XXXII.

Gauß in seiner Vorlesung über die Methode der kleinsten Quadrate.

[Festschrift zur Feier des hundertfünfzigjährigen Bestehens der Königlichen Gesellschaft der Wissenschaften zu Göttingen. Beiträge zur Gelehrtengeschichte Göttingens. S. 45—59 (1901).]

Als geborener Braunschweiger habe ich schon früh von Gauß sprechen hören, und ich glaubte gern an seine Größe, ohne zu wissen, worin sie bestand. Um so tieferen Eindruck machte es auf mich, als ich zuerst von seiner geometrischen Darstellung der imaginären oder, wie man zu jener Zeit wohl noch sagte, der unmöglichen Größen hörte. Ich war damals als Student auf dem Collegium Carolinum (der heutigen Technischen Hochschule) ein wenig in die höhere Mathematik eingedrungen, und bald darauf, als Gauß im Juli 1849 sein 50jähriges Doktor-Jubiläum feierte, sandte unser Lehrkörper einen von dem geistreichen Philologen Petri verfaßten Glückwunsch an ihn, worin mir der Passus, er habe das Unmögliche möglich gemacht, ganz besonders gefiel. Zu Ostern 1850 kam ich nach Göttingen, und hier wuchs mein Verständnis schon etwas mehr, als ich im Seminar durch eine kurze, aber sehr interessante Vorlesung von Stern in die Elemente der Zahlentheorie eingeführt wurde und den Reziprozitätssatz kennen lernte. Auf meinen Wegen nach oder von der Sternwarte, wo ich eine Vorlesung des trefflichen Professors Goldschmidt über populäre Astronomie hörte, begegnete ich zuweilen Gauß und erfreute mich des Anblicks seiner stattlichen, Ehrfurcht gebietenden Erscheinung, und sehr oft sah ich ihn in größter Nähe auf seinem festen Platze im Literarischen Museum, das er regelmäßig besuchte, um Zeitungen zu lesen.

Zu Anfang des folgenden Wintersemesters hielt ich mich für reif, seine Vorlesung über die Methode der kleinsten Quadrate zu hören, und so betrat ich, mit dem Testierbuch ausgerüstet und nicht ohne

Herzklopfen, zum ersten Male sein Wohnzimmer, wo ich ihn an seinem Schreibtisch sitzend fand. Meine Meldung schien ihn wenig zu erfreuen, ich hatte auch wohl gehört, daß er sich ungern entschloß, Vorlesungen zu halten; nachdem er seinen Namen in das Buch eingetragen hatte, sagte er nach kurzem Schweigen: „Sie wissen vielleicht, daß es immer sehr zweifelhaft ist, ob meine Vorlesungen zustande kommen; wo wohnen Sie? bei dem Barbier Vogel? Nun, das trifft sich ja glücklich, denn der ist auch mein Barbier, durch ihn werde ich Sie benachrichtigen."

Einige Tage darauf trat dann Vogel, eine stadtbekannte Persönlichkeit, ganz erfüllt von der Wichtigkeit seiner Mission, bei mir ein, um zu bestellen, daß sich noch mehrere Zuhörer gemeldet hätten, und daß Herr Geh. Hofrat Gauß die Vorlesung halten würde.

