den grössten Werth erlangt, und dass es deshalb aus den ν Gleichungen

$$\frac{d\Omega}{dp} = 0, \quad \frac{d\Omega}{dq} = 0, \quad \frac{d\Omega}{dr} = 0, \quad \frac{d\Omega}{ds} = 0 \text{ etc.}$$

zu ermitteln ist. Diese Gleichungen nehmen, wenn man $M - V = v$, $M' - V' = v'$, $M'' - V'' = v''$ etc. und $\frac{d\varphi(\Delta)}{\varphi(\Delta)\, d\Delta} = \varphi'(\Delta)$ setzt, folgende Form an:

$$\frac{dv}{dp}\varphi'(v) + \frac{dv'}{dp}\varphi'(v') + \frac{dv''}{dp}\varphi'(v'') + \text{etc.} = 0$$
$$\frac{dv}{dq}\varphi'(v) + \frac{dv'}{dq}\varphi'(v') + \frac{dv''}{dq}\varphi'(v'') + \text{etc.} = 0$$
$$\frac{dv}{dr}\varphi'(v) + \frac{dv'}{dr}\varphi'(v') + \frac{dv''}{dr}\varphi'(v'') + \text{etc.} = 0$$
$$\frac{dv}{ds}\varphi'(v) + \frac{dv'}{ds}\varphi'(v') + \frac{dv''}{ds}\varphi'(v'') + \text{etc.} = 0 \text{ etc.}$$

Hieraus wird man also durch Elimination die völlig bestimmte Lösung der Aufgabe ableiten können, sobald nur die Natur der Funktion φ' bekannt ist. Da diese aber a priori nicht definirt werden kann, so wollen wir die Sache von einer anderen Seite angreifen, und nachforschen, auf welcher stillschweigend gleichsam als Grundlage angenommenen Funktion ein landläufiges Princip eigentlich beruht, dessen Vortrefflichkeit allgemein anerkannt ist. Wie ein Axiom pflegt man nämlich die Hypothese zu behandeln, wenn irgend eine Grösse durch mehrere unmittelbare, unter gleichen Umständen und mit gleicher Sorgfalt angestellte Beobachtungen bestimmt worden ist, dass alsdann das arithmetische Mittel zwischen allen beobachteten Werthen, wenn auch nicht mit absoluter Strenge, so doch wenigstens sehr nahe den wahrscheinlichsten Werth gebe, so dass es immer das sicherste ist, an diesem festzuhalten. Setzen wir daher $V = V' = V''$ etc. $= p$, so wird allgemein

$$\varphi'(M - p) + \varphi'(M' - p) + \varphi'(M'' - p) + \text{etc.} = 0$$

sein müssen, wenn für p der Werth $\frac{1}{\mu}(M + M' + M'' + \text{etc.})$ eingesetzt wird, was für eine ganze positive Zahl μ auch sein möge. Nimmt man also $M' = M''$ etc. $= M - \mu N$, so wird allgemein, d. h. für jeden ganzen positiven Werth von μ,

$$\varphi'[(\mu - 1)N] = (1 - \mu)\,\varphi'(-N)$$

sein, woraus man leicht entnimmt, dass allgemein $\frac{\varphi'(\Delta)}{\Delta}$ eine constante Grösse sein müsse, welche wir mit k bezeichnen wollen. Hieraus folgt $\log \varphi(\Delta) = \frac{1}{2} k\Delta^2 + \text{Const.}$, oder wenn wir die Basis der hyperbolischen Logarithmen mit e bezeichnen, und Const. $= \log \varkappa$ setzen,

$$\varphi(\Delta) = \varkappa e^{\frac{1}{2} k\Delta^2}.$$

Ferner ist leicht einzusehen, dass k nothwendig negativ sein muss, damit Ω wirklich ein Maximum werden könne, weshalb wir

$$\frac{1}{2} k = -h^2$$

setzen; und da nach einem zuerst von *Laplace* gefundenen, eleganten Lehrsatz das von $\Delta = -\infty$ bis $\Delta = +\infty$ genommene Integral

$$\int e^{-h^2\Delta^2} d\Delta = \frac{\sqrt{\pi}}{h}$$

wird (wo π den halben Kreisumfang für den Radius 1 bezeichnet), so wird unsere Funktion

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

178.

Die soeben ermittelte Funktion kann zwar in aller Strenge die Wahrscheinlichkeiten der Fehler sicher nicht darstellen; denn da die möglichen Fehler immer in bestimmten Grenzen eingeschlossen sind, so müsste die Wahrscheinlichkeit grösserer Fehler sich immer $= 0$ ergeben, während unsere Formel immer einen endlichen Werth liefert. Jedoch ist dieser Mangel, mit dem jede analytische Funktion ihrer Natur nach behaftet sein muss, für alle praktischen Zwecke ohne alle Bedeutung, da, sobald nur erst $h\Delta$ einen beträchtlichen Werth erlangt hat, der Werth unserer Funktion so schnell abnimmt, dass man ihn sicher der Null gleichkommend annehmen darf. Die Fehlergrenzen selbst mit völliger Strenge anzugeben, wird überdies die Natur der Sache niemals gestatten.