Wir waren neun Studenten, von denen ich A. Ritter (später Professor der Mechanik in Hannover und Aachen) und Moritz Cantor (später Professor in Heidelberg) nach und nach näher kennen lernte; wir alle kamen sehr regelmäßig, es hat wohl selten einer von uns gefehlt, obgleich der Weg nach der Sternwarte im Winter bisweilen nicht angenehm war. Das Auditorium war durch ein Vorzimmer von Gauß' Arbeitszimmer getrennt und ziemlich klein. Wir saßen an einem Tisch, dessen Längsseiten für je drei, aber nicht für vier Personen bequemen Platz darboten. Der Tür gegenüber am oberen Ende saß Gauß in mäßiger Entfernung vom Tische, und wenn wir vollzählig waren, so mußten zwei von uns, die zuletzt kamen, ganz in seine Nähe rücken und ihr Heft auf den Schoß nehmen. Gauß trug ein leichtes schwarzes Käppchen, einen ziemlich langen braunen Gehrock, graue Beinkleider; er saß meist in bequemer Haltung, etwas gebeugt vor sich niedersehend, mit über dem Leib gefalteten Händen. Er sprach ganz frei, sehr deutlich, einfach und schlicht; wenn er aber einen neuen Gesichtspunkt hervorheben wollte, wobei er ein besonders charakteristisches Wort gebrauchte, so erhob er wohl plötzlich den Kopf, wandte sich zu einem seiner Nachbarn und blickte ihn während der nachdrücklichen Rede ernst mit seinen schönen, durchdringenden blauen Augen an. Das war unvergeßlich. Seine Sprache war fast ganz dialektfrei, nur bisweilen kamen Anklänge an unsere stadt-braunschweigische Mundart; beim Zählen z. B., wobei er auch den Gebrauch der Finger nicht verschmähte, sagte er nicht eins, zwei, drei, sondern eine, zweie, dreie usf., wie man es noch jetzt bei uns auf dem Markte hören kann. Ging er von einer prinzipiellen Erörterung zur Entwicklung mathematischer Formeln über, so erhob er sich, und in stattlicher, ganz aufrechter Haltung schrieb er an einer neben ihm stehenden Tafel mit der ihm eigenen schönen Handschrift, wobei es ihm immer durch Sparsamkeit und zweckmäßige Anordnung gelang, mit dem ziemlich kleinen Raume auszukommen. Für die Zahlenbeispiele, auf deren sorgfältige Durchführung er besonderen Wert legte, brachte er die erforderlichen Data auf kleinen Zetteln mit.

Indem ich nun zu dem Inhalt der (wöchentlich dreistündigen) Vorlesung übergehe, beziehe ich mich auf einen Brief von Gauß an Schumacher aus dem Jahre 1844, welcher im Band VIII von Gauß' Werken (S. 147—148) abgedruckt ist; er berichtet dort, daß seine Vorlesung aus drei Teilen besteht, von denen der erste eine nur auf Prinzipien der Zweckmäßigkeit basierte Begründungsart und die eigentliche praktische Anwendung der Methode der kleinsten Quadrate gibt, während der zweite und dritte Teil die beiden wesentlich verschiedenen Begründungsarten der Methode durch die Wahrscheinlichkeitsrechnung behandelt, wie sie in der Theoria motus corporum coelestium und in der Theoria combinationis observationum dargestellt sind. Denselben Weg schlug Gauß auch zu meiner Zeit ein, und ich möchte hier einiges aus dem ersten Teile mitteilen, was, wie ich glaube, weniger bekannt ist als der Inhalt der anderen Teile; freilich kann ich es auch hierbei nicht vermeiden, sehr bekannte Dinge mit einzuflechten. Den Zweck der Methode der kleinsten Quadrate und ihre elementare Begründung stellte Gauß ungefähr so dar:

Es liegt eine Reihe von Beobachtungen (Messungen) vor, die dazu dienen sollen, gewisse unbekannte Größen $x, y, z \cdots$ zahlenmäßig zu bestimmen; die unmittelbaren Gegenstände dieser Beobachtungen können diese Unbekannten selbst, allgemeiner aber gewisse Funktionen von ihnen sein, d. h. Größen $V, V', V'' \cdots$, welche sich streng berechnen lassen würden, wenn die Werte von $x, y, z \cdots$ schon bekannt wären. Werden nun für diese Funktionen durch unmittelbare Beobachtungen die Werte $M, M', M'' \cdots$ gefunden, so erhält man ein entsprechendes System von sogenannten Beobachtungsgleichungen $V = M$, $V' = M'$, $V'' = M'' \cdots$, aus denen die unbekannten Elemente $x, y, z \cdots$ ermittelt werden sollen; dies ist natürlich nur dann möglich, wenn die Anzahl der Beobachtungen mindestens ebenso groß ist, wie die der unbekannten Elemente. Da aber alle unsere Messungen nur einen begrenzten Genauigkeitsgrad besitzen, also Fehlern unterworfen sind, so sucht man deren schädlichen Einfluß durch Vermehrung der Anzahl der Beobachtungen zu bekämpfen, und dann fragt sich, wie soll man ein solches, mit unbekannten Fehlern $M - V, M' - V', M'' - V'' \cdots$ behaftetes, überzähliges System von Beobachtungsgleichungen $V = M, V' = M', V'' = M'' \cdots$ behandeln, um die unvermeidlichen Widersprüche zwischen ihnen auszugleichen und so die plausibelsten Werte der unbekannten Elemente $x, y, z \cdots$ zu finden? Hierbei wird vorausgesetzt, daß diese Beobachtungen gleiche Zuverlässigkeit besitzen, d. h., daß wir keinen Grund haben, einer von ihnen größeres Zutrauen zu schenken als den übrigen. Macht man eine bestimmte Hypothese über die Werte $x, y, z \cdots$ und berechnet daraus die entsprechenden Werte der Funktionen $V, V', V'' \cdots$, so erhält man ein entsprechendes System von Fehlern $M - V, M' - V', M'' - V'' \cdots$, aber es fragt sich, welchen Maßstab soll man zugrunde legen, um nach Beschaffenheit dieser Fehler einer solchen Hypothese den Vorzug vor einer anderen zuzuerkennen?

Hier könnte man, da die Beobachtungen wegen ihrer gleichen Zuverlässigkeit eine gleichmäßige Behandlung verdienen, zunächst an die algebraische Summe der Fehler denken, um nach ihrer Kleinheit die Brauchbarkeit einer Hypothese zu beurteilen, und dies würde geradezu dahin führen, die Hypothese für die beste zu erklären, für welche diese Summe gleich Null wird. Allein man sieht sofort, daß dies nicht als allgemeines Prinzip gelten kann, weil, sobald die Anzahl der unbekannten Elemente mindestens gleich zwei ist, unendlich viele verschiedene Hypothesen dieser Forderung genügen würden, und außerdem würde eine Hypothese, bei welcher große Fehler durch die entgegengesetzten Vorzeichen sich in der Summe aufheben, für ebenso gut gelten, als eine andere, in welcher die einzelnen Fehler absolut genommen kleiner sind.

Dies kann uns veranlassen, nur die absoluten Werte der Fehler zu betrachten und die Hypothese für die beste zu erklären, für welche deren Summe so klein wie möglich ausfällt. Allein auch gegen dieses Prinzip lassen sich mehrere triftige Einwände erheben, nämlich:

a) Es verstößt gegen den mathematischen Sinn, daß hierbei die negativen Fehler auf andere Weise in die Rechnung eintreten sollen als die positiven.

b) Die Behandlung wird bei einer größeren Anzahl unbekannter Elemente $x, y, z \cdots$ bald sehr verwickelt.

c) Selbst im einfachsten Falle, wo nur eine einzige unbekannte Größe auftritt, und diese zugleich der unmittelbare Gegenstand der sämtlichen Beobachtungen ist, führt dieses Prinzip zu Resultaten, denen wir unseren Beifall nicht schenken können. Hätte man z. B. vier Beobachtungen $x = 900$, $x = 903$, $x = 917$, $x = 921$ gemacht, so würde jeder zwischen 903 und 917 liegende Wert x für gleich brauchbar gelten müssen, weil für alle diese Werte x die Summe der absoluten Fehler denselben kleinsten Wert $(x - 900) + (x - 903) + (917 - x) + (921 - x) = 35$ erhalten würde; dieselbe Erscheinung tritt immer auf, wenn die Anzahl der Beobachtungen gerade ist, während bei einer ungeraden Anzahl immer der in der Mitte liegende Wert x für den besten gelten müßte, so daß die übrigen Beobachtungen auch hier gar keinen Einfluß auf die Bestimmung von x ausüben würden.

d) Von besonderem Gewicht ist endlich der Einwand, daß nach diesem Prinzip bei einer größeren Anzahl von Beobachtungen ein großer Fehler keinen stärkeren Einfluß auf das Resultat ausüben würde, als eine Reihe kleiner Fehler, deren absolute Werte dieselbe Summe besitzen, während doch die Hypothese, welcher die letztere Erscheinung entspricht, nach unserem Gefühl gewiß den Vorzug vor der ersteren verdient.