Uebrigens wird man die Constante h als das Maass für die Genauigkeit der Beobachtungen ansehen können. Wenn man nämlich annimmt, die Wahrscheinlichkeit eines Fehlers Δ werde in irgend einer Gruppe von Beobachtungen durch

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

in einer anderen Gruppe von genaueren oder ungenaueren Beobachtungen aber durch

$$\frac{h'}{\sqrt{\pi}} e^{-h'^2\Delta^2}$$

ausgedrückt, so wird die Erwartung, es sei bei irgend einer Beobachtung der ersteren Gruppe der Fehler in den Grenzen $-\delta$ und $+\delta$ enthalten, durch das von $\Delta = -\delta$ bis $\Delta = +\delta$ genommene Integral

$$\int \frac{h}{\sqrt{\pi}} e^{-h^2\Delta^2}\, d\Delta$$

ausgedrückt, und ebenso wird die Erwartung, dass der Fehler irgend einer Beobachtung der letzteren Gruppe die Grenzen $-\delta'$ und $+\delta'$ nicht überschreite, durch das von $\Delta = -\delta'$ bis $\Delta = +\delta'$ genommene Integral

$$\int \frac{h'}{\sqrt{\pi}} e^{-h'^2\Delta^2}\, d\Delta$$

ausgedrückt; beide Integrale werden aber offenbar einander gleich, sobald man $h\delta = h'\delta'$ hat. Wenn also z. B. $h' = 2h$ ist, so kann in der ersten Gruppe ebenso leicht ein doppelter Fehler begangen werden, wie in der zweiten ein einfacher, in welchem Falle man den letzteren Beobachtungen nach dem allgemeinen Sprachgebrauch die doppelte Genauigkeit zuschreibt.

179.

Nun werden wir Folgerungen aus diesem Gesetze ziehen. Damit das Produkt

$$\Omega = h^\mu \pi^{-\frac{1}{2}\mu} e^{-h^2(v^2 + v'^2 + v''^2 + \text{etc.})}$$

ein Maximum werde, muss offenbar die Summe $v^2 + v'^2 + v''^2 +$ etc. ein Minimum werden. *Das wahrscheinlichste Werthsystem der Unbekannten* p, q, r, s etc. *wird daher dasjenige sein, bei welchem die Quadrate der Differenzen zwischen den beobachteten und berechneten Werthen der Funktionen* V, V′, V″ etc. *die kleinste Summe ergeben*, wenn nur bei allen Beobachtungen der gleiche Grad von Genauigkeit vorausgesetzt werden darf.

Dies Princip, welches bei allen Anwendungen der Mathematik auf die Naturwissenschaften sehr häufig von Nutzen ist, muss überall mit demselben Recht als Axiom gelten, mit welchem das arithmetische Mittel zwischen mehreren beobachteten Werthen derselben Grösse als wahrscheinlichster Werth angenommen wird.

Auf Beobachtungen von *ungleicher* Genauigkeit kann unser Princip jetzt ohne jede Mühe ausgedehnt werden. Wenn nämlich das Maass der Genauigkeit der Beobachtungen, vermittelst deren $V = M$, $V' = M'$, $V'' = M''$ etc. gefunden ist, bezw. durch h, h', h'' etc. ausgedrückt wird, d. h. wenn man voraussetzt, dass Fehler, welche diesen Grössen umgekehrt proportional sind, bei jenen Beobachtungen gleich leicht begangen werden können, so wird dies offenbar auf dasselbe hinauskommen, als wenn durch Beobachtungen von gleicher Genauigkeit (deren Maass $= 1$ ist) die Werthe der Funktionen hV, $h'V'$, $h''V''$ etc. unmittelbar $= hM$, $h'M'$, $h''M''$ etc. gefunden worden wären; deshalb wird das wahrscheinlichste Werthsystem für die Grössen p, q, r, s etc. dasjenige sein, bei welchem die Summe $h^2v^2 + h'^2v'^2 + h''^2v''^2 +$ etc. d. h. *bei welchem die Summe der Quadrate der Differenzen zwischen den wirklich beobachteten und den berechneten Werthen multiplicirt mit den ihren Genauigkeitsgrad messenden Zahlen ein Minimum wird.* Hiernach ist es nicht einmal nothwendig, dass die Funktionen V, V′, V″ etc. sich auf homogene Grössen beziehen, sondern sie können auch heterogene (z. B. Bogen- und Zeitsecunden) darstellen, wenn man nur das Verhältniss der Fehler zu schätzen vermag, welche bei den einzelnen gleich leicht begangen werden konnten.

180.

Das in dem vorhergehenden Art. dargestellte Princip empfiehlt sich auch aus dem Grunde, weil die numerische Bestimmung der Unbekannten zu einem sehr bequemen Algorithmus führt, wenn die Funktionen V, V′, V″ etc. lineare sind. Wir nehmen an, es sei

$$\begin{aligned} M - V &= v = -m + ap + bq + cr + ds + \text{etc.} \\ M' - V' &= v' = -m' + a'p + b'q + c'r + d's + \text{etc.} \\ M'' - V'' &= v'' = -m'' + a''p + b''q + c''r + d''s + \text{etc.} \end{aligned}$$

etc., und setzen

$$\begin{aligned} av + a'v' + a''v'' + \text{etc.} &= P \\ bv + b'v' + b''v'' + \text{etc.} &= Q \\ cv + c'v' + c''v'' + \text{etc.} &= R \\ dv + d'v' + d''v'' + \text{etc.} &= S \end{aligned}$$

etc. Dann werden die ν Gleichungen des Art. 177., aus welchen die Werthe der Unbekannten zu bestimmen sind, offenbar folgende sein:

$$P = 0, \quad Q = 0, \quad R = 0, \quad S = 0 \quad \text{etc.},$$

wenigstens wenn wir die Beobachtungen als gleich gut voraussetzen, auf welchen Fall nach den Anweisungen des vorigen Art. die übrigen zurückgeführt werden können. Es sind also eben so viele lineare Gleichungen vorhanden, als Unbekannte zu bestimmen sind, woraus deren Werthe durch die gewöhnliche Elimination abgeleitet werden.