Aus allen diesen Gründen ist dieses Prinzip zu verwerfen, und wir müssen einen anderen Maßstab suchen, durch welchen diese Mängel beseitigt werden, zumal die Wahl dieses Maßstabes ganz unserer Willkür überlassen ist. Hierzu führt uns von selbst der letztgenannte Einwand; die Rücksicht darauf, daß ein Fehler, welcher a-mal so groß ist als ein Fehler, der a-mal vorkommt, stärker ins Gewicht fallen muß, als diese a einzelnen Fehler, veranlaßt uns, statt der Fehler selbst ihre Quadrate zu nehmen und die Brauchbarkeit einer Hypothese nach der Kleinheit der ihr entsprechenden Summe der Fehlerquadrate zu schätzen. So gewinnen wir den Grundsatz der sogenannten Methode der kleinsten Quadrate, nach welchem diejenige Hypothese über die Werte der unbekannten Größen $x, y, z \cdots$ als die beste gelten soll, für welche die Summe der Fehlerquadrate so klein wie möglich wird. Durch die Wahl dieses Maßstabes weichen wir den obigen Einwänden a) und d) aus, ebenso gestaltet sich der unter b) erwähnte Übelstand weit besser, und die unter c) bemerkte Erscheinung wird ganz unmöglich. Wollte man, um den Einwand d) zu beseitigen, eine noch höhere Potenz der Fehler einführen, so müßte dieselbe jedenfalls eine paare sein, um dem ersten Vorwurf zu begegnen; aber dann werden die Rechnungen so außerordentlich verwickelt, daß diese Behandlung die Mühe nicht lohnen würde.

Nach dieser sehr einleuchtenden, nur auf Prinzipien der Zweckmäßigkeit beruhenden Begründung der Methode ging Gauß sofort zur Bildung der sogenannten Normalgleichungen über, die durch partielle Differentiation der Summe

$$\Omega = (V - M)^2 + (V' - M')^2 + (V'' - M'')^2 + \cdots$$

der Fehlerquadrate in bezug auf jede der unbekannten Größen x, $y, z \cdots$ gewonnen werden. Der Kürze wegen will ich hier

$$d\Omega = 2(X dx + Y dy + Z dz + \cdots)$$

setzen, dann gibt die Forderung des Minimums von Ω die auf x, $y, z \cdots$ bezüglichen Normalgleichungen

$$X = 0, \quad Y = 0, \quad Z = 0 \cdots,$$

die von Gauß vollständig entwickelt dargestellt wurden.

Der zunächst behandelte und wichtigste Hauptfall ist der, wo die Größen $V, V', V'' \cdots$, also auch $X, Y, Z \cdots$ lineare Funktionen von $x, y, z \cdots$ sind. Zuerst wurde natürlich der spezielle Fall vorgeführt, wo eine einzige unbekannte Größe x durch wiederholte unmittelbare Messungen bestimmt werden soll, und wo die entsprechende Normalgleichung zu der altbekannten Regel des arithmetischen Mittels führt. Ein dazugehöriges Zahlenbeispiel — Bestimmung der Polhöhe von Lauenburg aus 11 Beobachtungen — benutzte Gauß, um uns auf gewisse Rechnungsvorteile aufmerksam zu machen. Da alle beobachteten Werte natürlich dieselbe Anzahl 53 der Grade aufwiesen und erst in den Minuten zwischen 21 und 22 schwankten, so wäre es ja töricht, bei der Bildung des arithmetischen Mittels diese so weit miteinander übereinstimmenden Werte wirklich zu addieren und ihre Summe nachher durch ihre Anzahl zu dividieren; statt dessen ist es offenbar vorteilhafter, etwa $53^0\,21'$ oder $53^0\,22'$ als genäherten Wert von x anzusehen, also $x = 53^0\,21' + t''$ oder $x = 53^0\,22' + u''$ zu setzen, und nur das arithmetische Mittel dieser Korrektionen t oder u

in Sekunden zu berechnen. Das war freilich sehr einleuchtend, aber Gauß verschmähte es nicht, eine solche scheinbare Kleinigkeit hervorzuheben, weil in ihr der Keim eines allgemeinen Prinzips enthalten war, und ebenso verfuhr er in den folgenden Aufgaben mit einer oder mehreren Unbekannten; immer wurde die allgemeine Regel an bestimmten Zahlenbeispielen durchgeführt, die zu ähnlichen Bemerkungen Veranlassung gaben. Von allen diesen Aufgaben glaube ich hier die letzte (fünfte) mitteilen zu dürfen, weil sie wohl auch heute noch einiges Interesse darbietet.

Es handelt sich um die als nahezu gleich zuverlässig anzusehenden Messungen der Höhendifferenzen (in Metern) von den folgenden fünf Punkten: P (Boden der Göttinger Sternwarte), Q (Meridianzeichen der Wehnder Papiermühle), R (Fläche des Postamentes auf dem Hohenhagen), S (östlicher Abhang des Hils, eine Viertelstunde von Ammensen), T (Brocken, Marmorplatte des vormaligen, im Hause gelegenen Turms). Bedeuten diese Zeichen zugleich die Höhen der entsprechenden Punkte, so liegen folgende sieben Beobachtungen vor:

$$\begin{aligned} Q &= P + 64{,}334 \\ R &= P + 349{,}366 \\ R &= Q + 283{,}596 \\ S &= Q + 206{,}580 \\ S &= R - 76{,}108 \\ T &= R + 648{,}427 \\ T &= S + 719{,}612 \end{aligned}$$

In diesen Gleichungen treten tatsächlich nur vier Unbekannte auf, nämlich die relativen Höhen $Q-P$, $R-P$, $S-P$, $T-P$ über Göttingen; denn absolute Höhen können natürlich hieraus nicht gefunden werden. Nun wird man sich zunächst durch geschickte Kombinationen genäherte Werte für diese Unbekannten verschaffen und hierauf

$$\begin{aligned} Q &= P + 64{,}334 + q \\ R &= P + 348{,}648 + r \\ S &= P + 271{,}727 + s \\ T &= P + 994{,}207 + t \end{aligned}$$

setzen, wo q, r, s, t die Korrektionen dieser Näherungswerte bedeuten. Führt man sie als neue Unbekannte ein und drückt sie in Millimetern aus, so nehmen die obigen sieben Beobachtungsgleichungen folgende Form an:

$$\begin{aligned}
0 + q \quad &= 0\\
-\ 718 + r \quad &= 0\\
+\ 718 - q + r &= 0\\
+\ 813 - q + s &= 0\\
-\ 813 - r + s &= 0\\
-\ 2868 - r + t &= 0\\
+\ 2868 - s + t &= 0
\end{aligned}$$

und die vier Normalgleichungen lauten:

$$\begin{aligned}
0 &= -1531 + 3q - r - s\\
0 &= +3681 - q + 4r - s - t\\
0 &= -2868 - q - r + 3s - t\\
0 &= \qquad\qquad\quad - r - s + 2t
\end{aligned}$$

Nach einigen Bemerkungen über die direkte Auflösung dieser Gleichungen durch zweckmäßige Anordnung der sukzessiven Eliminationen teilte uns Gauß eine indirekte Auflösungsmethode mit, einen Kunstgriff, durch den man sich die beschwerliche Eliminationsarbeit, wie er sagte, oft erleichtern könne. Derselbe besteht wesentlich darin, die konstanten Glieder durch fortgesetzte Substitutionen auf immer kleinere absolute Werte herabzudrücken, und dieser Prozeß beginnt in unserem Beispiel auf folgende Weise. Ignoriert man in der zweiten Gleichung, welche das größte konstante Glied (3681) enthält, die Unbekannten q, s, t neben dem Gliede $4r$, welches den größten Koeffizienten (4) hat, und vernachlässigt Bruchteile, so erhält man für r den Wert -920; man betrachtet ihn als eine Annäherung und führt eine Korrektion r' als neue Unbekannte ein, indem man $r = -920 + r'$ setzt; die unbekannten Glieder werden hierdurch nur insofern berührt, daß r' an Stelle von r tritt, während die konstanten Glieder in -611, $+1$, -1948, $+920$ übergehen. Indem man nach derselben Regel fortfährt, wird man in der dritten Gleichung die Unbekannten q, r', t ignorieren und $s = +649 + s'$ setzen, wodurch die konstanten Glieder in -1260, -648, -1, $+271$ übergehen. Offenbar kommt es immer nur darauf an, die neue Substitution zu notieren, wobei man die Akzente der neuen Unbekannten füglich unterdrücken darf, und die neuen Werte der konstanten Glieder zu berechnen; den ganzen

Mechanismus kann man in leicht verständlicher Weise durch eine Tabelle darstellen, deren Anfang hier folgen mag:

	r	s	q	r	s
	-920	$+649$	$+420$	$+267$	$+229$
-1531	-611	-1260	0	-267	-496
$+3681$	$+1$	-648	-1068	0	-229
-2868	-1948	-1	-421	-688	-1
0	$+920$	$+271$	$+221$	$+4$	-225