Wir wollen jetzt nachsehen, ob diese Elimination immer möglich ist, oder ob jemals die Lösung unbestimmt oder sogar unmöglich werden kann. Aus der Theorie der Elimination weiss man, dass der zweite oder dritte Fall dann eintreten werde, wenn aus den Gleichungen $P = 0$, $Q = 0$, $R = 0$, $S = 0$ etc., bei Auslassung von einer derselben, eine Gleichung gebildet werden kann, die entweder mit der ausgelassenen identisch ist, oder ihr widerstreitet, oder was auf dasselbe hinaus kommt, wenn man eine lineare Funktion $\alpha P + \beta Q + \gamma R + \delta S +$ etc. angeben kann, welche entweder identisch $= 0$ ist, oder wenigstens keine einzige der Unbekannten p, q, r, s etc. enthält. Wir nehmen also an, es werde

$$\alpha P + \beta Q + \gamma R + \delta S + \text{etc.} = \varkappa.$$

Man erhält ohne weiteres die identische Gleichung

$$(v + m)\, v + (v' + m')\, v' + (v'' + m'')\, v'' + \text{etc.} = pP + qQ + rR + sS + \text{etc.}$$

Wenn wir demnach annehmen, dass die Funktionen v, v', v'' etc. durch die Substitutionen $p = \alpha x$, $q = \beta x$, $r = \gamma x$, $s = \delta x$ etc. bezw. in $-m + \lambda x$, $-m' + \lambda' x$, $-m'' + \lambda'' x$ etc. übergehen, so wird offenbar die identische Gleichung entstehen:

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

d. h. es wird

$$\lambda^2 + \lambda'^2 + \lambda''^2 + \text{etc.} = 0, \quad \varkappa + \lambda m + \lambda' m' + \lambda'' m'' + \text{etc.} = 0;$$

hiernach wird aber nothwendigerweise $\lambda = 0$, $\lambda' = 0$, $\lambda'' = 0$ etc. und $\varkappa = 0$ sein müssen. Hieraus erhellt, dass alle Funktionen V, V', V'' etc. so beschaffen sind, dass sich ihre Werthe nicht ändern, wenn die Grössen p, q, r, s etc. um beliebige Grössen zunehmen oder abnehmen, welche den Zahlen α, β, γ, δ etc. proportional sind; dass aber solche Fälle, in welchen die Bestimmung der Unbekannten offenbar auch dann nicht einmal möglich wäre, wenn selbst die wahren Werthe der Funktionen V, V', V'' etc. gegeben sein würden, nicht hierher gehören, daran haben wir schon oben erinnert.

Uebrigens lassen sich auf den hier betrachteten Fall alle übrigen, wo die Funktionen V, V′, V″ etc. nicht linear sind, leicht zurückführen. Bezeichnen wir nämlich mit π, χ, ϱ, σ etc. angenäherte Werthe der Unbekannten p, q, r, s etc. (welche wir leicht ableiten können, indem wir von den μ Gleichungen $V = M$, $V' = M'$, $V'' = M''$ etc. zunächst nur ν benutzen), und führen wir an Stelle der Unbekannten andere p', q', r', s' etc. ein, indem wir $p = \pi + p'$, $q = \chi + q'$, $r = \varrho + r'$, $s = \sigma + s'$ etc. setzen, so werden offenbar die Werthe dieser neuen Unbekannten so klein sein, dass man ihre Quadrate und Produkte vernachlässigen darf, wodurch die Gleichungen von selbst linear werden. Wenn aber alsdann nach beendigter Rechnung die Werthe der Unbekannten p', q', r', s' etc. wider Erwarten sich so gross ergeben, dass die Vernachlässigung der Quadrate und Produkte nicht gefahrlos erschiene, so wird eine Wiederholung derselben Operation (indem man an Stelle der π, χ, ϱ, σ etc. die verbesserten Werthe der p, q, r, s etc. nimmt) schnell Abhülfe schaffen.

181.

So oft demnach nur eine einzige Unbekannte p vorhanden ist, zu deren Bestimmung die Werthe der Funktionen $ap + n$, $a'p + n'$, $a''p + n''$ etc. bezw. $= M, M', M''$ etc., und zwar durch gleich genaue Beobachtungen gefunden sind, so wird der wahrscheinlichste Werth des p

$$= \frac{am + a'm' + a''m'' + \text{etc.}}{a^2 + a'^2 + a''^2 + \text{etc.}} = A$$

sein, wenn man m, m', m'' etc. bezw. für $M - n$, $M' - n'$, $M'' - n''$ etc. schreibt.