Hat man 73 solche Operationen gemacht, so sind die konstanten Teile so klein geworden, daß sie durch eine neue Substitution nicht mehr vermindert werden können, und es genügt dann, jede der ursprünglichen Unbekannten durch Addition ihrer sukzessiven Näherungswerte zu berechnen:

$$\begin{aligned} q &= +420 \cdots \\ r &= -920 + 267 \cdots \\ s &= +649 + 229 \cdots \\ t &= \cdots \end{aligned}$$

Noch viel kürzer wird die Arbeit, wenn man mit diesem Kunstgriff einen zweiten verbindet, welcher im folgenden besteht. Man führt noch eine neue Unbekannte p ein, indem man den Anfangspunkt verlegt und q, r, s, t durch $q-p$, $r-p$, $s-p$, $t-p$ ersetzt; behandelt man die hierdurch umgeformten Beobachtungsgleichungen wieder nach der Methode der kleinsten Quadrate, so erhält man die folgenden fünf Normalgleichungen:

$$\begin{aligned} 0 &= +718 + 2p - q - r \\ 0 &= -1531 - p + 3q - r - s \\ 0 &= +3681 - p - q + 4r - s - t \\ 0 &= -2868 - q - r + 3s - t \\ 0 &= - r - s + 2t, \end{aligned}$$

von denen die vier letzten in die früheren übergehen, wenn $p = 0$ gesetzt wird. Sie haben zwei merkwürdige Eigenschaften, deren Grund man leicht erkennt: erstens ist die Summe der konstanten Glieder, und ebenso die Summe der Koeffizienten jeder einzelnen Unbekannten gleich Null, und zweitens ist auch die Summe der Koeffizienten aller Unbekannten in jeder einzelnen Gleichung gleich Null. Zufolge der ersten Eigenschaft ist jede Gleichung eine identische Folge der vier anderen, und es ist daher unmöglich, bestimmte Werte der fünf Unbekannten aus ihnen abzuleiten. Wendet man aber die oben beschriebene indirekte Auflösungsmethode an, so wird man finden, daß dieselbe jetzt viel schneller zum Abschluß kommt, und außerdem ergibt sich aus dem Umstande, daß die Summe der konstanten Glieder stets gleich Null bleiben muß, eine überaus angenehme Kontrolle der fortlaufenden Rechnung, deren Anfang wieder durch die folgende Tabelle dargestellt werden mag:

	r — 920	s + 649	p — 819	q + 147
+ 718	+ 1638	+ 1638	0	— 147
— 1531	— 611	— 1260	— 441	0
+ 3681	+ 1	— 648	+ 171	+ 24
— 2868	— 1948	— 1	— 1	— 148
0	+ 920	+ 271	+ 271	+ 271

Nach etwa 20 Operationen schließt diese Rechnung ab und liefert bestimmte Werte der fünf Unbekannten, aus denen sich schließlich die eigentlichen Unbekannten $q-p$, $r-p$, $s-p$, $t-p$ ergeben. In der Vorlesung wurden natürlich nur die ersten Schritte dieses wie des früheren Prozesses wirklich ausgeführt. Über die Vorzüge und Nachteile dieser indirekten Auflösungsmethode gegenüber der gewöhnlichen durch sukzessive Elimination der Unbekannten muß ich mich jeder Bemerkung enthalten; es genügt mir, durch die Mitteilung dieses Beispiels wieder darauf hinzuweisen, wie unablässig Gauß bemüht war, auch bei dem praktischen Rechnen sinnreiche Kunstgriffe zu erfinden.

Hierauf folgte die Behandlung des Falles, wo die Beobachtungsgleichungen $V = M \cdots$ die Unbekannten nicht mehr, wie bisher, in linearer Form enthalten. Die Einführung genäherter Werte, welche früher nur als ein die Zahlenrechnungen vereinfachender Kunstgriff auftrat, wird hier zu dem Prinzip ausgebildet, durch welches dieser Fall auf den früheren der linearen Gleichungen zurückzuführen ist, indem man die kleinen Korrektionen als neue Unbekannte behandelt und deren Produkte bei der Entwicklung nach dem Satze von Taylor vernachlässigt. Als Beispiel diente die Pothenot'sche Aufgabe in der praktischen Geometrie, speziell die Ortsbestimmung von Rosdorf bei Göttingen aus sechs Einschnitten.

Nun ging Gauß zu einer ebenfalls elementar gehaltenen Entwicklung des Begriffs der Präzision einer Beobachtungsmethode über. Wenn die vorliegenden Beobachtungen $V = M$, $V' = M' \cdots$ nicht mehr, wie bisher vorausgesetzt wurde, gleiche Zuverlässigkeit besitzen, so muß man sie als auf verschiedene Maßstäbe bezogen ansehen und jede Gleichung mit einem entsprechenden, die gleiche Zuverlässigkeit wiederherstellenden Koeffizienten k multiplizieren. Die hieraus nach der Methode der kleinsten Quadrate abgeleiteten Normalgleichungen enthalten diese (relativen) Präzisionen k nur in ihren Quadraten, und diese heißen die entsprechenden Gewichte p der Beobachtungen.

Diese Betrachtung gibt zugleich ein Mittel an die Hand, die Zuverlässigkeit der durch die Methode der kleinsten Quadrate gewonnenen Resultate im Vergleich mit der Zuverlässigkeit der gegebenen Beobachtungen zu bestimmen. Das Prinzip, auf welches sich diese Ableitung gründet, ist das der Konsequenz. Man denkt sich zu der ursprünglich vorhandenen Gruppe von Beobachtungen, die wieder als gleich zuverlässig vorausgesetzt werden, und deren Präzision $= 1$ angenommen wird, eine beliebige Anzahl anderer Beobachtungen von derselben Präzision hinzu, wodurch die zuerst gefundenen plausibelsten Werte der Unbekannten $x, y, z \cdots$ in andere Werte übergehen werden. Um diese zu finden, kann man nun zwei verschiedene Wege einschlagen, welche aber notwendig zu denselben Resultaten führen müssen. Der erste Weg besteht darin, daß man nach der Methode der kleinsten Quadrate sämtliche Beobachtungsgleichungen beider Gruppen gleichzeitig behandelt, der zweite darin, daß man die allein aus der ersten Gruppe (der wirklich gegebenen Beobachtungen) abgeleiteten Resultate mit gewissen, noch unbestimmt gelassenen Präzisions-Koeffizienten k multipliziert und hierauf mit der zweiten, hinzugedachten Gruppe von Beobachtungen kombiniert. Durch die Forderung, daß die auf diesen beiden verschiedenen Wegen erhaltenen Schlußresultate miteinander übereinstimmen müssen, ergeben sich dann die Werte der Präzisionen k und ihrer Quadrate, der Gewichte p.

Diese allgemeine Anleitung zur Gewichtsbestimmung der durch die Methode der kleinsten Quadrate gewonnenen Resultate wurde zunächst an den einfachsten Fällen durchgeführt, wo die (immer als linear vorausgesetzten) Beobachtungsgleichungen nur eine Unbekannte enthalten. Um aber für eine beliebige Anzahl n von Unbekannten $x, y \cdots$, deren letzte z sein möge, die Regel allgemein auszudrücken (wobei die uns damals noch wenig geläufige Sprache der Determinanten-Theorie gänzlich vermieden wurde), beschrieb Gauß zunächst ein Verfahren zur Auflösung der auf $x, y \cdots z$ bezüglichen Normalgleichungen $X = 0$, $Y = 0 \cdots Z = 0$, welches er mit dem Namen der rechten Elimination belegte (vgl. Disquisitiones circa elementa elliptica Palladis). Man eliminiere die erste Unbekannte x aus allen n Gleichungen, indem man nur die erste, auf x bezügliche Normalgleichung $X = 0$ mit geeigneten Koeffizienten multipliziert und von den folgenden, unveränderten Normalgleichungen abzieht, welche dadurch in $Y' = 0 \cdots Z' = 0$ übergehen und frei von x sind; nun eliminiere man ebenso die zweite Unbekannte y aus allen diesen $(n-1)$ Gleichungen, indem man nur die erste Gleichung $Y' = 0$ mit geeigneten Koeffizienten multipliziert und von allen folgenden abzieht; fährt man so fort, so gelangt man schließlich zu einer Gleichung von der Form $H(z - C) = 0$, in welcher nur noch die letzte Unbekannte z auftritt, und wo H, C bekannte Werte bedeuten, die sich aus dem Verlauf dieser rechten Elimination mit Bestimmtheit ergeben. Die Auflösung der Normalgleichungen liefert also für z den plausibelsten Wert C, und die Regel von Gauß besteht darin, daß der Koeffizient H zugleich das Gewicht dieses Resultats $z = C$ darstellt, d. h. also: In der Methode der kleinsten Quadrate wird das Gewicht jedes einzelnen Resultats für eine Unbekannte durch den Koeffizienten dargestellt, welchen diese Unbekannte bei rechter Elimination in der letzten Gleichung erhält.