Um nun den Grad der Genauigkeit zu schätzen, die wir bei diesem Werthe anzunehmen haben, setzen wir voraus, die Wahrscheinlichkeit eines Fehlers Δ bei den Beobachtungen werde durch

$$\frac{h}{\sqrt{\pi}} e^{-h^2\Delta^2}$$

ausgedrückt. Hiernach wird die Wahrscheinlichkeit, dass der wahre Werth des p gleich $A + p'$ sei, der Funktion

$$e^{-h^2[(ap - m)^2 + (a'p - m')^2 + (a''p - m'')^2 + \text{etc.}]}$$

proportional sein, wenn man $A + p'$ für p einsetzt. Der Exponent dieser Funktion kann auf die Form

$$-h^2(a^2 + a'^2 + a''^2 + \text{etc.})(p^2 - 2pA + B)$$

gebracht werden, wo B von p unabhängig ist; deshalb wird die Funktion selbst zu

$$e^{-h^2(a^2+a'^2+a''^2+\text{etc.})p'^2}$$

proportional sein. Offenbar ist also dem Werthe A derselbe Grad der Genauigkeit zuzuschreiben, als wenn er durch eine unmittelbare Beobachtung gefunden wäre, deren Genauigkeit sich zur Genauigkeit der ursprünglichen Beobachtungen verhält wie

$h\sqrt{a^2+a'^2+a''^2+\text{etc.}}$ zu h, oder wie $\sqrt{a^2+a'^2+a''^2+\text{etc.}}$ zu 1.

182.

Der Untersuchung über den Grad der Genauigkeit, welchen man den Werthen der Unbekannten, wenn ihrer mehrere vorhanden sind, beilegen muss, ist eine genauere Betrachtung der Funktion $v^2+v'^2+v''^2+\text{etc.}$, die wir mit W bezeichnen wollen, vorauszuschicken.

I. Setzen wir

$$\frac{1}{2}\frac{d\mathrm{W}}{dp} = p' = \lambda + \alpha p + \beta q + \gamma r + \delta s + \text{etc.},$$

und

$$\mathrm{W} - \frac{p'^2}{\alpha} = \mathrm{W}',$$

so wird offenbar $p' = \mathrm{P}$, und da

$$\frac{d\mathrm{W}'}{dp} = \frac{d\mathrm{W}}{dp} - \frac{2p'}{\alpha}\frac{dp'}{dp} = 0$$

ist, so muss die Funktion W′ von p unabhängig sein. Der Coefficient $\alpha = a^2+a'^2+a''^2+\text{etc.}$ wird offenbar immer eine positive Grösse sein.

II. Setzen wir ebenso

$$\frac{1}{2}\frac{d\mathrm{W}'}{dq} = q' = \lambda' + \beta' q + \gamma' r + \delta' s + \text{etc.},$$

und

$$\mathrm{W}' - \frac{q'^2}{\beta'} = \mathrm{W}'',$$

so wird

$$q' = \frac{1}{2}\frac{d\mathrm{W}}{dq} - \frac{p'}{\alpha}\frac{dp'}{dq} = \mathrm{Q} - \frac{\beta}{\alpha}p', \text{ und } \frac{d\mathrm{W}''}{dq} = 0$$

sein, wonach offenbar die Funktion W″ von p und q unabhängig sein wird. Dies würde nicht stattfinden, wenn $\beta' = 0$ werden könnte. Offenbar ergiebt sich aber W' aus $v^2 + v'^2 + v''^2 +$ etc., indem man die Grösse p aus v, v', v'' etc. mit Hülfe der Gleichung $p' = 0$ eliminirt; sonach wird β' die Summe der Coefficienten von q^2 in v^2, v'^2, v''^2 etc. nach jener Elimination sein; diese einzelnen Coefficienten selbst sind aber Quadrate und können nicht alle zugleich verschwinden, abgesehen von dem oben ausgeschlossenen Fall, in welchem die Unbekannten unbestimmt bleiben. Es muss β' deshalb offenbar eine positive Grösse sein.

III. Setzen wir ferner

$$\frac{1}{2}\frac{dW''}{dr} = r' = \lambda'' + \gamma'' r + \delta'' s + \text{etc.},$$

und

$$W'' - \frac{r'^2}{\gamma''} = W''',$$

so wird

$$r' = R - \frac{\gamma}{\alpha} p' - \frac{\gamma'}{\beta'} q',$$

und W''' sowohl von p, als von q und r unabhängig sein. Dass übrigens der Coefficient γ'' nothwendig positiv ist, wird analog wie in II. bewiesen. Man sieht nämlich leicht ein, dass γ'' die Summe der Coefficienten von r^2 in v^2, v'^2, v''^2 etc. ist, nachdem man die Grössen p und q mit Hülfe der Gleichungen $p' = 0$, $q' = 0$ aus v, v', v'' etc. eliminirt hat.

IV. Auf dieselbe Weise wird, wenn wir

$$\frac{1}{2}\frac{dW'''}{ds} = s' = \lambda''' + \delta''' s + \text{etc., und } W'''' = W''' - \frac{s'^2}{\delta'''}$$

setzen,

$$s' = S - \frac{\delta}{\alpha} p' - \frac{\delta'}{\beta'} q' - \frac{\delta''}{\gamma''} r',$$

W'''' von p, q, r, s unabhängig und δ''' eine positive Grösse sein.

V. Man kann, wenn es ausser p, q, r, s noch andere Unbekannte giebt, ebenso weiter gehen, so dass man endlich

$$W = \frac{1}{\alpha} p'^2 + \frac{1}{\beta'} q'^2 + \frac{1}{\gamma''} r'^2 + \frac{1}{\delta'''} s'^2 + \text{etc.} + \text{Const.}$$

erhält, wo alle Coefficienten α, β', γ'', δ''' etc. positive Grössen sein werden.