Bei dem Beweise dieses Satzes, den ich hier des Raumes wegen mit einer kleinen Änderung der Bezeichnung wiedergebe, beschränkte sich Gauß auf den Fall von drei Unbekannten x, y, z. Durch die rechte Elimination von x, y werden offenbar zwei Koeffizienten α, β gewonnen, welche bewirken, daß identisch

$$\alpha X + \beta Y + Z = H\ (z - C)$$

wird. Denkt man sich nun zu den wirklich vorhandenen m Beobachtungsgleichungen, denen die Normalgleichungen $X = 0$, $Y = 0$, $Z = 0$ mit dem Resultat $z = C$ entsprechen, noch eine Beobachtung $z = D$ von derselben Präzision hinzu und schlägt den oben beschriebenen ersten Weg ein, bei welchem alle $(m + 1)$ Beobachtungen gleichzeitig behandelt werden, so bleiben offenbar die auf x, y bezüglichen Normalgleichungen $X = 0$, $Y = 0$ ungeändert bestehen, während die dritte $Z = 0$ in $Z + (z - D) = 0$ übergeht; zufolge der obigen Identität ergibt sich daher z jetzt aus der Gleichung

$$H(z - C) + (z - D) = 0.$$

Schlägt man aber den zweiten Weg ein, indem man das aus den m wirklich vorhandenen Beobachtungen durch die Methode der kleinsten Quadrate gewonnene Resultat $z = C$ mit einer noch unbestimmten Präzision k multipliziert und nun nach derselben Methode mit der hinzugedachten Beobachtung $z = D$ kombiniert, so erhält man, wenn das unbekannte Gewicht $kk = p$ gesetzt wird, die einzige Normalgleichung

$$p(z - C) + (z - D) = 0,$$

und durch den Vergleich mit dem Resultate des ersten Weges folgt $p = H$, w. z. b. w.

In nahem Zusammenhang mit der eben beschriebenen rechten Elimination steht die sukzessive identische Umformung der Summe Ω der Fehlerquadrate in eine Reihe von Quadraten linearer Funktionen $A, B, C \cdots$, die so gewählt werden, daß x nur in A, y nur in A und B, z nur in A, B, C usf. auftritt. Der zuletzt verbleibende konstante Bestandteil stellt dann das Minimum von Ω dar, und die plausibelsten Werte der Unbekannten $x, y, z \cdots$ ergeben sich aus den Gleichungen $A = 0$, $B = 0$, $C = 0 \cdots$ in umgekehrter Folge.

Mit dieser Darstellung schloß Gauß am 24. Januar 1851 den ersten Teil seiner Vorlesung, durch den er uns mit dem Wesen der Methode der kleinsten Quadrate vollkommen vertraut gemacht hatte. Es folgte nun eine überaus klare und durch originelle Beispiele erläuterte Entwicklung der Grundbegriffe und der Hauptsätze der Wahrscheinlichkeitsrechnung, die als Einleitung zu der zweiten und dritten Begründungsart der Methode diente, worauf ich hier nicht mehr eingehen darf. Ich kann nur sagen, daß wir diesem ausgezeichneten Vortrage, in welchem auch einige Beispiele aus der Theorie der bestimmten Integrale behandelt wurden, mit immer steigendem Interesse gefolgt sind. Aber es schien uns auch, als ob Gauß selbst, der vorher wenig Neigung gezeigt hatte, die Vorlesung zu halten, im Laufe derselben doch einige Freude an seiner Lehrtätigkeit empfand. So kam es am 13. März zum Schluß, Gauß erhob sich, wir alle mit ihm, und er entließ uns mit den freundlichen Abschiedsworten: „Es bleibt mir nur noch übrig, Ihnen zu danken für die große Regelmäßigkeit und Aufmerksamkeit, mit der Sie meinem, doch wohl recht trocken zu nennenden Vortrage gefolgt sind.“ Seitdem ist nun ein halbes Jahrhundert verflossen, aber dieser angeblich trockene Vortrag steht mir unvergeßlich in der Erinnerung als einer der schönsten, die ich je gehört habe.