VI. Die Wahrscheinlichkeit irgend eines Systems von bestimmten Werthen der Grössen p, q, r, s etc. ist nun der Funktion

$e^{-h^2 W}$ proportional, so dass, wenn der Werth der Grösse p unbestimmt bleibt, die Wahrscheinlichkeit eines Systems bestimmter Werthe der übrigen dem von $p = -\infty$ bis $p = +\infty$ ausgedehnten Integral

$$\int e^{-h^2 W} dp$$

proportional sein wird, welches nach dem Theorem von *Laplace*

$$= h^{-1} \alpha^{-\frac{1}{2}} \pi^{\frac{1}{2}} e^{-h^2\left[\frac{1}{\beta'} q'^2 + \frac{1}{\gamma''} r'^2 + \frac{1}{\delta'''} s'^2 + \text{etc.}\right]}$$

wird; es wird also die obige Wahrscheinlichkeit der Funktion $e^{-h^2 W'}$ proportional sein. Wenn überdies q ebenso als Variable angesehen wird, dann ist die Wahrscheinlichkeit eines Systems bestimmter Werthe für r, s etc. dem von $q = -\infty$ bis $q = +\infty$ genommenen Integral

$$\int e^{-h^2 W'} dq$$

proportional, welches

$$= h^{-1} \beta'^{-\frac{1}{2}} \pi^{\frac{1}{2}} e^{-h^2\left[\frac{1}{\gamma''} r'^2 + \frac{1}{\delta'''} s'^2 + \text{etc.}\right]}$$

wird, oder proportional der Funktion $e^{-h^2 W''}$. Wenn man auch r als Variable behandelt, so wird ganz analog die Wahrscheinlichkeit bestimmter Werthe der übrigen s etc. der Funktion $e^{-h^2 W'''}$ proportional sein, und so weiter. Wir wollen annehmen, die Anzahl der Unbekannten steige bis auf vier; dann wird derselbe Schluss auch gelten, wenn sie grösser oder kleiner ist. Der wahrscheinlichste Werth von s wird hier $= -\frac{\lambda'''}{\delta'''}$ sein, und die Wahrscheinlichkeit, dass derselbe sich von dem wahren um die Differenz σ unterscheide, wird der Funktion $e^{-h^2 \delta''' \sigma^2}$ proportional sein, woraus wir schliessen, dass das Maass der jener Bestimmung beizulegenden relativen Genauigkeit durch $\sqrt{\delta'''}$ ausgedrückt wird, wenn das Maass der den ursprünglichen Beobachtungen beizulegenden Genauigkeit $= 1$ gesetzt wird.

183.

Durch die Methode des vorhergehenden Art. wird das Genauigkeitsmaass nur für diejenige Unbekannte bequem ausgedrückt, der beim Geschäft der Elimination der letzte Platz angewiesen ist; um nun diese Unbequemlichkeit zu vermeiden, empfiehlt es sich, den Coefficienten δ''' auf andere Weise auszudrücken. Aus den Gleichungen

$$\mathrm{P} = p'$$

$$\mathrm{Q} = q' + \frac{\beta}{\alpha} p'$$

$$\mathrm{R} = r' + \frac{\gamma'}{\beta'} q' + \frac{\gamma}{\alpha} p'$$

$$\mathrm{S} = s' + \frac{\delta''}{\gamma''} r' + \frac{\delta'}{\beta'} q' + \frac{\delta}{\alpha} p'$$

folgt, dass die p', q', r', s' durch P, Q, R, S folgendermaassen ausgedrückt werden können:

$$\begin{aligned} p' &= \mathrm{P} \\ q' &= \mathrm{Q} + \mathfrak{A}\mathrm{P} \\ r' &= \mathrm{R} + \mathfrak{B}'\mathrm{Q} + \mathfrak{A}'\mathrm{P} \\ s' &= \mathrm{S} + \mathfrak{C}''\mathrm{R} + \mathfrak{B}''\mathrm{Q} + \mathfrak{A}''\mathrm{P}, \end{aligned}$$

so dass $\mathfrak{A}$, $\mathfrak{A}'$, $\mathfrak{B}'$, $\mathfrak{A}''$, $\mathfrak{B}''$, $\mathfrak{C}''$ bestimmte Grössen sind. Es wird daher (wenn wir die Anzahl der Unbekannten auf vier einschränken)

$$s = -\frac{\lambda'''}{\delta'''} + \frac{\mathfrak{A}''}{\delta'''}\mathrm{P} + \frac{\mathfrak{B}''}{\delta'''}\mathrm{Q} + \frac{\mathfrak{C}''}{\delta'''}\mathrm{R} + \frac{1}{\delta'''}\mathrm{S}.$$

Hieraus leiten wir folgenden Schluss ab. Die wahrscheinlichsten Werthe der Unbekannten p, q, r, s etc., welche durch Elimination aus den Gleichungen $\mathrm{P} = 0$, $\mathrm{Q} = 0$, $\mathrm{R} = 0$, $\mathrm{S} = 0$ etc. abzuleiten sind, werden offenbar, wenn man für den Augenblick P, Q, R, S etc. als Variable betrachtet, demselben Eliminationsverfahren gemäss in linearer Form durch P, Q, R, S etc. ausgedrückt, so dass man erhält

$$\begin{aligned} p &= \mathrm{L} + \mathrm{AP} + \mathrm{BQ} + \mathrm{CR} + \mathrm{DS} + \text{etc.} \\ q &= \mathrm{L}' + \mathrm{A'P} + \mathrm{B'Q} + \mathrm{C'R} + \mathrm{D'S} + \text{etc.} \\ r &= \mathrm{L}'' + \mathrm{A''P} + \mathrm{B''Q} + \mathrm{C''R} + \mathrm{D''S} + \text{etc.} \\ s &= \mathrm{L}''' + \mathrm{A'''P} + \mathrm{B'''Q} + \mathrm{C'''R} + \mathrm{D'''S} + \text{etc.} \\ &\text{etc.} \end{aligned}$$

Hiernach werden die wahrscheinlichsten Werthe der p, q, r, s etc. offenbar bezw. L, L′, L″, L‴ etc. sein, und das diesen Bestimmungen zuzuschreibende Genauigkeitsmaass wird bezw. durch

$$\sqrt{\frac{1}{\mathrm{A}}},\ \sqrt{\frac{1}{\mathrm{B}'}},\ \sqrt{\frac{1}{\mathrm{C}''}},\ \sqrt{\frac{1}{\mathrm{D}'''}} \text{ etc.}$$

ausgedrückt, wenn die Genauigkeit der ursprünglichen Beobachtungen $= 1$ gesetzt ist. Was wir nämlich in Betreff der Bestimmung der Unbekannten s vorher gezeigt haben (bei welcher δ''' dem $\frac{1}{D'''}$ entspricht), lässt sich durch blosse Permutation der Unbekannten auf alle übrigen übertragen.

184.

Um die vorhergehenden Untersuchungen an einem Beispiel zu erläutern, wollen wir annehmen, es sei durch Beobachtungen, bei denen man eine gleiche Genauigkeit voraussetzen darf, gefunden

$$\begin{aligned} p - q + 2r &= 3 \\ 3p + 2q - 5r &= 5 \\ 4p + q + 4r &= 21, \end{aligned}$$

durch eine vierte aber, der nur die halbe Genauigkeit zuzuschreiben ist, habe sich ergeben

$$-2p + 6q + 6r = 28.$$

An Stelle der letzten Gleichung führen wir daher die folgende ein:

$$-p + 3q + 3r = 14,$$

und nehmen an, diese sei aus einer den früheren an Genauigkeit gleichen Beobachtung hervorgegangen.

Hiernach wird

$$\begin{aligned} P &= 27p + 6q - 88 \\ Q &= 6p + 15q + r - 70 \\ R &= q + 54r - 107, \end{aligned}$$

und hieraus durch Elimination

$$\begin{aligned} 19899\,p &= 49154 + 809\,P - 324\,Q + 6\,R \\ 19899\,q &= 70659 - 324\,P + 1458\,Q - 27\,R \\ 19899\,r &= 38121 + 6\,P - 27\,Q + 369\,R. \end{aligned}$$

Die wahrscheinlichsten Werthe der Unbekannten werden daher

$$\begin{aligned} p &= 2{,}470 \\ q &= 3{,}551 \\ r &= 1{,}916 \end{aligned}$$

sein, und die diesen Bestimmungen zuzuschreibende relative Genauigkeit, wenn die Genauigkeit der ursprünglichen Beobachtungen $= 1$ gesetzt ist, wird

$$\text{für } p \ldots \sqrt{\frac{19899}{809}} = 4{,}96$$

$$\text{für } q \ldots \sqrt{\frac{19899}{1458}} = 3{,}69$$

$$\text{für } r \ldots \sqrt{\frac{19899}{369}} = 7{,}34\,.$$

185.

Der bisher behandelte Gegenstand könnte zu mehreren eleganten, analytischen Untersuchungen Veranlassung geben, bei denen wir uns jedoch hier nicht aufhalten wollen, um uns nicht zu weit von unserem Vorhaben zu entfernen. Aus demselben Grunde müssen wir uns die Auseinandersetzung der Kunstgriffe, durch welche die numerische Rechnung auf einen schneller zum Ziele führenden Algorithmus gebracht werden kann, für eine andere Gelegenheit vorbehalten. Eine einzige Bemerkung wollen wir uns hier anzufügen gestatten. Falls die Anzahl der Funktionen oder der vorgelegten Gleichungen beträchtlich ist, wird die Rechnung deshalb hauptsächlich ein wenig lästiger, weil die Coefficienten, mit welchen die ursprünglichen Gleichungen zu multipliciren sind, um P, Q, R, S etc. abzuleiten, meistens wenig bequeme Decimalbrüche enthalten. Wenn es in einem solchen Falle nicht die Mühe zu lohnen scheint, diese Multiplicationen mit Hülfe der Logarithmentafeln so genau wie möglich auszuführen, so wird es in den meisten Fällen ausreichen, an Stelle dieser Multiplicatoren andere für die Rechnung bequemere anzuwenden, welche von jenen wenig verschieden sind. Diese Vernachlässigung kann keine merklichen Fehler erzeugen, abgesehen von dem einen Falle, wo sich das Genauigkeitsmaass für die Bestimmung der Unbekannten viel geringer ergiebt, als die Genauigkeit der ursprünglichen Beobachtungen war.

186.

Uebrigens wird das Princip, dass die Quadrate der Differenzen zwischen den beobachteten und berechneten Grössen die allerkleinste Summe ergeben müssen, auch unabhängig von der Wahrscheinlichkeitsrechnung auf folgende Weise erwogen werden können.

Wenn die Anzahl der Unbekannten der Anzahl der beobachteten und von ihnen abhängigen Grössen gleich ist, so kann man jene so bestimmen, dass diesen genau Genüge geschieht. Wenn aber jene Anzahl kleiner als diese ist, so kann man ein absolut genaues Zusammenstimmen nicht erlangen, insoweit sich die Beobachtungen nicht absoluter Genauigkeit erfreuen. In diesem Falle muss man sich daher Mühe geben, die möglichst beste Uebereinstimmung herzustellen, oder die Differenzen so viel wie möglich zu verkleinern. Diese Forderung enthält aber ihrer Natur nach etwas Unbestimmtes. Wenn nämlich auch ein Werthsystem der Unbekannten, welches *alle* Differenzen bezw. kleiner als ein anderes ergiebt, diesem letzteren zweifelsohne vorzuziehen ist, so bleibt nichtsdestoweniger die Wahl zwischen zwei Systemen, von denen das eine für einige Beobachtungen eine bessere Uebereinstimmung erzeugt, das andere aber für andere, in gewisser Hinsicht unserem Ermessen überlassen, und es können offenbar unzählige Principien vorgeschlagen werden, durch welche die obige Bedingung erfüllt wird. Bezeichnet man die Differenzen zwischen den Beobachtungen und der Rechnung mit Δ, Δ', Δ'' etc., so wird der obigen Bedingung nicht nur genügt, wenn $\Delta^2 + \Delta'^2 + \Delta''^2 +$ etc. ein Minimum wird (was unserem Princip entspricht), sondern auch wenn $\Delta^4 + \Delta'^4 + \Delta''^4 +$ etc. oder $\Delta^6 + \Delta'^6 + \Delta''^6 +$ etc. oder allgemein die Summe der Potenzen mit irgend einem beliebigen geraden Exponenten zu einem Minimum wird. Von allen diesen Principien ist aber das unsrige das einfachste, während wir bei den übrigen zu den verwickeltsten Rechnungen geführt werden. Uebrigens ist unser Princip, dessen wir uns schon seit dem Jahre 1795 bedient haben, kürzlich auch von *Legendre* in dem Werke „Nouvelles méthodes pour la détermination des orbites des comètes, Paris 1806“ aufgestellt worden, woselbst auch mehrere andere Eigenthümlichkeiten dieses Princips auseinandergesetzt sind, welche wir hier der Kürze wegen unterdrücken.

Wenn wir eine Potenz mit einem unendlich grossen, geraden Exponenten annehmen würden, so würden wir auf dasjenige System geführt werden, bei welchem die grössten Differenzen so klein wie möglich werden.

Laplace bedient sich zur Auflösung linearer Gleichungen, deren Anzahl grösser ist als die Anzahl der unbekannten Grössen, eines anderen Princips, welches seiner Zeit schon von *Boscovich* vorgeschlagen war, dass nämlich die Differenzen selbst, aber alle positiv genommen, eine möglichst kleine Summe erzeugen sollen. Es lässt sich leicht zeigen, dass das System der Werthe der Unbekannten, welches aus diesem Princip allein ermittelt ist, nothwendig*) so vielen Gleichungen aus der Anzahl der vorgelegten genau genügen muss, als Unbekannte vorhanden sind, so dass die übrigen Gleichungen nur in so weit in Betracht kommen, als sie *zur entscheidenden Wahl beitragen;* wenn daher z. B. die Gleichung V = M zu der Anzahl derer gehört, welchen nicht genügt wird, so würde an dem System der nach jenem Princip gefundenen Werthe nichts geändert, wenn man auch an Stelle von M irgend einen anderen beliebigen Werth N beobachtet hätte, wenn nur die Differenzen $M - n$ und $N - n$, wo mit n der berechnete Werth bezeichnet ist, mit demselben Vorzeichen behaftet sind. Uebrigens regulirt *Laplace* jenes Princip in gewisser Hinsicht durch Hinzufügung einer neuen Bedingung: er fordert nämlich, dass die Summe der Differenzen selbst, ohne Aenderung der Vorzeichen, = 0 wird. Hierdurch wird bewirkt, dass die Anzahl der genau dargestellten Gleichungen um eine Einheit kleiner wird als die Anzahl der unbekannten Grössen; gleichwohl wird aber unsere obige Bemerkung immer noch statthaben, wenn nur wenigstens zwei Unbekannte vorhanden waren.

187.

Nach diesen allgemeinen Untersuchungen wenden wir uns wieder zu unserem eigentlichen Vorhaben, um dessentwillen jene angestellt worden sind. Bevor wir zur möglichst genauen Bestimmung der Bahn aus mehr Beobachtungen als den nothwendigerweise erforderlichen schreiten können, muss schon eine angenäherte Bestimmung vorhanden sein, welche von allen gegebenen Beobachtungen nicht allzuviel abweicht. Die Verbesserungen, welche an diese angenäherten Elemente noch anzubringen sind, um einen möglichst genauen Anschluss zu erreichen, sollen als die gesuchten Grössen der Aufgabe angesehen werden. Da man annehmen kann, dass diese sich so klein ergeben werden, dass die Quadrate und Produkte vernachlässigt werden dürfen, so werden die Aenderungen, welche die berechneten geocentrischen Oerter des Himmelskörpers hierdurch erlangen, nach den im zweiten Abschnitt des ersten Buches gegebenen Differentialformeln berechnet werden können. Die gemäss den gesuchten, verbesserten Elementen berechneten

*) Abgesehen von besonderen Fällen, wo die Lösung in gewisser Beziehung unbestimmt bleibt.

Oerter werden daher durch lineare Funktionen der Verbesserungen der Elemente dargestellt, und die Vergleichung jener mit den beobachteten Oertern führt nach den oben auseinandergesetzten Principien zur Bestimmung der wahrscheinlichsten Werthe. Diese Operationen erfreuen sich einer so grossen Einfachheit, dass sie einer weiteren Erläuterung nicht bedürfen, und es ist von selbst einleuchtend, dass beliebig viele und beliebig weit von einander entfernte Beobachtungen zur Benutzung herangezogen werden können. — Derselben Methode kann man sich auch zur Verbesserung der *parabolischen* Bahnen der Cometen bedienen, wenn zufällig eine längere Beobachtungsreihe vorhanden ist, und die möglichst beste Uebereinstimmung verlangt wird.

188.

Die vorstehende Methode ist vorzüglich den Fällen angepasst, wo die höchste Genauigkeit gewünscht wird, sehr häufig aber treten Fälle ein, wo man ohne Bedenken ein wenig von jener aufgeben darf, wenn man auf diese Weise die Weitschichtigkeit der Rechnung wesentlich einschränken kann, hauptsächlich wenn die Beobachtungen noch keinen grossen Zeitraum einschliessen, und deshalb an eine sozusagen definitive Bestimmung der Bahn noch nicht gedacht wird. In solchen Fällen kann die nachfolgende Methode mit bemerkenswerthem Vortheil in Benutzung genommen werden.

Es mögen aus der ganzen Menge der Beobachtungen zwei vollständige Oerter L und L′ ausgewählt, und für die entsprechenden Zeiten aus den angenäherten Elementen die Entfernungen des Himmelskörpers von der Erde berechnet werden. Darauf bilde man in Hinblick auf diese Entfernungen drei Hypothesen, indem man bei der ersten die berechneten Werthe beibehält, bei der zweiten Hypothese die erste Entfernung ändert und die zweite bei der dritten Hypothese; beide Aenderungen können nach Maassgabe der Unsicherheit, welche man bei jenen Entfernungen als übrigbleibend voraussetzt, nach Belieben angenommen werden. Gemäss diesen drei Hypothesen, welche wir in folgendem Schema darstellen,

	Hyp. I.	Hyp. II.	Hyp. III.
Dem ersten Orte entsprechende Entfernung*)	D	$D + \delta$	D
Dem zweiten Orte entsprechende Entfernung	D'	D'	$D' + \delta'$

*) Noch bequemer wird es sein, an Stelle der Entfernungen selbst die Logarithmen der curtirten Distanzen zu benutzen.

werden nach den im ersten Buch auseinandergesetzten Methoden aus den beiden Oertern L und L′ drei Elementensysteme berechnet und darauf aus allen diesen die geocentrischen Oerter des Himmelskörpers, welche den Zeiten aller übrigen Beobachtungen entsprechen. Diese seien (indem man die einzelnen Längen und Breiten oder Rectascensionen und Declinationen besonders aufschreibt)

im ersten System	M,	M',	M''	etc.
im zweiten System . . .	$M + \alpha$,	$M' + \alpha'$,	$M'' + \alpha''$	etc.
im dritten System . . .	$M + \beta$,	$M' + \beta'$,	$M'' + \beta''$	etc.
Ferner seien die beobachteten Oerter bezw. . .	N,	N',	N''	etc.

Insoweit nun kleinen Aenderungen der Entfernungen D, D′ proportionale Aenderungen der einzelnen Elemente und der aus ihnen berechneten geocentrischen Oerter entsprechen, wird man voraussetzen dürfen, dass die aus einem vierten Elementensystem berechneten geocentrischen Oerter, welches auf die Entfernungen von der Erde $D + x\delta$, $D' + y\delta'$ begründet ist, bezw. $M + \alpha x + \beta y$, $M' + \alpha' x + \beta' y$, $M'' + \alpha'' x + \beta'' y$ etc. sein werden. Hieraus werden darauf nach den vorhergehenden Untersuchungen die Grössen x und y so bestimmt, dass jene Grössen bezw. mit N, N′, N″ etc. möglichst gut übereinstimmen (indem man der relativen Genauigkeit der Beobachtungen Rechnung trägt). Das verbesserte Elementensystem selbst, wird entweder ebenso aus L, L′ und den Entfernungen $D + x\delta$, $D' + y\delta'$, oder nach bekannten Regeln aus den drei ersten Elementensystemen durch einfache Interpolation abgeleitet werden können.

189.

Diese Methode weicht von der vorhergehenden nur insofern ab, dass sie zwei geocentrische Oerter genau und darauf die übrigen so genau wie möglich darstellt, während nach der anderen Methode keine Beobachtung den übrigen vorgezogen wird, sondern die Fehler möglichst auf alle vertheilt werden. Die Methode des vorhergehenden Art. ist daher der früheren nur insofern hintanzusetzen, als man dadurch, dass die Oerter L, L′ einen gewissen Theil der Fehler aufnehmen, die Fehler in den übrigen Oertern erheblich verkleinern kann; meistens kann man sich jedoch durch eine zweckmässige Auswahl der Beobachtungen L, L′ leicht davor bewahren, dass dieser Unterschied von grosser Bedeutung werden kann. Man muss sich nämlich Mühe geben, für L, L′ solche Beobachtungen auszuwählen, welche sich nicht nur einer ausgesuchten Genauigkeit erfreuen, sondern auch so beschaffen sind, dass die aus ihnen und den Entfernungen abgeleiteten Elemente durch kleine Variationen der geocentrischen Positionen selbst nicht allzusehr beeinflusst werden. Man würde daher thöricht handeln, wenn man um einen geringen Zeitraum von einander entfernte Beobachtungen auswählen würde, oder solche, denen sehr nahe gegenüberliegende oder zusammenfallende heliocentrische Oerter entsprechen.